Some topics in functional analysis

Stephen Semmes
Rice University

Preface

These informal notes are intended to be at least somewhat introductory, particularly near the beginning. We shall also be concerned with connections with other aspects of analysis, or other areas of mathematics, even if we do not get into it too much here. Of course, there are many texts with more information, some of which may be found in the bibliography.

The main prerequisites, at least near the beginning, are familiarity with basic analysis and linear algebra, including abstract metric spaces and vector spaces. In some places, it may be helpful for the reader to be familiar with some general topology, measure theory and integration, or complex analysis. Some familiarity with abstract algebra, including groups and rings, would be helpful in some places as well. We shall sometimes consider simpler versions of some basic notions or results, which could be treated more fully with more familiarity or review of related facts.

Of course, linear algebra deals with vector spaces and linear mappings between them. This is already very interesting and important in the finite-dimensional case, and one may consider infinite-dimensional vector spaces as well. One may consider vector spaces over arbitrary fields too. Here we deal with vector spaces over the real and complex numbers, and norms and inner products on these vector spaces. The corresponding metrics determine topologies on these vector spaces, and we may also consider other types of topologies on them.

Beginning in Part II, we shall consider algebras over the real and complex numbers. These are basically vector spaces on which a bilinear operation of multiplication is defined. Some basic examples include algebras of functions on some nonempty set, and algebras of linear mappings from a vector space into itself. Complex analysis can play a large role in this, and we shall often try to include real versions where possible.

The definition of a norm uses an absolute value function on the corresponding field of scalars. Here we normally use the standard absolute value functions on the real and complex numbers, but one can consider absolute value functions on arbitrary fields, as in [44, 80]. This will be discussed in Appendix A, for which the reader should have some familiarity with some basic notions related to fields. Some topics related to algebras over fields with non-archimedean absolute value functions may be found in [60, 61]. Some related references include [11, 41, 53, 95, 136, 145, 147, 149, 193].

Contents

Part I

Norms and bounded linear mappings

Chapter 1

Norms and Banach Spaces

1.1 Norms on vector spaces

Let V be a vector space over the real or complex numbers. A nonnegative real-valued function N on V is said to be a *norm* on V if it satisfies the following three conditions. First, for each $v \in V$,

$$N(v) = 0 \text{ if and only if } v = 0. \tag{1.1.1}$$

Second,

$$N(t\,v) = |t|\,N(v) \tag{1.1.2}$$

for every $v \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. Here $|t|$ denotes the usual absolute value of the real or complex number t. Third,

$$N(v + w) \le N(v) + N(w) \tag{1.1.3}$$

for every $v, w \in V$. This is known as the *triangle inequality* for N.

Of course, the real line $\mathbf{R}$ and complex plane $\mathbf{C}$ may be considered as one-dimensional vector spaces over themselves. The usual absolute value functions on $\mathbf{R}$ and $\mathbf{C}$ may be considered as norms on these spaces. More precisely, the only norms on $\mathbf{R}$ or $\mathbf{C}$, as vector spaces over themselves, are positive constant multiples of the usual absolute value functions.

If V is a vector space over the complex numbers, then V may be considered as a vector space over the real numbers, by simply restricting scalar multiplication on V to real numbers. We may use

$$V_{\mathbf{R}} \tag{1.1.4}$$

to denote V considered as a vector space over $\mathbf{R}$ in this way. If N is a norm on V as a vector space over $\mathbf{C}$, then N may also be considered as a norm on $V_{\mathbf{R}}$. However, a norm on $V_{\mathbf{R}}$ may not be a norm on V, as a vector space over the complex numbers.

Let X be a nonempty set, and note that the spaces of all real or complex-valued functions on X are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions on X. Also let

$$\ell^\infty(X, \mathbf{R}),\ \ell^\infty(X, \mathbf{C}) \tag{1.1.5}$$

be the spaces of bounded real and complex-valued functions on X, respectively. It is well known and not difficult to check that these are linear subspaces of the spaces of all real and complex-valued functions on X, respectively. In particular, $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$ are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions on X.

If f is a bounded real or complex-valued function on X, then put

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}. \tag{1.1.6}$$

One can verify that this defines a norm on each of $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$, as vector spaces over the real and complex numbers, respectively. This is known as the *supremum norm*, and we may sometimes use the notation

$$\|f\|_{sup} \tag{1.1.7}$$

instead. We may also use the notation

$$\|f\|_{\ell^\infty(X,\mathbf{R})},\ \|f\|_{\ell^\infty(X,\mathbf{C})} \tag{1.1.8}$$

as well, to be more precise.

Let V be a vector space over the real or complex numbers again, and let N be a norm on V. It is easy to see that

$$d_N(v, w) = N(v - w) \tag{1.1.9}$$

defines a metric on V. The metric on $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$ associated to the supremum norm in this way is the usual *supremum metric*, for instance.

In particular, all of the usual notions and results about metric spaces can be used in this situation. Of course, there are a lot of additional properties in this case.

1.2 Seminorms and convex sets

Let V be a vector space over the real numbers. A subset E of V is said to be *convex* if for every $v, w \in E$ and real number t with $0 \le t \le 1$, we have that

$$(1 - t)\, v + t\, w \in E. \tag{1.2.1}$$

If V is a vector space over the complex numbers, then one can use the same definition, which basically corresponds to considering V as a vector space over the real numbers.

Now let V be a vector space over the real or complex numbers. A nonnegative real-valued function N on V is said to be a *seminorm* or *pseudonorm* if it satisfies the homogeneity condition (1.1.2) and triangle inequality (1.1.3). Note that (1.1.2) implies that $N(0) = 0$, by taking $t = 0$. Thus a seminorm N on V is a norm if and only if $N(v) > 0$ for every $v \in V$ with $v \neq 0$.

Let N be a seminorm on V. If $v \in V$ and r is a positive real number, then the *open ball* in V centered at v with radius r may be defined by

$$B(v, r) = B_N(v, r) = \{w \in V : N(v - w) < r\}. \tag{1.2.2}$$

Similarly, if r is a nonnegative real number, then the *closed ball* in V centered at v with radius r may be defined by

$$\overline{B}(v, r) = \overline{B}_N(v, r) = \{w \in V : N(v - w) \le r\}. \tag{1.2.3}$$

One can check that these are convex sets in V.

If N is a norm on V, then (1.2.2) and (1.2.3) are the same as the open and closed balls in V determined by the metric $d_N(\cdot, \cdot)$ on V associated to N. If N is a seminorm on V, then (1.1.9) defines a *semimetric* or *pseudometric* on V. This means that it satisfies the same conditions as a metric, except that $d_N(v, w) = 0$ does not necessarily imply that $v = w$.

Let N be a nonnegative real-valued function on V that satisfies the homogeneity condition (1.1.2). We may still use the notation (1.2.2) and (1.2.3) in this case, even if (1.1.9) is not necessarily a metric or semimetric on V. If the open unit ball $B_N(0, 1)$ in V with respect to N is convex, then it is well known that N satisfies the triangle inequality, and is thus a seminorm on V.

To see this, let $v, w \in V$ be given, and let r, t be real numbers with

$$N(v) < r, \; N(w) < t. \tag{1.2.4}$$

This implies that

$$r^{-1}\, v, \, t^{-1}\, w \in B_N(0, 1). \tag{1.2.5}$$

If $B_N(0, 1)$ is convex, then it follows that

$$(r + t)^{-1}\, (v + w) = (r\, (r + t)^{-1})\, (r^{-1}\, v) + (t\, (r + t)^{-1})\, (t^{-1}\, w) \tag{1.2.6}$$

is an element of $B_N(0, 1)$ too. This means that

$$N(v + w) < r + t. \tag{1.2.7}$$

One can use this to get (1.1.3).

Similarly, if the closed unit ball $\overline{B}_N(0, 1)$ is convex, then N satisfies the triangle inequality on V. More precisely, it would be enough to know that (1.2.5) implies that (1.2.6) is an element of $\overline{B}_N(0, 1)$. In this case, the inequality in (1.2.7) might not be strict, but the non-strict version would be sufficient to get (1.1.3).

If N is a seminorm on V, then

$$N(v) \le N(w) + N(v - w) \tag{1.2.8}$$

for every $v, w \in V$, and similarly with the roles of v and w interchanged. One can use this to check that

$$|N(v) - N(w)| \le N(v - w) \tag{1.2.9}$$

for all $v, w \in V$, using the ordinary absolute value of a real number on the left side. If N is a norm on V, then this implies that N is uniformly continuous as a real-valued function on V, with respect to the metric $d_N(\cdot, \cdot)$ associated to N on V, and the standard Euclidean metric on $\mathbf{R}$.

1.3 Some norms on $\mathbf{R}^n$, $\mathbf{C}^n$

Let n be a positive integer, and let $\mathbf{R}^n$, $\mathbf{C}^n$ be the usual spaces of n-tuples of real and complex numbers, respectively. These are vector spaces over the real and complex numbers, respectively, with respect to coordinatewise addition and scalar multiplication.

If $v \in \mathbf{R}^n$ or $\mathbf{C}^n$, then put

$$\|v\|_\infty = \max_{1 \le j \le n} |v_j|. \tag{1.3.1}$$

One can check that this defines a norm on each of $\mathbf{R}^n$ and $\mathbf{C}^n$. This corresponds to the suremum norm defined in Section 1.1, with X equal to the set of integers from 1 to n.

Let p be a positive real number, and put

$$\|v\|_p = \Big(\sum_{j=1}^n |v_j|^p\Big)^{1/p} \tag{1.3.2}$$

for each $v \in \mathbf{R}^n$ or $\mathbf{C}^n$. Observe that

$$\|v\|_p = 0 \text{ if and only if } v = 0. \tag{1.3.3}$$

We also have that

$$\|t\, v\|_p = |t|\, \|v\|_p \tag{1.3.4}$$

for each $v \in \mathbf{R}^n$ or $\mathbf{C}^n$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate.

If $p \ge 1$, then it is well known that (1.3.2) satisfies the triangle inequality on $\mathbf{R}^n$ and $\mathbf{C}^n$, and thus defines a norm. This is *Minkowski's inequality* for finite sums. Of course, this follows easily from the triangle inequality for the standard absolute value functions on $\mathbf{R}$ and $\mathbf{C}$ when $p = 1$.

If $p > 1$, then the triangle inequality can be obtained from the convexity of the open or closed unit ball, as in the previous section. This also uses the convexity of r^p as a function of a nonnegative real number r. More precisely, if x, y, and t are nonnegative real numbers, and $t \le 1$, then

$$((1-t)\, x + t\, y)^p \le (1-t)\, x^p + t\, y^p. \tag{1.3.5}$$

If $p = 2$, then (1.3.2) is the *standard Euclidean norm* on $\mathbf{R}^n$, $\mathbf{C}^n$. One can also look at this in terms of inner products, and we shall return to that in Section 1.9.

If $n = 1$, then (1.3.2) is the same as the standard asolute value function on $\mathbf{R}$, $\mathbf{C}$. If $n \geq 2$ and $0 < p < 1$, then one can check that (1.3.2) does not satisfy the triangle inequality on $\mathbf{R}^n$ or $\mathbf{C}^n$. However, it is well known that

$$\|v + w\|_p^p \leq \|v\|_p^p + \|w\|_p^p \tag{1.3.6}$$

for all $v, w \in \mathbf{R}^n$ or $\mathbf{C}^n$ when $0 < p \leq 1$. This implies that

$$\|v - w\|_p^p \tag{1.3.7}$$

defines a metric on each of $\mathbf{R}^n$ and $\mathbf{C}^n$ when $0 < p \leq 1$.

Clearly

$$\|v\|_\infty \leq \|v\|_p \tag{1.3.8}$$

for every $v \in \mathbf{R}^n$ or $\mathbf{C}^n$ and positive real number p. It is also easy to see that

$$\|v\|_p \leq n^{1/p}\,\|v\|_\infty \tag{1.3.9}$$

for all $v \in \mathbf{R}^n$ or $\mathbf{C}^n$ and $0 < p < \infty$. It is well known that $n^{1/p} \to 1$ as $p \to \infty$. This implies that

$$\lim_{p \to \infty} \|v\|_p = \|v\|_\infty \tag{1.3.10}$$

for all $v \in \mathbf{R}^n$ or $\mathbf{C}^n$.

1.4 Linear subspaces and continuous functions

Let V be a vector space over the real or complex numbers, and let W be a linear subspace of V. If N is a norm on V, then the restriction of N to W is a norm on W. Similarly, if N is a seminorm on V, then the restriction of N to W is a seminorm on W.

Let N be a norm on V, and let $d_N(u, v)$ be the corresponding metric on V. Note that the restriction of $d_N(u, v)$ to $u, v \in W$ is the same as the metric on W associated to the restriction of N to W.

One can check that the closure $\overline{W}$ of W in V with respect to $d_N(\cdot, \cdot)$ is a linear subspace of V as well. This can be viewed in terms of convergent sequences, which will be discussed further in the next section.

Let X be a nonempty metric space, or topological space. Consider the spaces $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ of continuous real and complex-valued functions on X. More precisely, this uses the standard metrics or topologies on $\mathbf{R}$, $\mathbf{C}$, respectively. It is well known that $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions. Equivalently, these are linear subspaces of the corresponding spaces of all real and complex-valued functions on X, respectively.

Consider the spaces

(1.4.1) $C_b(X, \mathbf{R}) = C(X, \mathbf{R}) \cap \ell^\infty(X, \mathbf{R}),\ C_b(X, \mathbf{C}) = C(X, \mathbf{C}) \cap \ell^\infty(X, \mathbf{C})$

of all bounded continuous real and complex-valued functions on X, respectively. These are linear subspaces of the spaces of all continuous real and complex-valued functions on X, respectively, as well as the spaces of all bounded real and complex-valued functions on X. In particular, the restriction of the supremum norm to each of these spaces is a norm, that we shall normally denote $\|\cdot\|_{sup}$.

It is well known that

(1.4.2) $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ are closed subsets of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$,

respectively, with respect to the supremum metric. This basically corresponds to the fact that if $\{f_j\}_{j=1}^\infty$ is a sequence of continuous real or complex-valued functions on X that converges uniformly to a real or complex-valued function f on X, then f is continuous on X.

If X is compact, then every continuous real or complex-valued function f on X is bounded. More precisely, $f(X)$ is a compact subset of the real line or complex plane, as appropriate, with respect to the standard Euclidean metric in this case. In particular, this implies that $f(X)$ is a bounded subset of $\mathbf{R}$ or $\mathbf{C}$, as appropriate.

If X is equipped with the discrete metric or topology, then every function on X is continuous. This means that $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ are the same as $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively, in this case.

1.5 Sequences, series, and Banach spaces

Let V be a vector space over the real or complex numbers, let N be a norm on V, and let $d_N(\cdot, \cdot)$ be the corresponding metric on V. Suppose that $\{v_j\}_{j=1}^\infty$ and $\{w_j\}_{j=1}^\infty$ are sequences of elements of V that converge to elements v and w of V, respectively, with respect to $d_N(\cdot, \cdot)$. Under these conditions, it is easy to see that $\{v_j + w_j\}_{j=1}^\infty$ converges to $v + w$ with respect to $d_N(\cdot, \cdot)$, so that

$$\lim_{j \to \infty} (v_j + w_j) = \Big(\lim_{j \to \infty} v_j\Big) + \Big(\lim_{j \to \infty} w_j\Big). \tag{1.5.1}$$

This is analogous to the corresponding statement for convergent sequences of real or complex numbers. Similarly, if $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $\{t\, v_j\}_{j=1}^\infty$ converges to $t\, v$ with respect to $d_N(\cdot, \cdot)$ on V, so that

$$\lim_{j \to \infty} (t\, v_j) = t\,\Big(\lim_{j \to \infty} v_j\Big). \tag{1.5.2}$$

If $\{t_j\}_{j=1}^\infty$ is a sequence of real or complex numbers that converges to $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then one can check that $\{t_j\, v_j\}_{j=1}^\infty$ converges to $t\, v$ with respect to $d_N(\cdot, \cdot)$ on V, so that

$$\lim_{j \to \infty} (t_j\, v_j) = \Big(\lim_{j \to \infty} t_j\Big)\,\Big(\lim_{j \to \infty} v_j\Big). \tag{1.5.3}$$

More precisely, one can verify that

$$\lim_{j \to \infty} (t_j\, v_j) = 0 \tag{1.5.4}$$

when $\{t_j\}_{j=1}^\infty$ is a bounded sequence of real or complex numbers, as appropriate, and $\{v_j\}_{j=1}^\infty$ converges to 0 in V. Similarly, (1.5.4) holds when $\{v_j\}_{j=1}^\infty$ is a bounded sequence in V, and $\{t_j\}_{j=1}^\infty$ converges to 0 in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. One can obtain (1.5.3) from these other properties of limits in V, in essentially the same way as for products of convergent sequences of real and complex numbers.

An infinite series $\sum_{j=1}^\infty v_j$ with terms in V is said to *converge* in V if the corresponding sequence of partial sums $\sum_{j=1}^n v_j$ converges in V with respect to $d_N(\cdot, \cdot)$. In this case, the value of the sum is defined by

$$\sum_{j=1}^\infty v_j = \lim_{n \to \infty} \sum_{j=1}^n v_j, \tag{1.5.5}$$

as usual. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then it is easy to see that $\sum_{j=1}^\infty t\, v_j$ converges, with

$$\sum_{j=1}^\infty (t\, v_j) = t \sum_{j=1}^\infty v_j. \tag{1.5.6}$$

If $\sum_{j=1}^\infty w_j$ is another convergent series with terms in V, then $\sum_{j=1}^\infty (v_j + w_j)$ converges, with

$$\sum_{j=1}^\infty (v_j + w_j) = \sum_{j=1}^\infty v_j + \sum_{j=1}^\infty w_j. \tag{1.5.7}$$

It is sometimes convenient to consider infinite series starting at $j = 0$, for which there are analogous statements.

It is well known that convergent sequences in metric spaces are Cauchy sequences. Let $\sum_{j=1}^\infty v_j$ be an infinite series with terms in V again. One can check that the corresponding sequence of partial sums is a Cauchy sequence if and only if for every $\epsilon > 0$ there is a positive integer $L(\epsilon)$ such that

$$N\Big(\sum_{j=l}^n v_j\Big) < \epsilon \tag{1.5.8}$$

for all positive integers l, n with $n \ge l \ge L(\epsilon)$. In particular, this implies that

$$N(v_l) < \epsilon \tag{1.5.9}$$

for all $l \ge L(\epsilon)$, by taking $n = l$. This means that

$$\lim_{l \to \infty} v_l = 0 \tag{1.5.10}$$

in V, with respect to $d_N(\cdot, \cdot)$.

If V is complete as a metric space with respect to $d_N(\cdot,\cdot)$, then V is said to be a *Banach space* with respect to N. In this case, the criterion for the sequence of partial sums to be a Cauchy sequence in the preceding paragraph implies the convergence of the infinite series. This is analogous to another classical fact about convergent series of real and complex numbers.

1.6 Completeness and subsets

Let $(M, d(\cdot,\cdot))$ be a metric space, and let E be a subset of M. The restriction of $d(\cdot,\cdot)$ to E defines a metric on E, so that E may be considered as a metric space as well. If M is complete as a metric space with respect to $d(\cdot,\cdot)$, and if E is a closed set in M, then

(1.6.1) $$E \text{ is complete as a metric space,}$$

with respect to the restriction of $d(\cdot,\cdot)$ to E. This uses the fact that a sequence of elements of E is a Cauchy sequence in E if and only if it is a Cauchy sequence as a sequence of elements of M.

If E is complete as a metric space, then one can check that

(1.6.2) $$E \text{ is a closed set in } M.$$

More precisely, this does not use completeness of M. It suffices to show that if a sequence of elements of E converges to an element of M, then the limit of the sequence is in E. Such a sequence may be considered as a Cauchy sequence in E, which converges to an element of E, by hypothesis. This is the same as the limit of the sequence in M, by uniqueness of the limit.

Let V be a real or complex vector space with a norm N, and let W be a linear subspace of V. If V is a Banach space with respect to N, and if W is a closed set in V with respect to the metric $d_N(\cdot,\cdot)$ associated to N, then it follows that

(1.6.3) $$W \text{ is a Banach space,}$$

with respect to the restriction of N to W. If W is a Banach space with respect to the restriction of N to W, then W is a closed set in V with respect to $d_N(\cdot,\cdot)$, as in the preceding paragraph.

If X is a nonempty set, then it is well known that

(1.6.4) $$\ell^\infty(X, \mathbf{R}) \text{ and } \ell^\infty(X, \mathbf{C}) \text{ are Banach spaces,}$$

with respect to the supremum norm. Indeed, let $\{f_j\}_{j=1}^\infty$ be a Cauchy sequence of bounded real or complex-valued functions on X with respect to the supremum metric. It is easy to see that for each $x \in X$,

(1.6.5) $$\{f_j(x)\}_{j=1}^\infty \text{ is a Cauchy sequence}$$

in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to the standard Euclidean metric. This implies that $\{f_j(x)\}_{j=1}^\infty$ converges in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, because the real line and complex plane are complete with respect to their standard metrics. Put

$$f(x) = \lim_{j \to \infty} f_j(x) \tag{1.6.6}$$

for each $x \in X$, which defines f as a real or complex-valued function on X, as appropriate. Thus $\{f_j\}_{j=1}^{\infty}$ converges to f pointwise on X, by construction. One can use this and the fact that $\{f_j\}_{j=1}^{\infty}$ is a Cauchy sequence with respect to the supremum metric to get that

$$\{f_j\}_{j=1}^{\infty} \text{ converges to } f \text{ uniformly on } X. \tag{1.6.7}$$

Using this, it is easy to see that f is bounded on X, because f_j is bounded on X for each j, by hypothesis. This means that $f \in \ell^{\infty}(X, \mathbf{R})$ or $\ell^{\infty}(X, \mathbf{C})$, as appropriate, and one can verify that $\{f_j\}_{j=1}^{\infty}$ converges to f with respect to the supremum metric.

Suppose now that X is a nonempty metric space, or topological space. Remember that $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ are closed sets in $\ell^{\infty}(X, \mathbf{R})$, $\ell^{\infty}(X, \mathbf{C})$, respectively, with respect to the supremum metric, as in Section 1.4. It follows that

$$C_b(X, \mathbf{R}) \text{ and } C_b(X, \mathbf{C}) \text{ are Banach spaces,} \tag{1.6.8}$$

with respect to the supremum norm too.

1.7 Absolute convergence

It is well known that an infinite series $\sum_{j=1}^{\infty} a_j$ of nonnegative real numbers converges if and only if the corresponding sequence of partial sums is bounded. In this case,

$$\sum_{j=1}^{\infty} a_j = \sup_{n \ge 1} \sum_{j=1}^{n} a_j. \tag{1.7.1}$$

Let V be a vector space over the real or complex numbers, and let N be a norm on V. An infinite series $\sum_{j=1}^{\infty} v_j$ with terms in V is said to *converge absolutely* with respect to N if

$$\sum_{j=1}^{\infty} N(v_j) \tag{1.7.2}$$

converges as an infinite series of nonnegative real numbers.

If l, n are positive integers with $l \le n$, then

$$N\Big(\sum_{j=l}^{n} v_j\Big) \le \sum_{j=l}^{n} N(v_j), \tag{1.7.3}$$

by the triangle inequality. If $\sum_{j=1}^{\infty} v_j$ converges absolutely with respect to N, then the corresponding sequence of partial sums is a Cauchy sequence with respect to the metric d_N on V associated to N. This follows from the characterization of the Cauchy condition mentioned in Section 1.5.

If V is a Banach space with respect to N, then we get that $\sum_{j=1}^{\infty} v_j$ converges in V. One can also check that

$$N\Big(\sum_{j=1}^{\infty} v_j\Big) \le \sum_{j=1}^{\infty} N(v_j) \tag{1.7.4}$$

under these conditions.

More precisely, it is well known that the completeness of V with respect to d_N is characterized by the condition that absolute convergence of an infinite series with respect to N imply convergence. To see this, let $\{w_j\}_{j=1}^{\infty}$ be a Cauchy sequence of elements of V with respect to d_N. One can find a subsequence $\{w_{j_l}\}_{l=1}^{\infty}$ of $\{w_j\}_{j=1}^{\infty}$ such that

$$\sum_{l=1}^{\infty} N(w_{j_l} - w_{j_{l+1}}) \tag{1.7.5}$$

converges as an infinite series of nonnegative real numbers. Indeed, one can choose the j_l's so that

$$N(w_{j_l} - w_{j_{l+1}}) \le 2^{-l} \tag{1.7.6}$$

for each l, for instance.

Thus we get that

$$\sum_{l=1}^{\infty} (w_{j_l} - w_{j_{l+1}}) \tag{1.7.7}$$

converges in V, by hypothesis. Note that

$$\sum_{l=1}^{n} (w_{j_l} - w_{j_{l+1}}) = w_{j_1} - w_{j_{n+1}} \tag{1.7.8}$$

for every positive integer n. This means that the convergence of (1.7.7) is equivalent to the convergence of $\{w_{j_l}\}_{l=1}^{\infty}$ to an element of V.

It is well known and not difficult to check that if a Cauchy sequence of elements of a metric space has a subsequence that converges to an element of the metric space, then the whole Cauchy sequence converges to the same limit. It follows that $\{w_j\}_{j=1}^{\infty}$ converges to an element of V with respect to d_N.

Let X be a nonempty set, and suppose that V is the space of bounded real or complex-valued functions on X, equipped with the supremum norm. In this case, the absolute convergence of an infinite series with terms in V with respect to the supremum norm corresponds to a well-known criterion of Weierstrass for the uniform convergence of the corresponding sequence of partial sums.

1.8 Holomorphic functions

Let U be a nonempty open subset of the complex plane, with respect to the standard Euclidean metric on $\mathbf{C}$. Consider the space $\mathcal{H}(U)$ of all complex-valued functions on U that are complex-analytic, or holomorphic. This is a linear subspace of the space $C(U, \mathbf{C})$ of all continuous complex-valued functions on U, with respect to the restriction of the standard Euclidean metric on $\mathbf{C}$ to U.

If $\{f_j\}_{j=1}^\infty$ is a sequence of holomorphic functions on U that converges uniformly to a complex-valued function f on U, then it is well known that f is holomorphic on U as well. Indeed, one can use the Cauchy integral formula for the f_j's to get that f can be expressed in the same way. This also works when $\{f_j\}_{j=1}^\infty$ converges to f uniformly on compact subsets of U.

Let

$$H^\infty(U) = \mathcal{H}(U) \cap C_b(U, \mathbf{C}) \tag{1.8.1}$$

be the space of all bounded holomorphic functions on U. This is a linear subspace of each of the spaces of all holomorphic functions on U, and all bounded continuous complex-valued functions on U. In fact, $H^\infty(U)$ is a closed set in $C_b(U, \mathbf{C})$, with respect to the supremum metric, as in the preceding paragraph. Thus $H^\infty(U)$ is a Banach spaces with respect to the supremum norm.

Let $\overline{U}$ be the closure of U in $\mathbf{C}$, with respect to the standard Euclidean metric. Consider the space $A(U)$ of continuous complex-valued functions on $\overline{U}$ that are holomorphic on U. This is a linear subspace of the space $C(\overline{U}, \mathbf{C})$ of all continuous complex-valued functions on $\overline{U}$. Similarly, let

$$A_b(U) = A(U) \cap C_b(\overline{U}, \mathbf{C}) \tag{1.8.2}$$

be the space of bounded continuous complex-valued functions on $\overline{U}$ that are holomorphic on U. Of course, if U is a bounded open set in $\mathbf{C}$, then $\overline{U}$ is compact, and $A_b(U) = A(U)$. As before, $A_b(U)$ is a closed linear subspace of $C_b(\overline{U}, \mathbf{C})$, with respect to the supremum metric. This means that $A_b(U)$ is a Banach space too, with respect to the supremum norm.

There are many examples of Banach spaces related to complex analysis like these. There are also analogous spaces of harmonic functions on open sets in $\mathbf{R}^n$, and spaces of holomorphic functions of several complex variables, for instance. Of course, Banach spaces and related notions are involved in the study of complex analysis and differential equations in other ways too. See [112, 141, 153] for some nice introductions to several complex variables, as well as [113] for more information.

1.9 Inner products and Hilbert spaces

Let V and W be vector spaces over the complex numbers. It is sometimes convenient to refer to a linear mapping from V into W, as vector spaces over $\mathbf{C}$, as being *complex-linear*. Similarly, it is sometimes convenient to refer to a linear mapping from V into W, considered as vector spaces over the real numbers, as being *real-linear*. Thus a real-linear mapping T from V into W is complex-linear if and only if

$$T(i\,v) = i\,T(v) \tag{1.9.1}$$

for every $v \in V$.

A real-linear mapping T from V into W is said to be *conjugate-linear* if

$$T(i\,v) = -i\,T(v) \tag{1.9.2}$$

for every $v \in V$. This implies that

$$T(a\,v) = \overline{a}\,T(v) \tag{1.9.3}$$

for every $v \in V$ and $a \in \mathbf{C}$. Here $\overline{a}$ is the complex-conjugate of $a \in \mathbf{C}$, as usual.

Now let V be a vector space over the real or complex numbers. An *inner product* on V is a real or complex-valued function $\langle v, w\rangle = \langle v, w\rangle_V$, as appropriate, defined for $v, w \in V$, that satisfies the following three conditions. First, for each $w \in V$, $\langle v, w\rangle$ should be linear as a function of v. Second, for each $v, w \in V$, we should have that

$$\langle v, w\rangle = \langle w, v\rangle \tag{1.9.4}$$

in the real case, and

$$\langle v, w\rangle = \overline{\langle w, v\rangle} \tag{1.9.5}$$

in the complex case. This implies that for each $v \in V$, $\langle v, w\rangle$ is linear as a function of w in the real case, and conjugate-linear as a function of w in the complex case. This also implies that

$$\langle v, v\rangle \in \mathbf{R} \tag{1.9.6}$$

for each $v \in V$ in the complex case. The third condition is that

$$\langle v, v\rangle > 0 \tag{1.9.7}$$

for each $v \in V$ with $v \neq 0$. Of course, $\langle v, v\rangle = 0$ when $v = 0$.

If $\langle \cdot, \cdot\rangle$ is an inner product on V, then we put

$$\|v\| = \|v\|_V = \langle v, v\rangle^{1/2} \tag{1.9.8}$$

for each $v \in V$, using the nonnegative square root on the right side. Note that $\|v\| = 0$ if and only if $v = 0$, and that

$$\|t\,v\| = |t|\,\|v\| \tag{1.9.9}$$

for each $v \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. The *Cauchy–Schwarz inequality* states that

$$|\langle v, w\rangle| \leq \|v\|\,\|w\| \tag{1.9.10}$$

for all $v, w \in V$. This can be obtained from the fact that

$$\langle v + t\,w, v + t\,w \rangle \geq 0 \tag{1.9.11}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. One can use the Cauchy–Schwarz inequality to get that

$$\|v + w\| \leq \|v\| + \|w\| \tag{1.9.12}$$

for all $v, w \in V$, so that (1.9.8) defines a norm on V.

If V is a complex vector space, then

$$\langle v, w \rangle_{V_{\mathbf{R}}} = \operatorname{Re}\langle v, w \rangle_V \tag{1.9.13}$$

defines an inner product on V as a vector space over the real numbers, where $\operatorname{Re} a$ is the real part of a complex number a. Clearly the norm associated to (1.9.13) is the same as before.

If n is a positive integer, then the standard inner product on $\mathbf{R}^n$ is defined by

$$\langle v, w \rangle = \langle v, w \rangle_{\mathbf{R}^n} = \sum_{j=1}^{n} v_j\, w_j. \tag{1.9.14}$$

Similarly, the standard inner product on $\mathbf{C}^n$ is defined by

$$\langle v, w \rangle = \langle v, w \rangle_{\mathbf{C}^n} = \sum_{j=1}^{n} v_j\, \overline{w_j}. \tag{1.9.15}$$

The norms associated to these inner products are the same as the standard Euclidean norms, as in Section 1.3.

Let V be a real or complex vector space with an inner product $\langle \cdot, \cdot \rangle$ again. If V is complete with respect to the metric associated to the corresponding norm $\| \cdot \|$, then V is said to be a *Hilbert space*. In particular, this means that V is a Banach space with respect to $\| \cdot \|$.

1.10 Sums of orthogonal vectors

Let $\sum_{j=1}^{\infty} a_j$ be a convergent series of nonnegative real numbers. It is well known that $\sum_{j=1}^{\infty} a_j^2$ also converges under these conditions. This uses the fact that the sequence of a_j's is bounded in this case. More precisely, we have that

$$\sum_{j=1}^{\infty} a_j^2 \leq \Big(\sup_{l \geq 1} a_l\Big) \sum_{j=1}^{\infty} a_j. \tag{1.10.1}$$

Let V be a vector space over the real or complex numbers with an inner product $\langle \cdot, \cdot \rangle$, and let $\| \cdot \|$ be the corresponding norm on V, as in the previous section. We say that $v, w \in V$ are *orthogonal* with respect to $\langle \cdot, \cdot \rangle$ if

$$\langle v, w \rangle = 0. \tag{1.10.2}$$

Note that this condition is symmetric in v and w. In this case, we get that

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2. \tag{1.10.3}$$

Suppose that $\sum_{j=1}^{\infty} v_j$ is an infinite series of pairwise-orthogonal vectors in V. This means that

$$\langle v_j, v_l \rangle = 0 \tag{1.10.4}$$

when $j \neq l$. If l, n are positive integers with $l \leq n$, then

$$\Big\|\sum_{j=l}^{n} v_j\Big\|^2 = \sum_{j=l}^{n} \|v_j\|^2. \tag{1.10.5}$$

If we take $l = 1$ in (1.10.5), then we get that the sequence of partial sums

$$\sum_{j=1}^{n} v_j \tag{1.10.6}$$

is bounded in V with respect to $\|\cdot\|$ if and only if the sequence of partial sums

$$\sum_{j=1}^{n} \|v_j\|^2 \tag{1.10.7}$$

is bounded in $\mathbf{R}$. Of course, this happens if and only if

$$\sum_{j=1}^{\infty} \|v_j\|^2 \tag{1.10.8}$$

converges as an infinite series of nonnegative real numbers. If this series converges, then the right side of (1.10.5) is as small as we want when l is large enough. This implies that the sequence of partial sums (1.10.6) is a Cauchy sequence in V with respect to the metric associated to $\|\cdot\|$, as in Section 1.5.

Suppose now that V is a Hilbert space. If (1.10.8) converges, then it follows that $\sum_{j=1}^{\infty} v_j$ converges in V. In this case, we also get that

$$\Big\|\sum_{j=1}^{\infty} v_j\Big\|^2 = \sum_{j=1}^{\infty} \|v_j\|^2. \tag{1.10.9}$$

Of course, if the sequence of partial sums (1.10.6) is a Cauchy sequence in V with respect to the metric associated to $\|\cdot\|$, then it is bounded. In particular, this happens when $\sum_{j=1}^{\infty} v_j$ converges in V.

1.11 Arbitrary norms on $\mathbf{R}^n$, $\mathbf{C}^n$

Let n be a positive integer, and let N be a seminorm on $\mathbf{R}^n$ or $\mathbf{C}^n$. Also let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$ or $\mathbf{C}^n$, so that the lth coordinate of e_j is equal to 1 when $j = l$, and to 0 otherwise. It is easy to see that

$$N(v) \leq \sum_{j=1}^{n} N(e_j)\, |v_j| \tag{1.11.1}$$

for all $v \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, by expressing v as a linear combination of the e_j's, and using the triangle inequality. This implies that

$$N(v) \leq \Big(\sum_{j=1}^{n} N(e_j)^2\Big)^{1/2} \|v\|_2 \tag{1.11.2}$$

for each $v \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, where $\|v\|_2$ is the standard Euclidean norm of v, as in Section 1.3. This also uses the Cauchy–Schwarz inequality.

It follows that

$$|N(v) - N(w)| \leq N(v - w) \leq \Big(\sum_{j=1}^{n} N(e_j)^2\Big)^{1/2} \|v - w\|_2 \tag{1.11.3}$$

for all $v, w \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, using (1.2.9) in the first step. In particular, this means that N is uniformly continuous as a real-valued function on $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, with respect to the corresponding Euclidean metric on the domain and range.

Suppose now that N is a norm on $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. The extreme value theorem implies that N attains its minimum on the unit sphere in $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, with respect to the standard Euclidean metric. The minimum of N on the unit sphere is positive, because N is a norm. This implies that there is a positive real number c such that

$$N(u) \geq c \tag{1.11.4}$$

for each $v \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, with $\|v\|_2 = 1$. Using this, one can check that

$$c\,\|v\|_2 \leq N(v) \tag{1.11.5}$$

for all $v \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate.

It is well known that $\mathbf{R}^n$ and $\mathbf{C}^n$ are complete with respect to their standard Euclidean metrics. One can use this and (1.11.2), (1.11.5) to get the completeness of $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, with respect to the metric associated to N.

If V is any vector space over the real or complex numbers of positive finite dimension n, then V is isomorphic to $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, as a vector space. If N_V is a norm on V, then it follows that V is complete with respect to the metric associated to N_V, as before.

1.12 Functions with finite support

Let X be a nonempty set, and let f be a real or complex-valued function on X. The *support* of f in X may be defined by

$$\operatorname{supp} f = \{x \in X : f(x) \neq 0\}. \tag{1.12.1}$$

Let $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ be the spaces of real and complex-valued functions f on X such that

$$\operatorname{supp} f \text{ has only finitely many elements,} \tag{1.12.2}$$

respectively. These are linear subspaces of the spaces of all real and complex-valued functions on X, as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively, with respect to pointwise addition and scalar multiplication. Of course, if X has only finitely many elements, then every real or complex-valued function on X has finite support.

A real or complex-valued function on X with finite support is obviously bounded, so that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ may be considered as linear subspaces of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively. In particular, the supremum norm defines a norm on each of $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$.

If f is a real or complex-valued function on X with finite support, then

$$\sum_{x \in X} f(x) \tag{1.12.3}$$

may be defined as a real or complex number, as appropriate. This is the same as the sum of $f(x)$ over any nonempty finite subset of X that contains the support of f. Similarly, if p is a positive real number and f is a real or complex-valued function on X with finite support, then $|f(x)|^p$ is a nonnegative real-valued function on X with finite support, so that

$$\sum_{x \in X} |f(x)|^p \tag{1.12.4}$$

may be defined as a nonnegative real number, as before. Under these conditions, we put

$$\|f\|_p = \Big(\sum_{x \in X} |f(x)|^p \Big)^{1/p}. \tag{1.12.5}$$

This corresponds to the definition of $\| \cdot \|_p$ on $\mathbf{R}^n$, $\mathbf{C}^n$ in Section 1.3 when X is the set of integers from 1 to n.

Clearly $\|f\|_p = 0$ if and only if $f = 0$ on X. It is easy to see that

$$\|t\, f\|_p = |t| \, \|f\|_p \tag{1.12.6}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. If $p \geq 1$, then

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \tag{1.12.7}$$

for all real and complex-valued functions f, g on X with finite support. This can be obtained from the analogous statement on $\mathbf{R}^n$, $\mathbf{C}^n$ in Section 1.3, or shown in essentially the same way. Thus (1.12.5) defines a norm on each of $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ when $p \geq 1$.

As before,

$$\|f\|_\infty \leq \|f\|_p \tag{1.12.8}$$

for every $p > 0$. If $0 < p_1 \le p_2 < +\infty$, then

$$\begin{aligned}(1.12.9) \qquad \|f\|_{p_2}^{p_2} = \sum_{x \in X} |f(x)|^{p_2} \quad &\le \quad \|f\|_\infty^{p_2 - p_1} \sum_{x \in X} |f(x)|^{p_1} \\ &= \quad \|f\|_\infty^{p_2 - p_1} \, \|f\|_{p_1}^{p_1} \le \|f\|_{p_1}^{p_2}.\end{aligned}$$

This means that

$$(1.12.10) \qquad \|f\|_{p_2} \le \|f\|_{p_1}.$$

If $0 < p \le 1$, then it is well known that

$$(1.12.11) \qquad (a + b)^p \le a^p + b^p$$

for all nonnegative real numbers a, b. This can be obtained from (1.12.10), with $p_1 = p$, and $p_2 = 1$. Using this, one can check that

$$(1.12.12) \qquad \|f + g\|_p^p \le \|f\|_p^p + \|g\|_p^p$$

for all real or complex-valued functions on on X finite support. This implies that

$$(1.12.13) \qquad \|f - g\|_p^p$$

defines a metric on each of $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$.

It is easy to see that

$$(1.12.14) \qquad \langle f, g \rangle = \langle f, g \rangle_{c_{00}(X, \mathbf{R})} = \sum_{x \in X} f(x)\, g(x)$$

and

$$(1.12.15) \qquad \langle f, g \rangle = \langle f, g \rangle_{c_{00}(X, \mathbf{C})} = \sum_{x \in X} f(x)\, \overline{g(x)}$$

define inner products on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$, respectively. The norms associated to these inner products are equal to (1.12.5), with $p = 2$. These inner products correspond to the standard inner products on $\mathbf{R}^n$ and $\mathbf{C}^n$ when X is the set of integers from 1 to n.

If $x \in X$, then let δ_x be the real-valued function on X defined by

$$\begin{aligned}(1.12.16) \qquad \delta_x(y) \quad &= \quad 1 \quad \text{when } y = x \\ &= \quad 0 \quad \text{when } y \ne x.\end{aligned}$$

It is easy to see that the collection of δ_x's, $x \in X$, is a basis for each of $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$, as vector spaces over the real and complex numbers, respectively. Of course,

$$(1.12.17) \qquad \|\delta_x\|_p = 1$$

for every $x \in X$ and $0 < p \le \infty$. The δ_x's are also pairwise orthogonal with respect to the inner products defined in the preceding paragraph.

1.13 Vanishing at infinity

Let X be a nonempty set, and let f be a real or complex-valued function on X again. We say that f *vanishes at infinity* on X if for every $\epsilon > 0$,

$$|f(x)| < \epsilon \tag{1.13.1}$$

for all but finitely many $x \in X$. This holds automatically when f has finite support in X, and in particular when X has only finitely many elements. If X is the set $\mathbf{Z}_+$ of all positive integers, then f vanishes at infinity if and only if

$$\lim_{j \to \infty} f(j) = 0. \tag{1.13.2}$$

If f is a real or complex-valued function on X that vanishes at infinity, then

$$\operatorname{supp} f \text{ has at most finitely or countably many elements.} \tag{1.13.3}$$

Indeed, for each positive integer n,

$$\{x \in X : |f(x)| \ge 1/n\} \tag{1.13.4}$$

has only finitely many elements, and the support of f is the same as the union of these sets.

Suppose that E is a countably infinite subset of X, and let $\{x_j\}_{j=1}^\infty$ be a sequence of elements of E such that every element of E occurs in the sequence exactly once. Let f be a real or complex-valued function on X with support contained in E. Under these conditions, f vanishes at infinity on X if and only if

$$\lim_{j \to \infty} f(x_j) = 0. \tag{1.13.5}$$

Let

$$c_0(X, \mathbf{R}),\ c_0(X, \mathbf{C}) \tag{1.13.6}$$

be the spaces of real and complex-valued functions on X that vanish at infinity, respectively. If f is a real or complex-valued function on X that vanishes at infinity, then it is easy to see that f is bounded on X. Thus

$$c_{00}(X, \mathbf{R}) \subseteq c_0(X, \mathbf{R}) \subseteq \ell^\infty(X, \mathbf{R}) \tag{1.13.7}$$

and

$$c_{00}(X, \mathbf{C}) \subseteq c_0(X, \mathbf{C}) \subseteq \ell^\infty(X, \mathbf{C}). \tag{1.13.8}$$

More precisely, one can check that $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$ are linear subspaces of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively.

One can also verify that

$$c_0(X, \mathbf{R}),\ c_0(X, \mathbf{C}) \text{ are closed sets in } \ell^\infty(X, \mathbf{R}),\ \ell^\infty(X, \mathbf{C}), \tag{1.13.9}$$

respectively, with respect to the supremum metric. Equivalently, if a real or complex-valued function f on X can be approximated uniformly on X by functions that vanish at infinity on X, then f vanishes at infinity on X as well. Thus

$$c_0(X, \mathbf{R}),\ c_0(X, \mathbf{C}) \text{ are Banach spaces} \tag{1.13.10}$$

with respect to the supremum norm.

If f is a real or complex-valued function on X that vanishes at infinity, then it is easy to see that f can be approximated by functions with finite support on X, uniformly on X. This means that $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$ are the same as the closures of $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ in $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively, with respect to the supremum metric.

Suppose that $\{x_j\}_{j=1}^\infty$ is a sequence of distinct elements of X, and let δ_{x_j} be as in (1.12.16) for each j. Also let $a_1, a_2, a_3, \ldots$ be an infinite sequence of real or complex numbers. One can check that

$$\sum_{j=1}^\infty a_j\, \delta_{x_j} \tag{1.13.11}$$

converges with respect to the supremum metric if and only if $\{a_j\}_{j=1}^\infty$ converges to 0. More precisely, it is easy to see that the sequence of partial sums

$$\sum_{j=1}^n a_j\, \delta_{x_j} \tag{1.13.12}$$

converges pointwise on X to the function that is equal to a_j at x_j for every j, and to 0 at all other points in X. The previous statement is basically the same as saying that this sequence of partial sums converges uniformly on X to the function just mentioned if and only if $\{a_j\}_{j=1}^\infty$ converges to 0.

1.14 Separability

A metric or topological space X is said to be *separable* if there is a dense set E in X such that E has only finitely or countably many elements. It is well known that $\mathbf{R}^n$ and $\mathbf{C}^n$ are separable with respect to their standard Euclidean metrics for each positive integer n.

Let V be a vector space over the real or complex numbers with a norm N. If V has finite dimension, then one can check that V is separable, with respect to the metric associated to N.

Similarly, suppose that $v_1, v_2, v_3, \ldots$ is a sequence of elements of V whose linear span is dense in V, with respect to the metric associated to N. One can verify that

$$V \text{ is separable with respect to the metric associated to } N \tag{1.14.1}$$

under these conditions. More precisely, one can get a countable dense set in V by taking linear combinations of the v_j's with rational coefficients in the real case, and with coefficients that have rational real and imaginary parts in the complex case.

If X is an infinite set, then it is well known and not too difficult to show that $\ell^\infty(X, \mathbf{R})$ is not separable with respect to the supremum metric. If X is countably infinite, then $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are separable with respect to the metric associated to any norm, as in the preceding paragraph. Similarly, $c_0(X, \mathbf{R})$ and $c_0(X, \mathbf{C})$ are separable with respect to the supremum metric in this case.

A collection $\mathcal{B}$ of open subsets of a metric or topological space X is said to be a *base for the topology of* X if every open set in X can be expressed as a union of elements of $\mathcal{B}$. If there is a base $\mathcal{B}$ for the topology of X such that $\mathcal{B}$ has only finitely or countably many elements, then X is said to satisfy the *second countability condition*. It is well known that X is separable in this case. Topological spaces that satisfy the second countability condition are also sometimes said to be *completely separable*. It is well known that separable metric spaces satisfy the second countability condition.

Let Y be a subset of X, equipped with the induced topology. If X satisfies the second countability condition, then it is easy to see that Y satisfies the second countability condition too. In particular, this implies that Y is separable, as before.

If $d(\cdot, \cdot)$ is a metric on X, then the restriction of $d(\cdot, \cdot)$ to Y is a metric on Y. It is well known that the topology determined on Y by the restriction of $d(\cdot, \cdot)$ to Y is the same as the topology induced on Y by the topology determined on X by $d(\cdot, \cdot)$. If X is separable, then one can show more directly that Y is separable in this case.

1.15 Some continuous extensions

Let (X, d_X) and (Y, d_Y) be metric spaces. Suppose for the moment that f is a uniformly continuous mapping from X into Y. If $\{x_j\}_{j=1}^\infty$ is a Cauchy sequence of elements of X, then one can check that

$$\{f(x_j)\}_{j=1}^\infty \text{ is a Cauchy sequence in } Y. \tag{1.15.1}$$

Of course, if Y is complete, then it follows that $\{f(x_j)\}_{j=1}^\infty$ converges to an element of Y.

Suppose now that E is a dense set in X, and that f is a uniformly continuous mapping from E into Y, with respect to the restriction of $d_X(\cdot, \cdot)$ to E. If Y is complete, then it is well known that

$$\begin{aligned} &f \text{ has a unique extension to a uniformly} \\ &\text{continuous mapping from } X \text{ into } Y. \end{aligned} \tag{1.15.2}$$

More precisely, uniqueness of the extension only requires ordinary continuity on X.

To get the existence of the extension, let $x \in X$ be given, and let $\{x_j\}_{j=1}^\infty$ be a sequence of elements of E that converges to x. Thus $\{x_j\}_{j=1}^\infty$ may be considered as a Cauchy sequence in E, so that

$$\{f(x_j)\}_{j=1}^\infty \text{ converges to an element of } Y, \tag{1.15.3}$$

as before. One can check that the limit only depends on x, and not the particular sequence $\{x_j\}_{j=1}^\infty$. One can also verify that this defines a uniformly continuous extension of f to X.

A mapping ϕ from X into Y is said to be an *isometry* if

$$d_Y(\phi(x), \phi(w)) = d_X(x, w) \tag{1.15.4}$$

for every $x, w \in X$. If ϕ is an isometry from X into Y, and X is complete, then

$$\phi(X) \text{ is a closed set in } Y. \tag{1.15.5}$$

More precisely, $\phi(X)$ is complete with respect to the restriction of d_Y to $\phi(X)$ in this case. This implies that $\phi(X)$ is a closed set in Y, as in Section 1.6.

A *completion* of X as a metric space may be defined as an isometric embedding of X onto a dense subset of a complete metric space. If X is already complete, then one can simply use the identity mapping on X. It is well known that every metric space has a completion. One can also show that completions are unique, up to a sutiable isometric equivalence, using the previous remarks.

Let V be a vector space over the real or complex numbers with a norm N. If V is not already complete with respect to the metric associated to N, then it is well known that V has a completion that is a Banach space. More precisely, this may be considered as an isometric linear mapping from V onto a dense linear subspace of a Banach space. If one uses a standard abstract construction of the completion, then one can check that the completion has these additional properties in this case. Alternatively, one can show that the vector space operations and norm on V can be extended to any completion of V as a metric space in a natural way.

Similarly, if V is an inner product space that it not already complete with respect to the metric associated to the corresponding norm, then it is well known that V has a completion that is a Hilbert space. As before, this can be shown using a standard abstract construction of the completion, or by extending the inner product to any completion of V in a natural way.

Chapter 2

Lipschitz and bounded linear mappings

2.1 Lipschitz mappings

Let (X, d_X) and (Y, d_Y) be metric spaces. A mapping f from X into Y is said to be *Lipschitz* if

$$d_Y(f(x), f(w)) \le C\, d_X(x, w) \tag{2.1.1}$$

for some nonnegative real number C and all $x, w \in X$. In this case, we may say that f is Lipschitz with constant C, to be more precise. Note that

(2.1.2) Lipschitz mappings are uniformly continuous.

Clearly a mapping f from X into Y is Lipschitz with constant $C = 0$ if and only if f is constant on X.

One can check that a real-valued function f on X is Lipschitz with constant $C \ge 0$ if and only if

$$f(x) \le f(w) + C\, d_X(x, w) \tag{2.1.3}$$

for every $x, w \in X$. This uses the standard Euclidean metric on the real line, as the range of f. In particular, if $p \in X$, then

$$f_p(x) = d_X(x, p) \tag{2.1.4}$$

is Lipschitz with constant 1 on X.

Let f be a continuous mapping from X into Y, and let E be a dense set in X. Suppose that f is Lipschitz with constant C on E, with respect to the restriction of d_X to E. Under these conditions, one can check that

(2.1.5) f is Lipschitz with constant C on all of X.

The space of all Lipschitz mappings from X into Y may be denoted

$$\mathrm{Lip}(X, Y). \tag{2.1.6}$$

Suppose that $X \neq \emptyset$, and that f is a Lipschitz mapping from X into Y. If X has at least two elements, then put

$$\mathrm{Lip}(f) = \mathrm{Lip}_{X,Y}(f) = \sup\left\{ \frac{d_Y(f(x), f(w))}{d_X(x,w)} : x, w \in X,\, x \neq w \right\}. \tag{2.1.7}$$

Otherwise, if X has only one element, then this may be interpreted as being equal to 0. It is easy to see that f is Lipschitz with constant $\mathrm{Lip}(f)$. More precisely, $\mathrm{Lip}(f)$ is the smallest nonnegative real number C such that f is Lipschitz with constant C.

Let (Z, d_Z) be another metric space, and suppose that f is a Lipschitz mapping from X into Y, and that g is a Lipschitz mapping from Y into Z. It is easy to see that the compostion $g \circ f$ of f and g is a Lipschitz mapping from X into Z, with

$$\mathrm{Lip}_{X,Z}(g \circ f) \le \mathrm{Lip}_{X,Y}(f)\, \mathrm{Lip}_{Y,Z}(g). \tag{2.1.8}$$

Suppose now that Y is a vector space over the real or complex numbers. The space of all functions on X with values in Y is a vector space over $\mathbf{R}$ or $\mathbf{C}$ as well, as appropriate, with respect to pointwise addition and scalar multiplication of functions. Let N_Y be a norm on Y, and suppose that $d_Y = d_{N_Y}$ is the metric on Y associated to N_Y. One can check that

$$\begin{aligned} &\mathrm{Lip}(X, Y) \text{ is a linear subspace of the space} \\ &\text{of all functions on } X \text{ with values in } Y, \end{aligned} \tag{2.1.9}$$

and that

$$\mathrm{Lip}(f) \text{ is a seminorm on } \mathrm{Lip}(X, Y). \tag{2.1.10}$$

Let us take $Y = \mathbf{R}$ with the standard Euclidean metric again, and let f_p be as in (2.1.4) for each $p \in X$. One can verify that $f_p - f_q$ is bounded on X for every $p, q \in X$, with

$$\sup_{x \in X} |f_p(x) - f_q(x)| = d_X(p, q). \tag{2.1.11}$$

If X is bounded with respect to d_X, then f_p is bounded on X for each $p \in X$. This means that

$$p \mapsto f_p \tag{2.1.12}$$

defines an isometric embedding of X into the space $C_b(X, \mathbf{R})$ of all bounded continuous real-valued functions on X, with respect to the supremum metric.

If X is not bounded, then let us fix a basepoint $p_0 \in X$. Thus $f_p - f_{p_0}$ is bounded on X for each $p \in X$, as in the preceding paragraph. It is easy to see that

$$p \mapsto f_p - f_{p_0} \tag{2.1.13}$$

is an isometric embedding of X into $C_b(X, \mathbf{R})$, with respect to the supremum metric.

2.2 Bounded linear mappings

Let V, W be vector spaces, both real or both complex. The space of all linear mappings from V into W may be denoted

$$\mathcal{L}(V, W). \tag{2.2.1}$$

This is a linear subspace of the space of all functions on V with values in W.

Let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively. A linear mapping T from V into W is said to be *bounded* with respect to these norms if

$$\|T(v)\|_W \le C\,\|v\|_V \tag{2.2.2}$$

for some nonnegative real number C and all $v \in V$. This implies that

$$\|T(u) - T(v)\|_W = \|T(u - v)\|_W \le C\,\|u - v\|_V \tag{2.2.3}$$

for all $u, v \in V$, so that T is Lipschitz with constant C with respect to the metrics on V, W associated to the norms.

If T is a linear mapping from V into W, and if $\|T(v)\|_W$ is bounded on a ball of positive radius in V centered at 0 with respect to the metric associated to $\|\cdot\|_V$, then one can check that

$$T \text{ is a bounded linear mapping.} \tag{2.2.4}$$

In particular, this holds when T is continuous at 0 with respect to the metrics on V, W associated to the norms.

Let

$$\mathcal{BL}(V, W) \tag{2.2.5}$$

be the space of all bounded linear mappings from V into W. One can verify that this is a linear subspace of $\mathcal{L}(V, W)$.

If T is a bounded linear mapping from V into W, then put

$$\|T\|_{op} = \|T\|_{op,VW} = \sup\{\|T(v)\|_W : v \in V,\ \|v\|_V \le 1\}. \tag{2.2.6}$$

This is the same as the smallest nonnegative real number C such that (2.2.2) holds. This is also the same as the Lipschitz constant $\mathrm{Lip}(T)$ of T as in the previous section, with respect to the metrics on V, W associated to their norms. One can check that

$$\|\cdot\|_{op} \text{ defines a norm on } \mathcal{BL}(V, W), \tag{2.2.7}$$

which is the *operator norm* associated to the norms on V, W.

Let Z be another vector space over the real or complex numbers, as appropriate, and with a norm $\|\cdot\|_Z$. If T_1 is a bounded linear mapping from V into W, and T_2 is a bounded linear mapping from W into Z, then their composition $T_2 \circ T_1$ is a bounded linear mapping from V into Z, with

$$\|T_2 \circ T_1\|_{op,VZ} \le \|T_1\|_{op,VW}\,\|T_2\|_{op,WZ}. \tag{2.2.8}$$

If W is complete with respect to the metric associated to $\| \cdot \|_W$, then it is well known that

$$\mathcal{BL}(V, W) \text{ is complete} \tag{2.2.9}$$

with respect to the metric associated to the operator norm. To see this, let $\{T_j\}_{j=1}^\infty$ be a Cauchy sequence in $\mathcal{BL}(V, W)$ with respect to the metric associated to the operator norm. This means that for each $\epsilon > 0$ there is a positive integer $L(\epsilon)$ such that

$$\|T_j - T_l\|_{op} < \epsilon \tag{2.2.10}$$

for every $j, l \geq L(\epsilon)$. It follows that

$$\|T_j(v) - T_l(v)\|_W \leq \epsilon \, \|v\|_V \tag{2.2.11}$$

for every $v \in V$ and $j, l \geq L(\epsilon)$. In particular, this implies that

$$\{T_j(v)\}_{j=1}^\infty \text{ is a Cauchy sequence in } W \tag{2.2.12}$$

for each $v \in V$, with respect to the metric associated to $\| \cdot \|_W$. Thus $\{T_j(v)\}_{j=1}^\infty$ converges to an element of W for each $v \in V$, by completeness. If we put

$$T(v) = \lim_{j \to \infty} T_j(v) \tag{2.2.13}$$

for each $v \in V$, then it is easy to see that T defines a linear mapping from V into W.

It is well known and easy to see that a Cauchy sequence in any metric space is bounded. In this case, this means that $\{T_j\}_{j=1}^\infty$ is bounded with respect to the operator norm, so that $\{\|T_j\|_{op}\}_{j=1}^\infty$ is a bounded sequence of nonnegative real numbers. If $v \in V$, then one can check that

$$\|T(v)\|_W \leq \sup_{j \geq 1} \|T_j(v)\|_W \leq \Big(\sup_{j \geq 1} \|T_j\|_{op}\Big) \, \|v\|_V. \tag{2.2.14}$$

This means that T is a bounded linear mapping from V into W, with

$$\|T\|_{op} \leq \sup_{j \geq 1} \|T_j\|_{op}. \tag{2.2.15}$$

Similarly, one can verify that

$$\|T(v) - T_l(v)\|_W \leq \epsilon \, \|v\|_V \tag{2.2.16}$$

for every $v \in V$ and $l \geq L(\epsilon)$. This implies that

$$\|T - T_l\|_{op} \leq \epsilon \tag{2.2.17}$$

for every $l \geq L(\epsilon)$. This shows that $\{T_l\}_{l=1}^\infty$ converges to T with respect to the metric on $\mathcal{BL}(V, W)$ associated to the operator norm, as desired.

Let V_0 be a linear subspace of V that is dense in V, with respect to the metric associated to $\| \cdot \|_V$. Also let T_0 be a bounded linear mapping from V_0

into W, with respect to the restriction of $\|\cdot\|_V$ to V_0. If W is complete with respect to the metric associated to $\|\cdot\|_W$, then there is a unique extension of T_0 to a uniformly continuous mapping T from V into W, as in Section 1.15. In this case, one can check that

$$T \text{ is a linear mapping from } V \text{ into } W. \tag{2.2.18}$$

More precisely,

$$T \text{ is a bounded linear mapping from } V \text{ into } W, \tag{2.2.19}$$

with operator norm equal to the operator norm of T_0 on V_0.

2.3 Some linear mappings

Let X be a nonempty set, and let W be a vector space over the real or complex numbers. If f is a function on X with values in W, then the *support* of f may be defined as the set $\operatorname{supp} f$ of $x \in X$ such that

$$f(x) \neq 0, \tag{2.3.1}$$

just as for real or complex-valued functions on X, as in Section 1.12. Let $c_{00}(X, W)$ be the space of W-valued functions f on X such that

$$\operatorname{supp} f \text{ has only finitely many elements.} \tag{2.3.2}$$

This is a linear subspace of the space of all W-valued functions on X.

If $f \in c_{00}(X, W)$, then

$$\sum_{x \in X} f(x) \tag{2.3.3}$$

may be defined as an element of W. This is the same as the sum of $f(x)$ over any nonempty finite subset of X that contains the support of f, as before. This defines a linear mapping from $c_{00}(X, W)$ into W.

It is convenient to put

$$V = c_{00}(X, \mathbf{R}) \text{ or } c_{00}(X, \mathbf{C}) \tag{2.3.4}$$

for the moment, depending on whether W is a vector space over $\mathbf{R}$ or $\mathbf{C}$. Let a be any W-valued function on X. If $f \in V$, then

$$a\, f \in c_{00}(X, W), \tag{2.3.5}$$

and we put

$$T_a(f) = \sum_{x \in X} a(x)\, f(x). \tag{2.3.6}$$

This defines a linear mapping from V into W.

If $y \in X$, then let δ_y be the real-valued function on X equal to 1 at x and 0 otherwise, as in Section 1.12. Note that

$$T_a(\delta_y) = a(y) \tag{2.3.7}$$

for every $y \in X$. It is easy to see that

(2.3.8) every linear mapping T from V into W is of the form T_a

for a unique W-valued function a on X. This uses the fact that the collection of δ_y's, $y \in X$, is a basis for V as a vector space over **R** or **C**, as appropriate, as in Section 1.12. More precisely, this holds with

$$a(y) = T(\delta_y) \tag{2.3.9}$$

for every $y \in X$.

Let $\| \cdot \|_W$ be a norm on W. A W-valued function a on X is said to be *bounded* on X with respect to $\| \cdot \|_W$ if

(2.3.10) $\|a(x)\|_W$ is bounded as a real-valued function on X.

Let

$$\ell^\infty(X, W) \tag{2.3.11}$$

be the space of W-valued functions on X that are bounded with respect to $\| \cdot \|_W$. This is a linear subspace of the space of all W-valued functions on X.

If $a \in \ell^\infty(X, W)$, then put

$$\|a\|_\infty = \|a\|_{sup} = \|a\|_{\ell^\infty(X,W)} = \sup\{\|a(x)\|_W : x \in X\}. \tag{2.3.12}$$

One can check that this defines a norm on $\ell^\infty(X, W)$, which is another version of the *supremum norm.* The metric associated to the supremum norm on $\ell^\infty(X, W)$ is the same as the *supremum metric* corresponding to the metric on W associated to $\| \cdot \|_W$. If W is a Banach space with respect to $\| \cdot \|_W$, then

(2.3.13) $\ell^\infty(X, W)$ is a Banach space

with respect to (2.3.12), by standard arguments. This is analogous to the cases of bounded real and complex-valued functions on X, as in Section 1.6.

Let a be a W-valued function on X again, and let T_a be as in (2.3.6). Also let $\| \cdot \|_1$ be the norm defined on V as in Section 1.12, with $p = 1$. Note that

$$\|\delta_y\|_1 = 1 \tag{2.3.14}$$

for every $y \in X$. If T_a is bounded as a linear mapping from V into W, with respect to $\| \cdot \|_1$ and $\| \cdot \|_W$, respectively, then

$$\|a(y)\|_W = \|T_a(\delta_y)\|_W \le \|T_a\|_{op} \, \|\delta_y\|_1 = \|T_a\|_{op} \tag{2.3.15}$$

for every $y \in X$. This means that $a \in \ell^\infty(X, W)$, with

$$\|a\|_\infty \le \|T_a\|_{op}. \tag{2.3.16}$$

If a is any element of $\ell^\infty(X, W)$ and $f \in V$, then

$$\|T_a(f)\|_W \le \sum_{x \in X} \|a(x)\|_W \, |f(x)| \le \|a\|_\infty \, \|f\|_1. \tag{2.3.17}$$

This implies that T_a is a bounded linear mapping from V into W, with respect to $\|\cdot\|_1$ and $\|\cdot\|_W$, respectively, with

$$\|T_a\|_{op} \le \|a\|_\infty. \tag{2.3.18}$$

It follows that

$$\|T_a\|_{op} = \|a\|_\infty \tag{2.3.19}$$

under these conditions.

2.4 Nonnegative sums

Let X be a nonempty set, and let f be a nonnegative real-valued function on X. If A is a nonempty finite subset of X, then the sum

$$\sum_{x \in A} f(x) \tag{2.4.1}$$

is defined as a nonnegative real number in the usual way. This may be interpreted as being equal to 0 when $A = \emptyset$.

The sum

$$\sum_{x \in X} f(x) \tag{2.4.2}$$

may be defined as a nonnegative extended real number as the supremum of the finite subsums (2.4.1) over all nonempty finite subsets A of X. Of course, if X has only finitely many elements, then this is the same as the usual sum over X. This definition of the sum over an arbitrary nonempty set X is the same as the Lebesgue integral of f over X, with respect to counting measure on X.

If t is a positive real number, then

$$\sum_{x \in X} t\, f(x) = t \sum_{x \in X} f(x), \tag{2.4.3}$$

where the right side is interpreted as being equal to $+\infty$ when (2.4.2) is equal to $+\infty$, as usual. This also works when $t = 0$, with the right side interpreted as being equal to 0 even when (2.4.2) is $+\infty$.

If g is another nonnegative real-valued function on X, then one can check that

$$\sum_{x \in X} (f(x) + g(x)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x). \tag{2.4.4}$$

The right side should be interpreted as being equal to $+\infty$ when either of the sums is $+\infty$.

If (2.4.2) is finite, then f is said to be *summable* on X. It is easy to see that

$$f \text{ vanishes at infinity on } X \tag{2.4.5}$$

in this case, in the sense of Section 1.13. In particular, this implies that

(2.4.6) the support of f has only finitely or countably many elements,

as before.

Let $\{x_j\}_{j=1}^\infty$ be a sequence of distinct elements of X, and suppose for the moment that the support of f is contained in the set of x_j's. Under these conditions, one can verify that

$$\sum_{x \in X} f(x) = \sum_{j=1}^\infty f(x_j). \tag{2.4.7}$$

More precisely, the sum on the right should be interpreted as being $+\infty$ when the series does not converge.

If E is any subset of X, then $\sum_{x \in E} f(x)$ may be defined as a nonnegative extended real number, as before. If E_1 and E_2 are disjoint subsets of X, then

$$\sum_{x \in E_1 \cup E_2} f(x) = \sum_{x \in E_1} f(x) + \sum_{x \in E_2} f(x), \tag{2.4.8}$$

as in (2.4.4).

Suppose for the moment that f is summable on X, and let $\epsilon > 0$ be given. Observe that there is a finite subset $A(\epsilon)$ of X such that

$$\sum_{x \in X} f(x) < \sum_{x \in A(\epsilon)} f(x) + \epsilon, \tag{2.4.9}$$

by the definition of the sum (2.4.2). This means that

$$\sum_{x \in X \setminus A(\epsilon)} f(x) < \epsilon. \tag{2.4.10}$$

Let $\{f_j\}_{j=1}^\infty$ be a sequence of nonnegative real-valued functions on X that converges to f pointwise on X. Suppose that there is a nonnegative real-number C such that

$$\sum_{x \in X} f_j(x) \le C \tag{2.4.11}$$

for each j. Under these conditions, it is well known that

$$\sum_{x \in X} f(x) \le C. \tag{2.4.12}$$

Indeed, if A is any finite subset of X, then

$$\sum_{x \in A} f(x) = \lim_{j \to \infty} \sum_{x \in A} f_j(x) \le C. \tag{2.4.13}$$

2.5 p-Summability

Let X be a nonempty set, and let p be a positive real number. A nonnegative real-valued function f on X is said to be *p-summable* on X if

$$f(x)^p \text{ is summable on } X. \tag{2.5.1}$$

If g is another nonnegative real-valued function on X, and f, g are both p-summable on X, then one can check that

$$f + g \text{ is } p\text{-summable on } X \tag{2.5.2}$$

too.

More precisely, if $0 < p \le 1$, then

$$\sum_{x \in X} (f(x) + g(x))^p \le \sum_{x \in X} (f(x)^p + g(x)^p) = \sum_{x \in X} f(x)^p + \sum_{x \in X} g(x)^p, \tag{2.5.3}$$

using (1.12.11) in the first step. If $p \ge 1$, and f, g are p-summable on X, then

$$\Big(\sum_{x \in X} (f(x) + g(x))^p\Big)^{1/p} \le \Big(\sum_{x \in X} f(x)^p\Big)^{1/p} + \Big(\sum_{x \in X} g(x)^p\Big)^{1/p}. \tag{2.5.4}$$

This is *Minkowski's inequality* for arbitrary sums. This can be obtained from Minkowski's inequality for finite sums, as in Section 1.3, or shown using similar arguments.

If f is p-summable on X for some $p > 0$, then f is bounded on X, with

$$\sup_{x \in X} f(x) \le \Big(\sum_{x \in X} f(x)^p\Big)^{1/p}. \tag{2.5.5}$$

We also have that f vanishes at infinity on X, because $f(x)^p$ vanishes at infinity on X, as in the previous section. In fact, for each $\epsilon > 0$, we have that

$$\epsilon^p \, (\#\{x \in X : f(x) \ge \epsilon\}) \le \sum_{x \in X} f(x)^p. \tag{2.5.6}$$

Here $\#E$ denotes the number of elements of a set E.

Suppose that $0 < p_1 \le p_2 < \infty$, and that f is p_1-summable on X. Under these conditions, f is p_2-summable on X, with

$$\sum_{x \in X} f(x)^{p_2} \le \Big(\sup_{x \in X} f(x)\Big)^{p_2 - p_1} \sum_{x \in X} f(x)^{p_1} \le \Big(\sum_{x \in X} f(x)^{p_1}\Big)^{p_2/p_1}, \tag{2.5.7}$$

using (2.5.5) with $p = p_1$ in the second step. This implies that

$$\begin{aligned} \Big(\sum_{x \in X} f(x)^{p_2}\Big)^{1/p_2} &\le \Big(\sup_{x \in X} f(x)\Big)^{1 - (p_1/p_2)} \Big(\sum_{x \in X} f(x)^{p_1}\Big)^{1/p_2} \\ &\le \Big(\sum_{x \in X} f(x)^{p_1}\Big)^{1/p_1}. \end{aligned} \tag{2.5.8}$$

One can use (2.5.5) and the first inequality in (2.5.8) to get that

$$\lim_{p\to\infty} \Big(\sum_{x\in X} f(x)^p\Big)^{1/p} = \sup_{x\in X} f(x) \tag{2.5.9}$$

in this case.

2.6 ℓ^p Spaces

Let X be a nonempty set again, and let p be a positive real number. Also let W be a vector space over the real or complex numbers, with a norm $\|\cdot\|_W$. Consider the space

$$\ell^p(X, W) \tag{2.6.1}$$

of W-valued functions f on X such that

(2.6.2) $\|f(x)\|_W$ is p-summable as a nonnegative real-valued function on X.

In particular, we can take $W = \mathbf{R}$ or $\mathbf{C}$, considered as one-dimensional vector spaces over themselves, with the usual absolute value function as the norm.

One can check that $\ell^p(X, W)$ is a linear subspace of the space of all W-valued functions on X. If $f \in \ell^p(X, W)$, then put

$$\|f\|_p = \|f\|_{\ell^p(X,W)} = \Big(\sum_{x\in X} \|f(x)\|_W^p\Big)^{1/p}. \tag{2.6.3}$$

This defines a norm on $\ell^p(X, W)$ when $p \ge 1$. If $0 < p \le 1$, then this satisfies the usual homogeneity property of a norm, as well as

$$\|f + g\|_p^p \le \|f\|_p^p + \|g\|_p^p \tag{2.6.4}$$

for all $f, g \in \ell^p(X, W)$. This implies that

$$\|f - g\|_p^p \tag{2.6.5}$$

defines a metric on $\ell^p(X, W)$ when $p \le 1$, as usual.

If $0 < p_1 \le p_2 \le +\infty$, then

$$\ell^{p_1}(X, W) \subseteq \ell^{p_2}(X, W), \tag{2.6.6}$$

as in the previous section. We also have that

$$\|f\|_{p_2} \le \|f\|_{p_1} \tag{2.6.7}$$

for all $f \in \ell^{p_1}(X, W)$ in this case.

Let us say that a W-valued function f on X *vanishes at infinity* on X with respect to $\|\cdot\|_W$ on W if

(2.6.8) $\|f(x)\|_W$ vanishes at infinity as a nonnegative real-valued function on X.

Let $c_0(X, W)$ be the space of W-valued functions on X that vanish at infinity on X with respect to $\|\cdot\|_W$ on W. It is easy to see that

$$c_0(X, W) \text{ is a linear subspace of } \ell^\infty(X, W). \tag{2.6.9}$$

More precisely, one can check that

$$c_0(X, W) \text{ is the same as the closure of } c_{00}(X, W) \text{ in } \ell^\infty(X, W) \tag{2.6.10}$$

with respect to the supremum metric, as in Section 1.13.

If $0 < p < \infty$, then

$$c_{00}(X, W) \subseteq \ell^p(X, W) \subseteq c_0(X, W). \tag{2.6.11}$$

One can verify that

$$c_{00}(X, W) \text{ is dense in } \ell^p(X, W) \tag{2.6.12}$$

when $p < \infty$, using (2.4.10). This uses the metric on $\ell^p(X, W)$ associated to (2.6.3) when $p \geq 1$, and the metric $\|f - g\|_p^p$ when $p \leq 1$.

If W is complete with respect to the metric associated to $\|\cdot\|_W$, then

$$\ell^p(X, W) \text{ is complete} \tag{2.6.13}$$

with respect to the metric associated to (2.6.3) when $p \geq 1$, and with respect to the metric $\|f - g\|_p^p$ when $p \leq 1$. To see this, let $\{f_j\}_{j=1}^\infty$ be a Cauchy sequence in $\ell^p(X, W)$ with respect to the appropriate metric. If $x \in X$, then $\{f_j(x)\}_{j=1}^\infty$ is a Cauchy sequence in W, with respect to the metric associated to $\|\cdot\|_W$. This sequence converges in W, by hypothesis. This means that $\{f_j\}_{j=1}^\infty$ converges to a W-valued function f pointwise on X.

It is well known that a Cauchy sequence in any metric space is bounded. One can use this and (2.4.12) to get that $f \in \ell^p(X, W)$. Similarly, one can use the Cauchy condition for $\{f_j\}_{j=1}^\infty$ in $\ell^p(X, W)$ and (2.4.12) to get that the distance from f_j to f is small when j is sufficiently large. This means that $\{f_j\}_{j=1}^\infty$ converges to f with respect to the appropriate metric on $\ell^p(X, W)$.

In particular, $\ell^p(X, \mathbf{R})$ and $\ell^p(X, \mathbf{C})$ are complete with respect to their appropriate metrics.

2.7 Generalized convergence of sums

Let X be a nonempty set, and let W be a vector space over the real or complex numbers with a norm $\|\cdot\|_W$. Also let f be a function on X with values in W. We say that that the sum

$$\sum_{x \in X} f(x) \tag{2.7.1}$$

converges in the generalized sense with respect to the metric on W associated to $\|\cdot\|_W$ if there is a $w \in W$ with the following property: for each $\epsilon > 0$ there is a finite set $A(\epsilon) \subseteq X$ such that for every finite set $A \subseteq X$ with

$$A(\epsilon) \subseteq A, \tag{2.7.2}$$

we have that

$$\Big\|\sum_{x\in A} f(x) - w\Big\|_W < \epsilon. \tag{2.7.3}$$

One can check that such a $w \in W$ is unique when it exists, in which case it is considered to be the value of the sum (2.7.1).

If X has only finitely many elements, then this is the same as treating (2.7.1) as a finite sum. If the support of f in X has only finitely many elements, then this is the same as reducing (2.7.1) to a finite sum in W, as in Section 2.3.

Of course, if A, B are finite subsets of X, then $A \cup B$ is a finite subset of X that contains A and B. This means that the collection of all finite subsets of X is a directed system, as a partially-ordered set with respect to inclusion. The sums

$$\sum_{x\in A} f(x) \tag{2.7.4}$$

of f over finite subsets A of X may be considered as a net in W, indexed by this directed system. The convergence of the sum (2.7.1) is the same as the convergence of this net in W, with respect to the metric associated to $\|\cdot\|_W$.

Let us say that the sum (2.7.1) satisfies *the generalized Cauchy condition* if for every $\epsilon > 0$ there is a finite subset $A_0(\epsilon)$ of X with the following property: if A, B are finite subsets of X such that

$$A_0(\epsilon) \subseteq A, B, \tag{2.7.5}$$

then

$$\Big\|\sum_{x\in A} f(x) - \sum_{x\in B} f(x)\Big\|_W < \epsilon. \tag{2.7.6}$$

This is the same as saying that the net of finite sums (2.7.4) is a Cauchy net with respect to the metric on W associated to $\|\cdot\|_W$. If (2.7.1) converges in the generalized sense, then it is easy to see that it satisfies the generalized Cauchy condition, with

$$A_0(\epsilon) = A(\epsilon/2) \tag{2.7.7}$$

for each $\epsilon > 0$.

Equivalently, (2.7.1) satisfies the generalized Cauchy condition if and only if for every $\epsilon > 0$ there is a finite subset $A_1(\epsilon)$ of X with the following property: if C is a finite subset of X such that

$$A_1(\epsilon) \cap C = \emptyset, \tag{2.7.8}$$

then

$$\Big\|\sum_{x\in C} f(x)\Big\|_W < \epsilon. \tag{2.7.9}$$

More precisely, the previous version implies this one, with

$$A_1(\epsilon) = A_0(\epsilon), \tag{2.7.10}$$

by taking $A = A_0(\epsilon) \cup C$ and $B = A_0(\epsilon)$ in (2.7.6). Conversely, this version implies the previous one, with

$$A_0(\epsilon) = A_1(\epsilon/2) \tag{2.7.11}$$

for each $\epsilon > 0$.

If (2.7.1) satisfies the generalized Cauchy condition, then f vanishes at infinity on X with respect to $\|\cdot\|_W$ on W. Indeed, (2.7.9) implies that

$$\|f(x)\|_W < \epsilon \tag{2.7.12}$$

when $x \in X \setminus A_1(\epsilon)$, by taking $C = \{x\}$. In particular, this means that the support of f has only finitely or countably many elements, as in Section 1.13.

Let $\{x_j\}_{j=1}^\infty$ be a sequence of distinct elements of X, and suppose for the moment that the support of f is contained in the set of x_j's. If (2.7.1) converges in the generalized sense, then it is easy to see that

$$\sum_{j=1}^\infty f(x_j) \tag{2.7.13}$$

converges in W, and that the value of the sum is the same as for (2.7.1).

Similarly, if (2.7.1) satisfies the generalized Cauchy condition, and (2.7.13) converges in W, then one can check that (2.7.1) converges in the generalized sense, and with the same value of the sum.

If (2.7.1) satisfies the generalized Cauchy condition, then the sequence of partial sums

$$\sum_{j=1}^n f(x_j) \tag{2.7.14}$$

of (2.7.13) is a Cauchy sequence in W with respect to the metric associated to $\|\cdot\|_W$. If W is complete with respect to this metric, then (2.7.13) converges in W. This implies that (2.7.1) converges in the generalized sense, with sum equal to (2.7.13), as in the preceding paragraph.

If W is a Banach space, and (2.7.1) satisfies the generalized Cauchy condition, then one can use the remarks in the prevous paragraphs to get that (2.7.1) converges in the generalized sense. Note that if (2.7.1) converges in the generalized sense, then every rearrangement of (2.7.13) converges, and with the same value of the sum. Conversely, if every rearrangement of (2.7.13) converges, then it is well known that (2.7.1) satisfies the generalized Cauchy condition, as in Proposition 1.c.1 on p15 of [131].

If f is any W-valued function on X, then

$$\left\|\sum_{x\in C} f(x)\right\|_W \le \sum_{x\in C} \|f(x)\|_W \tag{2.7.15}$$

for every finite subset C of X, by the triangle inequality. If $f \in \ell^1(X, W)$, then one can use this to get that (2.7.1) satisfies the generalized Cauchy condition.

If (2.7.1) converges in the generalized sense, then we have that

$$\left\| \sum_{x \in X} f(x) \right\|_W \le \sum_{x \in X} \|f(x)\|_W \tag{2.7.16}$$

in this case.

Suppose now that $\|\cdot\|_W$ is the norm associated to an inner product $\langle \cdot, \cdot \rangle_W$ on W, and that the values of f on X are pairwise-orthogonal in W. This means that

$$\langle f(x), f(y) \rangle_W = 0 \tag{2.7.17}$$

for every $x, y \in X$ with $x \ne y$. If C is a finite subset of X, then it follows that

$$\left\| \sum_{x \in C} f(x) \right\|_W^2 = \sum_{x \in C} \|f(x)\|_W^2. \tag{2.7.18}$$

If $f \in \ell^2(X, W)$, then one can use this to get that (2.7.1) satisfies the generalized Cauchy condition. If (2.7.1) converges in the generalized sense, then one can verify that

$$\left\| \sum_{x \in X} f(x) \right\|_W^2 = \sum_{x \in X} \|f(x)\|_W^2 \tag{2.7.19}$$

under these conditions.

2.8 More on generalized convergence

Let X be a nonempty set again, and let W be a vector space over the real or complex numbers with a norm $\|\cdot\|_W$. Consider the space

$$Sum(X, W) \tag{2.8.1}$$

of W-valued functions f on X such that the sum (2.7.1) converges in the generalized sense. It is easy to see that this is a linear subspace of the space of all W-valued functions on X, and that

$$f \mapsto \sum_{x \in X} f(x) \tag{2.8.2}$$

is a linear mapping from $Sum(X, W)$ into W.

Similarly, let

$$GCC(X, W) \tag{2.8.3}$$

be the space of W-valued functions f on X such that (2.7.1) satisfies the generalized Cauchy condition. This is also a linear subspace of the space of all W-valued functions on X, with

$$c_{00}(X, W) \subseteq Sum(X, W) \subseteq GCC(X, W) \subseteq c_0(X, W), \tag{2.8.4}$$

as in the previous section. If W is a Banach space, then

$$Sum(X, W) = GCC(X, W), \tag{2.8.5}$$

as before.

Remember that

$$\ell^1(X, W) \subseteq GCC(X, W), \tag{2.8.6}$$

so that

$$\ell^1(X, W) \subseteq Sum(X, W) \tag{2.8.7}$$

when W is a Banach space. In this case, (2.8.2) is a bounded linear mapping from $\ell^1(X, W)$ into W, with operator norm equal to 1.

Let f be a nonnegative real-valued function on X. If f is summable on X, then one can check that the sum (2.7.1) converges in the generalized sense, with the same value of the sum as in Section 2.4.

Similarly, if f is a real or complex-valued function on X such that $|f|$ is summable as a nonnegative real-valued function on X, then (2.7.1) converges in the generalized sense. This can be seen by expressing f as a linear combination of summable nonnegative real-valued functions on X, and using the remark in the preceding paragraph. Alternatively, this follows from (2.8.7), with $W = \mathbf{R}$ or $\mathbf{C}$, as appropriate.

Let f, g be real or complex-valued functions on X such that $|f|$, $|g|$ are 2-summable on X. One can check that $|f|\,|g|$ is summable on X. More precisely, one can use the Cauchy–Schwarz inequality for finite sums to get that

$$\sum_{x \in X} |f(x)|\,|g(x)| \le \Big(\sum_{x \in X} |f(x)|^2\Big)^{1/2} \Big(\sum_{x \in X} |g(x)|^2\Big)^{1/2}. \tag{2.8.8}$$

This could also be obtained more directly, using the same type of arguments as for finite sums.

Suppose that $\|\cdot\|_W$ is the norm associated to an inner product $\langle\cdot,\cdot\rangle_W$ on W. If f, g are W-valued functions on X, then

$$|\langle f(x), g(x)\rangle_W| \le \|f(x)\|_W\,\|g(x)\|_W \tag{2.8.9}$$

for every $x \in X$, by the Cauchy–Schwarz inequality. If $f, g \in \ell^2(X, W)$, then we get that

$$|\langle f(x), g(x)\rangle_W| \tag{2.8.10}$$

is summable as a nonnegative real-valued function on X, as in the preceding paragraph. In this case, we put

$$\langle f, g\rangle_{\ell^2(X,W)} = \sum_{x \in X} \langle f(x), g(x)\rangle_W. \tag{2.8.11}$$

where the right side converges in the generalized sense, as before. One can check that this defines an inner product on $\ell^2(X, W)$, for which the corresponding norm is the usual norm on $\ell^2(X, W)$ associated to $\|\cdot\|_W$. If W is a Hilbert space, then $\ell^2(X, W)$ is complete with respect to the metric associatived to this norm, as in Section 2.6. This means that $\ell^2(X, W)$ is a Hilbert space under these conditions.

Of course, the standard absolute value functions on $\mathbf{R}$ and $\mathbf{C}$ may be considered as norms associated to inner products. These are the same as the standard inner products on $\mathbf{R}^n$ and $\mathbf{C}^n$, respectively, as in Section 1.9, with $n = 1$. This leads to inner products

$$\langle f, g\rangle_{\ell^2(X,\mathbf{R})} = \sum_{x \in X} f(x)\, g(x) \tag{2.8.12}$$

on $\ell^2(X, \mathbf{R})$ and

$$\langle f, g\rangle_{\ell^2(X,\mathbf{C})} = \sum_{x \in X} f(x)\, \overline{g(x)} \tag{2.8.13}$$

on $\ell^2(X, \mathbf{C})$ as in the preceding paragraph. Thus $\ell^2(X, \mathbf{R})$ and $\ell^2(X, \mathbf{C})$ are Hilbert spaces over the real and complex numbers, respectively, as before.

2.9 Bounded finite sums

Let X be a nonempty set, and let W be a vector space over the real or complex numbers with a norm $\|\cdot\|_W$. Let us say that a W-valued function f on X has *bounded finite sums* on X with respect to $\|\cdot\|_W$ if the norms

$$\Big\|\sum_{x \in A} f(x)\Big\|_W \tag{2.9.1}$$

of the sums of f over all finite subsets A of X are bounded. It is easy to see that the space

$$BFS(X, W) \tag{2.9.2}$$

of W-valued functions on X with bounded finite sums is a linear subspace of the space of all W-valued functions on X.

Let us check that

$$GCC(X, W) \subseteq BFS(X, W) \subseteq \ell^\infty(X, W). \tag{2.9.3}$$

If $f \in GCC(X, W)$, then there is a finite set $A_1(1) \subseteq X$ such that

$$\Big\|\sum_{x \in C} f(x)\Big\|_W < 1 \tag{2.9.4}$$

for all finite sets $C \subseteq X$ such that $A_1(1) \cap C = \emptyset$, as in (2.7.9). If A is a finite subset of X, then we can take $C = A \setminus A_1(1)$ in (2.9.4), and use the triangle inequality to get that

$$\begin{aligned} \Big\|\sum_{x \in A} f(x)\Big\|_W &\le \Big\|\sum_{x \in A \cap A_1(1)} f(x)\Big\|_W + \Big\|\sum_{x \in A \setminus A_1(1)} f(x)\Big\|_W \\ &< \sum_{x \in A_1(1)} \|f(x)\|_W + 1. \end{aligned} \tag{2.9.5}$$

This shows that $f \in BFS(X, W)$. The second inclusion in (2.9.3) corresponds to taking A to be a subset of X with only one element in (2.9.1).

If $f \in BFS(X, W)$, then put

$$\|f\|_{BFS} = \|f\|_{BFS(X,W)} = \sup\Big\{\Big\|\sum_{x \in A} f(x)\Big\|_W : A \text{ is a finite subset of } X\Big\}. \tag{2.9.6}$$

This defines a norm on $BFS(X, W)$, with

$$\|f\|_\infty \le \|f\|_{BFS} \tag{2.9.7}$$

for every $f \in BFS(X, W)$. If $f \in \ell^1(X, W)$, then $f \in BFS(X, W)$, with

$$\|f\|_{BFS} \le \|f\|_1, \tag{2.9.8}$$

by the triangle inequality. If $f \in Sum(X, W)$, then $f \in BFS(X, W)$, by (2.8.4) and (2.9.3), and one can check that

$$\Big\|\sum_{x \in X} f(x)\Big\|_W \le \|f\|_{BFS}. \tag{2.9.9}$$

One can also verify that

$$GCC(X, W) \text{ is the same as the closure of } c_{00}(X, W) \text{ in } BFS(X, W). \tag{2.9.10}$$

Let $\{f_j\}_{j=1}^\infty$ be a sequence of W-valued functions on X that converges pointwise to a W-valued function f on X, with respect to the metric on W associated to $\|\cdot\|_W$. If A is a finite subset of X, then it follows that

$$\lim_{j \to \infty} \sum_{x \in A} f_j(x) = \sum_{x \in A} f(x). \tag{2.9.11}$$

Suppose that $f_j \in BFS(X, W)$ for each j, and that there is a nonnegative real number C such that

$$\|f_j\|_{BFS} \le C \tag{2.9.12}$$

for all j. This implies that $f \in BFS(X, W)$, with

$$\|f\|_{BFS} \le C, \tag{2.9.13}$$

because of (2.9.11).

If W is complete with respect to the metric associated to $\|\cdot\|_W$, then

$$BFS(X, W) \text{ is complete} \tag{2.9.14}$$

with respect to the metric associated to (2.9.6). Indeed, if $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $BFS(X, W)$, then $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to the supremum metric, and in particular $\{f_j\}_{j=1}^\infty$ converges pointwise to a W-valued function f on X. One can check that f is an element of $BFS(X, W)$, because Cauchy sequences are bounded. One can also verify that $\{f_j\}_{j=1}^\infty$ converges to f with respect to the metric associated to (2.9.6), using the fact that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to this metric again.

Suppose for the moment that $W = \mathbf{R}$, as a one-dimensional vector space over itself, with the usual absolute value function as the norm. If $f \in BFS(X, \mathbf{R})$, then $\max(f, 0)$ and $\max(-f, 0)$ are summable on X, with

$$\sum_{x \in X} \max(f(x), 0), \ \sum_{x \in X} \max(-f(x), 0) \le \|f\|_{BFS}. \tag{2.9.15}$$

This implies that $f \in \ell^1(X, \mathbf{R})$, with

$$\|f\|_1 \le 2 \, \|f\|_{BFS}. \tag{2.9.16}$$

Thus

$$BFS(X, \mathbf{R}) = \ell^1(X, \mathbf{R}), \tag{2.9.17}$$

as linear subspaces of the space of all real-valued functions on X.

Suppose now that $W = \mathbf{C}$, as a one-dimensional vector space over itself, and with the usual absolute value function as the norm. If $f \in BFS(X, \mathbf{C})$, then the real and imaginary parts of f are in $BFS(X, \mathbf{R})$, with BFS norms less than or equal to the BFS norm of f. This implies that $f \in \ell^1(X, \mathbf{C})$, with

$$\|f\|_1 \le 4 \, \|f\|_{BFS}, \tag{2.9.18}$$

using (2.9.16) for the real and imaginary parts of f. It follows that

$$BFS(X, \mathbf{C}) = \ell^1(X, \mathbf{C}), \tag{2.9.19}$$

as linear subspaces of the space of all complex-valued functions on X.

Let W be a vector space over the real or complex numbers again, and let $\langle \cdot, \cdot \rangle_W$ be an inner product on W, with associated norm $\|\cdot\|_W$. Also let f be a W-valued function on X with pairwise-orthogonal values, so that $\langle f(x), f(y) \rangle_W = 0$ for every $x, y \in X$ with $x \ne y$. Under these conditions,

$$f \in BFS(X, W) \text{ if and only if } f \in \ell^2(X, W), \tag{2.9.20}$$

because of (2.7.18). In this case, we have that

$$\|f\|_{BFS} = \|f\|_2. \tag{2.9.21}$$

2.10 Some related linear mappings

Let X be a nonempty set, and let W, Z be vector spaces, both real or both complex. Also let T be a linear mapping from W into Z. If f is a W-valued function on X, then $T(f(x))$ is a Z-valued function on X. More precisely,

$$f \mapsto T \circ f \tag{2.10.1}$$

is a linear mapping from the space of all W-valued functions on X into the space of all Z-valued functions on X.

Suppose that $\|\cdot\|_W$, $\|\cdot\|_Z$ are norms on W, Z, respectively, and that T is bounded with respect to these norms. If f is a W-valued function on X, then

$$\|T(f(x))\|_Z \le \|T\|_{op,WZ}\,\|f(x)\|_W \tag{2.10.2}$$

for every $x \in X$. If $f \in \ell^p(X,W)$ for some $0 < p \le +\infty$, then it follows that

$$T \circ f \in \ell^p(X,Z), \tag{2.10.3}$$

with

$$\|T \circ f\|_{\ell^p(X,Z)} \le \|T\|_{op,WZ}\,\|f\|_{\ell^p(X,W)}. \tag{2.10.4}$$

Similarly, if $f \in c_0(X,W)$, then

$$T \circ f \in c_0(X,Z). \tag{2.10.5}$$

If $f \in BFS(X,W)$, then it is easy to see that

$$T \circ f \in BFS(X,Z), \tag{2.10.6}$$

with

$$\|T \circ f\|_{BFS(X,Z)} \le \|T\|_{op,WZ}\,\|f\|_{BFS(X,W)}. \tag{2.10.7}$$

If $f \in GCC(X,W)$, then one can check that

$$T \circ f \in GCC(X,Z). \tag{2.10.8}$$

If $f \in Sum(X,W)$, then one can verify that

$$T \circ f \in Sum(X,Z), \tag{2.10.9}$$

with

$$\sum_{x \in X} T(f(x)) = T\Big(\sum_{x \in X} f(x)\Big). \tag{2.10.10}$$

If $\sum_{j=1}^\infty w_j$ is a convergent series in W, then

$$\sum_{j=1}^\infty T(w_j) \text{ is a convergent series in } Z, \tag{2.10.11}$$

with

$$\sum_{j=1}^\infty T(w_j) = T\Big(\sum_{j=1}^\infty w_j\Big). \tag{2.10.12}$$

If $\sum_{j=1}^\infty w_j$ converges absolutely with respect to $\|\cdot\|_W$, then

$$\sum_{j=1}^\infty T(w_j) \text{ converges absolutely with respect to } \|\cdot\|_Z. \tag{2.10.13}$$

This corresponds to the fact that (2.10.1) maps $\ell^1(\mathbf{Z}_+, W)$ into $\ell^1(\mathbf{Z}_+, Z)$, as before.

If

$$\|T(w)\|_Z = \|w\|_W \tag{2.10.14}$$

for every $w \in W$, then T is said to be an *isometric linear mapping* from W into Z. This implies that T is an isometry from W into Z, with respect to the metrics associated to the norms. In particular, this means that T is injective.

Suppose now that $\|\cdot\|_W$, $\|\cdot\|_Z$ are associated to inner products $\langle\cdot,\cdot\rangle_W$, $\langle\cdot,\cdot\rangle_Z$ on W, Z, respectively. If

$$\langle T(u), T(w)\rangle_Z = \langle u, w\rangle_W \tag{2.10.15}$$

for every $u, w \in W$, then (2.10.14) holds for every $w \in W$, by taking $u = w$ in (2.10.15).

Conversely, if (2.10.14) holds for every $w \in W$, then it is well known that (2.10.15) holds for every $u, w \in W$. This can be obtained from suitable *polarization identities*, which can be used to express an inner product in terms of the corresponding norm. If $u, w \in W$, then

$$(1/2)\,(\|u + w\|_W^2 - \|u\|_W^2 - \|w\|_W^2) \tag{2.10.16}$$

is equal to $\langle u, w\rangle_W$ in the real case. In the complex case, this is equal to the real part of the inner product. To get the imaginary part, one can multiply u by $-i$.

If T maps W onto Z and satisfies (2.10.15) for every $u, w \in W$, then T is said to be a *unitary mapping* from W onto Z. In the real case, one may say that T is an *orthogonal transformation*.

2.11 Distances to sets

Let (X, d) be a metric space, and let A be a nonempty subset of X. If $x \in X$, then put

$$\operatorname{dist}(x, A) = \inf_{a \in A} d(x, a). \tag{2.11.1}$$

If $y \in X$ and $a \in A$, then

$$\operatorname{dist}(x, A) \le d(x, a) \le d(x, y) + d(y, a). \tag{2.11.2}$$

This means that

$$\operatorname{dist}(x, A) - d(x, y) \le d(y, a), \tag{2.11.3}$$

so that

$$\operatorname{dist}(x, A) - d(x, y) \le \operatorname{dist}(y, A). \tag{2.11.4}$$

It follows that $\operatorname{dist}(x, A)$ is Lipschitz with constant 1 as a real-valued function on X, as in Section 2.1.

It is easy to see that

$$\operatorname{dist}(x, A) = 0 \tag{2.11.5}$$

if and only if x is an element of the closure $\overline{A}$ of A in X. If $A \subseteq B \subseteq X$, then

$$\mathrm{dist}(x, B) \le \mathrm{dist}(x, A) \tag{2.11.6}$$

for every $x \in X$. One can check that

$$\mathrm{dist}(x, A) = \mathrm{dist}(x, \overline{A}) \tag{2.11.7}$$

for every $x \in A$.

Let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$, and let V_0 be a linear subspace of V. If $v \in V$, then put

$$N_{V_0}(v) = \inf_{w \in V_0} \|v - w\|_V. \tag{2.11.8}$$

Equivalently,

$$N_{V_0}(v) = \mathrm{dist}(v, V_0), \tag{2.11.9}$$

where the right side is defined as in (2.11.1), using the metric on V associated to $\|\cdot\|_V$. Thus

$$N_{V_0}(v) = 0 \tag{2.11.10}$$

if and only if v is an element of the closure $\overline{V_0}$ of V_0 in V, as in the preceding paragraph. Similarly,

$$N_{V_0} = N_{\overline{V_0}} \tag{2.11.11}$$

on V, as in (2.11.7).

One can check that

$$N_{V_0}(t\, v) = |t|\, N_{V_0}(v) \tag{2.11.12}$$

for all $v \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. If $v_1, v_2 \in V$ and $w_1, w_2 \in V_0$, then

$$\begin{aligned} N_{V_0}(v_1 + v_2) &\le \|(v_1 + v_2) - (w_1 + w_2)\|_V \\ &\le \|v_1 - w_1\|_V + \|v_2 - w_2\|_V. \end{aligned} \tag{2.11.13}$$

One can use this to verify that

$$N_{V_0}(v_1 + v_2) \le N_{V_0}(v_1) + N_{V_0}(v_2). \tag{2.11.14}$$

Thus N_{V_0} is a seminorm on V.

Note that

$$N_{V_0}(v) \le \|v\|_V \tag{2.11.15}$$

for every $v \in V$. This implies that

$$|N_{V_0}(v) - N_{V_0}(u)| \le N_{V_0}(u - v) \le \|u - v\|_V \tag{2.11.16}$$

for all $u, v \in V_0$, where the first step is as in Section 1.2. This is another way to look at the Lipschitz condition for N_{V_0} as a real-valued function on V, with respect to the metric associated to $\|\cdot\|_V$.

2.12 Distances and orthogonality

Let V be a vector space over the real or complex numbers with an inner product $\langle \cdot, \cdot \rangle_V$ and the associated norm $\| \cdot \|_V$. Also let W be a linear subspace of V, and let $v \in V$ be given. We shall sometimes be interested in trying to get a $u \in W$ such that

$$\langle v - u, w \rangle_V = 0 \tag{2.12.1}$$

for every $w \in W$. Let us first check that such an element of W is unique. Suppose that $u' \in W$ satisfies

$$\langle v - u', w \rangle_V = 0 \tag{2.12.2}$$

for every $w \in W$ as well. This implies that

$$\langle u - u', w \rangle_V = 0 \tag{2.12.3}$$

for every $w \in W$. It follows that

$$u = u', \tag{2.12.4}$$

because we can take $w = u - u'$ in (2.12.3).

If $u \in W$ satisfies (2.12.1), then we can take $w = u$ to get that

$$\|v\|_V^2 = \|v - u\|_V^2 + \|u\|_V^2. \tag{2.12.5}$$

Similarly, if $y \in W$, then

$$\|v - y\|_V^2 = \|(v - u) + (u - y)\|_V^2 = \|v - u\|_V^2 + \|u - y\|_V^2, \tag{2.12.6}$$

because of (2.12.1), with $w = u - y$. In particular,

$$\|v - u\|_V \le \|v - y\|_V, \tag{2.12.7}$$

with equality if and only if $u = y$. This implies that

$$\mathrm{dist}(v, W) = \|v - u\|_V, \tag{2.12.8}$$

where the left side is as in the previous section, using the metric on V associated to $\| \cdot \|_V$. Note that

$$\|u\|_V \le \|v\|_V, \tag{2.12.9}$$

with equality if and only if $u = v$, by (2.12.5).

Suppose that $v_1, \dots, v_n$ are finitely many *orthonormal vectors* in V, so that

$$\|v_j\|_V = 1 \tag{2.12.10}$$

for each $j = 1, \dots, n$, and

$$\langle v_j, v_l \rangle_V = 0 \tag{2.12.11}$$

when $j \neq l$. Let W be the linear subspace of V spanned by $v_1, \ldots, v_n$. If $v \in V$, then put

$$P_W(v) = \sum_{j=1}^{n} \langle v, v_j \rangle_V \, v_j. \tag{2.12.12}$$

This defines a linear mapping from V into W. This is known as the *orthogonal projection of V onto W*, because of the properties mentioned in the next two paragraphs.

It is easy to see that

$$P_W(v_l) = v_l \tag{2.12.13}$$

for each $l = 1, \ldots, n$. This implies that

$$P_W(v) = v \tag{2.12.14}$$

for every $v \in W$, so that P_W maps V onto W in particular. If $v \in V$, then

$$\langle P_W(v), v_l \rangle_V = \langle v, v_l \rangle_V \tag{2.12.15}$$

for each $l = 1, \ldots, n$. This means that

$$\langle v - P_W(v), v_l \rangle_V = 0 \tag{2.12.16}$$

for each $l = 1, \ldots, n$, so that

$$\langle v - P_W(v), w \rangle_V = 0 \tag{2.12.17}$$

for every $w \in W$. Note that $P_W(v) = 0$ if and only if $\langle v, w \rangle_V = 0$ for every $w \in W$, by the definition (2.12.12) of $P_W(v)$ and (2.12.15).

Thus $u = P_W(v) \in W$ satisfies (2.12.1), and is uniquely determined by these properties, as before. It follows that $P_W(v)$ does not depend on the particular orthonormal vectors $v_1, \ldots, v_n$ spanning W. Observe that

$$\|P_W(v)\|_V^2 = \sum_{j=1}^{n} |\langle v, v_j \rangle_V|^2. \tag{2.12.18}$$

Combining this with (2.12.5), we get that

$$\|v\|_V^2 = \|v - P_W(v)\|_V^2 + \sum_{j=1}^{n} |\langle v, v_j \rangle_V|^2. \tag{2.12.19}$$

We also have that

$$\operatorname{dist}(v, W) = \|v - P_W(v)\|_V, \tag{2.12.20}$$

as in (2.12.8).

2.13 Orthonormal families of vectors

Let V be a vector space over the real or complex numbers with an inner product $\langle \cdot, \cdot \rangle_V$ and the associated norm $\| \cdot \|_V$ again. Also let A be a nonempty set, and let $\{v_\alpha\}_{\alpha \in A}$ be an *orthonormal family of vectors* in V indexed by A. This means that

$$\|v_\alpha\|_V = 1 \tag{2.13.1}$$

for all $\alpha \in A$, and that

$$\langle v_\alpha, v_\beta \rangle_V = 0 \tag{2.13.2}$$

for all $\alpha, \beta \in A$ with $\alpha \neq \beta$.

If E is a nonempty subset of A, then

$$\begin{gathered} \text{let } W_E \text{ be the linear subspace of } V \\ \text{spanned by the } v_\alpha\text{'s with } \alpha \in E. \end{gathered} \tag{2.13.3}$$

This may be interpreted as being $\{0\}$ when $E = \emptyset$. If $E_1 \subseteq E_2 \subseteq A$, then

$$W_{E_1} \subseteq W_{E_2}. \tag{2.13.4}$$

If E is a nonempty finite subset of A, then let

$$P_{W_E}(v) = \sum_{\alpha \in E} \langle v, v_\alpha \rangle_V \, v_\alpha \tag{2.13.5}$$

be the corresponding orthogonal projection of V onto W_E, as in the previous section. Thus

$$\|v\|_V^2 = \|v - P_{W_E}(v)\|_V^2 + \sum_{\alpha \in E} |\langle v, v_\alpha \rangle_V|^2 \tag{2.13.6}$$

and

$$\operatorname{dist}(v, W_E) = \|v - P_{W_E}(v)\|_V \tag{2.13.7}$$

for every $v \in V$, as before. If $E_1 \subseteq E_2$ are nonempty finite subsets of A, then

$$\|v - P_{W_{E_2}}(v)\|_V = \operatorname{dist}(v, W_{E_2}) \leq \operatorname{dist}(v, W_{E_1}) = \|v - P_{W_{E_1}}(v)\|_V, \tag{2.13.8}$$

because of (2.13.4). This also corresponds to the fact that the sum on the right side of (2.13.6) can only increase as E gets larger.

If $v \in V$, then put

$$f_v(\alpha) = \langle v, v_\alpha \rangle_V \tag{2.13.9}$$

for every $\alpha \in A$. This defines a real or complex-valued function on A, as appropriate. If E is a nonempty finite subset of A, then

$$\sum_{\alpha \in E} |f_v(\alpha)|^2 \leq \|v\|_V^2, \tag{2.13.10}$$

by (2.13.6). This means that $f_v \in \ell^2(A, \mathbf{R})$ or $\ell^2(A, \mathbf{C})$, as appropriate, with

$$\|f_v\|_2^2 = \sum_{\alpha \in A} |f_v(\alpha)|^2 \leq \|v\|_V^2. \tag{2.13.11}$$

Note that

$$v \mapsto f_v \tag{2.13.12}$$

is a linear mapping from V into $\ell^2(A, \mathbf{R})$ or $\ell^2(A, \mathbf{C})$, as appropriate.

Every element of W_A is an element of W_E for some finite set $E \subseteq A$. Using this, one can check that

$$\begin{aligned} \operatorname{dist}(v, W_A) &= \inf\{\operatorname{dist}(v, W_E) : E \subseteq A,\ E \neq \emptyset,\ E \text{ has only} \\ &\qquad \text{finitely many elements}\}. \end{aligned} \tag{2.13.13}$$

Note that

$$\|v\|_V^2 = \operatorname{dist}(v, W_E)^2 + \sum_{\alpha \in E} |\langle v, v_\alpha \rangle_V|^2 \tag{2.13.14}$$

for every nonempty finite subset E of A, by (2.13.6) and (2.13.7). One can use (2.13.13) and (2.13.14) to get that

$$\|v\|_V^2 = \operatorname{dist}(v, W_A)^2 + \|f_v\|_2^2. \tag{2.13.15}$$

Let

$$W \text{ be the closure } \overline{W_A} \text{ of } W_A \text{ in } V, \tag{2.13.16}$$

with respect to the metric associated to $\|\cdot\|_V$. It follows from (2.13.15) that $v \in W$ if and only if

$$\|v\|_V = \|f_v\|_2. \tag{2.13.17}$$

Observe that

$$\sum_{\alpha \in A} f_v(\alpha)\, v_\alpha \tag{2.13.18}$$

satisfies the generalized Cauchy condition, as in Section 2.7. If E is a nonempty finite subset of A, then

$$\sum_{\alpha \in E} f_v(\alpha)\, v_\alpha = P_{W_E}(v), \tag{2.13.19}$$

by (2.13.5) and (2.13.9). If $v \in W_A$, then (2.13.9) is equal to 0 for all but finitely many $\alpha \in A$, so that (2.13.18) reduces to a finite sum, which is equal to v. If $v \in W$, then one can check that (2.13.18) converges in the generalized sense, with

$$\sum_{\alpha \in A} f_v(\alpha)\, v_\alpha = v, \tag{2.13.20}$$

using (2.13.7), (2.13.8), and (2.13.19).

Suppose for the rest of the section that V is a Hilbert space. This implies that (2.13.18) converges in V in the generalized sense for every $v \in V$, as in Section 2.7. More precisely, the sum is an element of W, and we put

$$P_W(v) = \sum_{\alpha \in A} f_v(\alpha)\, v_\alpha \tag{2.13.21}$$

for each $v \in V$. This defines a linear mapping from V into W, with

$$P_W(v) = v \tag{2.13.22}$$

for every $v \in W$, as in (2.13.20). Of course, this is the same as in the previous section when A has only finitely many elements.

If $v \in V$ and $\beta \in A$, then it is easy to see that

$$\langle P_W(v), v_\beta \rangle_V = f_v(\beta) = \langle v, v_\beta \rangle_V, \tag{2.13.23}$$

using a remark in Section 2.10. Equivalently,

$$\langle v - P_W(v), v_\beta \rangle_V = 0 \tag{2.13.24}$$

for every $\beta \in A$. This implies that

$$\langle v - P_W(v), w \rangle_V = 0 \tag{2.13.25}$$

for every $w \in W_A$. It follows that this holds for every $w \in W$.

Note that $P_W(v) \in W$ is uniquely determined by (2.13.25), as in the previous section. Thus $P_W(v)$ depends only on v and W, and not on the particular orthonormal family of vectors whose closed linear span is W. As before, P_W is known as the *orthogonal projection of V onto W*.

2.14 Orthonormal bases

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a Hilbert space over the real or complex numbers, with the corresponding norm $\| \cdot \|_V$. A collection of orthonormal vectors in V is said to be an *orthonormal basis* for V if

$$\text{its linear span is dense in } V, \tag{2.14.1}$$

with respect to the metric associated to $\| \cdot \|_V$.

If V has finite dimension as a vector space over $\mathbf{R}$ or $\mathbf{C}$, then V is spanned by finitely many vectors. In this case, one can use the Gram–Schmidt process to get finitely many orthonormal vectors in V that span V.

Similarly, for any sequence of vectors in V, one can use the Gram–Schmidt process to get a finite or infinite sequence of orthonormal vectors in V with the same linear span. If the linear span of the initial sequence is dense in V, then we get an orthonormal basis for V.

Let A be a nonempty set, let $\{v_\alpha\}_{\alpha \in A}$ be an orthonormal family of vectors in V indexed by A, and let W be the closure of the linear span of the v_α's, $\alpha \in A$, as in the previous section. Suppose that

$$W \neq V, \tag{2.14.2}$$

and let u be an element of V not in W. If P_W is the orthogonal projection of V onto W, as in the previous section, then

$$u \neq P_W(u), \tag{2.14.3}$$

because $P_W(u) \in W$. Put

$$y = \|u - P_W(u)\|_V^{-1} \, (u - P_W(u)), \tag{2.14.4}$$

so that $\|y\|_V = 1$ and

$$\langle y, w \rangle_V = 0 \tag{2.14.5}$$

for every $w \in W$, by construction. This implies that the collection of v_α's, $\alpha \in A$, together with y, is an orthonormal collection of vectors in V.

The remarks in the preceding paragraph imply that if a collection of orthonormal vectors in V is maximal with respect to inclusion, then it is an orthonormal basis for V. There are well-known arguments to get a maximal orthonormal collection of vectors using the axiom of choice, based on Zorn's lemma or Hausdorff's maximality principle.

Suppose now that $\{v_\alpha\}_{\alpha \in A}$ is an orthonormal basis for V. If $f \in \ell^2(A, \mathbf{R})$ or $\ell^2(A, \mathbf{C})$, as appropriate, then

$$\sum_{\alpha \in A} f(\alpha) \, v_\alpha \tag{2.14.6}$$

converges in the generalized sense in V, as in Section 2.7. Put

$$T(f) = \sum_{\alpha \in A} f(\alpha) \, v_\alpha, \tag{2.14.7}$$

which defines a linear mapping from $\ell^2(A, \mathbf{R})$ or $\ell^2(A, \mathbf{C})$, as appropriate, into V. More precisely,

$$\begin{aligned} &\text{this defines an isometric linear mapping from} \\ &\ell^2(A, \mathbf{R}) \text{ or } \ell^2(A, \mathbf{C}), \text{ as appropriate, into } V, \end{aligned} \tag{2.14.8}$$

as in Section 2.7 again. This implies that T preserves the appropriate inner products, as in Section 2.10, which could also be verified more directly in this case.

In fact,

$$T \text{ maps } \ell^2(A, \mathbf{R}) \text{ or } \ell^2(A, \mathbf{C}), \text{ as appropriate, onto } V \tag{2.14.9}$$

under these conditions. This follows from the remarks in the previous section, because $W = V$ in this case. Alternatively, one can observe that

$$\text{the image of } T \text{ is a closed set in } V, \tag{2.14.10}$$

because of the completeness of $\ell^2(A, \mathbf{R})$ or $\ell^2(A, \mathbf{C})$, as appropriate.

If V_0 is a closed linear subspace of V, then V_0 may be considered as a Hilbert space, using the restriction of $\langle \cdot, \cdot \rangle_V$ to V_0. One can use an orthonormal basis for V_0 to get the orthogonal projection of V onto V_0, as in the previous section.

2.15 Minimizing distances

Let $(V, \langle \cdot, \cdot \rangle_V)$ be an inner product space over the real or complex numbers, with the corresponding norm $\| \cdot \|_V$. If $v, w \in V$, then it is easy to see that

$$\|v + w\|_V^2 + \|v - w\|_V^2 = 2\,\|v\|_V^2 + 2\,\|w\|_V^2. \tag{2.15.1}$$

This is known as the *parallelogram law.*

Let E be a nonempty subset of V, and let $y \in V$ be given. Also let $\{z_j\}_{j=1}^\infty$ be a sequence of elements of E such that

$$\lim_{j \to \infty} \|y - z_j\|_V = \operatorname{dist}(y, E), \tag{2.15.2}$$

where the distance from y to E is defined using the metric on V associated to $\| \cdot \|_V$. Using the parallelogram law, we get that

$$\|(y - z_j) + (y - z_l)\|_V^2 + \|z_j - z_l\|_V^2 = 2\,\|y - z_j\|_V^2 + 2\,\|y - z_l\|_V^2 \tag{2.15.3}$$

for all $j, l \geq 1$.

Suppose now that E is convex, so that $(z_j + z_l)/2 \in E$ for all j, l. This implies that

$$\|(y - z_j) + (y - z_l)\|_V = 2\,\|y - ((z_j + z_l)/2)\|_V \geq 2 \operatorname{dist}(y, E) \tag{2.15.4}$$

for all j, l. It follows that $\{z_j\}_{j=1}^\infty$ is a Cauchy sequence in V with respect to the metric associated to $\| \cdot \|_V$.

Suppose that V is a Hilbert space, and that E is a closed set in V, with respect to the metric associated to $\| \cdot \|_V$. Under these conditions, $\{z_j\}_{j=1}^\infty$ converges to an element z of E, and

$$\|y - z\|_V = \operatorname{dist}(y, E). \tag{2.15.5}$$

Suppose that E is a closed linear subspace of V, so that

$$\|y - z + w\|_V \geq \operatorname{dist}(y, E) \tag{2.15.6}$$

for every $w \in E$. Equivalently, this means that

$$\|y - z + t\,w\|_V^2 \geq \|y - z\|_V^2 \tag{2.15.7}$$

for every $w \in E$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. One can use this to get that

$$\langle y - z, w \rangle_V = 0 \tag{2.15.8}$$

for every $w \in E$.

Remember that $z \in E$ is uniquely determined by (2.15.8), as in Section 2.12. Thus we get a mapping P_E from V into E which sends $y \in V$ to the corresponding $z \in E$. It is easy to see that P_E is a linear mapping, because of this characterization of z. Of course, $z = y$ when $y \in E$. This is another way to get the orthogonal projection of V onto a closed linear subspace E.

If V is any inner product space and V_0 is a linear subspace of V, then put

$$V_0^\perp = \{u \in V : \langle u, w\rangle_V = 0 \text{ for every } w \in V_0\}. \tag{2.15.9}$$

This is called the *orthogonal complement* of V_0 in V, and it is a closed linear subspace of V. It is easy to see that

$$V_0 \cap V_0^\perp = \{0\}. \tag{2.15.10}$$

We also have that

$$(\overline{V_0})^\perp = V_0^\perp, \tag{2.15.11}$$

where $\overline{V_0}$ is the closure of V_0 in V with respect to the metric associated to $\|\cdot\|_V$. The orthogonal complement of $V_0^\perp$ in V can be defined in the same way, and we have that

$$V_0 \subseteq (V_0^\perp)^\perp. \tag{2.15.12}$$

If V is a Hilbert space and V_0 is a closed linear subspace of V, then

$$\begin{aligned}&V_0^\perp \text{ is the same as the kernel}\\ &\text{of the orthogonal projection } P_{V_0} \text{ of } V \text{ onto } V_0.\end{aligned} \tag{2.15.13}$$

In this case,

$$\begin{aligned}&\text{every element of } V \text{ can be expressed}\\ &\text{as a sum of elements of } V_0 \text{ and } V_0^\perp \text{ in a unique way,}\end{aligned} \tag{2.15.14}$$

as before. One can use this to check that

$$(V_0^\perp)^\perp = V_0. \tag{2.15.15}$$

More precisely, let $v \in (V_0^\perp)^\perp$ be given, and let us express v as

$$v = w + u, \tag{2.15.16}$$

with $w \in V_0$ and $u \in V_0^\perp$. This implies that

$$v - w = u \tag{2.15.17}$$

is an element of $V_0^\perp$ and $(V_0^\perp)^\perp$, and is thus equal to 0, as in (2.15.10).

Chapter 3

Bounded linear functionals

3.1 Dual spaces

Let V a vector space over the real or complex numbers. A *linear functional* on V is a linear mapping from V into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, considered as a one-dimensional vector space over itself. We shall call the space V^{alg} of all linear functionals on V the *algebraic dual* of V. This is a linear subspace of the space of all real or complex-valued functions on V, as appropriate.

Let $\|\cdot\|_V$ be a norm on V. A *bounded linear functional* is a linear functional on V that is bounded as a linear mapping, using the standard absolute value function on $\mathbf{R}$ or $\mathbf{C}$ as the norm, as appropriate. Let V' be the *dual space* of all bounded linear functionals on V, which is a linear subspace of the algebraic dual V^{alg}.

If $\lambda \in V'$, then put

$$\|\lambda\|_{V'} = \sup\{|\lambda(v)| : v \in V,\ \|v\|_V \leq 1\}. \tag{3.1.1}$$

This is the same as the operator norm of λ, as a bounded linear mapping from V into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Thus $\|\cdot\|_{V'}$ defines a norm on V', which is known as the *dual norm* associated to $\|\cdot\|_V$ on V. Note that

$$V' \text{ is complete} \tag{3.1.2}$$

with respect to the metric associated to the dual norm, as in Section 2.2.

Suppose now that $\|\cdot\|_V$ is the norm associated to an inner product $\langle\cdot,\cdot\rangle_V$ on V. If $w \in V$, then

$$\lambda_w(v) = \langle v, w\rangle_V \tag{3.1.3}$$

defines a linear functional on V. More precisely, this is a bounded linear functional on V, because of the Cauchy–Schwarz inequality. In fact,

$$\|\lambda_w\|_{V'} = \|w\|_V, \tag{3.1.4}$$

because $\lambda_w(w) = \|w\|_V^2$. Note that

$$w \mapsto \lambda_w \tag{3.1.5}$$

is linear in the real case, and conjugate-linear in the complex case.

Suppose that V is a Hilbert space. If λ is a bounded linear functional on V, then the *Riesz representation theorem* states that

$$\lambda = \lambda_w \tag{3.1.6}$$

for a unique $w \in V$. The uniqueness of w follows from the remarks in the preceding paragraph. To get existence, we may suppose that $\lambda \neq 0$, since otherwise we could take $w = 0$. If $\lambda \neq 0$, then the kernel of λ is a proper closed linear subspace of V. In this case, one can use the orthogonal projection from V onto the kernel of λ to get $y \in V$ such that $y \neq 0$ and

(3.1.7) y is in the orthogonal complement of the kernel of λ.

One can check that (3.1.6) holds with w equal to a suitable multiple of y.

3.2 Duals of ℓ^1 spaces

Let X be a nonempty set, and let f be a real or complex-valued function on X such that $|f|$ is summable on X. If g is a bounded real or complex-valued function on X, as appropriate, then it is easy to see that $|f|\,|g|$ is summable on X, with

$$\sum_{x \in X} |f(x)|\,|g(x)| \le \|f\|_1\,\|g\|_\infty. \tag{3.2.1}$$

Put

$$\lambda_g(f) = \sum_{x \in X} f(x)\,g(x), \tag{3.2.2}$$

where the sum on the right converges in the generalized sense. Observe that

$$|\lambda_g(f)| \le \|f\|_1\,\|g\|_\infty, \tag{3.2.3}$$

by (3.2.1).

This defines a bounded linear functional on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$, as appropriate. One can check that

(3.2.4) the dual norm of λ_g with respect to the usual ℓ^1 norm is equal to the supremum norm of g.

More precisely, one can verify that $\|g\|_\infty$ is less than or equal to the dual norm of λ_g, by considering functions f on X that are equal to 1 at one point, and equal to 0 at all other points in X.

Note that

$$g \mapsto \lambda_g \tag{3.2.5}$$

defines a linear mapping from each of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$ into the duals of $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$, respectively. We would like to show that this mapping is surjective in both cases.

If g is any real or complex-valued function on X, then (3.2.2) defines a linear functional on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate. It is easy to see that every linear functional on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ corresponds to a unique real or complex-valued function g on X, as appropriate, in this way.

Let λ be a bounded linear functional on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$, with respect to the usual ℓ^1 norm. In particular,

(3.2.6) the restriction of λ to $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate, is of the form λ_g as in (3.2.2)

for some real or complex-valued function g on X, as appropriate, as in the preceding paragraph. One can use the boundedness of λ with respect to the ℓ^1 norm to get that

$$g \text{ is bounded on } X, \tag{3.2.7}$$

with supremum norm bounded by the dual norm of λ, as before. Thus λ_g also defines a bounded linear functional on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$, as appropriate, as before.

By construction,

$$\lambda = \lambda_g \tag{3.2.8}$$

on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate. One can use this to get that they are the same on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$, as appropriate. This uses the fact that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$, respectively, with respect to the metrics associated to the ℓ^1 norms.

Of course, if X has only finitely many elements, then $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are the same as $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$, respectively. This means that every linear functional on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$ may be expressed as in (3.2.2), with dual norm with respect to the ℓ^1 norm corresponding to $\|g\|_\infty$, as before.

3.3 Duals of c_0 spaces

Let X be a nonempty set, and let f be a bounded real or complex-valued function on X. If g is a real or complex-valued function on X such that $|g|$ is summable on X, then $|f|\,|g|$ is summable on X, with

$$\sum_{x \in X} |f(x)|\,|g(x)| \le \|f\|_\infty\,\|g\|_1, \tag{3.3.1}$$

as in the previous section. Put

$$\lambda_g(f) = \sum_{x \in X} f(x)\,g(x), \tag{3.3.2}$$

where the sum on the right converges in the generalized sense, as before. We also have that

$$|\lambda_g(f)| \le \|f\|_\infty \, \|g\|_1, \tag{3.3.3}$$

by (3.3.1). This defines a bounded linear functional on $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, as appropriate.

It is easy to see that

(3.3.4) the dual norm of λ_g with respect to the supremum norm is equal to $\|g\|_1$.

More precisely, one can do this using a bounded real or complex-valued function f on X, as appropriate, such that $\|f\|_\infty = 1$ and

$$f(x)\, g(x) = |g(x)| \tag{3.3.5}$$

for every $x \in X$. Of course,

$$g \mapsto \lambda_g \tag{3.3.6}$$

defines a linear mapping from each of $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$ into the duals of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively.

We may also consider λ_g as a bounded linear functional on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, with respect to the supremum norm. One can check that (3.3.4) holds in this case too. Indeed, if A is a nonempty finite subset of X, then one can verify that

$$\sum_{x \in A} |g(x)| \tag{3.3.7}$$

is less than or equal to the dual norm of λ_g, as a bounded linear functional on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, with respect to the supremum norm. This uses a real or complex-valued function f on X supported on A such that $\|f\|_\infty = 1$ and (3.3.5) holds for every $x \in A$.

Let λ be a bounded linear functional on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, with respect to the supremum norm. As in the previous section, the restriction of λ to $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate, is of the form λ_g as in (3.3.2) for some real or complex-valued function g on X, as appropriate. If A is a nonempty finite subset of X, then (3.3.7) is less than or equal to the dual norm of λ, as in the preceding paragraph. This implies that

$$g \in \ell^1(X, \mathbf{R}) \text{ or } \ell^1(X, \mathbf{C}), \tag{3.3.8}$$

as appropriate, with ℓ^1 norm less than or equal to the dual norm of λ.

Thus λ_g defines a bounded linear functional on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, with respect to the supremum norm, as before. One can check that

$$\lambda = \lambda_g \tag{3.3.9}$$

on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate. This uses the density of $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ in $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$, respectively, with respect to the supremum metric. If X has only finitely many elements, then $c_{00}(X, \mathbf{R})$, $c_0(X, \mathbf{R})$, and $\ell^\infty(X, \mathbf{R})$ are all the same, as are $c_{00}(X, \mathbf{C})$, $c_0(X, \mathbf{C})$, and $\ell^\infty(X, \mathbf{C})$.

Suppose that X has infinitely many elements. Let us say that a real or complex-valued function f on X has a *limit at infinity* on X if there is a real or complex number t, as appropriate, such that

$$f(x) - t \text{ vanishes at infinity on } X, \tag{3.3.10}$$

as a real or complex-valued function of $x \in X$, as appropriate. It is easy to see that t is unique when it exists, in which case t is called the limit at infinity of f on X. Under these conditions, f is bounded on X, and

$$|t| \le \|f\|_\infty. \tag{3.3.11}$$

If X is countably infinite, and $\{x_j\}_{j=1}^\infty$ is a sequence of elements of X in which every element of X occurs exactly once, then f has a limit at infinity on X if and only if $\{f(x_j)\}_{j=1}^\infty$ converges as a sequence of real or complex numbers, as appropriate.

Let $c^{lai}(X, \mathbf{R})$, $c^{lai}(X, \mathbf{C})$ be the spaces of real and complex-valued functions on X, respectively, with a limit at infinity. Observe that these are linear subspaces of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively. One can check that these are also closed sets in $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively, with respect to the supremum metric. More precisely, let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on X with limits at infinity such that $\{f_j\}_{j=1}^\infty$ converges uniformly to a real or complex-valued function f on X, as appropriate. If t_j is the limit of f_j at infinity on X for each j, then one can verify that

$$\{t_j\}_{j=1}^\infty \text{ is a Cauchy sequence} \tag{3.3.12}$$

in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to the standard Euclidean metric. This implies that $\{t_j\}_{j=1}^\infty$ converges to a real or complex number t, as appropriate. One can use this to get that (3.3.10) holds.

Of course, constant functions on X have a limit at infinity. The mapping

$$f \mapsto \text{ the limit of } f \text{ at infinity on } X \tag{3.3.13}$$

is a bounded linear functional on each of $c^{lai}(X, \mathbf{R})$ and $c^{lai}(X, \mathbf{C})$, with respect to the supremum norm, and with dual norm equal to 1. If λ is any linear functional on $c^{lai}(X, \mathbf{R})$ or $c^{lai}(X, \mathbf{C})$ that is equal to 0 on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, then λ is a constant multiple of (3.3.13). This constant is equal to the value of λ at the function on X equal to 1 at every $x \in X$.

3.4 Hölder's inequality for sums

Let p, q be real numbers with $1 < p, q < \infty$. Suppose that

$$1/p + 1/q = 1, \tag{3.4.1}$$

in which case p and q are said to be *conjugate exponents*. One may also allow $p = 1$, $q = \infty$ or $p = \infty$, $q = 1$, which were basically considered already in the previous two sections. Note that (3.4.1) implies that

$$p = q/(q-1), \ q = p/(p-1). \tag{3.4.2}$$

Let a, b be nonnegative real numbers. It is well known that

$$a\,b \le a^p/p + b^q/q. \tag{3.4.3}$$

Of course, this is clear when a or b is 0, and so we may as well suppose that $a, b > 0$. This can be obtained from the convexity of the exponential function on the real line. Alternatively, the derivative of

$$a^p/p - a\,b + b^q/q, \tag{3.4.4}$$

as a function of a, is equal to

$$a^{p-1} - b. \tag{3.4.5}$$

It is easy to see that this is equal to 0 if and only if (3.4.6) holds, using (3.4.2).

If

$$a^p = b^q, \tag{3.4.6}$$

then

$$a\,b = a^p = b^q, \tag{3.4.7}$$

and equality holds in (3.4.3). More precisely, the strict convexity of the exponential function on $\mathbf{R}$ implies that equality only holds in (3.4.3) when (3.4.6) holds. This can also be obtained by considering (3.4.5).

Let X be a nonempty set, and let f, g be real or complex-valued functions on X. If $|f|$ is p-summable on X, and $|g|$ is q-summable on X, then $|f|\,|g|$ is summable on X, with

$$\sum_{x \in X} |f(x)|\,|g(x)| \le (1/p) \sum_{x \in X} |f(x)|^p + (1/q) \sum_{x \in X} |g(x)|^q, \tag{3.4.8}$$

because of (3.4.3). Under these conditions, *Hölder's inequality* states that

$$\sum_{x \in X} |f(x)|\,|g(x)| \le \|f\|_p \, \|g\|_q. \tag{3.4.9}$$

This follows from (3.4.8) when $\|f\|_p = \|g\|_q = 1$. Otherwise, one can reduce to that case, using scalar multiplication.

Let A be a nonempty subset of X, and suppose for the moment that

$$f(x)\,g(x) = |f(x)|^p = |g(x)|^q \tag{3.4.10}$$

for every $x \in A$. This implies that

$$\begin{aligned} \sum_{x \in A} f(x)\,g(x) &= \sum_{x \in A} |f(x)|^p = \sum_{x \in A} |g(x)|^q \\ &= \Big(\sum_{x \in A} |f(x)|^p\Big)^{1/p} \Big(\sum_{x \in A} |g(x)|^q\Big)^{1/q}. \end{aligned} \tag{3.4.11}$$

Of course, if A has only finitely many elements, then this works without any summability conditions on f or g.

If $g \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, then put

$$\lambda_g(f) = \sum_{x \in X} f(x)\, g(x) \tag{3.4.12}$$

for all $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, where the sum on the right converges in the generalized sense, as usual. This defines a bounded linear functional on $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate. One can check that

(3.4.13) the dual norm of λ_g with respect to the ℓ^p norm is equal to $\|g\|_q$,

using (3.4.11) with $A = X$. More precisely, one can choose f so that (3.4.10) holds for every $x \in X$. As before,

$$g \mapsto \lambda_g \tag{3.4.14}$$

defines a linear mapping from each of $\ell^q(X, \mathbf{R})$, $\ell^q(X, \mathbf{C})$ into the duals of $\ell^p(X, \mathbf{R})$, $\ell^p(X, \mathbf{C})$, respectively.

Let λ be a bounded linear functional on $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$. As in the previous two sections, the restriction of λ to $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate, is of the form λ_g as in (3.4.12) for some real or complex-valued function g on X, as appropriate. If A is a finite subset of X, then one can verify that

$$\Big(\sum_{x \in A} |g(x)|^q\Big)^{1/q} \tag{3.4.15}$$

is less than or equal to the dual norm of λ, using (3.4.11). Here we choose f so that (3.4.10) holds when $x \in A$, and $f = 0$ on $X \setminus A$.

This implies that

$$g \in \ell^q(X, \mathbf{R}) \text{ or } \ell^q(X, \mathbf{C}), \tag{3.4.16}$$

as appropriate, with ℓ^q norm less than or equal to the dual norm of λ. It follows that λ_g defines a bounded linear functional on $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, as before. Of course,

$$\lambda = \lambda_g \tag{3.4.17}$$

on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate, by construction. Remember that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $\ell^p(X, \mathbf{R})$, $\ell^p(X, \mathbf{C})$, respectively, with respect to the metrics associated to the ℓ^p norms when $p < \infty$. One can use this to get that (3.4.17) holds on $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, as before.

3.5 Hilbert space adjoints

Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be Hilbert spaces, both real or both complex, and let $\|\cdot\|_V$, $\|\cdot\|_W$ be the norms corresponding to these inner products, respectively.

Also let T be a bounded linear mapping from V into W. One can check that the operator norm of T may be given equivalently by

$$\|T\|_{op,VW} = \sup\{|\langle T(v), w\rangle_W| : v \in V,\, w \in W,\, \|v\|_V, \|w\|_W \le 1\}. \tag{3.5.1}$$

More precisely, right side is less than or equal to the operator norm of T, because of the Cauchy–Schwarz inequality. To get the opposite inequality, one can use the fact that

$$\|y\|_W = \sup\{|\langle y, w\rangle_W| : w \in W,\, \|w\|_W \le 1\} \tag{3.5.2}$$

for every $y \in W$.

If $w \in W$, then

$$\mu_w(v) = \langle T(v), w\rangle_W \tag{3.5.3}$$

defines a bounded linear functional on V. More precisely,

$$|\mu_w(v)| \le \|T(v)\|_W\, \|w\|_W \le \|T\|_{op,VW}\, \|w\|_W\, \|v\|_V \tag{3.5.4}$$

for every $v \in V$. It follows that there is a unique element of V that we shall denote $T^*(w)$ such that

$$\langle T(v), w\rangle_W = \langle v, T^*(w)\rangle_V \tag{3.5.5}$$

for every $v \in V$, as in Section 3.1. Note that

$$|\langle v, T^*(w)\rangle_V| = |\langle T(v), w\rangle_W| \le \|T\|_{op,VW}\, \|w\|_W\, \|v\|_V \tag{3.5.6}$$

for all $v \in V$, as in (3.5.4). One can use this to get that

$$\|T^*(w)\|_V \le \|T\|_{op,VW}\, \|w\|_W, \tag{3.5.7}$$

as in (3.5.2).

It is easy to see that T^* is a linear mapping from W into V. Using (3.5.7), we get that T^* is a bounded linear mapping, with

$$\|T^*\|_{op,WV} \le \|T\|_{op,VW}. \tag{3.5.8}$$

As in (3.5.1), we have that

$$\|T^*\|_{op,WV} = \sup\{|\langle v, T^*(w)\rangle_V| : v \in V,\, w \in W,\, \|v\|_V, \|w\|_W \le 1\}. \tag{3.5.9}$$

One can use this, (3.5.1), and (3.5.5) to get that

$$\|T^*\|_{op,WV} = \|T\|_{op,VW}. \tag{3.5.10}$$

The operator T^* is called the *adjoint* of T.

Similarly, the adjoint $(T^*)^*$ of T^* is a bounded linear mapping from V into W. One can check that

$$(T^*)^* = T. \tag{3.5.11}$$

One can also use this to get (3.5.10) from (3.5.8) and its analogue for T^*.

Note that

$$T \mapsto T^* \tag{3.5.12}$$

is a linear mapping from $\mathcal{BL}(V, W)$ into $\mathcal{BL}(W, V)$ in the real case. In the complex case, this mapping is conjugate-linear.

Let $(Z, \langle \cdot, \cdot \rangle_Z)$ be another Hilbert space, which is real or complex, depending on whether V, W are real or complex. If T_1 is a bounded linear mapping from V into W, and T_2 is a bounded linear mapping from W into Z, then their composition $T_2 \circ T_1$ is a bounded linear mapping from V into Z. Observe that

$$\langle T_2(T_1(v)), z\rangle_Z = \langle T_1(v), T_2^*(z)\rangle_W = \langle v, T_1^*(T_2^*(z))\rangle_Z \tag{3.5.13}$$

for every $v \in V$ and $z \in Z$. This implies that

$$(T_2 \circ T_1)^* = T_1^* \circ T_2^*, \tag{3.5.14}$$

as bounded linear mappings from Z into V.

If T is any bounded linear mapping from V into W, then $T^* \circ T$ is a bounded linear mapping from V into itself. Note that

$$\langle T^*(T(u)), v\rangle_V = \langle T(u), T(v)\rangle_W \tag{3.5.15}$$

for every $u, v \in V$. It follows that T is an isometric linear mapping from V into W if and only if

$$\langle T^*(T(u)), v\rangle_V = \langle u, v\rangle_V \tag{3.5.16}$$

for every $u, v \in V$. One can check that this happens if and only if

$$T^* \circ T \text{ is the identity mapping } I_V \text{ on } V. \tag{3.5.17}$$

If we take $u = v$ in (3.5.15), then we get that

$$\|T(v)\|_W^2 = \langle T^*(T(v)), v\rangle_V \tag{3.5.18}$$

for every $v \in V$. This implies that

$$\|T(v)\|_V^2 \le \|T^*(T(v))\|_V \, \|v\|_V \le \|T^* \circ T\|_{op,VV} \, \|v\|_V^2 \tag{3.5.19}$$

for every $v \in V$, using the Cauchy–Schwarz inequality in the first step. It follows that

$$\|T\|_{op,VW}^2 \le \|T^* \circ T\|_{op,VV}. \tag{3.5.20}$$

It is easy to see that the opposite inequality holds too, so that

$$\|T\|_{op,VW}^2 = \|T^* \circ T\|_{op,VV}. \tag{3.5.21}$$

This is the C^* *identity* for the operator norm of a bounded linear mapping between Hilbert spaces.

3.6 Sublinear functions

Let V be a vector space over the real numbers, and let p be a real-valued function on V. If

$$p(v + w) \le p(v) + p(w) \tag{3.6.1}$$

for every $v, w \in V$, then p is said to be *subadditive* on V. If

$$p(t\, v) = t\, p(v) \tag{3.6.2}$$

for every $v \in V$ and nonnegative real number t, then p is said to be *homogeneous of degree* 1 on V. Of course, this implies that

$$p(0) = 0, \tag{3.6.3}$$

by taking $t = 0$. If p is both subadditive and homogeneous of degree 1, then p is said to be *sublinear* on V.

Let us say that p is *symmetric* on V if

$$p(-v) = p(v) \tag{3.6.4}$$

for every $v \in V$. If p is subadditive on V, then it is easy to see that

$$p(0) \ge 0. \tag{3.6.5}$$

We also have that

$$p(0) \le p(v) + p(-v) \tag{3.6.6}$$

for every $v \in V$. If p is symmetric on V as well, then we get that

$$0 \le p(0)/2 \le p(v) \tag{3.6.7}$$

for every $v \in V$. Thus a seminorm on V is the same as a symmetric sublinear function on V.

Note that linear functionals on V are sublinear. More precisely, if p is sublinear on V, then p is linear on V if and only if

$$p(-v) = -p(v) \tag{3.6.8}$$

for every $v \in V$.

Let p_1, p_2 be real-valued functions on V. If p_1, p_2 are subadditive on V, then it is easy to see that

$$\max(p_1, p_2) \text{ is subadditive on } V. \tag{3.6.9}$$

If p_1, p_2 are both homogeneous of degree 1, or both symmetric on V, then $\max(p_1, p_2)$ has the same property.

If p is sublinear on V, then clearly

$$p(-v) \text{ is sublinear on } V. \tag{3.6.10}$$

It follows that

$$\max(p(v), p(-v)) \text{ is a seminorm on } V. \tag{3.6.11}$$

Note that

$$\max(p(v), 0) \text{ is sublinear on } V \tag{3.6.12}$$

in this case too.

If p is subadditive on V, then

$$p(v) \leq p(w) + p(v - w) \tag{3.6.13}$$

for every $v, w \in V$. Similarly,

$$p(w) \leq p(v) + p(w - v) \tag{3.6.14}$$

for every $v, w \in V$. It follows that

$$|p(v) - p(w)| \leq \max(p(v - w), p(w - v)) \tag{3.6.15}$$

for every $v, w \in V$. If p is symmetric on V, then this reduces to

$$|p(v) - p(w)| \leq p(v - w) \tag{3.6.16}$$

for every $v, w \in V$. Of course, the analogous statement for seminorms was mentioned in Section 1.2.

If p is sublinear on V, then

$$\{v \in V : p(v) < r\} \tag{3.6.17}$$

is a convex set in V for every real number r. Similarly,

$$\{v \in V : p(v) \leq r\} \tag{3.6.18}$$

is a convex set in V for every $r \in \mathbf{R}$. Note that 0 is an element of (3.6.17) when $r > 0$, and an element of (3.6.18) when $r \geq 0$. Analogous statements for seminorms were mentioned in Section 1.2.

If E is any subset of V, then put

$$-E = \{-v : v \in E\}. \tag{3.6.19}$$

If

$$-E = E, \tag{3.6.20}$$

then E is said to be *symmetric* about 0 in V. If p is symmetric in V, then (3.6.17) and (3.6.18) are symmetric about 0 in V for every $r \in \mathbf{R}$.

3.7 Convex cones

Let V be a vector space over the real numbers again. A subset C of V is said to be a *convex cone* in V if it satisfies the following two conditions. First, if $v \in C$ and t is a positive real number, then

$$t\,v \in C. \tag{3.7.1}$$

Second, if $v, w \in C$, then

$$v + w \in C. \tag{3.7.2}$$

It is easy to see that

$$\text{convex cones are convex sets.} \tag{3.7.3}$$

If a subset C of V satisfies (3.7.1) and is convex, then (3.7.2) holds.

Linear subspaces of V are convex cones. If a nonempty convex cone C in V is symmetric about 0 in V, then

$$C \text{ is a linear subspace of } V. \tag{3.7.4}$$

If C is a convex cone in V, and if

$$C \cap (-C) \tag{3.7.5}$$

is nonempty, then (3.7.5) is a linear subspace of V.

If p is a sublinear real-valued function on V, then

$$\{v \in V : p(v) \le 0\} \tag{3.7.6}$$

is a convex cone in V. If p is a seminorm on V, then this is the same as

$$\{v \in V : p(v) = 0\}, \tag{3.7.7}$$

which is a linear subspace of V.

Let C be a nonempty convex cone in V, and let p be a nonnegative real-valued function on V. If $v \in V$, then put

$$p_C(v) = \inf\{p(v - w) : w \in C\}. \tag{3.7.8}$$

If $p(0) = 0$, then

$$p_C(v) = 0 \tag{3.7.9}$$

for every $v \in C$. If p is a norm on V, then p_C is the same as the distance to C with respect to the metric associated to C, as in Section 2.11.

If p is homogeneous of degree 1 on V, then one can check that p_C is homogeneous of degree 1 on V as well. More precisely, one can verify that

$$p_C(t\,v) = t\,p_C(v) \tag{3.7.10}$$

for every $v \in V$ and positive real number t. We also have that

$$p_C(0) = 0, \tag{3.7.11}$$

so that (3.7.10) holds when $t = 0$ too.

Suppose that p is subadditive on V. If $v_1, v_2 \in V$ and $w_1, w_2 \in C$, then

$$p_C(v_1 + v_2) \le p(v_1 + v_2 - w_1 - w_2) \le p(v_1 - w_1) + p(v_2 - w_2). \tag{3.7.12}$$

One can use this to get that

$$p_C(v_1 + v_2) \le p_C(v_1) + p_C(v_2). \tag{3.7.13}$$

3.8 Minkowski functionals

Let V be a vector space over the real numbers, and let E be a subset of V. If $t \in \mathbf{R}$, then put

$$t\,E = \{t\,v : v \in E\}. \tag{3.8.1}$$

We say that E is *starlike about* 0 if

$$t\,E \subseteq E \tag{3.8.2}$$

for every $t \in [0, 1]$. Of course, this implies that $0 \in E$ when $E \ne \emptyset$. If E is convex and $0 \in E$, then E is starlike about 0.

Let us say that a subset A of V is *radial at* 0 if for every $v \in V$ there is a positive real number $r(v)$ such that

$$r\,v \in A \tag{3.8.3}$$

for every $r \in \mathbf{R}$ with $0 \le r \le r(v)$, as on p14 of [105]. Note that this implies that $0 \in A$. This is a type of "absorbing" property, although that term is used in slightly different ways as well. A convex set $A \subseteq V$ may be said to be *absorbing* if

$$\text{every } v \in V \text{ is an element of } t\,A \text{ for some } t > 0, \tag{3.8.4}$$

as on p24 of [162]. This implies that $0 \in A$, and that A is radial at 0, which would also work when A is starlike about 0 instead of convex.

Let p be a real-valued function on V that is homogeneous of degree 1. If r is a positive real number, then (3.6.17) and (3.6.18) are starlike about 0 and radial at 0.

Suppose that $A \subseteq V$ satisfies (3.8.4). The corresponding *Minkowski functional* on V is defined by

$$\rho_A(v) = \inf\{t > 0 : t^{-1}\,v \in A\} = \inf\{t > 0 : v \in t\,A\}, \tag{3.8.5}$$

as on p15 of [105], and p24 of [162]. Clearly

$$\rho_A(0) = 0, \tag{3.8.6}$$

because $0 \in A$. It is easy to see that

$$\rho_A(r\,v) = r\,\rho_A(v) \tag{3.8.7}$$

for every $v \in V$ and $r > 0$, so that ρ_A is homogeneous of degree 1 on V. If A is symmetric about 0 in V, then ρ_A is symmetric on V.

Put

$$B = \{v \in V : \rho_A(v) < 1\} \tag{3.8.8}$$

and

$$C = \{v \in V : \rho_A(v) \le 1\}. \tag{3.8.9}$$

By construction,

$$A \subseteq C. \tag{3.8.10}$$

If A is starlike about 0, then

$$B \subseteq A. \tag{3.8.11}$$

If A is convex, then one can check that

$$\rho_A \text{ is subadditive on } V, \tag{3.8.12}$$

using the argument in Section 1.2.

3.9 Some one-step extensions

Let V be a vector space over the real numbers, and let p be a sublinear real-valued function on V. Also let V_0 be a linear subspace of V, and let λ_0 be a linear functional on V_0. Suppose that

$$\lambda_0(v_0) \le p(v_0) \tag{3.9.1}$$

for every $v_0 \in V_0$.

If $V_0 \ne V$, then let u_1 be an element of V not in V_0. Consider

$$V_1 = \{v_0 + t\,u_1 : v_0 \in V_0,\ t \in \mathbf{R}\}, \tag{3.9.2}$$

which is the linear subspace of V spanned by V_0 and u_1. Under these conditions, there is an extension λ_1 of λ_0 to a linear functional on V_1 such that

$$\lambda_1(v_1) \le p(v_1) \tag{3.9.3}$$

for every $v_1 \in V_1$. This is part of the theorem of Hahn and Banach.

One can check that every element of V_1 can be expressed in a unique way as

$$v_0 + t\,u_1 \tag{3.9.4}$$

for some $v_0 \in V_0$ and $t \in \mathbf{R}$, because $u_1 \notin V_0$. If α_1 is a real number, then we can get an extension λ_1 of λ_0 to a linear functional on V_1 by putting

$$\lambda_1(v_0 + t\,u_1) = \lambda_0(v_0) + t\,\alpha_1 \tag{3.9.5}$$

for every $v_0 \in V_0$ and $t \in \mathbf{R}$. In fact, every such extension λ_1 of λ_0 corresponds to a unique $\alpha_1 \in \mathbf{R}$ in this way.

Thus we would like to choose α_1 so that

$$\lambda_0(v_0) + t\,\alpha_1 \le p(v_0 + t\,u_1) \tag{3.9.6}$$

for every $v_0 \in V_0$ and $t \in \mathbf{R}$. This condition holds when $t = 0$, by (3.9.1). In order to get that this condition holds when $t \neq 0$, one can reduce to the cases where $t = \pm 1$, because p is homogeneous of degree 1 on V, and V_0 is a linear subspace of V. This means that it suffices to choose α_1 so that

$$\lambda_0(v_0) + \alpha_1 \le p(v_0 + u_1) \tag{3.9.7}$$

and

$$\lambda_0(v_0) - \alpha_1 \le p(v_0 - u_1) \tag{3.9.8}$$

for every v_0. Equivalently, we would like to choose α_1 so that

$$\lambda_0(v_0) - p(v_0 - u_1) \le \alpha_1 \le p(w_0 + u_1) - \lambda_0(w_0) \tag{3.9.9}$$

for every $v_0, w_0 \in V_0$.

Let us check that

$$\lambda_0(v_0) - p(v_0 - u_1) \le p(w_0 + u_1) - \lambda_0(w_0) \tag{3.9.10}$$

for every $v_0, w_0 \in V_0$. This is the same as saying that

$$\lambda_0(v_0) + \lambda_0(w_0) \le p(v_0 - u_1) + p(w_0 + u_1) \tag{3.9.11}$$

for all $v_0, w_0 \in V_0$. If $v_0, w_0 \in V_0$, then

$$\begin{aligned} \lambda_0(v_0) + \lambda_0(w_0) = \lambda_0(v_0 + w_0) &\le p(v_0 + w_0) \\ &\le p(v_0 - u_1) + p(w_0 + u_1). \end{aligned} \tag{3.9.12}$$

This uses (3.9.1) in the second step, and the subadditivity of p on V in the third step. This shows that (3.9.10) holds, which implies that there is an $\alpha_1 \in \mathbf{R}$ such that (3.9.9) holds.

More precisely,

$$\begin{aligned} &\sup\{\lambda_0(v_0) - p(v_0 - u_1) : v_0 \in V_0\} \\ &\qquad \le \inf\{p(w_0 + u_1) - \lambda_0(w_0) : w_0 \in V_0\}, \end{aligned} \tag{3.9.13}$$

because of (3.9.10). In order to get (3.9.9), one can take any $\alpha_1 \in \mathbf{R}$ such that

$$\begin{aligned} &\sup\{\lambda_0(v_0) - p(v_0 - u_1) : v_0 \in V_0\} \\ &\qquad \le \alpha_1 \le \inf\{p(w_0 + u_1) - \lambda_0(w_0) : w_0 \in V_0\}. \end{aligned} \tag{3.9.14}$$

Of course, (3.9.1) is the same as saying that

$$-\lambda_0(v_0) = \lambda_0(-v_0) \le p(-v_0) \tag{3.9.15}$$

for every $v_0 \in V_0$. This means that

$$-p(-v_0) \le \lambda_0(v_0) \tag{3.9.16}$$

for every $v_0 \in V_0$. Thus (3.9.1) implies that

$$|\lambda_0(v_0)| \le \max(p(v_0), p(-v_0)) \tag{3.9.17}$$

for every $v_0 \in V_0$.

Let $\| \cdot \|_V$ be a norm on V, and suppose that

$$p(v) \le C\, \|v\|_V \tag{3.9.18}$$

for some nonnegatve real number C and every $v \in V$. This implies that

$$p(-v) \le C\, \| -v\|_V = C\, \|v\|_V \tag{3.9.19}$$

for every $v \in V$ as well. It follows that

$$|p(v) - p(w)| \le \max(p(v-w), p(w-v)) \le C\, \|v - w\|_V \tag{3.9.20}$$

for every $v, w \in V$, using (3.6.15) in the first step. In particular, this means that p is continuous as a real-valued function on V, with respect to the metric associated to $\| \cdot \|_V$.

Similarly,

$$|\lambda_0(v_0)| \le C\, \|v_0\|_V \tag{3.9.21}$$

for every $v_0 \in V_0$, by (3.9.17). Under these conditions,

(3.9.22) λ_0 has a unique extension to a bounded linear functional on the closure $\overline{V_0}$ of V_0 in V,

with respect to the metric associated to $\| \cdot \|_V$, as in Section 2.2. This extension satisfies the analogue of (3.9.1) on $\overline{V_0}$, because p is continuous on V, as in the preceding paragraph.

3.10 The Hahn–Banach theorem

Let us continue with the same notation and hypotheses as at the beginning of the previous section. Under these conditions, the *Hahn–Banach theorem* states that there is an extension λ of λ_0 to a linear functional on V that satisfies

$$\lambda(v) \le p(v) \tag{3.10.1}$$

for every $v \in V$.

If V is spanned by V_0 and finitely many additional vectors, then λ can be obtained by repeating the argument in the previous section finitely many times. If V is spanned by V_0 and a sequence of additional vectors, then one can continue to repeat the argument to get λ.

Otherwise, there is an argument based on the axiom of choice using Zorn's lemma or Hausdorff's maximality principle to get a maximal extension of λ_0 to a linear subspace of V on which the extension is less than or equal to p. The argument in the previous section implies that such a maximal extension is defined on all of V.

Let $\|\cdot\|_V$ be a norm on V again, and suppose that p is less than or equal to a nonnegative real number C times $\|\cdot\|_V$ on V, as in (3.9.18). In this case, it suffices to find an extension of λ_0 to a dense linear subspace of V, with respect to the metric associated to $\|\cdot\|_V$, on which the extension is less than or equal to p, as in the previous section. If the linear span of V_0 and a sequence of additional vectors is dense in V, then such an extension of λ_0 can be obtained by repeating the argument in the previous section, as before.

If p is a seminorm on V, then (3.9.1) is the same as saying that

$$|\lambda_0(v_0)| \le p(v_0) \tag{3.10.2}$$

for every $v_0 \in V_0$, as in the previous section. Similarly, (3.10.1) is the same as saying that

$$|\lambda(v)| \le p(v) \tag{3.10.3}$$

for every $v \in V$. This corresponds to Theorem 3.3 on p57 of [162] in the real case.

In particular, we can take

$$p(v) = C\,\|v\|_V \tag{3.10.4}$$

on V for some $C \ge 0$. Using this, we get that the Hahn–Banach theorem implies that if λ_0 is a bounded linear functional on V_0 with respect to the restriction of $\|\cdot\|_V$ to V_0, then λ_0 has an extension to a bounded linear functional on V, with the same dual norm with respect to $\|\cdot\|_V$.

Let $u_0 \in V$ with $u_0 \ne 0$ be given, and let V_0 be the linear span of u_0 in V. Consider the linear functional λ_0 defined on V_0 by

$$\lambda_0(t\,u_0) = t\,\|u_0\|_V \tag{3.10.5}$$

for every $t \in \mathbf{R}$. The Hahn–Banach theorem implies that λ_0 has an extension to a bounded linear functional on V with dual norm equal to 1. This corresponds to the corollary on p58 of [162] in the real case.

Let W_0 be a closed linear subspace of V, with respect to the metric associated to $\|\cdot\|_V$, and suppose that $u_0 \in V \setminus W_0$. Thus there is a positive real number r such that

$$\|u_0 - w_0\|_V \ge r \tag{3.10.6}$$

for every $w_0 \in W_0$. Let V_0 be the linear subspace of V spanned by W_0 and u_0, so that every element of V_0 can be expressed in a unique way as

$$w_0 + t\,u_0 \tag{3.10.7}$$

for some $w_0 \in W_0$ and $t \in \mathbf{R}$. This uses the fact that $t\,u_0 \in W_0$ only when $t = 0$.

Consider the linear functional λ_0 defined on V_0 by

$$\lambda_0(w_0 + t\, u_0) = t \tag{3.10.8}$$

for every $w_0 \in W_0$ and $t \in \mathbf{R}$. One can check that

$$|\lambda_0(w_0 + t\, u_0)| = |t| \le r^{-1}\, \|w_0 + t\, u_0\|_V \tag{3.10.9}$$

for every $w_0 \in W_0$ and $t \in \mathbf{R}$, using (3.10.6) in the second step. The Hahn–Banach theorem implies that λ_0 can be extended to a bounded linear functional λ on V, with dual norm less than or equal to $1/r$. Note that $\lambda(u_0) = 1$, and that $\lambda(w_0) = 0$ for every $w_0 \in W_0$. This corresponds to a simplification of Theorem 3.5 on p59 of [162] in the real case.

3.11 Some separation results

Let V be a vector space over the real numbers, and let $\|\cdot\|_V$ be a norm on V. Also let A be a nonempty convex open subset of V, with respect to the metric associated to $\|\cdot\|_V$, and let v be an element of V not in A. Under these conditions, there is a bounded linear functional λ on V such that

$$\lambda(a) < \lambda(v) \tag{3.11.1}$$

for every $a \in A$. This corresponds to a simplification of part (a) of Theorem 3.4 on p58 of [162] in the real case.

To see this, we start by reducing to the case where $0 \in A$, which we can do using a translation on V. It is easy to see that A is radial at 0, or equivalently that A has the absorbing property mentioned in Section 3.8, because A is an open set in V. Let ρ_A be the Minkowski functional on V associated to A, as in Section 3.8. Remember that ρ_A is sublinear on V under these conditions.

Let V_0 be the linear subspace of V spanned by v, and let λ_0 be the linear functional on V_0 defined by

$$\lambda_0(t\, v) = t\, \rho_A(v) \tag{3.11.2}$$

for every $t \in \mathbf{R}$. Observe that

$$\lambda_0(t\, v) \le \rho_A(t\, v) \tag{3.11.3}$$

for every $t \in \mathbf{R}$, with equality when $t \ge 0$. More precisely, if $t < 0$, then the right side is greater than or equal to 0, and the left side is less than or equal to 0. The Hahn–Banach theorem implies that there is an extension λ of λ_0 to a linear functional on V that satisifes

$$\lambda(w) \le \rho_A(w) \tag{3.11.4}$$

for every $w \in V$.

It is easy to see that

$$\rho_A(v) \geq 1, \tag{3.11.5}$$

because $v \notin A$, by hypothesis. If $a \in A$, then one can check that

$$\rho_A(a) < 1, \tag{3.11.6}$$

because A is an open set in V. This implies (3.11.1).

Because A is an open set in V that contains 0, there is a positive real number r such that $a \in A$ when $a \in V$ and $\|a\|_V < r$. Using this, one can verify that

$$\rho_A(w) \leq r^{-1}\, \|w\|_V \tag{3.11.7}$$

for every $w \in V$. It follows from this and (3.11.4) that λ is a bounded linear functional on V, as in Section 3.9.

If $E_1, E_2 \subseteq V$, then put

$$E_1 + E_2 = \{w_1 + w_2 : w_1 \in E_1,\, w_2 \in E_2\}. \tag{3.11.8}$$

It is easy to see that this is an open set in V, with respect to the metric associated to $\|\cdot\|_V$, when E_1 or E_2 is an open set. If E_1 is an open set in V, for instance, then $E_1 + E_2$ is the same as the union of the translates of E_1 by elements of E_2, and each of these translates of E_1 is an open set in V. If E_1 and E_2 are both convex sets in V, then $E_1 + E_2$ is convex in V as well.

Now let B be a nonempty closed convex set in V, with respect to the metric associated to $\|\cdot\|_V$, and suppose that $v \notin B$. Thus there is a positive real number r_1 such that

$$\|v - b\|_V \geq r_1 \tag{3.11.9}$$

for every $b \in B$. Put

$$A_1 = B + B(0, r_1), \tag{3.11.10}$$

where $B(0, r_1)$ is the usual open ball in V centered at 0 with radius r_1 with respect to the metric associated to $\|\cdot\|_V$. This is a convex open subset of V, as in the preceding paragraph. We also have that $v \notin A_1$, by construction.

It follows that there is a bounded linear functional λ on V such that

$$\lambda(a_1) < \lambda(v) \tag{3.11.11}$$

for every $a_1 \in A_1$, as before. Equivalently, this means that

$$\lambda(b) + \lambda(w) = \lambda(b + w) < \lambda(v) \tag{3.11.12}$$

for every $b \in B$ and $w \in V$ with $\|w\|_V < r_1$. Of course, $\lambda \not\equiv 0$ on V, so that

$$\lambda(w) > 0 \tag{3.11.13}$$

for some $w \in V$ with $\|w\|_V < r_1$. Using this, we get that

$$\sup_{b \in B} \lambda(b) \leq \lambda(v) - \lambda(w) < \lambda(v). \tag{3.11.14}$$

This corresponds to a simplification of part (b) of Theorem 3.4 on p58 of [162] in the real case.

3.12 Complex linear functionals

Let V be a vector space over the complex numbers, and let $V_{\mathbf{R}}$ be V considered as a vector space over the real numbers, as before. If λ is a linear functional on V, then

$$\mu = \operatorname{Re}\lambda \tag{3.12.1}$$

is a linear functional on $V_{\mathbf{R}}$. It is easy to see that λ is uniquely determined by μ. More precisely,

$$\lambda(v) = \mu(v) - i\,\mu(i\,v) \tag{3.12.2}$$

for every $v \in V$, as in (1) on p56 of [162]. Conversely, if μ is a linear functional on $V_{\mathbf{R}}$, then one can check that (3.12.2) defines a linear functional on V, as in [162].

Let $\|\cdot\|_V$ be a norm on V, which may also be considered as a norm on $V_{\mathbf{R}}$. It is easy to see that λ is bounded as a linear functional on V with respect to $\|\cdot\|_V$ if and only if μ is bounded as a linear functional on $V_{\mathbf{R}}$ with respect to $\|\cdot\|_V$. In fact, the corresponding dual norms of λ and μ are the same, which is to say that

$$\|\lambda\|_{V'} = \|\mu\|_{V'_{\mathbf{R}}}. \tag{3.12.3}$$

Indeed, if $v \in V$, then

$$\begin{aligned} |\lambda(v)| &= \sup\{\operatorname{Re}(t\lambda(v)) : t \in \mathbf{C},\ |t| = 1\} \\ &= \sup\{\mu(t\,v) : t \in \mathbf{C},\ |t| = 1\}. \end{aligned} \tag{3.12.4}$$

Let N be a seminorm on V, which may be considered as a seminorm on $V_{\mathbf{R}}$ too. Also let V_0 be a linear subspace of V, and let λ_0 be a linear functional on V_0 such that

$$|\lambda_0(v_0)| \le N(v_0) \tag{3.12.5}$$

for every $v_0 \in V_0$. Thus $\mu_0 = \operatorname{Re}\lambda_0$ is a linear functional on V_0, as a linear subspace of $V_{\mathbf{R}}$, with

$$\mu_0(v_0) \le N(v_0) \tag{3.12.6}$$

for every $v_0 \in V_0$. It follows that μ_0 can be extended to a linear functional μ on $V_{\mathbf{R}}$ such that

$$\mu(v) \le N(v) \tag{3.12.7}$$

for every $v \in V_{\mathbf{R}}$, as in Section 3.10. Let λ be the linear functional on V whose real part is equal to μ, as before. Observe that

$$|\lambda(v)| \le N(v) \tag{3.12.8}$$

for every $v \in V$, because of (3.12.4). We also have that $\lambda = \lambda_0$ on V_0, because their real parts are the same on V_0, by construction. This corresponds to Theorem 3.3 on p57 of [162] in the complex case.

If $u_0 \in V$ and $u_0 \neq 0$, then there is a bounded linear functional λ on V such that

$$\lambda(u_0) = \|u_0\|_V \text{ and } \|\lambda\|_{V'} = 1. \tag{3.12.9}$$

This can be obtained from the remarks in the preceding paragraph in the same way as in Section 3.10. This corresponds to the corollary on p58 of [162] in the complex case.

Similarly, let W_0 be a closed linear subapace of V, with respect to the metric associated to $\|\cdot\|_V$, and let u_0 be an element of $V \setminus W_0$. Under these conditions, there is a bounded linear function λ on V such that

$$\lambda(u_0) = 1 \text{ and } \lambda(w_0) = 0 \text{ for every } w_0 \in W_0, \tag{3.12.10}$$

as in Section 3.10 again. This corresponds to a simplification of Theorem 3.5 on p59 of [162] in the complex case.

3.13 Dual linear mappings

Let V, W be vector spaces, both real or both complex. Also let T be a linear mapping from V into W. If λ is a linear functional on W, then

$$\lambda \circ T \tag{3.13.1}$$

is a linear functional on V. Put

$$T^{\text{alg}}(\lambda) = \lambda \circ T, \tag{3.13.2}$$

which defines a linear mapping from W^{alg} into V^{alg}. This is the *dual linear mapping* associated to T between the corresponding algebraic dual spaces.

Observe that

$$T \mapsto T^{\text{alg}} \tag{3.13.3}$$

defines a linear mapping from $\mathcal{L}(V, W)$ into $\mathcal{L}(W^{\text{alg}}, V^{\text{alg}})$. Let Z be another vector space over the real or complex numbers, depending on whether V, W are real or complex. If T_1 is a linear mapping from V into W, and T_2 is a linear mapping from W into Z, then one can check that

$$(T_2 \circ T_1)^{alg} = T_1^{\text{alg}} \circ T_2^{\text{alg}}, \tag{3.13.4}$$

as linear mappings from Z^{alg} into V^{alg}. More precisely, if $\mu \in Z^{\text{alg}}$, then

$$((T_2 \circ T_1)^{\text{alg}})(\mu) = \mu \circ (T_2 \circ T_1) = (\mu \circ T_2) \circ T_1 = T_1^{\text{alg}}(T_2^{\text{alg}}(\mu)). \tag{3.13.5}$$

Let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively. If T is a bounded linear mapping from V into W, and λ is a bounded linear functional on W, then (3.13.1) is a bounded linear functional on V. Put

$$T'(\lambda) = \lambda \circ T, \tag{3.13.6}$$

which defines a linear mapping from W' into V'. Of course, this is the same as the restriction of T^{alg} to W'. This is the *dual linear mapping* of T between the dual spaces W', V' of W, V associated to the norms $\|\cdot\|_W$, $\|\cdot\|_V$, respectively.

Note that

$$\|T'(\lambda)\|_{V'} \le \|T\|_{op,VW}\, \|\lambda\|_{W'} \tag{3.13.7}$$

for every $\lambda \in W'$. This implies that T' is a bounded linear mapping from W' into V', with respect to the dual norms corresponding to $\|\cdot\|_W$, $\|\cdot\|_V$, with

$$\|T'\|_{op,W'V'} \le \|T\|_{op,VW}. \tag{3.13.8}$$

We also have that

$$T \mapsto T' \tag{3.13.9}$$

is a linear mapping from $\mathcal{BL}(V, W)$ into $\mathcal{BL}(W', V')$.

If $v \in V$ and $\lambda \in W'$, then

$$\begin{aligned} |\lambda(T(v))| &= |(T'(\lambda))(v)| \le \|T'(\lambda)\|_{V'}\, \|v\|_V \\ &\le \|T'\|_{op.W'V'}\, \|\lambda\|_{W'}\, \|v\|_V. \end{aligned} \tag{3.13.10}$$

This implies that

$$\|T(v)\|_W \le \|T'\|_{op,W'V'}\, \|v\|_V, \tag{3.13.11}$$

because of (3.12.9) and its analogue in the real case. It follows that

$$\|T\|_{op,VW} \le \|T'\|_{op,W'V'}. \tag{3.13.12}$$

This means that

$$\|T'\|_{op,W'V'} = \|T\|_{op,VW}, \tag{3.13.13}$$

because of (3.13.8).

Suppose that T is an isometric linear mapping from V into W, as in Section 2.10. In particular, this implies that the operator norm of T is equal to 1, at least if $V \ne \{0\}$. If V is a linear subspace of W, for instance, and $\|\cdot\|_V$ is the same as the restriction of $\|\cdot\|_W$ to V, then one can take T to be the obvious inclusion mapping from W into V. This means that for each $v \in V$,

$$T(v) = v, \tag{3.13.14}$$

considered as an element of W.

In this case, T' is the obvious restriction mapping from W' into V', which sends $\lambda \in W'$ to its restriction to V, as an element of V'. Note that

$$T'(W') = V', \tag{3.13.15}$$

by the Hahn–Banach theorem. If T is any isometric linear mapping into W, then it is easy to see that T' is surjective, for essentially the same reason.

3.14 Second dual spaces

Let V be a vector space over the real or complex numbers. Consider the algebraic dual

$$(V^{\mathrm{alg}})^{\mathrm{alg}} \tag{3.14.1}$$

of the algebraic dual V^{alg} of V. If $v \in V$ and $\lambda \in V^{\mathrm{alg}}$, then put

$$\widehat{L}_v(\lambda) = \lambda(v). \tag{3.14.2}$$

This defines a linear functional on V^{alg}, and thus an element of $(V^{\mathrm{alg}})^{\mathrm{alg}}$. In fact,

$$v \mapsto \widehat{L}_v \tag{3.14.3}$$

is a linear mapping from V into $(V^{\mathrm{alg}})^{\mathrm{alg}}$.

If $v \neq 0$, then it is well known that there is a $\lambda \in V^{\mathrm{alg}}$ such that $\lambda(v) \neq 0$. This means that $\widehat{L}_v \neq 0$, so that (3.14.3) is one-to-one. If V has finite dimension, then V^{alg} has the same dimension, by standard arguments. This implies that $(V^{\mathrm{alg}})^{\mathrm{alg}}$ has the same dimension in this case, using the same argument. It follows that (3.14.3) maps V onto $(V^{\mathrm{alg}})^{\mathrm{alg}}$ under these conditions.

Let $\|\cdot\|_V$ be a norm on V, so that the corresponding dual norm $\|\cdot\|_{V'}$ can be defined on the dual space V' of bounded linear functionals on V in the usual way. Thus the dual space

$$V'' = (V')' \tag{3.14.4}$$

of bounded linear functionals on V' can be defined in the usual way too, with the dual norm

$$\|\cdot\|_{V''} = \|\cdot\|_{(V')'} \tag{3.14.5}$$

associated to $\|\cdot\|_{V'}$.

If $v \in V$ and $\lambda \in V'$, then put

$$L_v(\lambda) = \lambda(v). \tag{3.14.6}$$

Equivalently, L_v is the same as the restriction of $\widehat{L}_v$ to V'. Of course,

$$|L_v(\lambda)| = |\lambda(v)| \leq \|v\|_V \, \|\lambda\|_{V'}. \tag{3.14.7}$$

This implies that L_v is a bounded linear functional on V', with

$$\|L_v\|_{V''} \leq \|v\|_V. \tag{3.14.8}$$

More precisely,

$$\|L_v\|_{V''} = \|v\|_V, \tag{3.14.9}$$

because of (3.12.9) and its analogue in the real case.

Note that

$$v \mapsto L_v \tag{3.14.10}$$

is a linear mapping from V into V'', as before. This mapping is one-to-one, because of (3.14.9).

If (3.14.10) maps V onto V'', then V is said to be *reflexive*. This can only happen when V is a Banach space, because V'' is automatically a Banach space.

Of course,

$$\{L_v : v \in V\} \tag{3.14.11}$$

is a linear subspace of V''. If V is a Banach space, then (3.14.11) is a closed set in V'', with respect to the metric associated to $\|\cdot\|_{V''}$. This follows from (3.14.9), and a remark about isometries in Section 1.15.

3.15 Second duals of linear mappings

Let V, W be vector spaces, both real or both complex, and let T be a linear mapping from V into W. Remember that T^{alg} is the dual mapping from W^{alg} into V^{alg} corresponding to T, as in Section 3.13. Using this, we get the corresponding second dual mapping

$$(T^{\mathrm{alg}})^{\mathrm{alg}} \tag{3.15.1}$$

from $(V^{\mathrm{alg}})^{\mathrm{alg}}$ into $(W^{\mathrm{alg}})^{\mathrm{alg}}$. Note that

$$T \mapsto (T^{\mathrm{alg}})^{\mathrm{alg}} \tag{3.15.2}$$

is a linear mapping from $\mathcal{L}(V, W)$ into $\mathcal{L}((V^{\mathrm{alg}})^{\mathrm{alg}}, (W^{\mathrm{alg}})^{\mathrm{alg}})$, because of the analogous statement in Section 3.13.

If $v \in V$, then let $\widehat{L}_v^V = \widehat{L}_v \in (V^{\mathrm{alg}})^{\mathrm{alg}}$ be as in (3.14.2). Similarly, if $w \in W$ and $\mu \in W^{\mathrm{alg}}$, then

$$\widehat{L}_w^W(\mu) = \mu(w) \tag{3.15.3}$$

defines an element of $(W^{\mathrm{alg}})^{\mathrm{alg}}$. Remember that $v \mapsto \widehat{L}_v^V$ and $w \mapsto \widehat{L}_w^W$ define linear mappings from V and W into $(V^{\mathrm{alg}})^{\mathrm{alg}}$ and $(W^{\mathrm{alg}})^{\mathrm{alg}}$, respectively, as in the previous section.

Let us check that

$$(T^{\mathrm{alg}})^{\mathrm{alg}}(\widehat{L}_v^V) = \widehat{L}_{T(v)}^W. \tag{3.15.4}$$

Of course,

$$(T^{\mathrm{alg}})^{\mathrm{alg}}(\widehat{L}_v^V) = \widehat{L}_v^V \circ T^{\mathrm{alg}}, \tag{3.15.5}$$

by construction. If $\mu \in W^{\mathrm{alg}}$, then

$$(\widehat{L}_v^V \circ T^{\mathrm{alg}})(\mu) = \widehat{L}_v^V(T^{\mathrm{alg}}(\mu)) = \widehat{L}_v^V(\mu \circ T) = \mu(T(v)) = \widehat{L}_{T(v)}^W(\mu). \tag{3.15.6}$$

This means that

$$\widehat{L}_v^V \circ T^{\mathrm{alg}} = \widehat{L}_{T(v)}^W, \tag{3.15.7}$$

so that (3.15.4) holds.

Let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively, and suppose now that T is a bounded linear mapping from V into W. Using the dual mapping T' from

W' into V', we get the second dual mapping $T'' = (T')'$ from V'' into W''. We also have that

$$\|T''\|_{op,V''W''} = \|T'\|_{op,W'V'} = \|T\|_{op,VW}, \tag{3.15.8}$$

by (3.13.13). As before,

$$T \mapsto T'' \tag{3.15.9}$$

is a linear mapping from $\mathcal{BL}(V, W)$ into $\mathcal{BL}(V'', W'')$, because of the analogous statement in Section 3.13.

If $v \in V$, then let $L_v^V = L_v \in V''$ be as in (3.14.6). Similarly, if $w \in W$ and $\mu \in W'$, then

$$L_w^W(\mu) = \mu(w) \tag{3.15.10}$$

defines an element of W'', which is the same as the restriction of $\widehat{L}_w^W$ to W'. Of course, $v \mapsto L_v^V$ and $w \mapsto L_w^W$ are linear mappings from V and W into V'' and W'', respectively, as before. Note that

$$L_v^V \circ T' = L_{T(v)}^W \tag{3.15.11}$$

on W', as in (3.15.7). This implies that

$$T''(L_v^V) = L_v^V \circ T' = L_{T(v)}^W, \tag{3.15.12}$$

as in (3.15.4).

Chapter 4

Uniform boundedness and related topics

4.1 The Baire category theorem

Let X be a metric space, or a topological space, and let E be a subset of X. As usual, we say that E is *dense* in X if every element of X is an element of E, a limit point of E, or both. Equivalently, this means that

$$\text{the closure } \overline{E} \text{ of } E \text{ in } X \text{ is equal to } X. \tag{4.1.1}$$

It is easy to see that E is dense in X if and only if for every nonempty open subset V of X, we have that

$$E \cap V \neq \emptyset. \tag{4.1.2}$$

Sometimes this is used as the definition of a dense set, particularly in arbitrary topological spaces.

If E is a dense set in X and U is a dense open set in X, then one can use this characterization of density to check that

$$E \cap U \text{ is dense in } X. \tag{4.1.3}$$

If $U_1, \ldots, U_n$ are finitely many dense open sets in X, then one can use the previous statement to get that

$$\bigcap_{j=1}^{n} U_j \text{ is a dense open set in } X \tag{4.1.4}$$

too. If $U_1, U_2, U_3, \ldots$ is a sequence of dense open sets in X, then the *Baire category theorem* gives conditions under which

$$\bigcap_{j=1}^{\infty} U_j \text{ is dense in } X. \tag{4.1.5}$$

More precisely, this holds when X is a complete metric space, and also when X is a locally compact Hausdorff topological space.

It is easy to see that $A \subseteq X$ has empty interior if and only if

$$X \setminus A \text{ is dense in } X. \tag{4.1.6}$$

Thus A is a closed set in X with empty interior if and only if

$$X \setminus A \text{ is a dense open set in } X. \tag{4.1.7}$$

If $A_1, \ldots, A_n$ are finitely many closed sets in X with empty interior, then it follows that

$$\bigcup_{j=1}^{n} A_j \text{ is a closed set in } X \text{ with empty interior} \tag{4.1.8}$$

as well. The conclusion of the Baire category theorem is equivalent to saying that if $A_1, A_2, A_3, \ldots$ is a sequence of closed sets in X with empty interior, then

$$\bigcup_{j=1}^{\infty} A_j \text{ has empty interior in } X. \tag{4.1.9}$$

A subset A of X is said to be *nowhere dense* in X if

$$\text{the closure } \overline{A} \text{ of } A \text{ in } X \text{ has empty interior.} \tag{4.1.10}$$

If $A_1, \ldots, A_n$ are finitely many nowhere dense sets in X, then

$$\bigcup_{j=1}^{n} A_j \text{ is nowhere dense in } X. \tag{4.1.11}$$

This uses the well-known fact that

$$\overline{\Big(\bigcup_{j=1}^{n} A_j\Big)} = \bigcup_{j=1}^{n} \overline{A_j}. \tag{4.1.12}$$

A subset E of X is said to be of *first category* or *meager* in X if E can be expressed as

$$E = \bigcup_{j=1}^{\infty} A_j, \tag{4.1.13}$$

where $A_1, A_2, A_3, \ldots$ is a sequence of nowehere dense sets in X. Otherwise, E is said to be of *second category* or *non-meager* in X.

The conclusion of the Baire category theorem is the same as saying that if E is of first category in X, then

$$\text{the interior of } E \text{ is empty.} \tag{4.1.14}$$

This means that a subset of X with nonempty interior is of second category.

4.2 Pointwise and uniform boundedness

Let X be a nonempty set, let (Y, d_Y) be a nonempty metric space, and let $\mathcal{E}$ be a nonempty collection of mappings from X into Y. If $x \in X$, then put

$$\mathcal{E}(x) = \{f(x) : f \in \mathcal{E}\}. \tag{4.2.1}$$

We say that $\mathcal{E}$ is *pointwise bounded* on a subset A of X if for each $x \in A$,

$$\mathcal{E}(x) \text{ is a bounded subset of } Y, \tag{4.2.2}$$

with respect to d_Y.

Similarly, put

$$\mathcal{E}(A) = \{f(x) : x \in A,\ f \in \mathcal{E}\} \tag{4.2.3}$$

for each $A \subseteq X$. If this is a bounded set in Y, then we say that $\mathcal{E}$ is *uniformly bounded* on A. Of course, this implies that $\mathcal{E}$ is pointwise bounded on A.

Suppose for the moment that Y is the real line, equipped with the standard Euclidean metric. If n is a positive integer, then put

$$E_n = \{x \in X : |f(x)| \le n \text{ for every } f \in \mathcal{E}\}. \tag{4.2.4}$$

Thus

$$\bigcup_{n=1}^{\infty} E_n = \{x \in X : \mathcal{E}(x) \text{ is bounded in } \mathbf{R}\}. \tag{4.2.5}$$

Similarly, $\mathcal{E}$ is uniformly bounded on $A \subseteq X$ if and only if $A \subseteq E_n$ for some n.

Now let (X, d_X) be a nonempty metric space, or even a topological space. Suppose that each $f \in \mathcal{E}$ is a continuous real-valued function on X. This implies that

$$E_n \text{ is a closed set in } X \tag{4.2.6}$$

for each n. If (4.2.5) is of second category in X, then

$$E_n \text{ has nonempty interior} \tag{4.2.7}$$

for some n. This means that

$$\mathcal{E} \text{ is uniformly bounded on a nonempty open subset of } X. \tag{4.2.8}$$

Let Y be any nonempty metric space again, and let y_0 be an element of Y. Remember that $d_Y(y, y_0)$ is Lipschitz with constant 1 as a function of $y \in Y$, as in Section 2.1. If f is a continuous mapping from X into Y, then it follows that

$$F(x) = d_Y(f(x), y_0) \tag{4.2.9}$$

is a continuous real-valued function on X.

Let $\mathcal{E}_0$ be the collection of real-valued functions on X of the form (4.2.9), with $f \in \mathcal{E}$. If $x \in X$, then it is easy to see that $\mathcal{E}(x)$ is a bounded subset of Y if and only if

$$\mathcal{E}_0(x) \text{ is a bounded set in } \mathbf{R}. \tag{4.2.10}$$

Similarly, if $A \subseteq X$, then

$$\mathcal{E}(A) \text{ is bounded in } Y \tag{4.2.11}$$

if and only if

$$\mathcal{E}_0(A) \text{ is bounded in } \mathbf{R}. \tag{4.2.12}$$

If the set of $x \in X$ such that $\mathcal{E}_0(x)$ is bounded in $\mathbf{R}$ is of second category in X, then

$$\mathcal{E}_0 \text{ is uniformly bounded on a nonempty open subset of } X, \tag{4.2.13}$$

as before. This means that if the set of $x \in X$ such that $\mathcal{E}(x)$ is bounded in Y is of second category in X, then $\mathcal{E}$ is uniformly bounded on a nonempty open subset of X.

4.3 The Banach–Steinhaus theorem

Let V, W be vector spaces, both real or both complex, and with norms $\|\cdot\|_V$, $\|\cdot\|_W$, respectively. Also let $\mathcal{E}$ be a collection of bounded linear mappings from V into W. If $v \in V$, then put

$$\mathcal{E}(v) = \{T(v) : T \in \mathcal{E}\}, \tag{4.3.1}$$

as in the previous section.

Suppose that

$$\{v \in V : \mathcal{E}(v) \text{ is bounded in } W\} \tag{4.3.2}$$

is of second category in V, using the metrics on V, W associated to the norms. Note that this holds when V is a Banach space, and (4.3.2) is equal to V. This implies that $\mathcal{E}$ is uniformly bounded on a nonempty open subset of V, as before.

Equivalently, this means that $\mathcal{E}$ is uniformly bounded on an open ball in V of positive radius. One can reduce to the case where the ball is centered at 0, using linearity of the elements of $\mathcal{E}$. Similarly, one can reduce to the case where the radius of the ball is one, using linearity.

It follows that the operator norms of the elements of $\mathcal{E}$ are bounded, so that

$$\|T\|_{op,VW} \le C \tag{4.3.3}$$

for some nonnegative real number C and every $T \in \mathcal{E}$. This is the *Banach–Steinhaus theorem*, which is also known as the *uniform boundedness principle*. Note that this condition implies that $\mathcal{E}$ is pointwise bounded on V, and uniformly bounded on bounded subsets of V.

One may also consider (4.3.3) as an equicontinuity property of $\mathcal{E}$. It is the same as saying that each $T \in \mathcal{E}$ is Lipschitz with constant C as a mapping from V into W, with respect to the metrics associated to their norms.

4.4 Pointwise convergence

Let V, W be vector spaces, both real or both complex, with norms $\|\cdot\|_V$, $\|\cdot\|_W$, respectively. Also let $\{T_j\}_{j=1}^\infty$ be a sequence of bounded linear mappings from V into W. Consider

$$\{v \in V : \{T_j(v)\}_{j=1}^\infty \text{ is a bounded sequence in } W\}. \tag{4.4.1}$$

It is easy to see that this is a linear subspace of V. If (4.4.1) is of second category in V, with respect to the metric associated to $\|\cdot\|_V$, then it follows that $\{T_j\}_{j=1}^\infty$ is bounded with respect to the operator norm, by the Banach–Steinhaus theorem.

Similarly, consider

$$\{v \in V : \{T_j(v)\}_{j=1}^\infty \text{ is a Cauchy sequence in } W\}. \tag{4.4.2}$$

If W is a Banach space, then this is the same as

$$\{v \in V : \{T_j(v)\}_{j=1}^\infty \text{ converges in } W\}. \tag{4.4.3}$$

It is well known and easy to see that Cauchy sequences in metric spaces are bounded, so that (4.4.2) is contained in (4.4.1). Convergent sequences in metric spaces are Cauchy sequences, so that (4.4.3) is contained in (4.4.2). Note that (4.4.2) and (4.4.3) are linear subspaces of V.

If $v \in V$ is an element of (4.4.3), then put

$$T(v) = \lim_{j\to\infty} T_j(v). \tag{4.4.4}$$

This defines a linear mapping from (4.4.3) into W.

Suppose that $\{T_j\}_{j=1}^\infty$ is bounded with respect to the operator norm, so that

$$\|T_j\|_{op,VW} \le C \tag{4.4.5}$$

for some nonnegative real number C and each $j \ge 1$. In this case, one can check that (4.4.2) is a closed set in V, with respect to the metric associated to $\|\cdot\|_V$. In particular, if (4.4.2) is dense in V, then it follows that (4.4.2) is equal to V.

Suppose for the moment that T is any linear mapping from V into W. Consider

$$\{v \in V : \{T_j(v)\}_{j=1}^\infty \text{ converges to } T(v) \text{ in } W\}. \tag{4.4.6}$$

This is another linear subspace of V. If $\{T_j\}_{j=1}^\infty$ is bounded with respect to the operator norm, and T is a bounded linear mapping from V into W, then one can verify that (4.4.6) is a closed set in V, with respect to the metric associated to $\|\cdot\|_V$.

Suppose again that $\{T_j\}_{j=1}^\infty$ is bounded with respect to the operator norm, so that (4.4.5) holds for some $C \ge 0$ and each j. Equivalently, this means that

$$\|T_j(v)\|_W \le C\,\|v\|_V \tag{4.4.7}$$

for each $v \in V$ and $j \geq 1$. Suppose also that (4.4.3) is equal to V, so that (4.4.4) defines a linear mapping from V into W. Under these conditions, we have that

$$\|T(v)\|_W \leq C \, \|v\|_V \tag{4.4.8}$$

for every $v \in V$. This is the same as saying that T is bounded as a linear mapping from V into W, with $\|T\|_{op,VW} \leq C$.

Of course, if $\{T_j\}_{j=1}^\infty$ converges to a bounded linear mapping T from V into W with respect to the metric associated to the operator norm, then $\{T_j\}_{j=1}^\infty$ is bounded with respect to the operator norm, and $\{T_j\}_{j=1}^\infty$ converges to T pointwise on V. Pointwise convergence of $\{T_j\}_{j=1}^\infty$ to a bounded linear mapping T from V into W is equivalent to the convergence of $\{T_j\}_{j=1}^\infty$ to T with respect to the *strong operator topology* on the space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W. We shall not discuss this topology in detail for the moment.

4.5 A sequential compactness theorem

Let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$, and let V' be the dual of V with respect to $\|\cdot\|_V$, as usual. Pointwise convergence of a sequence of bounded linear functionals $\{\lambda_j\}_{j=1}^\infty$ on V to another bounded linear functional λ on V is equivalent to the convergence of $\{\lambda_j\}_{j=1}^\infty$ to λ with respect to the *weak* topology* on V'. This may be described as the weakest topology on V' such that

$$\lambda \mapsto \lambda(v) \tag{4.5.1}$$

is continuous for every $v \in V$. This corresponds to taking $W = \mathbf{R}$ or $\mathbf{C}$, as appropriate, in the previous section, using the standard absolute value function as the norm.

This terminology is related to the common use of V^* for the dual of V with respect to $\|\cdot\|_V$, instead of V'. The notation V' is being used here to be consistent with the notation T' for the dual of a bounded linear mapping T, as compared to the notation T^* for the adjoint of a bounded linear mapping T between Hilbert spaces.

Let $\{\lambda_j\}_{j=1}^\infty$ be a bounded sequence in V' with respect to the dual norm $\|\cdot\|_{V'}$. If $v \in V$, then $\{\lambda_j(v)\}_{j=1}^\infty$ is a bounded sequence of real or complex numbers, as appropriate, and there is a subsequence of $\{\lambda_j(v)\}_{j=1}^\infty$ that converges, by a well-known result. Let E be a subset of V with only finitely or countably many elements. Under these conditions, there is a subsequence $\{\lambda_{j_l}\}_{l=1}^\infty$ of $\{\lambda_j\}_{j=1}^\infty$ such that

$$\{\lambda_{j_l}(v)\}_{l=1}^\infty \text{ converges in } \mathbf{R} \text{ or } \mathbf{C}, \tag{4.5.2}$$

as appropriate, for every $v \in E$. This can be obtained from the previous statement and standard arguments.

It follows that (4.5.2) holds for every element v of the linear span of E in V, by linearity. In fact, (4.5.2) holds for every element v of the closure of the linear span of E in V, with respect to the metric associated to $\|\cdot\|_V$, as in the previous section. Suppose now that the linear span of E is dense in V, so that (4.5.2) holds for every $v \in V$. This means that

(4.5.3) $\{\lambda_{j_l}\}_{l=1}^{\infty}$ converges pointwise to a bounded linear functional on V,

as before. Remember that there is a subset E of V with these properties exactly when V is separable with respect to the metric associated to $\|\cdot\|_V$, as in Section 1.14.

The *Banach–Alaoglu theorem* states that

(4.5.4) the closed unit ball in V' is compact with respect to the weak* topology.

The remarks in the previous paragraphs correspond to a variant of this, namely, the sequential compactness of the closed unit ball in V' with respect to the weak* topology when V is separable. Of course, the compactness or sequential compactness of the closed unit ball in V' with respect to the weak* topology implies the analogous property of any closed ball in V'. If V is separable, then it is not too difficult to show that the topology induced on the closed unit ball in V' by the weak* topology is determined by a metric. In this case, compactness and sequential compactness of the closed unit ball in V' with respect to the weak* topology are the same.

4.6 A criterion for density

Let W be a vector space over the real or complex numbers with a norm $\|\cdot\|_W$, and let W_0 be a linear subspace of W. Suppose that there is a real number a, $0 \le a < 1$, such that for every $w \in W$ there is a $w_0 \in W_0$ with

$$\|w - w_0\|_W \le a \|w\|_W. \tag{4.6.1}$$

It is not too difficult to show that W_0 is dense in W with respect to the metric associated to $\|\cdot\|_W$. More precisely, one can use the same condition with w replaced by $w - w_0$ to get a $w_1 \in W_0$ such that

$$\|w - w_0 - w_1\|_W \le a\,\|w - w_0\|_W \le a^2\,\|w\|_W. \tag{4.6.2}$$

Continuing in this way, we get that w is in the closure of W_0 in W.

Let $0 \le a < 1$ be given, and let E_a be a subset of W such that for every $u \in W$ with $\|u\|_W = 1$ there is a $u_0 \in E_a$ with

$$\|u - u_0\|_W \le a. \tag{4.6.3}$$

Let W_0 be the linear span of E_a in W, which may be interpreted as being $\{0\}$ when $W = \{0\}$. It is easy to see that W_0 satisfies the condition mentioned in the preceding paragraph. This implies that W_0 is dense in W, as before.

If W_0 is a finite-dimensional linear subspace of W, then W_0 is complete with respect to the metric associated to the restriction of $\|\cdot\|_W$ to W_0, as in Section

1.11. This implies that W_0 is a closed set in W, with respect to the metric associated to $\| \cdot \|_W$, as in Section 1.6. If W_0 is the linear span of a subset E_a of W as in the preceding paragraph, and if E_a has only finitely many elements, then W_0 has finite dimension, and it follows that $W_0 = W$. In particular, this means that

$$W \text{ has finite dimension} \tag{4.6.4}$$

under these conditions.

A subset A of a metric space X is said to be *totally bounded* if for every $r > 0$, A can be covered by finitely many balls of radius r. It is easy to see that compact subsets of X are totally bounded. If the closed unit sphere in W is totally bounded with respect to the metric associated to $\| \cdot \|_W$, then (4.6.4) holds, as before.

4.7 Pointwise convergence in ℓ^q spaces

Let X be a nonempty set, and let q be a positive extended real number. Suppose that $\{g_j\}_{j=1}^\infty$ is a bounded sequence in $\ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, so that

$$\|g_j\|_q \le C \tag{4.7.1}$$

for some nonnegative real number C and each $j \ge 1$. Suppose also that $\{g_j\}_{j=1}^\infty$ converges pointwise to a real or complex-valued function g on X, as appropriate. Under these conditions, we have that $g \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, as appropriate, with

$$\|g\|_q \le C. \tag{4.7.2}$$

More precisely, if $q = +\infty$, then (4.7.1) says that

$$|g_j(x)| \le C \tag{4.7.3}$$

for each $x \in X$ and $j \ge 1$. This implies that

$$|g(x)| \le C \tag{4.7.4}$$

for every $x \in X$, so that (4.7.2) holds. If $q < \infty$, then (4.7.1) is the same as saying that

$$\sum_{x \in X} |g_j(x)|^q \le C^q \tag{4.7.5}$$

for each j. It follows that

$$\sum_{x \in X} |g(x)|^q \le C^q, \tag{4.7.6}$$

as in Section 2.4.

Let $\{g_j\}_{j=1}^\infty$ be a bounded sequence in $\ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$ again. If $x \in X$, then $\{g_j(x)\}_{j=1}^\infty$ is a bounded sequence of real or complex numbers. If $A \subseteq X$

has only finitely or countably many elements, then it follows that there is a subsequence $\{g_{j_l}\}_{l=1}^\infty$ of $\{g_j\}_{j=1}^\infty$ such that

$$\{g_{j_l}(x)\}_{l=1}^\infty \text{ converges in } \mathbf{R} \text{ or } \mathbf{C}, \tag{4.7.7}$$

as appropriate, for every $x \in A$. This uses the same type of argument as for (4.5.2). Of course, if X has only finitely or countably many elements, then one can simply take $A = X$.

If $q < \infty$, then the support of g_j has only finitely or countably many elements for each j. This implies that the union of the supports of the g_j's has only finitely or countably many elements as well. If we take

$$A = \bigcup_{j=1}^\infty \operatorname{supp} g_j, \tag{4.7.8}$$

then we get a subsequence $\{g_{j_l}\}_{l=1}^\infty$ of $\{g_j\}_{j=1}^\infty$ such that (4.7.7) holds for every $x \in A$, as before. In this case, it follows that (4.7.7) holds for every $x \in X$.

The space of all real or complex-valued functions on X is the same as the Cartesian product of a family of copies of $\mathbf{R}$ or $\mathbf{C}$, as appropriate, indexed by X. Let us consider the corresponding product topology on this space, using the standard topology on each factor of $\mathbf{R}$ or $\mathbf{C}$. Closed balls in $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$ are compact sets with respect to the product topology, by Tychonoff's theorem. If $q < \infty$, then one can check that closed balls in $\ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$ are closed sets with respect to the product topology. This implies that they are compact sets with respect to the product topology as well, because closed sets contained in compact sets are compact.

4.8 Weak* convergence in ℓ^q spaces

Let X be a nonempty set, and let $1 \le p, q \le \infty$ be conjugate exponents, so that $1/p + 1/q = 1$. Also let $\{g_j\}_{j=1}^\infty$ be a bounded sequence in $\ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, so that (4.7.1) holds for some $C \ge 0$ and all $j \ge 1$. Note that

$$\lambda_{g_j}(f) = \sum_{x \in X} f(x)\, g_j(x) \tag{4.8.1}$$

defines a bounded linear functional on $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, for each j, with dual norm equal to $\|g_j\|_q$, as in Sections 3.2, 3.3, and 3.4.

Suppose that $\{g_j\}_{j=1}^\infty$ converges pointwise to a real or complex-valued function g on X, as appropriate. Thus $g \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, as appropriate, as in the previous section, so that

$$\lambda_g(f) = \sum_{x \in X} f(x)\, g(x) \tag{4.8.2}$$

defines a bounded linear functional on $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate. If f has finite support in X, then it is easy to see that

$$\lim_{j \to \infty} \lambda_{g_j}(f) = \lambda_g(f). \tag{4.8.3}$$

The set of $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, such that (4.8.3) holds is a closed set with respect to the metric associated to the ℓ^p norm, as in Section 4.4. If $q > 1$, so that $p < \infty$, then it follows that (4.8.3) holds for all $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate. This uses the fact that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $\ell^p(X, \mathbf{R})$, $\ell^p(X, \mathbf{C})$, respectively, when $p < \infty$, as in Section 2.6.

If $q = 1$, then (4.8.1) and (4.8.2) may also be considered as bounded linear functionals on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, with respect to the supremum norm, as in Section 3.3. In this case, (4.8.3) holds for all real or complex-valued functions f on X, as appropriate, that vanish at infinity. This is because $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$, respectively, with respect to the supremum metric, as in Section 1.13.

4.9 Weak convergence

Let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$, and let V' be the corresponding dual space. A sequence $\{v_j\}_{j=1}^\infty$ of elements of V is said to *converge weakly* to $v \in V$ if

$$\lim_{j \to \infty} \lambda(v_j) = \lambda(v) \tag{4.9.1}$$

for every $\lambda \in V'$. This is equivalent to the convergence of $\{v_j\}_{j=1}^\infty$ to v with respect to the *weak topology* on V. This may be described as the weakest topology on V with respect to which every $\lambda \in V'$ is continuous.

If $\{v_j\}_{j=1}^\infty$ converges to v with respect to the metric associated to the norm, then $\{v_j\}_{j=1}^\infty$ converges weakly to v. Similarly, the topology determined on V by the metric associated to the norm is at least as strong as the weak topology on V, because every $\lambda \in V'$ is continuous on V with respect to the metric associated to the norm.

Remember that V' separates points in V, as in Section 3.10. If a sequence in V has a weak limit, then one can use this to get that the limit is unique. Similarly, one can use this to get that the weak topology on V is Hausdorff.

Let $V'' = (V')'$ be the dual of V' with respect to the dual norm $\|\cdot\|_{V'}$, as in Section 3.14. If $v \in V$, then $L_v(\lambda) = \lambda(v)$ defines an element of V'', as before. Using this notation, (4.9.1) is the same as saying that

$$\lim_{j \to \infty} L_{v_j}(\lambda) = L_v(\lambda) \tag{4.9.2}$$

for every $\lambda \in V'$. This means that $\{v_j\}_{j=1}^\infty$ converges to v weakly in V if and only if $\{L_{v_j}\}_{j=1}^\infty$ converges to L_v in the weak* sense in V'', as the dual of V'.

Let E be a subset of V, and put

$$\mathcal{E}_E = \{L_v : v \in E\}, \tag{4.9.3}$$

which is a subset of V''. If $\lambda \in V'$, then put

$$\mathcal{E}_E(\lambda) = \{L_v(\lambda) : v \in E\}, \tag{4.9.4}$$

as in Section 4.3. Equivalently, $\mathcal{E}_E(\lambda) = \lambda(E)$.

Let us say that E is *weakly bounded* in V if $\lambda(E)$ is a bounded set in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for each $\lambda \in V'$. This is the same as saying that $\mathcal{E}_E$ is pointwise bounded on V'. Remember that V' is complete with respect to the metric associated to $\|\cdot\|_{V'}$, as in Sections 2.2 and 3.1. Let $\|\cdot\|_{V''}$ be the dual norm on V'' associated to $\|\cdot\|_{V'}$ on V', as in Section 3.14. If $\mathcal{E}_E$ is pointwise bounded on V', then it follows that $\mathcal{E}_E$ is bounded with respect to $\|\cdot\|_{V''}$, because of the Baire category theorem and the Banach–Steinhaus theorem.

Remember that $\|L_v\|_{V''} = \|v\|_V$ for every $v \in V$, as in Section 3.14. If E is weakly bounded in V, then we get that E is bounded with respect to $\|\cdot\|_V$, using the remarks in the preceding paragraph.

If $\{v_j\}_{j=1}^\infty$ is a sequence of elements of V that converges weakly in V, then it is easy to see that the set of v_j's, $j \geq 1$, is weakly bounded in V. This implies that $\{v_j\}_{j=1}^\infty$ is bounded with respect to $\|\cdot\|_V$, as before.

Suppose that

$$\|v_j\|_V \leq C \tag{4.9.5}$$

for some nonnegative real number C and each $j \geq 1$. This implies that

$$|\lambda(v_j)| \leq C\,\|\lambda\|_{V'} \tag{4.9.6}$$

for each $\lambda \in V'$ and $j \geq 1$. If $\{v_j\}_{j=1}^\infty$ converges weakly to $v \in V$, then it follows that

$$|\lambda(v)| \leq C\,\|\lambda\|_{V'} \tag{4.9.7}$$

for every $\lambda \in V'$. Using this, we get that

$$\|v\|_V \leq C, \tag{4.9.8}$$

as in Sections 3.10 and 3.12.

4.10 Some multiplication operators

Let X be a nonempty set, and let p, q, and r be positive extended real numbers. Suppose that

$$1/p + 1/q = 1/r, \tag{4.10.1}$$

with suitable interpretations when any of p, q, or r is $+\infty$. Let $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$ and $g \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$ be given. Under these conditions, $f\,g \in \ell^r(X, \mathbf{R})$ or $\ell^r(X, \mathbf{C})$, as appropriate, with

$$\|f\,g\|_r \leq \|f\|_p\,\|g\|_q. \tag{4.10.2}$$

This can be verified directly when p or q is $+\infty$.

Suppose that $p, q < +\infty$, which implies that $r < +\infty$. In this case, $p, q > r$, so that $p/r, q/r > 1$, and (4.10.1) is the same as saying that

$$r/p + r/q = 1. \tag{4.10.3}$$

Observe that $|f|^r$ is (p/r)-summable on X, and that $|g|^r$ is (q/r)-summable on X. Hölder's inequality implies that $|f|^r \, |g|^r$ is summable on X, with

$$\sum_{x \in X} |f(x)|^r \, |g(x)|^r \le \|(|f|^r)\|_{p/r} \, \|(|g|^r)\|_{q/r}. \tag{4.10.4}$$

This is the same as saying that $|f| \, |g|$ is r-summable on X, and that (4.10.2) holds.

Let a be a real or complex-valued function on X. If f is another real or complex-valued function on X, as appropriate, then

$$M_a(f) = a \, f \tag{4.10.5}$$

defines real or complex-valued function on X too. This defines a linear mapping from the space of real or complex-valued functions on X, as appropriate, into itself, which is the *multiplication operator* associated to a. Note that if f has finite support in X, then $M_a(f)$ has finite support as well. Thus the restriction of M_a to $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate, defines a linear mapping from that space into itself.

Suppose for the moment that $a \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$ for some $q > 0$. If $p, r > 0$ satisfy (4.10.1), then

$$M_a \text{ maps } \ell^p(X, \mathbf{R}) \text{ or } \ell^p(X, \mathbf{C}) \text{ into } \ell^r(X, \mathbf{R}) \text{ or } \ell^r(X, \mathbf{C}), \tag{4.10.6}$$

as appropriate, as before. We also have that

$$\|M_a(f)\|_r \le \|a\|_q \, \|f\|_p \tag{4.10.7}$$

for every $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, by (4.10.2). If $q = +\infty$, then it is easy to see that

$$M_a \text{ maps } c_0(X, \mathbf{R}) \text{ or } c_0(X, \mathbf{C}) \text{ into itself} \tag{4.10.8}$$

too, as appropriate.

Suppose now that a is a real or complex-valued function on X such that (4.10.6) holds, with

$$\|M_a(f)\|_r \le C \, \|f\|_p \tag{4.10.9}$$

for some $C \ge 0$ and all $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate. More precisely, if this holds for all real or complex-valued functions f on X with finite support, as appropriate, then it is easy to see that this holds for all $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate. We would like to check that $a \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, as appropriate, with

$$\|a\|_q \le C. \tag{4.10.10}$$

If $p = +\infty$, then $q = r$, and this can be verified directly, by taking $f \equiv 1$ on X. One can also consider functions f that are equal to 1 on a finite subset of X and to 0 elsewhere, so that f has finite support in X. If $q = +\infty$, then $p = r$, and this can be verified directly again, by considering functions f equal to 1 at one point in X, and to 0 elsewhere.

Suppose that $p, q < +\infty$, so that $r < +\infty$ and $p, q > r$, as before. Let A be a nonempty finite subset of X, and let f_A be a real or complex-valued function on X, as appropriate, such that

$$|f_A(x)| = |a(x)|^{(q-r)/r} \tag{4.10.11}$$

for each $x \in A$, and $f_A(x) = 0$ when $x \in X \setminus A$. Thus

$$|f_A(x)|^r \, |a(x)|^r = |a(x)|^q \tag{4.10.12}$$

for every $x \in A$. This implies that

$$\|M_a(f_A)\|_r = \|a\, f_A\|_r = \Big(\sum_{x \in A} |a(x)|^q\Big)^{1/r}. \tag{4.10.13}$$

Observe that

$$r/p = 1 - r/q = (q - r)/q, \tag{4.10.14}$$

so that $p/r = q/(q - r)$. It follows that

$$|f_A(x)|^p = |a(x)|^{((q-r)\,p)/r} = |a(x)|^q \tag{4.10.15}$$

for every $x \in A$. This means that

$$\|f_A\|_p = \Big(\sum_{x \in A} |a(x)|^q\Big)^{1/p}. \tag{4.10.16}$$

Combining this with (4.10.13), we get that

$$\Big(\sum_{x \in A} |a(x)|^q\Big)^{1/q} \le C. \tag{4.10.17}$$

This implies that $a \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, as appropriate, and that (4.10.10) holds.

If $r \ge 1$, then $p, q \ge 1$, and it follows that M_a defines a bounded linear mapping from $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$ into $\ell^r(X, \mathbf{R})$ or $\ell^r(X, \mathbf{C})$, as appropriate, with operator norm equal to $\|a\|_q$. There is an analogous statement when $r < 1$, using suitable versions of some of our previous definitions, and related facts.

4.11 Convergence of multiplication operators

Let X be a nonempty set, and remember that $\ell^p(X, \mathbf{R})$, $\ell^p(X, \mathbf{C})$ may be considered as metric spaces when $0 < p \le +\infty$. This uses the metric associated to the ℓ^p norm when $p \ge 1$, and the analogous metric mentioned in Section 2.6 when $p < 1$.

Let $\{a_j\}_{j=1}^{\infty}$ be a sequence of real or complex-valued functions on X that converges pointwise to a real or complex-valued function a on X, as appropriate. If f is any real or complex-valued function on X, as appropriate, then

$$\{M_{a_j}(f)\}_{j=1}^{\infty} \text{ converges to } M_a(f) \text{ pointwise on } X. \tag{4.11.1}$$

Let p, q, r be positive extended real numbers satisfying (4.10.1) again. Suppose that $a_j \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$ for each j, as appropriate, with

$$\|a_j\|_q \le C \tag{4.11.2}$$

for some $C \ge 0$ and each j. This implies that $a \in \ell^q(X, \mathbf{R})$ or $\ell^q(X, \mathbf{C})$, as appropriate, with

$$\|a\|_q \le C, \tag{4.11.3}$$

as in Section 4.7. If $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, then we would like to have conditions under which

$$\{M_{a_j}(f)\}_{j=1}^{\infty} \text{ converges to } M_a(f) \text{ in } \ell^r(X, \mathbf{R}) \text{ or } \ell^r(X, \mathbf{C}), \tag{4.11.4}$$

as appropriate. It is easy to see that this holds when

$$\{a_j\}_{j=1}^{\infty} \text{ converges to } a \text{ in } \ell^q(X, \mathbf{R}) \text{ or } \ell^q(X, \mathbf{C}), \tag{4.11.5}$$

as appropriate, because of (4.10.2).

Of course, the set of $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$ such that (4.11.4) holds is a linear subspace of $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{R})$, as appropriate. This is also a closed set in $\ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate, because of (4.11.2) and (4.11.3), as in Section 4.4.

Note that (4.11.4) holds when f has finite support in X. If $p < +\infty$, then it follows that (4.11.4) holds for every $f \in \ell^p(X, \mathbf{R})$ or $\ell^p(X, \mathbf{C})$, as appropriate. This uses the fact that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $\ell^p(X, \mathbf{R})$, $\ell^p(X, \mathbf{C})$, respectively, when $p < +\infty$, as in Section 2.6.

Suppose that $p = +\infty$, so that $q = r$. If f is equal to 1 at every point in X, then (4.11.4) implies (4.11.5).

If f vanishes at infinity on X, then (4.11.4) holds, with $q = r$. This is because $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$, respectively, with respect to the supremum metric, as in Section 1.13.

4.12 L^p Spaces

Let $(X, \mathcal{A}, \mu)$ be a (nonempty) measure space, so that X is a nonempty set, $\mathcal{A}$ is a σ-algebra of measurable subsets of X, and μ is a countably additive (nonnegative) measure on the measurable space $(X, \mathcal{A})$. If p is a positive real number, then let $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$ be the corresponding spaces of real and complex-valued measurable functions f on X, as appropriate, such that $|f|^p$ is integrable with respect to μ. More precisely, these spaces consist of equivalence classes of functions that are equal almost everywhere on X with respect to μ. If $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, then we put

$$\|f\|_p = \Big(\int_X |f(x)|^p \, d\mu(x) \Big)^{1/p}, \tag{4.12.1}$$

as usual.

Similarly, let $L^\infty(X, \mathbf{R})$ and $L^\infty(X, \mathbf{C})$ be the spaces of equivalence classes of real and complex-valued measurable functions f on X that are essentially bounded on X with respect to μ. In this case, $\|f\|_\infty$ is the essential supremum norm of f with respect to μ.

It is well known that $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$ are vector spaces over the real and complex numbers, respectively, for each $p > 0$. If $p \ge 1$, then $\|f\|_p$ defines a norm on each of $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$. In this case, the triangle inequality for $\| \cdot \|_p$ is known as *Minkowski's inequality* for integrals.

If $0 < p \le 1$, then $\|f\|_p$ satisfies the usual homogeneity property of a norm, and

$$\|f + g\|_p^p \le \|f\|_p^p + \|g\|_p^p \tag{4.12.2}$$

for all $f, g \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$. This means that

$$\|f - g\|_p^p \tag{4.12.3}$$

defines a metric on each of $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$ when $p \le 1$.

It is well known that $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$ are complete with respect to the appropriate metric for each $p > 0$. In particular, $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$ are Banach spaces when $p \ge 1$.

Of course, $\ell^p(X, \mathbf{R})$ and $\ell^p(X, \mathbf{C})$ are the same as $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$, respectively, when μ is counting measure on X. In this case, one can take all subsets of X to be measurable.

If $f, g \in L^2(X, \mathbf{R})$ or $L^2(X, \mathbf{C})$, then it is well known that $f\, g \in L^1(X, \mathbf{R})$ or $L^1(X, \mathbf{C})$, as appropriate, with

$$\|f\, g\|_1 \le \|f\|_2 \, \|g\|_2. \tag{4.12.4}$$

This is the integral version of the *Cauchy–Schwarz inequality*. Put

$$\langle f, g \rangle = \int_X f(x)\, g(x)\, d\mu(x) \tag{4.12.5}$$

in the real case, and

$$\langle f, g \rangle = \int_X f(x)\, \overline{g(x)}\, d\mu(x) \tag{4.12.6}$$

in the complex case. These define inner products on $L^2(X, \mathbf{R})$ and $L^2(X, \mathbf{C})$, respectively, for which the corresponding norm is $\| \cdot \|_2$. Thus $L^2(X, \mathbf{R})$ and $L^2(X, \mathbf{C})$ are Hilbert spaces over the real and complex numbers, respectively.

Suppose that $1 \leq p, q \leq \infty$ are conjugate exponents, so that $1/p + 1/q = 1$. If $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ and $g \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, then $f\,g \in L^1(X, \mathbf{R})$ or $L^1(X, \mathbf{C})$, as appropriate, with

$$\|f\,g\|_1 \leq \|f\|_p\,\|g\|_q. \tag{4.12.7}$$

This is the integral version of *Hölder's inequality*.

If $g \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, then put

$$\lambda_g(f) = \int_X f(x)\,g(x)\,d\mu(x) \tag{4.12.8}$$

for all $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate. Of course,

$$|\lambda_g(f)| \leq \|f\|_p\,\|g\|_q \tag{4.12.9}$$

for all $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate, by Hölder's inequality. Thus λ_g is a bounded linear functional on $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate, with dual norm less than or equal to $\|g\|_q$. Note that

$$g \mapsto \lambda_g \tag{4.12.10}$$

defines a linear mapping from each of $L^q(X, \mathbf{R})$, $L^q(X, \mathbf{C})$ into the duals of $L^p(X, \mathbf{R})$, $L^p(X, \mathbf{C})$, respectively.

If $g \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$ and $1 \leq q < \infty$, then one can check directly that

(4.12.11) the dual norm of λ_g with respect to the L^p norm is equal to $\|g\|_q$.

This also works when $q = \infty$, under the additional condition that

(4.12.12) every measurable subset of X of positive measure contain a measurable set of positive finite measure.

Note that this condition holds when X is σ-finite with respect to μ. This condition also holds when μ is counting measure on X, and all subsets of X are measurable.

More precisely, suppose that $g \in L^\infty(X, \mathbf{R})$ or $L^\infty(X, \mathbf{C})$, with $\|g\|_\infty > 0$, and let r be a nonnegative real number such that $\|g\|_\infty > r$. This implies that the set where $|g| > r$ has positive measure with respect to μ, by the definition of the essential supremum norm of g. If (4.12.12) holds, then there is a measurable subset of X with positive finite measure on which $|g| > r$. One can use this to get that (4.12.11) holds with $q = \infty$.

If $1 < p < \infty$, then it is well known that every bounded linear functional on $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ is of this form. This follows from the analogous statement for Hilbert spaces when $p = 2$. If $p = 1$, then the analogous statement holds when X is σ-finite with respect to μ.

If $\mu(X)$ is finite and $0 < r < p \leq \infty$, then it is well known that $L^p(X)$ is contained in $L^r(X)$. A more precise version of this will be mentioned in Section 4.15.

4.13 A measure-theoretic lemma

Let $(X, \mathcal{A}, \mu)$ be a measure space, and let E be a measurable subset of X. Suppose that every measurable subset of E of positive measure contains a measurable set of positive finite measure. We would like to check that

$$\mu(E) = \sup\{\mu(A) : A \subseteq E,\ A \text{ measurable},\ \mu(A) < \infty\}. \tag{4.13.1}$$

Of course, the supremum on the right is automatically less than or equal to $\mu(E)$. If $\mu(E) < \infty$, then (4.13.1) holds trivially.

Let $\{A_j\}_{j=1}^\infty$ be a sequence of measurable subsets of E with finite measure such that $\{\mu(A_j)\}_{j=1}^\infty$ converges to the supremum on the right side of (4.13.1). Put

$$B_l = \bigcup_{j=1}^{l} A_j \tag{4.13.2}$$

for each $l \geq 1$, which is a measurable subset of E with finite measure. Note that $\mu(B_l)$ is less than or equal to the right side of (4.13.1) for each l. We also have that

$$\mu(A_l) \leq \mu(B_l) \tag{4.13.3}$$

for each l, by construction. It follows that $\{\mu(B_l)\}_{l=1}^\infty$ converges to the supremum on the right side of (4.13.1) as well.

Put

$$B = \bigcup_{l=1}^{\infty} B_l, \tag{4.13.4}$$

which is a measurable subset of E. Of course, $B_l \subseteq B_{l+1}$ for each l, by construction. Thus

$$\mu(B_l) \to \mu(B) \text{ as } l \to \infty, \tag{4.13.5}$$

by a standard argument. This means that $\mu(B)$ is equal to the right side of (4.13.1).

If the supremum on the right side of (4.13.1) is $+\infty$, then equality holds trivially. Thus we may suppose that the supremum is finite, which means that $\mu(B) < +\infty$. In this case, we would like to check that

$$\mu(E \setminus B) = 0. \tag{4.13.6}$$

This will imply that

$$\mu(E) = \mu(B), \tag{4.13.7}$$

so that (4.13.1) holds.

Suppose for the same of a contradiction that

$$\mu(E \setminus B) > 0. \tag{4.13.8}$$

This implies that there is a measurable set $C \subseteq E \setminus B$ such that $\mu(C)$ is positive and finite, by hypothesis. It follows that $B \cup C$ is a measurable subset of E such that

$$\mu(B \cup C) = \mu(B) + \mu(C) > \mu(B). \tag{4.13.9}$$

This contradicts the fact that $\mu(B)$ is equal to the right side of (4.13.1), because $\mu(B \cup C) < \infty$.

4.14 Some more multiplication operators

Let $(X, \mathcal{A}, \mu)$ be a nonempty measure space, and let p, q, and r be positive extended real numbers such that $1/p + 1/q = 1/r$. If $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ and $g \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, then $f\,g \in L^r(X, \mathbf{R})$ or $L^r(X, \mathbf{C})$, as appropriate, with

$$\|f\,g\|_r \le \|f\|_p\,\|g\|_q. \tag{4.14.1}$$

This can be obtained from the integral version of Hölder's inequality, as in Section 4.10.

Let a be a real or complex-valued measurable function on X. If f is another real or complex-valued measurable function on X, as appropriate, then

$$M_a(f) = a\,f \tag{4.14.2}$$

defines a real or complex-valued measurable function on X as well. This defines a linear mapping from the space of real or complex-valued measurable functions on X, as appropriate, into itself, which is the *multiplication operator* associated to a. We shall use the same notation for the induced linear mapping on the corresponding space of equivalence classes of functions that are equal almost everywhere on X with respect to μ. Note that this induced linear mapping only depends on the analogous equivalence class that contains a.

Suppose for the moment then $a \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$. In this case,

(4.14.3) M_a maps $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ into $L^r(X, \mathbf{R})$ or $L^r(X, \mathbf{C})$,

as appropriate, with

$$\|M_a(f)\|_r \le \|a\|_q\,\|f\|_p \tag{4.14.4}$$

for every $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate, by (4.14.1).

Suppose now that a is a real or complex-valued measurable function on X such that (4.14.3) holds, with

$$\|M_a(f)\|_r \le C\,\|f\|_p \tag{4.14.5}$$

for some $C \ge 0$ and all $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate. Under suitable conditions, we would like to show that $a \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, as appropriate, with

$$\|a\|_q \le C. \tag{4.14.6}$$

If $p = +\infty$, then $q = r$, and this follows by taking $f \equiv 1$ on X.

Suppose from now on in this section that $p < +\infty$, and that every measurable subset of

$$\{x \in X : a(x) \ne 0\} \tag{4.14.7}$$

of positive measure contains a measurable set of positive finite measure. If $q = +\infty$, then $p = r$, and one can check directly that a is essentially bounded on X, with L^∞ norm less than or equal to C, as in (4.14.6).

Suppose now that $q < +\infty$ too, so that $p, q > r$. Let $\epsilon > 0$ be given, and let us show that

$$\mu(\{x \in X : |a(x)| \geq \epsilon\}) < \infty. \tag{4.14.8}$$

Note that every measurable subset of

$$\{x \in X : |a(x)| \geq \epsilon\} \tag{4.14.9}$$

of positive measure contains a measurable set of positive finite measure, by hypothesis.

If $A \subseteq X$, then let $\mathbf{1}_A$ be the corresponding *indicator function* on X, which is equal to 1 on A and to 0 on $X \setminus A$. Suppose that A is measurable, so that $\mathbf{1}_A$ is measurable on X. If $\mu(A) < \infty$, then $\mathbf{1}_A \in L^p(X, \mathbf{R})$, and we get that $M_a(\mathbf{1}_A) = a\,\mathbf{1}_A \in L^r(X, \mathbf{R})$ or $L^r(X, \mathbf{C})$, as appropriate. We also get that

$$\int_A |a(x)|^r \, d\mu(x) \leq C^r \, \|\mathbf{1}_A\|_p^r = C^r \, \mu(A)^{r/p}, \tag{4.14.10}$$

by (4.14.5).

If A is contained in (4.14.9), then it follows that

$$\epsilon^r \, \mu(A) \leq C^r \, \mu(A)^{r/p}. \tag{4.14.11}$$

This implies that

$$\mu(A)^{(p-r)/p} \leq (C/\epsilon)^r, \tag{4.14.12}$$

so that

$$\mu(A) \leq (C/\epsilon)^{p\,r/(p-r)}. \tag{4.14.13}$$

This means that

$$\mu(\{x \in X : |a(x)| \geq \epsilon\}) \leq (C/\epsilon)^{pr/(p-r)}, \tag{4.14.14}$$

as in the previous section. Note that $(p - r)/(pr) = 1/r - 1/p = 1/q$, so that $pr/(p - r) = q$.

Let b be a real or complex-valued measurable function on X such that

$$|b| \leq |a| \tag{4.14.15}$$

almost everywhere on X with respect to μ. This implies that b satisfies the analogues of (4.14.3) and (4.14.5) in place of a. If $b \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, as appropriate, then one can verify that

$$\|b\|_q \leq C. \tag{4.14.16}$$

This can be obtained by considering f such that

$$|f(x)| = |b(x)|^{(q-r)/r} \tag{4.14.17}$$

on X, so that

$$|f(x)|^r\,|b(x)|^r = |b(x)|^q \tag{4.14.18}$$

and

$$|f(x)|^p = |b(x)|^q \tag{4.14.19}$$

on X, as in Section 4.10.

Let us check that $|a|^q$ is integrable on (4.14.9) for each $\epsilon > 0$, with

$$\Big(\int_{\{x \in X : |a(x)| \ge \epsilon\}} |a(x)|^q \, d\mu(x)\Big)^{1/q} \le C. \tag{4.14.20}$$

If b is a bounded real or complex-valued measurable function on X that is equal to 0 when $|a| < \epsilon$, then $b \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, as appropriate, because of (4.14.8). If b also satisfies (4.14.15), then (4.14.16) holds, as before. One can approximate a on (4.14.9) by such functions b, to get that $|a|^q$ is integrable on (4.14.9), and that (4.14.20) holds.

Using this, it is easy to see that $|a|^q$ is integrable on X, and that (4.14.6) holds.

If $r \ge 1$, then $p, q \ge 1$, and we get that M_a is a bounded liinear mapping from $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ into $L^r(X, \mathbf{R})$ or $L^r(X, \mathbf{C})$, as appropriate, with operator norm equal to $\|a\|_q$ under the conditions mentioned earlier. There is an analogous statement when $r < 1$, using suitable versions of some of our previous definitions, as before.

4.15 Convergence in measure

Let $(X, \mathcal{A}, \mu)$ be a nonempty measure space, and let p be a positive extended real number. Remember that $L^p(X, \mathbf{R})$ and $L^p(X, \mathbf{C})$ may be considered as metric spaces, as in Section 4.12.

Let $0 < r < p$ be given, and suppose that $q > 0$ satisfies $1/q = 1/r - 1/p$. If $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, $A \subseteq X$ is measurable, and $\mu(A) < \infty$, then $f\,\mathbf{1}_A \in L^r(X, \mathbf{R})$ or $L^r(X, \mathbf{C})$, as appropriate, with

$$\|f\,\mathbf{1}_A\|_r \le \|f\|_p\,\|\mathbf{1}_A\|_q = \|f\|_p\,\mu(A)^{1/q}, \tag{4.15.1}$$

as in (4.14.1).

Let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued measurable functions on X, and let f be another real or complex-valued measurable function on X, as appropriate. This sequence is said to converge to f *in measure* on a measurable set $A \subseteq X$ with respect to μ if for every $\epsilon > 0$,

$$\lim_{j \to \infty} \mu(\{x \in A : |f_j(x) - f(x)| \ge \epsilon\}) = 0. \tag{4.15.2}$$

If $\mu(A) < \infty$, and $\{f_j\}_{j=1}^\infty$ converges to f pointwise almost everywhere on A with respect to μ, then it is well known that

$$\{f_j\}_{j=1}^\infty \text{ converges to } f \text{ in measure on } A. \tag{4.15.3}$$

If $\{f_j\}_{j=1}^\infty$ converges to f in measure on X, then it is well known that

$$\text{there is a subsequence of } \{f_j\}_{j=1}^\infty \text{ that converges to } f \text{ pointwise almost everywhere on } X \tag{4.15.4}$$

with respect to μ.

Suppose for the moment that $f_j \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ for some $p > 0$ and each j, and that $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate, too. If

$$\{f_j\}_{j=1}^\infty \text{ converges to } f \text{ with respect to the } L^p \text{ metric,} \tag{4.15.5}$$

then it is easy to see that

$$\{f_j\}_{j=1}^\infty \text{ converges to } f \text{ in measure on } X \tag{4.15.6}$$

with respect to μ.

Suppose that $\{f_j\}_{j=1}^\infty$ converges to f in measure on X, and that

$$\|f_j\|_p \le C \tag{4.15.7}$$

for some $C \ge 0$ and each j. One can use this to get that

$$\|f\|_p \le C, \tag{4.15.8}$$

by passing to a subsequence that converges to f pointwise almost everywhere with respect to μ and using Fatou's lemma when $p < \infty$. If $p = \infty$, then (4.15.7) can also be obtained more directly from convergence in measure and the definition of the assential supremum norm.

Suppose that $0 < r < p$, and that A is a measurable subset of X with finite measure. Observe that

$$\|(f_j - f)\,\mathbf{1}_A\|_r \le \|f_j - f\|_p\, \mu(A)^{1/q} \tag{4.15.9}$$

for each j, as in (4.15.1). It is easy to see that $\|f_j - f\|_p$ is uniformly bounded in this case, because of (4.15.7) and (4.15.8). If $\mu(X) < \infty$, then one can use (4.15.9) to get that

$$\{f_j\}_{j=1}^\infty \text{ converges to } f \text{ in } L^r(X, \mathbf{R}) \text{ or } L^r(X, \mathbf{C}), \tag{4.15.10}$$

as appropriate.

Let $\{a_j\}_{j=1}^\infty$ be a sequence of real or complex-valued measurable functions on X, and let a, f be real or complex-valued measurable functions on X as well. If

$$\{a_j\}_{j=1}^\infty \text{ converges to } a \text{ pointwise almost everywhere on } X \tag{4.15.11}$$

with respect to μ, then

$$\{M_{a_j}(f)\}_{j=1}^\infty \text{ converges to } M_a(f) \text{ pointwise almost everywhere on } X \tag{4.15.12}$$

with respect to μ too.

Let A be a measurable subset of X with finite measure again. It is well known that

$$\lim_{n\to\infty} \mu(\{x \in A : |f(x)| > n\}) = 0, \tag{4.15.13}$$

by a standard argument. If

$$\{a_j\}_{j=1}^{\infty} \text{ converges to } a \text{ on } A \text{ in measure} \tag{4.15.14}$$

with respect to μ, then one can use this to check that

$$\{M_{a_j}(f)\}_{j=1}^{\infty} \text{ converges to } M_a(f) \text{ on } A \text{ in measure} \tag{4.15.15}$$

with respect to μ.

Suppose now that $a_j \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, as appropriate, for some $q > 0$ and each j, with

$$\|a_j\|_q \le C \tag{4.15.16}$$

for some $C \ge 0$ and each j. Similarly, suppose that $a \in L^q(X, \mathbf{R})$ or $L^q(X, \mathbf{C})$, as appropriate. If $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate, for some $p > 0$, and if $r > 0$ satisfies $1/p + 1/q = 1/r$, then we would like to have conditions under which

$$\{M_{a_j}(f)\}_{j=1}^{\infty} \text{ converges to } M_a(f) \text{ in } L^r(X, \mathbf{R}) \text{ or } L^r(X, \mathbf{C}), \tag{4.15.17}$$

as appropriate. Of course, this holds when

$$\{a_j\}_{j=1}^{\infty} \text{ converges to } a \text{ in } L^q(X, \mathbf{R}) \text{ or } L^q(X, \mathbf{C}), \tag{4.15.18}$$

as appropriate, by (4.14.1).

The set of $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ such that (4.15.17) holds is a linear subspace of $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate. This is a closed set in $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$ too, as appropriate, because of (4.15.16), as in Section 4.4.

Suppose for the moment that $p = +\infty$, so that $q = r$. If $f \equiv 1$ on X, then (4.15.17) implies (4.15.18).

Suppose now that $p < \infty$, so that $q > r$. Let A be a measurable subset of X with finite measure. In this case, (4.15.17) holds with $f = \mathbf{1}_A$ if and only if

$$\{a_j\}_{j=1}^{\infty} \text{ converges to } a \text{ on } A \text{ in measure with respect to } \mu, \tag{4.15.19}$$

as before.

If this holds for every such A, then (4.15.17) holds for all simple functions in $L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate. This implies that (4.15.17) holds for all $f \in L^p(X, \mathbf{R})$ or $L^p(X, \mathbf{C})$, as appropriate, because simple functions are dense in these spaces.

Chapter 5

Some more spaces and mappings

5.1 Compact support and local compactness

Let X be a nonempty metric space, or topological space, and let f be a real or complex-valued function on X. The *support* of f in X is defined to be the closure of the set where $f \neq 0$,

$$\operatorname{supp} f = \overline{\{x \in X : f(x) \neq 0\}}. \tag{5.1.1}$$

The support of a function on a set was previously defined to simply be the set where the function is nonzero. That corresponds to taking the set to be equipped with the discrete metric or topology here.

We shall often be interested in functions f on X whose support is a compact set. Note that this happens when $\operatorname{supp} f$ is contained in a compact subset of X, because closed sets that are contained in compact sets are compact as well. If X is a metric space, or a Hausdorff topological space, then it is well known that compact sets in X are closed sets. In this case, f has compact support in X when the set where $f \neq 0$ is contained in a compact subset of X.

Let $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ be the spaces of continuous real and complex-valued functions on X with compact support, respectively. One can check that these are linear subspaces of the spaces $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ of continuous real and complex-valued functions on X, respectively. This uses the fact that the union of two compact subsets of X is compact as well. If X is compact, then $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ are the same as the spaces $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ of all continuous real or complex-valued functions on X, respectively.

Suppose for the moment that X is equipped with the discrete metric or topology. It is easy to see that only the finite subsets of X are compact in this case. This means that $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ are the same as the spaces $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$, respectively, defined in Section 1.12.

Let us say that X is *locally compact* at a point $x \in X$ if there are an open set $U \subseteq X$ and a compact set $K \subseteq X$ such that $x \in U$ and $U \subseteq K$. If X is a metric space or a Hausdorff topological space, then K is a closed set in X, so that $\overline{U} \subseteq K$. This implies that $\overline{U}$ is compact, and local compactness is sometimes defined in this way. If X is locally compact at every $x \in X$, then X is simply said to be *locally compact* as a metric or topological space.

Suppose that f is a continuous real or complex-valued function on X with compact support. If $x \in X$ and $f(x) \neq 0$, then it is easy to see that X is locally compact at x.

Suppose now that X is a locally compact Hausdorff space. Let $K \subseteq X$ be a compact set, and let $U \subseteq X$ be an open set, with $K \subseteq U$. Under these conditions, it is well known that there is a continuous real-valued function f on X with compact support contained in U such that $f = 1$ on K, and $0 \leq f \leq 1$ on X. This is a version of *Urysohn's lemma*, which can be shown using analogous arguments, or obtained from the usual formulation for normal topological spaces. If X is a locally compact metric space, then this can be verified more directly.

5.2 Vanishing at infinity and compactness

Let X be a nonempty metric space or topological space again, and let f be a real or complex-valued function on X. Let us say that f *vanishes at infinity* on X if for each $\epsilon > 0$,

$$\{x \in X : |f(x)| \geq \epsilon\} \tag{5.2.1}$$

is contained in a compact subset of X. Of course, this holds trivially when X is compact. If X is equipped with the discrete metric or topology, then this reduces to the analogous definition in Section 1.13.

If f is continuous on X, then (5.2.1) is a closed set in X for each $\epsilon > 0$. If (5.2.1) is contained in a compact subset of X for some $\epsilon > 0$, then it follows that (5.2.1) is compact. Note that f is bounded on X in this case, because f is bounded on compact subsets of X.

Suppose that f is continuous and vanishes at infinity on X. If $x \in X$ and $f(x) \neq 0$, then one can check that X is locally compact at x.

Let ϕ be a mapping from $\mathbf{R}$ or $\mathbf{C}$ into itself, as appropriate, such that $\phi(0) = 0$ and ϕ is continuous at 0. If f vanishes at infinity on X, then it is easy to see that

$$\phi \circ f \text{ vanishes at infinity on } X \tag{5.2.2}$$

as well. If $\phi \equiv 0$ on a neighborhood of 0 in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, then $f \equiv 0$ on the complement of a compact subset of X.

Let $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$ be the spaces of continuous real and complex-valued functions on X that vanish at infinity. It is easy to see that these are linear subspaces of the spaces $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ of bounded continuous real and complex-valued functions on X, respectively. One can verify that $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$ are also closed sets in $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$, respectively, with respect to the supremum metric. If X is compact, then $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$ are the same as the spaces $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ of all continuous real and complex-valued functions on X, respectively. If X is equipped with the discrete metric or topology, then $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$ are the same as the spaces $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$, respectively, defined in Section 1.13.

Of course, any real or complex-valued function on X with compact support vanishes at infinity. If f is any continuous real or complex-valued function on X that vanishes at infinity, then f can be approximated uniformly by continuous functions with compact support in X. This can be seen using suitable compositions with continuous functions on $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in (5.2.2). If X is a locally compact Hausdorff topological space, then one can also use the version of Urysohn's lemma mentioned in the previous section. This means that $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ are dense in $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$, respectively, with respect to the supremum metric.

Let X be a locally compact Hausdorff topological space, and let λ be a bounded linear functional on $C_0(X, \mathbf{R})$ or $C_0(X, \mathbf{C})$, with respect to the supremum norm. It is well known that λ can be represented in a unique way in terms of integration with respect to a real or complex Borel measure on X, with suitable regularity properties. If there is a base for the topology of X with only finitely or countably many elements, then these additional regularity properties hold automatically.

5.3 Uniform convergence on compact sets

Let X be a metric or topological space, and let (Y, d_Y) be a metric spaces. Also let $\{f_j\}_{j=1}^{\infty}$ be a sequence of mappings from X into Y, and let f be another mapping from X into Y. Suppose that

$$\{f_j\}_{j=1}^{\infty} \text{ converges to } f \text{ uniformly on compact subsets of } X, \tag{5.3.1}$$

so that for each compact set $K \subseteq X$, $\{f_j\}_{j=1}^{\infty}$ converges to f uniformly on K. If f_j is continuous on X for each j, then it follows that

$$f \text{ is continuous on } K \text{ for every compact } K \subseteq X. \tag{5.3.2}$$

More precisely, this uses the topology induced on K by the topology on X, or the restriction of the metric on X to K when X is a metric space.

Of course, the same conclusion holds when f_j is only asked to be continuous on compact subsets of X for each j. Let us now consider conditions under which (5.3.2) implies that f is continuous on X.

If X is locally compact at a point $x \in K$, then (5.3.2) implies that f is continuous at x. Thus (5.3.2) implies that f is continuous on X when X is locally compact.

We say that f is *sequentially continuous* at a point $x \in X$ if for every sequence $\{x_l\}_{l=1}^{\infty}$ of elements of X that converges to x,

$$\{f(x_l)\}_{l=1}^{\infty} \text{ converges to } f(x) \text{ in } Y. \tag{5.3.3}$$

If f is continuous at x, then f is sequentially continuous at x, by a standard argument. If there is a local base for the topology of X with only finitely or countably many elements, and if f is sequentially continuous at x, then one can check that f is continuous at x. In particular, metric spaces have this property.

Let $\{x_l\}_{l=1}^{\infty}$ be a sequence of elements of X that converges to an element x of X. One can check that

$$\{x_l : l \in \mathbf{Z}_+\} \cup \{x\} \tag{5.3.4}$$

is a compact subset of X. If the restriction of f to this compact set is continuous at x, then it follows that (5.3.3) holds.

If f satisfies (5.3.2), then we get that f is *sequentially continuous* on X, which is to say that f is sequentially continuous at every point in X.

If for every $x \in X$ there is a local base for the topology of X at x with only finitely or countably many elements, then X is said to satisfy the *first countability condition.* Metric spaces have this property, as before. In this case, (5.3.2) implies that f is continuous on X.

5.4 Some related multiplication operators

Let X be a nonempty metric or topological space, and let a be a continuous real or complex-valued function on X. If f is another continuous real or complex-valued function on X, as appropriate, then

$$M_a(f) = a\, f \tag{5.4.1}$$

defines a continuous real or complex-valued function on X as well. This defines a linear mapping from $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$ into itself, as appropriate, which is the *multiplication operator* associated to a.

If f has compact support in X, then $M_a(f)$ has compact support too. This means that the restriction of M_a to $C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$, as appropriate, defines a linear mapping from that space into itself.

If a is bounded on X, then

$$M_a \text{ maps } C_b(X, \mathbf{R}) \text{ or } C_b(X, \mathbf{C}) \text{ into itself,} \tag{5.4.2}$$

as appropriate. Similarly, one can check that

$$M_a \text{ maps } C_0(X, \mathbf{R}) \text{ or } C_0(X, \mathbf{C}) \text{ into itself,} \tag{5.4.3}$$

as appropriate, in this case.

Let $\|\cdot\|_{sup}$ be the usual supremum norm on $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$. This notation is often preferable, to avoid confusion with L^∞ norms. If a is bounded on X, then

$$\|M_a(f)\|_{sup} \le \|a\|_{sup}\, \|f\|_{sup} \tag{5.4.4}$$

for every $f \in C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate.

Suppose that (5.4.2) holds, and that

$$\|M_a(f)\|_{sup} \le C\,\|f\|_{sup} \tag{5.4.5}$$

for some $C \ge 0$ and all $f \in C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate. This implies that a is bounded on X, with

$$\|a\|_{sup} \le C, \tag{5.4.6}$$

by taking $f \equiv 1$ on X. This means that the operator norm of M_a on $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$ with respect to the supremum norm is equal to $\|a\|_{sup}$, because of (5.4.4).

Suppose now that (5.4.5) holds for some $C \ge 0$ and all $f \in C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$, as appropriate. If X is a locally compact Hausdorff topological space, then one can use the version of Urysohn's lemma mentioned in Section 5.1 to get that a is bounded on X, and that (5.4.6) holds.

Let $\{a_j\}_{j=1}^\infty$ be a sequence of continuous real or complex-valued functions on X, as appropriate. Suppose that

$$\{a_j\}_{j=1}^\infty \text{ converges to } a \text{ uniformly on compact subsets of } X. \tag{5.4.7}$$

If f is a continuous real or complex-valued function on X, then one can check that

$$\{M_{a_j}(f)\}_{j=1}^\infty \text{ converges to } M_a(f) \text{ uniformly on compact subsets of } X. \tag{5.4.8}$$

This uses the fact that f is bounded on compact subsets of X.

Note that

$$\operatorname{supp} M_a(f) \subseteq \operatorname{supp} f, \tag{5.4.9}$$

and similarly

$$\operatorname{supp} M_{a_j}(f) \subseteq \operatorname{supp} f \tag{5.4.10}$$

for each j. If f has compact support in X, then it follows that

$$\{M_{a_j}(f)\}_{j=1}^\infty \text{ converges to } M_a(f) \text{ uniformly on } X. \tag{5.4.11}$$

Of course, (5.4.11) holds with $f \equiv 1$ on X if and only if

$$\{a_j\}_{j=1}^\infty \text{ converges to } a \text{ uniformly on } X. \tag{5.4.12}$$

In this case, (5.4.11) holds when f is bounded on X.

Suppose that the a_j's are uniformly bounded on X, so that

$$\|a_j\|_{sup} \le C \tag{5.4.13}$$

for some $C \ge 0$ and each j. If $\{a_j\}_{j=1}^\infty$ converges to a pointwise on X, then a is bounded on X, and (5.4.6) holds. If (5.4.7) holds, then (5.4.11) holds when f vanishes at infinity on X. This uses the fact that $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ are dense in $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$, respectively, with respect to the supremum metric, as in Section 5.2.

5.5 Some separation conditions

Let X be a nonempty topological space. If continuous real-valued functions on X separate points in X, then X is said to be a *Urysohn space*. More precisely, this means that if x, y are distinct elements of X, then there is a continuous real-valued function f on X such that

$$f(x) \neq f(y). \tag{5.5.1}$$

In this case, one can choose f so that

$$f(x) = 0,\ f(y) = 1,\ \text{and } 0 \leq f \leq 1 \text{ on } X. \tag{5.5.2}$$

Alterntively, one might choose f so that

$$f(x) = -1,\ f(y) = 1,\ \text{and } |f| \leq 1 \text{ on } X. \tag{5.5.3}$$

It is easy to see that Urysohn spaces are Hausdorff. If the topology on X is determined by a metric, then one can check directly that X is a Urysohn space. Locally compact Hausdorff topological spaces are Urysohn spaces, by the version of Urysohn's lemma mentioned in Section 5.1.

Let us say that X is *completely regular in the strict sense* if for every $x \in X$ and closed set $E \subseteq X$ with $x \notin E$ there is a continuous real-valued function f on X such that

$$f \equiv 0 \text{ on } E \text{ and } f(x) \neq 0. \tag{5.5.4}$$

One can choose f so that

$$f(x) = 1 \text{ and } 0 \leq f \leq 1 \text{ on } X. \tag{5.5.5}$$

If X also satisfies the first or even zeroth separation condition, then we say that X is *completely regular in the strong sense*. This implies that X is a Urysohn space, and that X is Hausdorff in particular.

Sometimes one says that X is *completely regular* when X is completely regular in the strict sense. Similarly, if X is completely regular in the strong sense, then one may say that X satisfies separation condition number three and a half, or equivalently that X is a $T_{3\frac{1}{2}}$-space. However, the opposite convention is sometimes used too. Sometimes each of these names is used when X is completely regular in the strong sense, and one may refer to complete regularity in the strict sense in some other way.

If X is a metric space, then one can check directly that X is completely regular in the strong sense. If X is a locally compact Hausdorff topological space, then the version of Urysohn's lemma mentioned in Section 5.1 implies that X is completely regular in the strong sense.

Let us say that X is *normal in the strict sense* if for every pair A, B of disjoint closed subsets of X there are disjoint open sets $U, V \subseteq X$ such that

$$A \subseteq U,\ B \subseteq V. \tag{5.5.6}$$

If X satisfies the first separation condition too, then we say that X is *normal in the strong sense*. Note that this implies that X is Hausdorff. It is well known that metric spaces are normal in the strong sense.

Sometimes one says that X is *normal* when X is normal in the strict sense, and that X satisfies the fourth separation condition, or equivalently that X is a T_4 space, when X is normal in the strong sense. The opposite convention is sometimes used as well. Sometimes each of these names is used when X is normal in the strong sense, and normality in the strict sense may be described another way.

Suppose that X is normal in the strict sense, and that A, B are disjoint closed subsets of X. Under these conditions, *Urysohn's lemma* states that there is a continuous real-valued function f on X such that

$$f \equiv 0 \text{ on } A, \ f \equiv 1 \text{ on } B, \text{ and } 0 \le f \le 1 \text{ on } X. \tag{5.5.7}$$

If X is normal in the strong sense, then it follows that f is completely regular in the strong sense. Note that the conclusion of Urysohn's lemma can be obtained more directly for metric spaces, using distances to sets, as in Section 2.11.

Let us say that X is *regular in the strict sense* if for every $x \in X$ and closed set $E \subseteq X$ with $x \notin E$ there are disjoint open sets $U, V \subseteq X$ such that

$$x \in U \text{ and } E \subseteq V. \tag{5.5.8}$$

If X also satisfies the first or zeroth separation condition, then we say that X is *regular in the strong sense*. If X is completely regular in the strict or strong sense, then X is regular in the strict or strong sense, respectively. If X is regular in the strong sense, then X is Hausdorff.

As before, one sometimes says that X is *regular* when X is regular in the strict sense, and that X satisfies the third separation condition, or equivalently that X is a T_3 space, when X is regular in the strong sense. The opposite convention is sometimes used too, and sometimes each of these names is used when X is regular in the strong sense, and regularity in the strict sense may be described in some other way.

One can check that X is regular in the strict sense if and only if for every $x \in X$ and open set $W \subseteq X$ with $x \in W$, there is an open set $U \subseteq X$ such that

$$x \in U \text{ and } \overline{U} \subseteq W, \tag{5.5.9}$$

where $\overline{U}$ is the closure of U in X, as usual. If X is a metric space, or a locally compact Hausdorff space, then it is somewhat easier to verify directly that X is regular in the strong sense than complete regularity. Similarly, if X is normal in the strong sense, then it is easy to see directly that X is regular in the strong sense.

We say that X is *completely Hausdorff* if for every $x, y \in X$ with $x \neq y$ there are open sets $U, V \subseteq X$ such that

$$x \in U, \ y \in V, \text{ and } \overline{U} \cap \overline{V} = \emptyset. \tag{5.5.10}$$

Note that completely Hausdorff spaces are Hausdorff in particular. One can verify that X is completely Hausdorff when X is a Urysohn space, and when X is regular in the strong sense.

Suppose that X satisfies the second countability condition. If X is also normal in the strong sense, then *Urysohn's metrization theorem* states that there is a metric on X that determines the same topology. Tychonoff showed that normality in the strong sense can be replaced with regularity in the strong sense.

5.6 Point evaluations

Let X be a nonempty metric or topological space. If $x \in X$, then put

$$\delta_x(f) = f(x) \tag{5.6.1}$$

for every continuous real or complex-valued function f on X. This defines a bounded linear functional on each of $C_b(X, \mathbf{R})$ and $C_b(X, \mathbf{C})$, with respect to the supremum norm. It is easy to see that the dual norm of this linear functional is equal to 1.

The restriction of δ_x to each of $C_0(X, \mathbf{R})$ and $C_0(X, \mathbf{C})$ is a bounded linear functional with dual norm less than or equal to 1, with respect to the supremum norm. If X is a locally compact Hausdorff topological space, then one can use the version of Urysohn's lemma mentioned in Section 5.1 to get that the dual norm of δ_x on these spaces is also equal to 1.

If $x, y \in X$, then $\delta_x - \delta_y$ defines a bounded linear functional on each of $C_b(X, \mathbf{R})$ and $C_b(X, \mathbf{C})$, with dual norm with respect to the supremum norm less than or equal to 2. If X is a Urysohn space, then the dual norm of $\delta_x - \delta_y$ on each of these spaces if equal to 2, because of (5.5.3). Similarly, if X is a locally compact Hausdorff topological space, then one can check that the dual norm of $\delta_x - \delta_y$ on each of $C_0(X, \mathbf{R})$ and $C_0(X, \mathbf{C})$ is equal to 2, with respect to the supremum norm.

Suppose that $\{x_j\}_{j=1}^\infty$ is a sequence of elements of X that converges to $x \in X$. Observe that

$$\{\delta_{x_j}\}_{j=1}^\infty \text{ converges to } \delta_x \tag{5.6.2}$$

with respect to the weak* topology on the dual of each of $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$.

Consider the mapping

$$x \mapsto \delta_x \tag{5.6.3}$$

from X into the dual of $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$. This mapping is continuous with respect to the weak* topology on the dual space. If X is a Urysohn space, then this mapping is one-to-one. If X is completely regular in the strong sense, then this mapping is a homeomorphism onto its image, with respect to the topology induced on its iimage by the weak* topology on the dual space.

Now consider (5.6.3) as a mapping from X into the dual of $C_0(X, \mathbf{R})$ or $C_0(X, \mathbf{C})$. This mapping is continuous with respect to the weak* topology on the dual space, as before. If X is a locally compact Hausdorff space, then this mapping is a homeomorphism onto its image, with respect to the topology induced on its image by the weak* topology on the dual space.

Suppose that X is a locally compact Hausdorff topological space that is not compact. Let us say that a sequence $\{x_l\}_{l=1}^\infty$ of elements of X *tends to infinity* if for every compact set $K \subseteq X$, we have that

$$x_l \in X \setminus K \tag{5.6.4}$$

for all sufficiently large l. This implies that

$$\{\delta_{x_l}\}_{l=1}^\infty \text{ converges to } 0 \tag{5.6.5}$$

with respect to the weak* topology on the dual of each of $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$. It is easy to see that the converse holds as well, using the version of Urysohn's lemma mentioned in Section 5.1.

5.7 Some remarks about Borel measures

Let X be a nonempty metric or topological space. A subset of X is called an F_σ *set* if it can be expressed as the union of a sequence of closed sets in X. Similarly, a subset of X is said to be a G_δ *set* if it can be expressed as the intersection of a sequence of open sets. Thus a G_δ set is the same as the complement of an F_σ set in X.

Suppose for the moment that the topology on X is determined by a metric d. If $A \subseteq X$ and r is a positive real number, then put

$$A_r = \bigcup_{x \in A} B(x, r), \tag{5.7.1}$$

where $B(x, r)$ is the open ball in X centered at $x \in X$ of radius r. This is an open set in X that contains A, and one can check that

$$\overline{A} = \bigcap_{l=1}^\infty A_{1/l}. \tag{5.7.2}$$

It follows that

$$\text{every closed set in } X \text{ is a } G_\delta \text{ set.} \tag{5.7.3}$$

Equivalently, this means that

$$\text{every open set in } X \text{ is an } F_\sigma \text{ set.} \tag{5.7.4}$$

Let X be any nonempty metric or topological space again. The collection of *Borel sets* in X is the smallest σ-algebra of subsets of X that contains the open sets, or equivalently the closed sets. Thus F_σ sets and G_δ sets in X are Borel sets.

Let μ be a nonnegative *Borel measure* on X, which is to say a nonnegative countably-additive measure on the σ-algebra of Borel sets in X. Suppose that (5.7.3) holds, which is the same as saying that (5.7.4) holds. Supose also that

$$\mu(X) < +\infty, \tag{5.7.5}$$

for the moment. If $E \subseteq X$ is a Borel set, then it is well known that there are closed sets $A \subseteq X$ and open sets $U \subseteq X$ such that

$$A \subseteq E \subseteq U \tag{5.7.6}$$

and

$$\mu(U \setminus A) \text{ is as small as we like.} \tag{5.7.7}$$

This basically corresponds to parts of Proposition D.1 on p419 of [20], Theorem 2.2.2 on p60 of [62], and Theorem 1.10 on p11 of [142].

To see this, observe that open and closed sets in X have these properties, because of (5.7.3), (5.7.4). Thus it suffices to check that the collection $\mathcal{C}$ of Borel sets $E \subseteq X$ with these properties is a σ-algebra. It is easy to see that $\mathcal{C}$ is closed under taking complements. It follows that the closure of $\mathcal{C}$ under countable unions and intersections are equivalent to each other. It is not too difficult to verify either of these conditions.

Equivalently, (5.7.7) means that

$$\mu(U \setminus E) \text{ is as small as we like} \tag{5.7.8}$$

and

$$\mu(E \setminus A) \text{ is as small as we like.} \tag{5.7.9}$$

Suppose that

$$\mu(E) < +\infty, \tag{5.7.10}$$

in place of (5.7.5). Note that

$$\mu_E(B) = \mu(B \cap E) \tag{5.7.11}$$

is a nonnegative Borel measure on X such that $\mu_E(X) = \mu(E) < +\infty$. If (5.7.3) or equivalently (5.7.4) holds, then the previous argument implies that there are closed sets A in X such that

$$A \subseteq E \tag{5.7.12}$$

and $\mu_E(E \setminus A)$ is as small as we like. Of course, the latter is the same as saying that (5.7.9) holds.

Suppose now that $W_1, W_2, W_3, \ldots$ is a sequence of open subsets of X such that

$$\mu(W_j) < +\infty \tag{5.7.13}$$

for each j, and

$$E \subseteq \bigcup_{j=1}^{\infty} W_j. \tag{5.7.14}$$

As before,

$$\mu_j(B) = \mu(B \cap W_j) \tag{5.7.15}$$

is a nonnegative Borel measure on X for each j, with $\mu_j(X) = \mu(W_j) < +\infty$. If (5.7.3) or equivalently (5.7.4) holds, then the earlier argument implies that for each j there is an open set $U_j \subseteq X$ such that

$$E \cap W_j \subseteq U_j \tag{5.7.16}$$

and

$$\mu_j(U_j \setminus (E \cap W_j)) \text{ is as small as we like.} \tag{5.7.17}$$

We may as well take U_j so that

$$U_j \subseteq W_j \tag{5.7.18}$$

for each j, by replacing U_j with its intersection with W_j, if necessary. In this case, (5.7.17) is the same as saying that

$$\mu(U_j \setminus (E \cap W_j)) \text{ is as small as we like} \tag{5.7.19}$$

for each j.

If we put

$$U = \bigcup_{j=1}^{\infty} U_j, \tag{5.7.20}$$

then U is an open set in X such that

$$E \subseteq U. \tag{5.7.21}$$

We can also use (5.7.14) and (5.7.19) to get that (5.7.8) holds. These two variants of the earlier argument correspond to parts of Theorem 2.2.2 on p60 of [62], and Theorem 1.10 on p11 of [142].

5.8 Approximation by continuous functions

Let X be a nonempty metric or topological space, and let μ be a nonnegative Borel measure on X. Also let $A \subseteq X$ be a closed set and $U \subseteq X$ be an open set such that

$$A \subseteq U. \tag{5.8.1}$$

If X is normal in the strict sense, then there is a continuous real-valued function ϕ on X such that

$$\phi \equiv 1 \text{ on } A,\ \phi \equiv 0 \text{ on } X \setminus U, \text{ and } 0 \le \phi \le 1 \text{ on } X, \tag{5.8.2}$$

by Urysohn's lemma. This implies that

$$\mu(A) \le \int_X \phi\, d\mu \le \mu(U). \tag{5.8.3}$$

If $E \subseteq X$ is a Borel set such that

$$A \subseteq E \subseteq U, \tag{5.8.4}$$

then

$$\mu(A) \le \mu(E) \le \mu(U). \tag{5.8.5}$$

If $\mu(U) < \infty$, then one can use this and (5.8.3) to get that

$$\left| \mu(E) - \int_X \phi \, d\mu \right| \le \mu(U) - \mu(A) = \mu(U \setminus A). \tag{5.8.6}$$

If $\mu(X) < +\infty$, and X satisfies (5.7.3) or equivalently (5.7.4), then we can find A, U such that $\mu(U \setminus A)$ is as small as we like, as in the previous section. Note that

$$\{x \in X : \mathbf{1}_E(x) \ne \phi(x)\} \subseteq U \setminus A. \tag{5.8.7}$$

Let us say that a topological space X is *perfectly normal in the strict sense* if X is normal in the strict sense, and (5.7.3) or equivalently (5.7.4) holds. If X also satisfies the first separation condition, then we say that X is *perfectly normal in the strong sense*. As usual, one may say that X is *perfectly normal* when X is perfectly normal in the strict sense, and *perfectly* T_4 when X is perfectly normal in the strict sense, but the opposite convention may be used as well. Alternatively, both of these terms may be used when X is perfectly normal in the strong sense, and perfectly normal spaces in the strict sense may be described in other ways.

Let X be a nonempty topological space, and let μ_1, μ_2 be nonnegative finite Borel measures on X. If f is a bounded continuous real-valued function on X, then put

$$\lambda_j(f) = \lambda_{\mu_j}(f) = \int_X f \, d\mu_j \tag{5.8.8}$$

for $j = 1, 2$. These define bounded linear functionals on $C_b(X, \mathbf{R})$ with respect to the supremum norm, with dual norms equal to

$$\mu_j(X), \tag{5.8.9}$$

$j = 1, 2$.

Suppose from now on in this section that X is perfectly normal in the strict sense. If

$$\lambda_1 = \lambda_2 \tag{5.8.10}$$

on $C_b(X, \mathbf{R})$, then

$$\mu_1 = \mu_2 \tag{5.8.11}$$

on X. Of course, this means that

$$\mu_1(E) = \mu_2(E) \tag{5.8.12}$$

for all Borel sets $E \subseteq X$. To see this, one can try to approximate $\mu_1(E)$ and $\mu_2(E)$ by

$$\lambda_1(\phi) = \lambda_2(\phi) \tag{5.8.13}$$

for suitable $\phi \in C_b(X, \mathbf{R})$.

More precisely,

$$\mu = \mu_1 + \mu_2 \tag{5.8.14}$$

is another nonnegative finite Borel measure on X. Thus there are closed sets $A \subseteq X$ and open sets $U \subseteq X$ such that (5.8.4) holds and $\mu(U \setminus A)$ is as small as we like, as in the previous section. If ϕ is as in (5.8.2), then

$$\left|\mu_j(E) - \int_X \phi \, d\mu_j\right| \le \mu_j(U \setminus A), \tag{5.8.15}$$

$j = 1, 2$, as in (5.8.6).

Alternatively, one can look at $\mu_1 - \mu_2$ as a real-valued or signed measure on X. Before doing this, we shall review real and complex-valued measures in the next section.

5.9 Real and complex measures

Let $(X, \mathcal{A})$ be a (nonempty) measurable space, so that X is a nonempty set, and $\mathcal{A}$ is a σ-algebra of measurable subsets of X. A real or complex-valued function μ on $\mathcal{A}$ is said to be a *measure* on $(X, \mathcal{A})$ if

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j) \tag{5.9.1}$$

for every sequence $A_1, A_2, A_3, \ldots$ of pairwise-disjoint measurable subsets of X. In the real case, μ may also be called a *signed measure* on X. More precisely, the convergence of the series on the right side of (5.9.1) is considered to be part of the definition of a measure. It follows that the series should converge absolutely, because any rearrangement of the series converges too.

If $A \subseteq X$ is measurable, then put

$$\begin{aligned} |\mu|(A) \quad = \quad \sup\Big\{ \sum_{j=1}^{\infty} |\mu(A_j)| : A_1, A_2, A_3, \ldots & \text{ is a pairwise-disjoint sequence} \\ \text{of measurable subsets of } X \text{ such that } & \bigcup_{j=1}^{\infty} A_j = A \Big\}. \end{aligned} \tag{5.9.2}$$

It is well known that this defines a nonnegative finite measure on $(X, \mathcal{A})$, which is the *total variation measure* associated to μ. Note that

$$|\mu(A)| \le |\mu|(A) \tag{5.9.3}$$

for every measurable set $A \subseteq X$. The total variation measure is the smallest nonnegative measure on $(X, \mathcal{A})$ with this property.

Let $\mathcal{M}(X, \mathbf{R})$, $\mathcal{M}(X, \mathbf{C})$ be the spaces of real and complex measures on $(X, \mathcal{A})$, respectively. It is easy to see that these are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication on $\mathcal{A}$. One can check that

$$\|\mu\| = |\mu|(X) \tag{5.9.4}$$

defines a norm on each of $\mathcal{M}(X, \mathbf{R})$ and $\mathcal{M}(X, \mathbf{C})$. One can show that $\mathcal{M}(X, \mathbf{R})$ and $\mathcal{M}(X, \mathbf{C})$ are Banach spaces over the real and complex numbers, respectively, with respect to this norm.

If μ is a real measure on X, then

$$\mu^+ = (1/2)\,(|\mu| + \mu), \quad \mu^- = (1/2)\,(|\mu| - \mu) \tag{5.9.5}$$

are nonnegative finite measures on $(X, \mathcal{A})$. These are the *positive* and *negative variation measures* associated to μ on X, respectively. Of course,

$$\mu = \mu^+ - \mu^- \tag{5.9.6}$$

and

$$|\mu| = \mu^+ + \mu^-, \tag{5.9.7}$$

by construction. The former is known as the *Jordan decomposition* of μ.

The *Hahn decomposition theorem* implies that there is a measurable set $B \subseteq X$ such that

$$\mu(A \cap B) \geq 0 \tag{5.9.8}$$

and

$$\mu(A \cap (X \setminus B)) \leq 0 \tag{5.9.9}$$

for all measurable sets $A \subseteq X$. Using this, one can check that

$$|\mu|(A) = \mu(A \cap B) - \mu(A \cap (X \setminus B)) \tag{5.9.10}$$

for all measurable sets $A \subseteq X$. Similarly,

$$\mu^+(A) = \mu(A \cap B) \tag{5.9.11}$$

and

$$\mu^-(A) = -\mu(A \cap (X \setminus B)) \tag{5.9.12}$$

for all measurable sets $A \subseteq X$.

If μ is a complex measure on X, then it is well known that there is a measurable complex-valued function h on X such that

$$|h| = 1 \text{ on } X \tag{5.9.13}$$

and

$$\mu(A) = \int_A h\, d|\mu| \tag{5.9.14}$$

for all measurable sets $A \subseteq X$. This reduces to the Hahn decomposition theorem in the real case.

If f is a measurable complex-valued function on X that is integrable with respect to $|\mu|$, then the integral of f with respect to μ on X can be defined by

$$\int_X f\,d\mu = \int_X f\,h\,d|\mu|. \tag{5.9.15}$$

In this case, we get that

$$\left|\int_X f\,d\mu\right| \le \int_X |f|\,d|\mu|. \tag{5.9.16}$$

Note that f is integrable with respect to $|\mu|$ when f is bounded on X, because $|\mu|(X) < \infty$.

5.10 Real and complex Borel measures

Let X be a topological space, so that X may be considered as a measurable space too, using the σ-algebra of Borel sets in X. Also let μ be a real or complex Borel measure on X, which is to say a real or complex measure on X with respect to the Borel sets. If f is a bounded continuous real or complex-valued function on X, then put

$$\lambda_\mu(f) = \int_X f\,d\mu. \tag{5.10.1}$$

Observe that

$$|\lambda_\mu(f)| \le \int_X |f|\,d|\mu| \le \|f\|_{sup}\,|\mu|(X), \tag{5.10.2}$$

by (5.9.16). Thus λ_μ defines a bounded linear functional on $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate, with respect to the supremum norm, and with dual norm less than or equal to $|\mu|(X)$.

If X is perfectly normal in the strict sense, then

(5.10.3) the dual norm of λ_μ with respect to the supremum norm is equal to $|\mu|(X)$.

In particular, this implies that $\lambda_\mu = 0$ only when $\mu = 0$ on X.

Suppose that μ is a real Borel measure on X, and let $B \subseteq X$ be a Borel set as in the Hahn decomposition of μ. In this case,

$$\lambda_\mu(f) = \int_B f\,d|\mu| - \int_{X\setminus B} f\,d|\mu| \tag{5.10.4}$$

for all $f \in C_b(X, \mathbf{R})$. Let $C \subseteq X$ be a closed set and $V \subseteq X$ an open set such that

$$C \subseteq B \subseteq V. \tag{5.10.5}$$

Using Urysohn's lemma, we can get a continuous real-valued function ψ on X such that

$$\psi \equiv 1 \text{ on } C,\ \psi \equiv -1 \text{ on } X \setminus V, \text{ and } |\psi| \le 1 \text{ on } X. \tag{5.10.6}$$

Observe that

$$\lambda_\mu(\psi) - |\mu|(X) = \int_B (\psi - 1)\, d|\mu| - \int_{X \setminus B} (\psi + 1)\, d|\mu| = \int_{B \setminus C} (\psi - 1)\, d|\mu| - \int_{V \setminus B} (\psi + 1)\, d|\mu|. \tag{5.10.7}$$

This implies that

$$|\lambda_\mu(\psi) - |\mu|(X)| \le \int_{B \setminus C} |\psi - 1|\, d|\mu| + \int_{V \setminus B} |\psi + 1|\, d|\mu| \le 2\,|\mu|(B \setminus C) + 2\,|\mu|(V \setminus B) = 2\,|\mu|(V \setminus C). \tag{5.10.8}$$

The right side can be as small as we want, as in Section 5.7. One can use this to get (5.10.3) in the real case.

In the complex case, there is a complex-valued Borel measurable function h on X that satisfies (5.9.13) such that

$$\lambda_\mu(f) = \int_X f\, h\, d|\mu| \tag{5.10.9}$$

for all $f \in C_b(X, \mathbf{C})$, as in (5.9.15). Note that

$$|\lambda_\mu(f) - |\mu|(X)| = \left| \int_X (f\, h - 1)\, d|\mu| \right| \le \int_X |f\, h - 1|\, d|\mu|. \tag{5.10.10}$$

To get (5.10.3), one can find f so that $|f| \le 1$ on X, and

$$\int_X |f\, h - 1|\, d|\mu| = \int_X |f - \overline{h}|\, d|\mu| \tag{5.10.11}$$

is as small as we like. One way to do this is to approximate $\overline{h}$ by Borel measurable simple functions, and to approximate those simple functions by bounded continuous functions on X. The condition $|f| \le 1$ on X can be obtained by replacing f with $f/|f|$ on the set where $|f| \ge 1$.

5.11 Some remarks about product spaces

If X and Y are any two sets, then their *Cartesian product* $X \times Y$ is the set of all ordered pairs (x, y), with $x \in X$ and $y \in Y$. If X and Y are topological spaces, then the corresponding *product topology* on $X \times Y$ can be defined in a standard way. More precisely, if $U \subseteq X$ and $V \subseteq Y$ are open sets, then $U \times V$ is an open set in $X \times Y$, and the open subsets of $X \times Y$ of this type form a base for the product topology.

Suppose now that (X, d_X), (Y, d_Y) are metric spaces. There are various ways that one might try to define a metric on $X \times Y$, using d_X, d_Y. If $1 \le p < \infty$, then put

$$d_p((x, y), (x', y')) = d_{X \times Y, p}((x, y), (x', y')) = (d_X(x, x')^p + d_Y(y, y')^p)^{1/p} \tag{5.11.1}$$

for all $x, x' \in X$ and $y, y' \in Y$. One can check that this defines a metric on $X \times Y$, using Minkowski's inequality for finite sums, as in Section 1.3. Of course, this is much easier when $p = 1$.

Similarly, put

$$d_\infty((x,y),(x',y')) = d_{X\times Y,\infty}((x,y),(x',y')) = \max(d_X(x,x'), d_Y(y,y')) \tag{5.11.2}$$

for all $x, x' \in X$ and $y, y' \in Y$. One can verify that this defines a metric on $X \times Y$ as well. If $1 \le p < \infty$, then

$$d_\infty((x,y),(x',y')) \le d_p((x,y),(x',y')) \le 2^{1/p}\, d_\infty((x,y),(x',y')) \tag{5.11.3}$$

for all $x, x' \in X$ and $y, y' \in Y$.

If $x \in X$ and $y \in Y$, then let $B_X(x,r)$, $B_Y(y,r)$ be the open balls in X, Y centered at x, y, respectively, with radius $r > 0$, and let $\overline{B}_X(x,r)$, $\overline{B}_Y(y,r)$ be the corresponding closed balls of radius $r \ge 0$. Similarly, if $1 \le p \le \infty$, then let $B_p((x,y),r)$ be the open ball in $X \times Y$ with respect to d_p centered at (x,y) with radius $r > 0$, and let $\overline{B}_p((x,y),r)$ be the corresponding closed ball of radius $r \ge 0$. It is easy to see that

$$B_\infty((x,y),r) = B_X(x,r) \times B_Y(y,r) \tag{5.11.4}$$

for every $r > 0$, and that

$$\overline{B}_\infty((x,y),r) = \overline{B}_X(x,r) \times \overline{B}_Y(y,r) \tag{5.11.5}$$

for every $r \ge 0$.

One can use (5.11.4) to check that the topology determined on $X \times Y$ by d_∞ is the same as the product topology associated to the topologies determined on X, Y by d_X, d_Y, respectively. This is the same as the topology determined on $X \times Y$ by d_p for any $p \ge 1$, because of (5.11.3).

If $1 \le p \le \infty$, then a sequence $\{(x_j, y_j)\}_{j=1}^\infty$ of elements of $X \times Y$ is a Cauchy sequence with respect to d_p if and only if $\{x_j\}_{j=1}^\infty$ and $\{y_j\}_{j=1}^\infty$ are Cauchy sequences in X and Y, respectively. If X, Y are complete with respect to d_X, d_Y, respectively, then it follows that $X \times Y$ is complete with respect to d_p.

We may sometimes be interested in uniform continuity properties of a mapping from $X \times Y$ into another metric space Z. Uniform continuity with respect to any of the metrics d_p implies uniform continuity with respect to the other metrics of this type, because of (5.11.3).

Similarly, boundedness of a subset of $X \times Y$ with respect to any of the metrics d_p implies boundedness with respect to the other metrics of this type. We may also be interested in uniform continuity properties of mappings from $X \times Y$ into Z on bounded sets in $X \times Y$, for which the choice of metric d_p may not matter too much.

5.12 Some remarks about direct sums

Let V, W be vector spaces, both real or both complex. Their Cartesian product $V \times W$ may be considered as a vector space over the real or complex numbers too, as appropriate, with respect to coordinatewise addition and scalar multiplication. This is the *direct sum* of V and W, as vector spaces. One may wish to identify V, W with the linear subspaces $V \times \{0\}$, $\{0\} \times W$ of $V \times W$ in the obvious way.

Let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively. If $1 \le p < \infty$, then put

$$\|(v,w)\|_{V\times W,p} = (\|v\|_V^p + \|w\|_W^p)^{1/p} \tag{5.12.1}$$

for every $(v,w) \in V \times W$. It is easy to see that this is a norm on $V \times W$ when $p = 1$. If $p > 1$, then one can verify that this is a norm on $V \times W$ using Minkowski's inequality for finite sums, as in Section 1.3.

Similarly, put

$$\|(v,w)\|_{V\times W,\infty} = \max(\|v\|_V . \|w\|_W) \tag{5.12.2}$$

for every $(v,w) \in V \times W$. One can check that this defines a norm on $V \times W$ as well. We also have that

$$\|(v,w)\|_{V\times W,\infty} \le \|(v,w)\|_{V\times W,p} \le 2^{1/p}\, \|(v,w)\|_{V\times W,\infty} \tag{5.12.3}$$

for every $(v,w) \in V \times W$ and $1 \le p < \infty$.

If $1 \le p \le \infty$, then

$$\begin{aligned} d_{V\times W,p}((v,w),(v',w')) &= \|(v,w) - (v',w')\|_{V\times W,p} \\ &= \|(v-v', w-w')\|_{V\times W,p} \end{aligned} \tag{5.12.4}$$

defines a metric on $V \times W$, as usual. This is the same as the metric corresponding to p and the metrics on V, W associated to their norms as in the previous section. If V, W are Banach spaces, then $V \times W$ is complete with respect to this metric, as before. This means that $V \times W$ is a Banach space with respect to $\|\cdot\|_{V\times W,p}$.

Suppose that $\langle\cdot,\cdot\rangle_V$, $\langle\cdot,\cdot\rangle_W$ are inner products on V, W, respectively. It is easy to see that

$$\langle (v,w),(v',w')\rangle_{V\times W} = \langle v,v'\rangle_V + \langle w,w'\rangle_W \tag{5.12.5}$$

defines an inner product on $V \times W$. If $\|\cdot\|_V$, $\|\cdot\|_W$ are the norms on V, W associated to their inner products, respectively, then the norm on $V \times W$ associated to this inner product is the same as (5.12.1) with $p = 2$. If V, W are Hilbert spaces, then it follows that $V \times W$ is a Hilbert space with respect to this inner product, as in the preceding paragraph.

Let Z be another vector space over the real or complex numbers, as appropriate. If T_1, T_2 are linear mappings from V, W into Z, respectively, then

$$T((v,w)) = T_1(v) + T_2(w) \tag{5.12.6}$$

defines a linear mapping from $V \times W$ into Z. Note that T is uniquely determined by the conditions that

$$T((v, 0)) = T_1(v) \tag{5.12.7}$$

for every $v \in V$, and

$$T((0, w)) = T_2(w) \tag{5.12.8}$$

for every $w \in W$. Conversely, if T is any linear mapping from $V \times W$ into Z, then (5.12.7) and (5.12.8) define linear mappings T_1, T_2 from V, W into Z, respectively.

Let $1 \le p \le \infty$ be given, and suppose that T is a bounded linear mapping from $V \times W$ into Z, with respect to $\| \cdot \|_{V \times W, p}$ on $V \times W$. Let $\|T\|_{op,p}$ be the corresponding operator norm of T. It is easy to see that the associated linear mappings T_1, T_2 from V, W into Z, respectively, are bounded, with

$$\|T_1\|_{op,VZ}, \ \|T_2\|_{op,WZ} \le \|T\|_{op,p}. \tag{5.12.9}$$

Suppose now that T_1, T_2 are bounded linear mappings from V, W into Z, respectively. If $v \in V$ and $w \in W$, then

$$\begin{aligned} \|T((v,w))\|_Z &\le \|T_1(v)\|_Z + \|T_2(w)\|_Z \\ &\le \|T_1\|_{op,VZ} \, \|v\|_V + \|T_2\|_{op,WZ} \, \|w\|_W. \end{aligned} \tag{5.12.10}$$

This implies that T is bounded as a linear mapping from $V \times W$ into Z, with respect to $\| \cdot \|_{V \times W, p}$ on $V \times W$. More precisely, if $p = 1$, then we get that

$$\|T((v,w))\|_Z \le \max(\|T_1\|_{op,VZ}, \|T_2\|_{op,WZ}) \, \|(v,w)\|_{V \times W, 1}, \tag{5.12.11}$$

so that

$$\|T\|_{op,1} \le \max(\|T_1\|_{op,VZ}, \|T_2\|_{op,WZ}). \tag{5.12.12}$$

If $p > 1$, then let $1 \le q < \infty$ be the exponent conjugate to p. In this case,

$$\|T((v,w))\|_Z \le (\|T_1\|_{op,VZ}^q + \|T_2\|_{op,WZ}^q)^{1/q} \, \|(v,w)\|_{V \times W, p} \tag{5.12.13}$$

for every $v \in V$ and $w \in W$. This follows easily from (5.12.10) when $p = \infty$, so that $q = 1$. If $1 < p < \infty$, then one can use Hölder's inequality for functions on a set with two elements. Thus

$$\|T\|_{op,p} \le (\|T_1\|_{op,VZ}^q + \|T_2\|_{op,WZ}^q)^{1/q} \tag{5.12.14}$$

when $1 < p \le \infty$.

5.13 Bilinear mappings

Let V, W, and Z be vector spaces, all real or all complex. Also let B be a mapping from $V \times W$ into Z, so that $B(v,w)$ is defined as an element of Z for each $v \in V$ and $w \in W$. We say that B is *bilinear* on $V \times W$ if $B(v,w)$ is linear as a function of $v \in V$ for each $w \in W$, and linear as a function of $w \in W$ for each $v \in V$. Of course, if B is a bilinear mapping from $V \times W$ into Z, then

$$\widetilde{B}(w, v) = B(v, w) \tag{5.13.1}$$

defines a bilinear mapping from $W \times V$ into Z.

If $w \in W$, then put

$$B_{1,w}(v) = B(v, w) \tag{5.13.2}$$

for every $v \in V$. It is easy to see that B is bilinear on $V \times W$ if and only if $B_{1,w}$ is a linear mapping from V into Z for each $w \in W$, and

$$w \mapsto B_{1,w} \tag{5.13.3}$$

is a linear mapping from W into the space $\mathcal{L}(V, Z)$ of all linear mappings from V into Z. Similarly, if $v \in V$, then put

$$B_{2,v}(w) = B(v, w) \tag{5.13.4}$$

for every $w \in W$. As before, B is bilinear on $V \times W$ if and only if $B_{2,v}$ is a linear mapping from W into Z for each $v \in V$, and

$$v \mapsto B_{2,v} \tag{5.13.5}$$

is a linear mapping from V into $\mathcal{L}(W, Z)$.

Suppose that B is bilinear, and let $\|\cdot\|_V$, $\|\cdot\|_W$, and $\|\cdot\|_Z$ be norms on V, W, and Z, respectively. We say that B is *bounded* on $V \times W$ if there is a nonnegative real number C such that

$$\|B(v, w)\|_Z \le C\,\|v\|_V\,\|w\|_W \tag{5.13.6}$$

for every $v \in V$ and $w \in W$. This is the same as saying that $B_{1,w}$ is a bounded linear mapping from V into Z for each $w \in W$, with

$$\|B_{1,w}\|_{op,VZ} \le C\,\|w\|_W. \tag{5.13.7}$$

This means that (5.13.3) is a bounded linear mapping from W into the space $\mathcal{BL}(V, Z)$ of bounded linear mappings from V into Z, with operator norm less than or equal to C. Similarly, (5.13.6) is the same as saying that $B_{2,v}$ is a bounded linear mapping from W into Z for each $v \in V$, with

$$\|B_{2,v}\|_{op,WZ} \le C\,\|v\|_V. \tag{5.13.8}$$

Equivalently, this means that (5.13.5) is a bounded linear mapping from V into $\mathcal{BL}(W, Z)$, with operator norm less than or equal to C. Of course, (5.13.6) holds with $C = 0$ if and only if $B \equiv 0$ on $V \times W$.

Remember that $V \times W$ may be considered as a metric space, using a metric obtained from the metrics associated to the norms on V and W as in Section 5.11. If B is bounded on $V \times W$, then it is easy to see that B is continuous at $(0, 0)$, as a mapping from $V \times W$ into Z.

If $v, v' \in V$ and $w, w' \in W$, then

$$B(v,w) - B(v',w') = B(v - v', w) + B(v', w - w'). \tag{5.13.9}$$

If (5.13.6) holds for some $C \geq 0$, then we get that

$$\begin{aligned} \|B(v,w) - B(v',w')\|_Z &\leq \|B(v-v',w)\|_Z + \|B(v', w-w')\|_Z \\ &\leq C\,\|v - v'\|_V\,\|w\|_W + \|v'\|_V\,\|w - w'\|_W. \end{aligned} \tag{5.13.10}$$

It follows from (5.13.10) that B is uniformly continuous on bounded subsets of $V \times W$, with respect to the metric on Z associated to the norm. In particular, this implies that B is continuous on $V \times W$.

If B is continuous at $(0,0)$ on $V \times W$, then one can check that (5.13.6) holds for some $C \geq 0$. More precisely, it suffices to ask that B be bounded on a product of balls in V, W with positive radii centered at 0.

Let V_0, W_0 be dense linear subspaces of V, W, respectively, with respect to the metrics associated to their norms. Suppose that B_0 is a bounded bilinear mapping from $V_0 \times W_0$ into Z, using the restriction of the norms on V, W to V_0, W_0, respectively. If Z is complete with respect to the metric associated to its norm, then there is a unique extension of B_0 to a bounded bilinear mapping from $V \times W$ into Z. This can be obtained from the analogous statement for bounded linear mappings in Section 2.2, one variable at a time. This could also be obtained from the uniform continuity properties of B_0 on bounded subsets of $V_0 \times W_0$, as in the preceding paragraph, and the extension result for uniformly continuous mappings mentioned in Section 1.15.

5.14 Separate continuity

Let V, W and Z be metric or topological spaces, and let B be a mapping from $V \times W$ into Z. We say that B is *separately continuous* on $V \times W$ if $B(v,w)$ is continuous as a function of $v \in V$ for each $w \in W$, and continuous as a function of $w \in W$ for each $v \in V$. If B is continuous with respect to the product topology on $V \times W$, then B is sometimes said to be jointly continuous on $V \times W$. It is easy to see that this implies that B is separately continuous on $V \times W$.

Suppose now that V, W, and Z are vector spaces, all real or all complex, and equipped with norms $\|\cdot\|_V$, $\|\cdot\|_W$, and $\|\cdot\|_Z$, respectively. Suppose also that B is a bilinear mapping from $V \times W$ into Z, and let $B_{1,w}$, $B_{2,v}$ be as in (5.13.2), (5.13.4), respectively, for each $v \in V$, $w \in W$. In this case, separate continuity of B on $V \times W$ means that

$$B_{1,w} \text{ is a bounded linear mapping from } V \text{ into } Z \tag{5.14.1}$$

for every $w \in W$, and

$$B_{2,v} \text{ is a bounded linear mapping from } W \text{ into } Z \tag{5.14.2}$$

for every $v \in V$.

Suppose that B satisfies these two conditions, and let E be a nonempty subset of W. Thus

$$\mathcal{E}_{1,E} = \{B_{1,w} : w \in E\} \tag{5.14.3}$$

is a nonempty collection of bounded linear mappings from V into Z. If $v \in V$, then put

$$\mathcal{E}_{1,E}(v) = \{B_{1,w}(v) : w \in E\}, \tag{5.14.4}$$

as in Section 4.3. Equivalently,

$$\mathcal{E}_{1,E}(v) = B_{2,v}(E). \tag{5.14.5}$$

If E is a bounded subset of W, then we get that (5.14.5) is a bounded subset of Z, because of (5.14.2).

If V is a Banach space, then the Banach–Steinhaus theorem implies that the operator norms of $B_{1,w}$, $w \in E$, are uniformly bounded. If we take E to be the closed unit ball in W, then we get that B is bounded as a bilinear mapping from $V \times W$ into Z. This corresponds to a simplification of Theorem 2.17 on p51 of [162]. Of course, there is an analogous statement when W is a Banach space.

5.15 Bilinear and sesquilinear forms

Let V, W be vector spaces, both real or both complex. A bilinear mapping b from $V \times W$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, may be called a *bilinear functional* on $V \times W$, as on p106, 137 of [105]. The term *bilinear form* is used for this on p88 of [167], although this term is often used for the case where $V = W$.

If $w \in W$, then put $b_{1,w}(v) = b(v, w)$, as in Section 5.13. This is a linear functional on V, and

$$w \mapsto b_{1,w} \tag{5.15.1}$$

defines a linear mapping from W into the algebraic dual V^{alg} of V, as before. Similarly, if $v \in V$, then $b_{2,v}(w) = b(v, w)$ is a linear functional on W, and

$$v \mapsto b_{2,v} \tag{5.15.2}$$

defines a linear mapping from V into W^{alg}. Conversely, any linear mapping from W into V^{alg} or from V into W^{alg} corresponds to a bilinear functional on $V \times W$ in this way.

Now let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively. If b is bounded as a bilinear mapping on $V \times W$, with the standard absolute value function as the norm on $\mathbf{R}$ or $\mathbf{C}$, as appropriate, then $b_{1,w}$ is a bounded linear functional on V for each $w \in W$, and (5.15.1) defines a bounded linear mapping from W into the dual V' of V with respect to $\|\cdot\|_V$. Similarly, $b_{2,v}$ is a bounded linear functional on W for each $v \in V$, and (5.15.2) defines a bounded linear mapping from V into W'. Conversely, any bounded linear mapping from W into V' or from V into W' corresponds to a bounded bilinear functional on $V \times W$ in this way.

Let us take $V = W$ for the rest of the section. Suppose for the moment that V is a vector space over $\mathbf{R}$. Let $\langle \cdot, \cdot \rangle_V$ be an inner product on V, with associated norm $\| \cdot \|_V$. If A is a linear mapping from V into itself, then

$$b_A(v, w) = \langle A(v), w \rangle_V \tag{5.15.3}$$

defines a bilinear form on V. One can check that this is bounded as a bilinear mapping from $V \times V$ into $\mathbf{R}$ if and only if A is a bounded linear mapping from V into itself.

Conversely, if V is a Hilbert space, then it is well known that any bounded bilinear form b on V is of this form, for a unique bounded linear mapping A on V. More precisely, suppose that b is a bilinear form on V such that $b_{2,v}(w) = b(v, w)$ is a bounded linear functional on V, as a function of w, for each $v \in V$. This implies that for each $v \in V$ there is a unique $A(v) \in V$ such that

$$b_{2,v}(w) = \langle A(v), w \rangle_V \tag{5.15.4}$$

for every $w \in V$, as in Section 3.1. It is easy to see that A is a linear mapping from V into itself, because of uniqueness. If b is bounded as a bilinear form on V, then A is bounded as a linear mapping on V, as in the preceding paragraph.

Suppose that V is a vector space over $\mathbf{C}$ for the rest of the section. A mapping b from $V \times V$ into $\mathbf{C}$ is said to be a *sesquilinear form* on V if $b(v, w)$ is complex-linear in v for each $w \in V$, and conjugate-linear in w for each $v \in V$. In this case, if $v \in V$, then

$$\widetilde{b}_{2,v}(w) = \overline{b(v, w)} \tag{5.15.5}$$

is complex-linear in w. Note that b may be considered as a real-bilinear mapping from $V \times V$ into $\mathbf{C}$, which is to say that b is bilinear over $\mathbf{R}$, or equivalently that b is bilinear when V and $\mathbf{C}$ are considered as vector spaces over $\mathbf{R}$.

Let $\langle \cdot, \cdot \rangle_V$ be an inner product on V, as a complex vector space, and let $\| \cdot \|_V$ be the associated norm. In particular, $\langle \cdot, \cdot \rangle_V$ is sesquilinear on V. A sesquilinear form b on V is said to be *bounded* with respect to $\| \cdot \|_V$ if there is a nonnegative real number C such that

$$|b(v, w)| \le C \, \|v\|_V \, \|w\|_V \tag{5.15.6}$$

for every $v, w \in V$. This is the same as saying that b is bounded as a real-bilinear mapping from $V \times V$ into $\mathbf{C}$, where $\| \cdot \|_V$ is considered as a norm on V as a vector space over $\mathbf{R}$, and the standard absolute value function on $\mathbf{C}$ is considered as a norm on $\mathbf{C}$ as a vector space over $\mathbf{R}$.

If A is a linear mapping from V into itself, then (5.15.3) defines a sesquilinear form on V. One can check that this is a bounded sesquilinear form on V with respect to $\| \cdot \|_V$ if and only if A is a bounded linear mapping from V into itself, as before. If V is a complex Hilbert space, then it is well known that every bounded sesquilinear form b on V is of this form, for a unique bounded linear mapping A on V.

Indeed, suppose that b is a sesquilinear form on V such that (5.15.5) is a bounded linear functional on V, as a function of w, for each $v \in V$. This implies that for each $v \in V$ there is a unique $A(v) \in V$ such that

$$\widetilde{b}_{2,v}(w) = \langle w, A(v) \rangle_V \tag{5.15.7}$$

for every $w \in V$, as in Section 3.1 again. Equivalently, this means that

$$b(v, w) = \overline{\langle w, A(v) \rangle_V} = \langle A(v), w \rangle_V \tag{5.15.8}$$

for every $w \in V$. One can use uniqueness to get that A is a linear mapping from V into itself, as before. If b is bounded as a sesquilinear form on V, then A is bounded as a linear mapping on V, as in the previous paragraph.

Part II

Algebras, norms, and operators

Chapter 6

Algebras and norms

6.1 Algebras in the strict sense

Let $\mathcal{A}$ be a vector space over the real or complex numbers. If $\mathcal{A}$ is also equipped with a bilinear mapping from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$, then we say that $\mathcal{A}$ is an *algebra in the strict sense.* This bilinear mapping may be expressed as

$$(a, b) \mapsto a\, b, \tag{6.1.1}$$

although one may use other notation, depending on the situation.

If

$$(a\, b)\, c = a\, (b\, c) \tag{6.1.2}$$

for every $a, b, c \in \mathcal{A}$, then $\mathcal{A}$ is said to be an *associative algebra.* If

$$a\, b = b\, a \tag{6.1.3}$$

for every $a, b \in \mathcal{A}$, then $\mathcal{A}$ is said to be *commutative.* An element $e = e_{\mathcal{A}}$ of $\mathcal{A}$ is said to be the *multiplicative identity element* in $\mathcal{A}$ if

$$e_{\mathcal{A}}\, a = a\, e_{\mathcal{A}} = a \tag{6.1.4}$$

for every $a \in \mathcal{A}$. It is easy to see that the multiplicative identity element in $\mathcal{A}$ is unique when it exists.

If X is a nonempty set, then the spaces of all real and complex-valued functions on X are commutative associative algebras over $\mathbf{R}$ and $\mathbf{C}$, respectively, with respect to pointwise multiplication of functions. The function $\mathbf{1}_X$ equal to 1 at every point in X is the multiplicative identity element of each of these algebras.

Let V be a vector space over the real or complex numbers, and let $\mathcal{L}(V) = \mathcal{L}(V, V)$ be the space of linear mappings from V into itself. This is an associative algebra over the real or complex numbers, as appropriate, with respect to composition of linear mappings. The identity mapping $I = I_V$ on V is the multiplicative identity element in $\mathcal{L}(V)$.

Let $\mathcal{A}$ be an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$. Also let $\mathcal{A}_0$ be a linear subspace of $\mathcal{A}$ such that

$$a\,b \in \mathcal{A}_0 \tag{6.1.5}$$

for every $a, b \in \mathcal{A}_0$. Under these conditions, $\mathcal{A}_0$ is an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to the restriction of multiplication on $\mathcal{A}$ to $\mathcal{A}_0$. We say that $\mathcal{A}_0$ is a *subalgebra* of $\mathcal{A}$ in this case. Of course, if $\mathcal{A}$ is associative or commutative, then $\mathcal{A}_0$ has the same property.

Let X be a nonempty metric or topological space. The spaces $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ of continuous real and complex-valued functions on X are subalgebras of the algebras of all real or complex-valued functions on X, respectively. If X is equipped with the discrete metric or topology, then every function on X is continuous.

If U is a nonempty open subset of the complex plane, then the space $\mathcal{H}(U)$ of holomorphic functions on U is a subalgebra of $C(U, \mathbf{C})$. Similarly, the space $A(U)$ of continuous complex-valued functions on the closure $\overline{U}$ of U in $\mathbf{C}$ that are holomorphic on U is a subalgebra of $C(\overline{U}, \mathbf{C})$. There are analogous statements for holomorphic functions of several complex variables.

6.2 Norms on algebras

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers, and let $\|\cdot\|_{\mathcal{A}}$ be a norm on $\mathcal{A}$, as a vector space over $\mathbf{R}$ or $\mathbf{C}$. To say that multiplication on $\mathcal{A}$ is bounded as a bilinear mapping with respect to $\|\cdot\|_{\mathcal{A}}$ means that

$$\|a\,b\|_{\mathcal{A}} \le C\,\|a\|_{\mathcal{A}}\,\|b\|_{\mathcal{A}} \tag{6.2.1}$$

for some nonnegative real number C and all $a, b \in \mathcal{A}$, as in Section 5.13. This happens exactly when multiplication on $\mathcal{A}$ is continuous as a mapping from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$, with respect to the metric on $\mathcal{A}$ associated to the norm, and the corresponding product topology on $\mathcal{A} \times \mathcal{A}$, as before.

If (6.2.1) holds with $C = 1$, then $\|\cdot\|_{\mathcal{A}}$ is said to be *submultiplicative* on $\mathcal{A}$. If $\mathcal{A}$ has finite dimension as a vector space over $\mathbf{R}$ or $\mathbf{C}$, then it is easy to see that (6.2.1) holds for some $C \ge 0$, using the remarks in Section 1.11.

If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then (6.2.1)

$$\|e_{\mathcal{A}}\|_{\mathcal{A}} = \|e_{\mathcal{A}}\,e_{\mathcal{A}}\|_{\mathcal{A}} \le C\,\|e_{\mathcal{A}}\|_{\mathcal{A}}^2. \tag{6.2.2}$$

If $e_{\mathcal{A}} \ne 0$, so that $\|e_{\mathcal{A}}\|_{\mathcal{A}} > 0$, then it follows that

$$1 \le C\,\|e_{\mathcal{A}}\|_{\mathcal{A}}. \tag{6.2.3}$$

Note that $\mathcal{A} = \{0\}$ when $e_{\mathcal{A}} = 0$. Sometimes

$$\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1 \tag{6.2.4}$$

is included as a condition on a norm on $\mathcal{A}$.

If X is a nonempty set, then the spaces $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$ of bounded real and complex-valued functions on X are subalgebras of the spaces of all real and complex-valued functions on X, respectively. If f and g are bounded real or complex-valued functions on X, then it is easy to see that

$$\|f\,g\|_\infty \le \|f\|_\infty \, \|g\|_\infty. \tag{6.2.5}$$

Note that $\|\mathbf{1}_X\|_\infty = 1$.

If X is a nonempty metric or topological space, then the spaces $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ of bounded continuous real and complex-valued functions are subalgebras of $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ and of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively. If U is a nonempty open subset of $\mathbf{C}$, then the space $H^\infty(U)$ of bounded holomorphic functions on U is a subalgebra of $\mathcal{H}(U)$ and $C_b(U, \mathbf{C})$. Similarly, the space $A_b(U)$ of bounded continuous complex-valued functions on $\overline{U}$ that are holomorphic on U is a subalgebra of $A(U)$ and $C_b(\overline{U}, \mathbf{C})$.

Let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$. The space $\mathcal{BL}(V) = \mathcal{BL}(V, V)$ of all bounded linear mappings from V into itself with respect to $\|\cdot\|_V$ is a subalgebra of $\mathcal{L}(V)$. The corresponding operator norm $\|\cdot\|_{op} = \|\cdot\|_{op,VV}$ is submultiplicative on $\mathcal{BL}(V)$, as in Section 2.2. Note that I_V is automatically bounded on V, with

$$\|I_V\|_{op} = 1 \tag{6.2.6}$$

when $V \ne \{0\}$.

Let $\mathcal{A}$ be an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$ with a norm $\|\cdot\|_{\mathcal{A}}$ again. If $\mathcal{A}$ is not complete with respect to the metric associated to the norm, then one can pass to a completion to get a Banach space, as in Section 1.15. If (6.2.1) holds for some $C \ge 0$, then multiplication on $\mathcal{A}$ has a unique extension to the completion that satisfies the analogous conditions, as in Section 5.13.

Suppose that $\mathcal{A}$ is an associative algebra over $\mathbf{R}$ or $\mathbf{C}$. If $a \in \mathcal{A}$ and j is a positive integer, then a^j can be defined in $\mathcal{A}$ in the usual way. If $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$, then

$$\|a^j\|_{\mathcal{A}} \le \|a\|_{\mathcal{A}}^j. \tag{6.2.7}$$

If $\|a\|_{\mathcal{A}} < 1$, then it follows that

$$\|a^j\|_{\mathcal{A}} \to 0 \text{ as } j \to \infty. \tag{6.2.8}$$

6.3 Algebra homomorphisms

Let $\mathcal{A}$, $\mathcal{B}$ be algebras in the strict sense, both real or both complex. Also let ϕ be a linear mapping from $\mathcal{A}$ into $\mathcal{B}$, as vector spaces over $\mathbf{R}$ or $\mathbf{C}$. If

$$\phi(x\,y) = \phi(x)\,\phi(y) \tag{6.3.1}$$

for all $x, y \in \mathcal{A}$, then ϕ is said to be an *algebra homomorphism* from $\mathcal{A}$ into $\mathcal{B}$. If $\mathcal{A}$, $\mathcal{B}$ also have multiplicative identity elements $e_{\mathcal{A}}$, $e_{\mathcal{B}}$, respectively, then one may wish to ask that

$$\phi(e_{\mathcal{A}}) = e_{\mathcal{B}} \tag{6.3.2}$$

in addition. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$ and ϕ maps $\mathcal{A}$ onto $\mathcal{B}$, then (6.3.1) implies that $\phi(e_{\mathcal{A}})$ is the multiplicative identity element in $\mathcal{B}$.

Let $\mathcal{C}$ be another real or complex algebra in the strict sense, as appropriate. If ϕ is a homomorphism from $\mathcal{A}$ into $\mathcal{B}$, and ψ is a homomorphism from $\mathcal{B}$ into $\mathcal{C}$, then their composition $\psi \circ \phi$ is a homomorphism from $\mathcal{A}$ into $\mathcal{C}$.

A one-to-one homomorphism ϕ from $\mathcal{A}$ onto $\mathcal{B}$ is called an *algebra isomorphism*, as usual. In this case, ϕ^{-1} is an isomorphism from $\mathcal{B}$ onto $\mathcal{A}$. If ψ is an isomorphism from $\mathcal{B}$ onto $\mathcal{C}$, then it follows that $\psi \circ \phi$ is an isomorphism from $\mathcal{A}$ onto $\mathcal{C}$.

Let $\mathcal{A}$ be an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$ again. If $a, x \in \mathcal{A}$, then put

$$L_a(x) = a\,x. \tag{6.3.3}$$

This defines L_a as a linear mapping from $\mathcal{A}$ into itself, as a vector space over $\mathbf{R}$ or $\mathbf{C}$, which is the *left multiplication operator* associated to a. Note that

$$a \mapsto L_a \tag{6.3.4}$$

is linear as a mapping from $\mathcal{A}$ into the space $\mathcal{L}(\mathcal{A})$ of linear mappings from $\mathcal{A}$ into itself.

Suppose that $e_{\mathcal{A}}$ is a multiplicative identity element in $\mathcal{A}$. Thus

$$L_{e_{\mathcal{A}}} = I_{\mathcal{A}}, \tag{6.3.5}$$

the identity mapping on $\mathcal{A}$. We also have that

$$L_a(e_{\mathcal{A}}) = a \tag{6.3.6}$$

for every $a \in \mathcal{A}$. In particular, this implies that (6.3.4) is injective.

Suppose that $\mathcal{A}$ is an associative algebra. If $a, b, x \in \mathcal{A}$, then

$$L_a(L_b(x)) = a\,(b\,x) = (a\,b)\,x = L_{a\,b}(x). \tag{6.3.7}$$

Equivalently, this means that

$$L_a \circ L_b = L_{a\,b}. \tag{6.3.8}$$

Thus (6.3.4) is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$ in this case.

Let $\|\cdot\|_{\mathcal{A}}$ be a norm on $\mathcal{A}$, and suppose that (6.2.1) holds for some $C \geq 0$. If $a \in \mathcal{A}$, then we get that

$$\|L_a(x)\|_{\mathcal{A}} \leq C\,\|a\|_{\mathcal{A}}\,\|x\|_{\mathcal{A}} \tag{6.3.9}$$

for every $x \in \mathcal{A}$. This means that L_a is a bounded linear mapping from $\mathcal{A}$ into itself, with

$$\|L_a\|_{op} \leq C\,\|a\|_{\mathcal{A}}. \tag{6.3.10}$$

It follows that (6.3.4) is a bounded linear mapping from $\mathcal{A}$ into the space $\mathcal{BL}(\mathcal{A})$ of bounded linear mappings from $\mathcal{A}$ into itself.

If $e_{\mathcal{A}}$ is a multiplicative identity element in $\mathcal{A}$, then

$$\|a\|_{\mathcal{A}} \le \|L_a\|_{op}\, \|e_{\mathcal{A}}\|_{\mathcal{A}} \tag{6.3.11}$$

for each $a \in \mathcal{A}$, by (6.3.6). If $C = 1$ and (6.2.4) holds, then we get that (6.3.4) is an isometric linear mapping from $\mathcal{A}$ into $\mathcal{BL}(\mathcal{A})$.

6.4 Opposite algebra homomorphisms

Let $\mathcal{A}$, $\mathcal{B}$ be algebras in the strict sense, both real or both complex, and let ϕ be a linear mapping from $\mathcal{A}$ into $\mathcal{B}$, as vector spaces over $\mathbf{R}$ or $\mathbf{C}$. If

$$\phi(x\,y) = \phi(y)\,\phi(x) \tag{6.4.1}$$

for all $x, y \in \mathcal{A}$, then ϕ is said to be an *opposite algebra homomorphism* from $\mathcal{A}$ into $\mathcal{B}$. If $\mathcal{A}$, $\mathcal{B}$ have multiplicative identity elements $e_{\mathcal{A}}$, $e_{\mathcal{B}}$, then one may wish to ask that $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$ too. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$ and ϕ maps $\mathcal{A}$ onto $\mathcal{B}$, then (6.4.1) implies that $\phi(e_{\mathcal{A}})$ is the multiplicative identity element in $\mathcal{B}$, as before.

A one-to-one opposite algebra homomorphism ϕ from $\mathcal{A}$ onto $\mathcal{B}$ is called an *opposite algebra isomorphism*. This implies that ϕ^{-1} is an opposite algebra isomorphism from $\mathcal{B}$ onto $\mathcal{A}$.

Let ϕ be an opposite algebra isomorphism from $\mathcal{A}$ onto itself. If

$$\phi \circ \phi = I_{\mathcal{A}}, \tag{6.4.2}$$

then ϕ is called an *algebra involution* on $\mathcal{A}$.

If $\mathcal{A}$ is an algebra in the strict sense over $\mathbf{C}$, then $\mathcal{A}$ may also be considered as an algebra in the strict sense over $\mathbf{R}$. In this case, one may be interested in *conjugate-linear algebra involutions* on $\mathcal{A}$. These are algebra involutions on $\mathcal{A}$ as an algebra in the strict sense over $\mathbf{R}$ that are also conjugate-linear.

Let $(V, \langle\cdot,\cdot\rangle_V)$ be a real or complex Hilbert space, with the associated norm $\|\cdot\|_V$. If T is a bounded linear mapping from V into itself, then its adjoint T^* is a bounded linear mapping on V as well, as in Section 3.5. In fact, $T \mapsto T^*$ is an algebra involution on $\mathcal{BL}(V)$, which is conjugate-linear in the complex case.

Let $\mathcal{A}$ be any algebra in the strict sense over the real or complex numbers again. If $a, x \in \mathcal{A}$, then put

$$R_a(x) = x\,a. \tag{6.4.3}$$

This defines a linear mapping from $\mathcal{A}$ into itself, which is the *right multiplication operator* associated to a. Clearly

$$a \mapsto R_a \tag{6.4.4}$$

is a linear mapping from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then

$$R_{e_{\mathcal{A}}} = I_{\mathcal{A}} \tag{6.4.5}$$

and

$$R_a(e_{\mathcal{A}}) = a \tag{6.4.6}$$

for every $a \in \mathcal{A}$.

If $\mathcal{A}$ is an associative algebra, then

$$R_b(R_a(x)) = (x\,a)\,b = x\,(a\,b) = R_{a\,b}(x) \tag{6.4.7}$$

for every $a, b, x \in \mathcal{A}$. This means that

$$R_b \circ R_a = R_{a\,b}, \tag{6.4.8}$$

so that (6.4.4) is an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$.

Let $\|\cdot\|_{\mathcal{A}}$ be a norm on $\mathcal{A}$ that satisfies (6.2.1) for some $C \ge 0$. If $a \in \mathcal{A}$, then

$$\|R_a(x)\|_{\mathcal{A}} \le C\,\|a\|_{\mathcal{A}}\,\|x\|_{\mathcal{A}} \tag{6.4.9}$$

for every $x \in \mathcal{A}$, so that R_a is a bounded linear mapping on $\mathcal{A}$, with

$$\|R_a\|_{op} \le C\,\|a\|_{\mathcal{A}}. \tag{6.4.10}$$

If $e_{\mathcal{A}}$ is a multiplicative identity element in $\mathcal{A}$, then

$$\|a\|_{\mathcal{A}} \le \|R_a\|_{op}\,\|e_{\mathcal{A}}\|_{\mathcal{A}} \tag{6.4.11}$$

for every $a \in \mathcal{A}$. In particular, (6.4.4) is an isometric linear mapping from $\mathcal{A}$ into $\mathcal{BL}(\mathcal{A})$ when $C = 1$ and (6.2.4) holds.

6.5 Banach algebras

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers, with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $\mathcal{A}$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, then $\mathcal{A}$ is said to be a *Banach algebra* with respect to $\|\cdot\|_{\mathcal{A}}$. Sometimes one may also ask that $\mathcal{A}$ have a multiplicative identity element $e_{\mathcal{A}}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$.

Let $\mathcal{A}$ be an associative algebra over $\mathbf{R}$ or $\mathbf{C}$ with a multiplicative identity element $e_{\mathcal{A}}$. An element x of $\mathcal{A}$ is said to be *invertible* in $\mathcal{A}$ if there is an element x^{-1} of $\mathcal{A}$ such that

$$x\,x^{-1} = x^{-1}\,x = e_{\mathcal{A}}. \tag{6.5.1}$$

One can check that x^{-1} is unique when it exists, in which case it is called the *multiplicative inverse* of x in $\mathcal{A}$.

If $a \in \mathcal{A}$ and n is a nonnegative integer, then

$$(e_{\mathcal{A}} - a)\sum_{j=0}^{n} a^j = \Big(\sum_{j=0}^{n} a^j\Big)(e_{\mathcal{A}} - a) = e_{\mathcal{A}} - a^{n+1}, \tag{6.5.2}$$

by a standard argument. Here a^j is interpreted as being equal to $e_{\mathcal{A}}$ when $j = 0$. More precisely, the left side and middle part of (6.5.2) are equal to

$$\sum_{j=0}^{n} a^j - \sum_{j=0}^{n} a^{j+1} = \sum_{j=0}^{n} a^j - \sum_{j=1}^{n+1} a^j = e_{\mathcal{A}} - a^{n+1}. \tag{6.5.3}$$

Suppose now that $\mathcal{A}$ is a Banach algebra with respect to a norm $\|\cdot\|_{\mathcal{A}}$, and that $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Note that $\|a^j\|_{\mathcal{A}} \le \|a\|_{\mathcal{A}}^j$ for each $j \ge 0$, where $\|a\|_{\mathcal{A}}^j$ is interpreted as being equal to 1 when $j = 0$. If

$$\|a\|_{\mathcal{A}} < 1, \tag{6.5.4}$$

then $\sum_{j=0}^{\infty} a^j$ converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, with

$$\sum_{j=0}^{\infty} \|a^j\|_{\mathcal{A}} \le \sum_{j=0}^{\infty} \|a\|_{\mathcal{A}}^j = (1 - \|a\|_{\mathcal{A}})^{-1}. \tag{6.5.5}$$

This implies that $\sum_{j=0}^{\infty} a^j$ converges in $\mathcal{A}$, with

$$\Big\|\sum_{j=0}^{\infty} a^j\Big\|_{\mathcal{A}} \le \sum_{j=0}^{\infty} \|a^j\|_{\mathcal{A}} \le (1 - \|a\|_{\mathcal{A}})^{-1}, \tag{6.5.6}$$

as in Section 1.7.

Under these conditions, we have that

$$(e_{\mathcal{A}} - a) \sum_{j=0}^{\infty} a^j = \Big(\sum_{j=0}^{\infty} a^j\Big)(e_{\mathcal{A}} - a) = e_{\mathcal{A}}, \tag{6.5.7}$$

because of (6.5.2). This means that $e_{\mathcal{A}} - a$ is invertible in $\mathcal{A}$, with

$$(e_{\mathcal{A}} - a)^{-1} = \sum_{j=0}^{\infty} a^j. \tag{6.5.8}$$

We also have that

$$\|(e_{\mathcal{A}} - a)^{-1}\|_{\mathcal{A}} \le (1 - \|a\|_{\mathcal{A}})^{-1}, \tag{6.5.9}$$

by (6.5.6).

6.6 More on invertible elements

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. If x, y are invertible elements of $\mathcal{A}$, then it is easy to see that $x\,y$ is invertible in $\mathcal{A}$ too, with

$$(x\,y)^{-1} = y^{-1}\,x^{-1}. \tag{6.6.1}$$

Thus the collection $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$ is a group.

Suppose that $\mathcal{A}$ is a Banach algebra with respect to a norm $\|\cdot\|_{\mathcal{A}}$, with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Let $x \in G(\mathcal{A})$ and $y \in \mathcal{A}$ be given, and observe that

$$y = x - (x - y) = (e_{\mathcal{A}} - (x - y)\, x^{-1})\, x. \tag{6.6.2}$$

If

$$\|(x - y)\, x^{-1}\|_{\mathcal{A}} < 1, \tag{6.6.3}$$

then

$$e_{\mathcal{A}} - (x - y)\, x^{-1} \in G(A), \tag{6.6.4}$$

as in the previous section. This implies that

$$y \in G(\mathcal{A}), \tag{6.6.5}$$

by (6.6.2).

Note that (6.6.3) holds when

$$\|x - y\|_{\mathcal{A}}\, \|x^{-1}\|_{\mathcal{A}} < 1. \tag{6.6.6}$$

Of course, this is the same as saying that

$$\|x - y\|_{\mathcal{A}} < 1/\|x^{-1}\|_{\mathcal{A}}. \tag{6.6.7}$$

In particular,

$$G(\mathcal{A}) \text{ is an open set in } \mathcal{A}, \tag{6.6.8}$$

with respect to the metric on $\mathcal{A}$ associated to the norm.

If (6.6.3) holds, then we also get that

$$\|(e_{\mathcal{A}} - (x - y)\, x^{-1})^{-1}\|_{\mathcal{A}} \le (1 - \|(x - y)\, x^{-1}\|_{\mathcal{A}})^{-1}, \tag{6.6.9}$$

as in (6.5.9). It follows that

$$\begin{aligned} \|y^{-1}\|_{\mathcal{A}} &= \|x^{-1}\, (e - (x - y)\, x^{-1})^{-1}\|_{\mathcal{A}} \\ &\le \|x^{-1}\|_{\mathcal{A}}\, (1 - \|(x - y)\, x^{-1}\|_{\mathcal{A}})^{-1}, \end{aligned} \tag{6.6.10}$$

using (6.6.2) in the first step. If (6.6.7) holds, then we have that

$$\|y^{-1}\|_{\mathcal{A}} \le \|x^{-1}\|_{\mathcal{A}}\, (1 - \|x - y\|_{\mathcal{A}}\, \|x^{-1}\|_{\mathcal{A}})^{-1}. \tag{6.6.11}$$

If $y \in G(\mathcal{A})$, then

$$y^{-1} - x^{-1} = y^{-1}\, (x - y)\, x^{-1}, \tag{6.6.12}$$

so that

$$\|y^{-1} - x^{-1}\|_{\mathcal{A}} \le \|y^{-1}\|_{\mathcal{A}}\, \|x - y\|_{\mathcal{A}}\, \|x^{-1}\|_{\mathcal{A}}. \tag{6.6.13}$$

If (6.6.7) holds, then we can combine this with (6.6.11) to get that

$$\|y^{-1} - x^{-1}\|_{\mathcal{A}} \le \|x^{-1}\|_{\mathcal{A}}^2\, (1 - \|x - y\|_{\mathcal{A}}\, \|x^{-1}\|_{\mathcal{A}})^{-1}\, \|x - y\|_{\mathcal{A}}. \tag{6.6.14}$$

We may consider $G(A)$ as a metric space, using the restriction to $G(\mathcal{A})$ of the metric on $\mathcal{A}$ associated to the norm. The topology determined on $G(\mathcal{A})$ by this metric is the same as the topology induced on $G(\mathcal{A})$ by the topology determined on $\mathcal{A}$ by the metric associated to the norm, by standard arguments. The previous arguments imply that

$$w \mapsto w^{-1} \tag{6.6.15}$$

is continuous on $G(\mathcal{A})$. More precisely, this mapping is continuous at any x in $G(\mathcal{A})$, because of (6.6.14).

Let C_0 be a nonnegative real number, and suppose that $x, y \in G(A)$ satisfy

$$\|x^{-1}\|_{\mathcal{A}},\ \|y^{-1}\|_{\mathcal{A}} \le C_0. \tag{6.6.16}$$

In this case, we have that

$$\|y^{-1} - x^{-1}\|_{\mathcal{A}} \le C_0^2\, \|x - y\|_{\mathcal{A}}, \tag{6.6.17}$$

by (6.6.13).

Note that multiplication on $G(\mathcal{A})$ is continuous, as a mapping

$$\text{from } G(\mathcal{A}) \times G(\mathcal{A}) \text{ into } G(\mathcal{A}), \tag{6.6.18}$$

with respect to the appropriate product topology on $G(\mathcal{A}) \times G(\mathcal{A})$, or a suitable product metric. This follows from the analogous statement for multiplication on $\mathcal{A}$, as in Section 6.2. This implies that

$$G(\mathcal{A}) \text{ is a topological group,} \tag{6.6.19}$$

because of the continuity of (6.6.15) on $G(\mathcal{A})$.

6.7 Some additional properties of $G(\mathcal{A})$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. In the previous two sections, we used completeness of $\mathcal{A}$ to get the invertibility of elements of $\mathcal{A}$ under certain conditions. We can use slightly different arguments to get analogous properties of invertible elements of $\mathcal{A}$, even if $\mathcal{A}$ may not be complete.

Let $x, y \in G(\mathcal{A})$ be given, and note that

$$\|e_{\mathcal{A}}\|_{\mathcal{A}} \le \|x\|_{\mathcal{A}}\, \|x^{-1}\|_{\mathcal{A}}. \tag{6.7.1}$$

We also have that

$$\|y^{-1} - x^{-1}\|_{\mathcal{A}} \le \|y^{-1}\|_{\mathcal{A}}\, \|(x - y)\, x^{-1}\|_{\mathcal{A}}, \tag{6.7.2}$$

because of (6.6.12). It follows that

$$\begin{aligned} \|y^{-1}\|_{\mathcal{A}} &\le \|x^{-1}\|_{\mathcal{A}} + \|y^{-1} - x^{-1}\|_{\mathcal{A}} \\ &\le \|x^{-1}\|_{\mathcal{A}} + \|y^{-1}\|_{\mathcal{A}}\, \|(x - y)\, x^{-1}\|\mathcal{A}. \end{aligned} \tag{6.7.3}$$

This implies that

$$(1 - \|(x - y)\, x^{-1}\|_{\mathcal{A}})\, \|y^{-1}\|_{\mathcal{A}} \le \|x^{-1}\|_{\mathcal{A}}. \tag{6.7.4}$$

If (6.6.3) holds, then we get that (6.6.10) holds.

If (6.6.7) holds, then (6.6.11) holds, as before. We can combine this with (6.6.13) to get that (6.6.14) holds in this case. This implies that (6.6.15) is continuous on $G(\mathcal{A})$, as before. It follows that $G(\mathcal{A})$ is a topological group, as before.

Let $\{x_j\}_{j=1}^{\infty}$ be a sequence of elements of $G(\mathcal{A})$ that converges to $x \in \mathcal{A}$ with respect to the metric associated to the norm. Suppose that there is a nonnegative real number C_1 such that

$$\|x_j^{-1}\|_{\mathcal{A}} \le C_1 \tag{6.7.5}$$

for each j. This implies that

$$\|x_j^{-1} - x_l^{-1}\|_{\mathcal{A}} \le C_1\, \|x_j - x_l\|_{\mathcal{A}} \tag{6.7.6}$$

for all j, l, by (6.6.17). Of course, $\{x_j\}_{j=1}^{\infty}$ is a Cauchy sequence with respect to the metric on $\mathcal{A}$ associated to the norm, because it converges, by hypohtesis. It follows that $\{x_j^{-1}\}_{j=1}^{\infty}$ is a Cauchy sequence in $\mathcal{A}$ too, because of (6.7.6).

If $\mathcal{A}$ is complete with respect to the metric associated to the norm, then $\{x_j^{-1}\}_{j=1}^{\infty}$ converges to an element z of $\mathcal{A}$. One can check that

$$x\, z = z\, x = e_{\mathcal{A}}, \tag{6.7.7}$$

so that $x \in G(\mathcal{A})$, with $x^{-1} = z$. Note that

$$\|x^{-1}\|_{\mathcal{A}} = \|z\|_{\mathcal{A}} \le C_1. \tag{6.7.8}$$

This shows that

$$\{y \in G(\mathcal{A}) : \|y^{-1}\|_{\mathcal{A}} \le C_1\} \tag{6.7.9}$$

is a closed set in $\mathcal{A}$, with respect to the metric associated to the norm. If $\mathcal{A}$ is not necessarily complete, then (6.7.9) is at least relatively closed in $G(\mathcal{A})$, which is to say that it is a closed set in $G(\mathcal{A})$ with respect to the restriction to $G(\mathcal{A})$ of the metric on $\mathcal{A}$ associated to the norm.

6.8 The spectrum of an element

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. If $x \in \mathcal{A}$, then the *spectrum* of x with respect to $\mathcal{A}$ is the set $\sigma_{\mathcal{A}}(x)$ of $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that

$$\lambda\, e_{\mathcal{A}} - x \notin G(\mathcal{A}). \tag{6.8.1}$$

The complementary set of $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that

$$\lambda\, e_{\mathcal{A}} - x \in G(\mathcal{A}) \tag{6.8.2}$$

is called the *resolvent set* of x with respect to $\mathcal{A}$.

Let X be a nonempty set, and suppose for the moment that $\mathcal{A}$ is the algebra of all real or complex-valued functions on X. If $a \in \mathcal{A}$, then it is easy to see that the spectrum of a with respect to $\mathcal{A}$ is equal to $a(X)$.

Suppose now that $\mathcal{A}$ is the algebra of all bounded real or complex-valued functions on X. If $a \in \mathcal{A}$, then one can check that the spectrum of a with respect to $\mathcal{A}$ is the closure of $a(X)$ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate.

Let V be a finite-dimensional vector space over the real or complex numbers, and let T be a linear mapping from V into itself. It is well known that T is one-to-one on V if and only if T maps V onto itself, in which case T is invertible on V. If T is not invertible on V, then it follows that the kernel of T is nontrivial. This implies that the spectrum of T with respect to $\mathcal{L}(V)$ is the usual set of eigenvalues of T, corresponding to nonzero eigenvectors in V.

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a real or complex Banach algebra, with a multiplicative identity element $e_{\mathcal{A}}$ with norm 1. If $x \in \mathcal{A}$, then the resolvent set of x is an open set in $\mathbf{R}$ or $\mathbf{C}$, because $G(\mathcal{A})$ is an open set in $\mathcal{A}$, as in Section 6.6. This implies that

$$\sigma_{\mathcal{A}}(x) \text{ is a closed set} \tag{6.8.3}$$

in $\mathbf{R}$ or $\mathbf{C}$, as appropriate.

If λ is a nonzero real or complex number, as appropriate, then

$$\lambda\, e_{\mathcal{A}} - x = \lambda\,(e_{\mathcal{A}} - \lambda^{-1}\, x). \tag{6.8.4}$$

If $|\lambda| > \|x\|_{\mathcal{A}}$, then

$$e_{\mathcal{A}} - \lambda^{-1}\, x \in G(\mathcal{A}), \tag{6.8.5}$$

as in Section 6.5. This means that (6.8.2) holds in this case. It follows that

$$|\lambda| \le \|x\|_{\mathcal{A}} \text{ when } \lambda \in \sigma_{\mathcal{A}}(x). \tag{6.8.6}$$

If $\mathcal{A}$ is a complex Banach algebra, then it is well known that

$$\sigma_{\mathcal{A}}(x) \ne \emptyset \tag{6.8.7}$$

for every $x \in \mathcal{A}$. The basic idea is that if $\sigma_{\mathcal{A}}(x) = \emptyset$, then

$$(\lambda\, e_{\mathcal{A}} - x)^{-1} \tag{6.8.8}$$

is a holomorphic function of λ on the complex plane with values in $\mathcal{A}$. One can also check that this tends to 0 as $|\lambda| \to \infty$. This would imply that (6.8.8) is a constant function of λ, which would have to be equal to 0.

More precisely, if μ is a bounded linear functional on $\mathcal{A}$, then

$$\mu((\lambda\, e_{\mathcal{A}} - x)^{-1}) \tag{6.8.9}$$

would be a complex-valued holomorphic function of λ on the complex plane. The same type of argument mentioned in the preceding paragraph implies that

(6.8.9) is equal to 0 on **C**, by standard results from complex analysis. This implies that (6.8.8) is equal to 0 on **C**, by the Hahn–Banach theorem.

Suppose that $\mathcal{A} = \mathbf{C}$, considered as a Banach algebra over the real numbers, using the standard absolute value function as the norm. It is easy to see that any element of $\mathcal{A}$ with nonzero imaginary part has empty spectrum with respect to $\mathcal{A}$.

Let $\mathcal{A}$ be the set of rational functions in a single variable with complex coefficients. This is a field that contains **C** as a subfield, and in particular $\mathcal{A}$ may be considered as a commutative associative algebra over **C** with a multiplicative identity element. One can check that the spectrum of any element of $\mathcal{A}$ that does not correspond to a constant function is the empty set.

6.9 Homomorphisms into R, C

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let h be a homomorphism from $\mathcal{A}$ into **R** or **C**, as appropriate, as an algebra over itself. Observe that

$$h(e_{\mathcal{A}})^2 = h(e_{\mathcal{A}}^2) = h(e_{\mathcal{A}}), \tag{6.9.1}$$

so that $h(e_{\mathcal{A}})\,(h(e_{\mathcal{A}}) - 1) = 0$. This means that $h(e_{\mathcal{A}}) = 0$ or

$$h(e_{\mathcal{A}}) = 1. \tag{6.9.2}$$

It is easy to see that $h \equiv 0$ on $\mathcal{A}$ when $h(e_{\mathcal{A}}) = 0$. Let us suppose from now on in this section that $h \not\equiv 0$ on $\mathcal{A}$, so that (6.9.2) holds.

Let X be a nonempty set, and suppose for the moment that $\mathcal{A}$ is an algebra of real or complex-valued functions on X, with respect to pointwise multiplication of functions. If $x \in X$, then

$$h_x(a) = a(x) \tag{6.9.3}$$

defines a homomorphism from $\mathcal{A}$ into **R** or **C**, as appropriate.

Using (6.9.2), we get that

$$h(a) \neq 0 \text{ when } a \in G(\mathcal{A}). \tag{6.9.4}$$

If a is any element of $\mathcal{A}$, then

$$h(h(a)\,e_{\mathcal{A}} - a) = h(a)\,h(e_{\mathcal{A}}) - h(a) = 0, \tag{6.9.5}$$

so that

$$h(a)\,e_{\mathcal{A}} - a \notin G(\mathcal{A}). \tag{6.9.6}$$

This means that

$$h(a) \in \sigma_{\mathcal{A}}(a). \tag{6.9.7}$$

Suppose now that $\mathcal{A}$ is a Banach algebra with respect to a norm $\|\cdot\|_{\mathcal{A}}$, with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $a \in \mathcal{A}$, then

$$|h(a)| \leq \|a\|_{\mathcal{A}}, \tag{6.9.8}$$

because of (6.8.6) and (6.9.7). This implies that h is a bounded linear functional on $\mathcal{A}$, with dual norm with respect to $\|\cdot\|_{\mathcal{A}}$ equal to 1, because of (6.9.2).

Let $\mathcal{A}'$ be the dual space of bounded linear functionals on $\mathcal{A}$, as usual. One can check that

(6.9.9) the set of nonzero algebra homomorphisms h from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$,

as appropriate, is a closed set in $\mathcal{A}'$, with respect to the weak* topology. It follows that

(6.9.10) (6.9.9) is compact with respect to the weak* topology,

by the Banach–Alaoglu theorem.

Suppose that $\mathcal{A}$ is a complex Banach algebra, and that every nonzero element of $\mathcal{A}$ is invertible. If $x \in \mathcal{A}$, then $\sigma_{\mathcal{A}}(x) \neq \emptyset$, as in Section 6.8. In this case, we get that $\sigma_{\mathcal{A}}(x)$ contains exactly one element. In fact,

$$x = \lambda\, e_{\mathcal{A}} \tag{6.9.11}$$

for some $\lambda \in \mathbf{C}$. This corresponds to a famous theorem of Gelfand and Mazur.

If $\mathcal{A}$ is commutative, then this can be used to get some homomorphisms from $\mathcal{A}$ into $\mathbf{C}$. This will be discussed further in Section 6.12.

6.10 Quotient spaces

Let V be a vector space over the real or complex numbers, and let W be a linear subspace of V. Under these conditions, the corresponding *quotient space* V/W may be defined as a vector space over the real or complex numbers, as appropriate, in a standard way. There is also a natural *quotient mapping* q from V onto V/W, which is a linear mapping with kernel equal to W.

Let N_V be a seminorm on V. We can use this to define a nonnegative real-valued function $N_{V/W}$ on V/W such that

$$N_{V/W}(q(v)) = \inf\{N_V(v - w) : w \in W\} \tag{6.10.1}$$

for every $v \in V$. More precisely, if $u, v \in V$ satisfy $q(u) = q(v)$, then $u - v$ is an element of W, and it is easy to see that (6.10.1) is equal to the analogous definition of $N_{V/W}(q(u))$. Thus (6.10.1) depends only on $q(v) \in V/W$, and not on the choice of $v \in V$, so that $N_{V/W}$ is well defined on V/W.

If $v \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then one can check that

$$N_{V/W}(t\, q(v)) = N_{V/W}(q(t\, v)) = |t|\, N_{V/W}(q(v)). \tag{6.10.2}$$

If $v_1, v_2 \in V$ and $w_1, w_2 \in W$, then

$$\begin{aligned} N_{V/W}(q(v_1) + q(v_2)) &= N_{V/W}(q(v_1 + v_2)) \\ &\le N_V((v_1 + v_2) - (w_1 + w_2)) \\ &\le N_V(v_1 - w_1) + N_V(v_2 - w_2). \end{aligned} \tag{6.10.3}$$

Using this, one can get that

$$N_{V/W}(q(v_1) + q(v_2)) \le N_{V/W}(q_1(v)) + N_{V/W}(q(v_2)). \tag{6.10.4}$$

This shows that $N_{V/W}$ is a seminorm on V/W.

Suppose now that N_V is a norm on V. Observe that $v \in V$ satisfies

$$N_{V/W}(q(v)) = 0 \tag{6.10.5}$$

if and only if v is in the closure of W in V, with respect to the metric on V associated to N_V. Suppose that

$$W \text{ is also a closed set in } V, \tag{6.10.6}$$

so that (6.10.5) holds if and only if $v \in W$. This means that (6.10.5) holds if and only if $q(v) = 0$ in V/W, so that

$$N_{V/W} \text{ defines a norm on } V/W. \tag{6.10.7}$$

Of course,

$$N_{V/W}(q(v)) \le N_V(v) \tag{6.10.8}$$

for every $v \in V$, by construction. This implies that q is bounded as a linear mapping from V onto V/W. More precisely, the operator norm of q is equal to 1 when W is a proper closed linear subspaces of V.

Suppose that V is a Banach space with respect to N_V, and let us check that V/W is a Banach space with respect to $N_{V/W}$. It suffices to verify that absolute convergence of a series in V/W implies convergence in V/W, as in Section 1.7. An infinite series in V/W may be expressed as

$$\sum_{j=1}^{\infty} q(v_j), \tag{6.10.9}$$

where $v_1, v_2, v_3, \ldots$ is a sequence of elements of V. We would like to show that this series converges in V/W when

$$\sum_{j=1}^{\infty} N_{V/W}(q(v_j)) \tag{6.10.10}$$

converges as an infinite series of nonnegative real numbers.

We can choose $w_j \in W$ such that $N_V(v_j - w_j)$ is as close as we like to $N_{V/W}(q(v_j))$ for each j, by the definition (6.10.1) of $N_{V/W}$. In particular, we can do this in such a way that

$$\sum_{j=1}^{\infty} N_V(v_j - w_j) \tag{6.10.11}$$

converges. This implies that

$$\sum_{j=1}^{\infty} (v_j - w_j) \tag{6.10.12}$$

converges in V, as in Section 1.7. It follows that

$$\sum_{j=1}^{\infty} q(v_j - w_j) \tag{6.10.13}$$

converges in V/W. This series is the same as (6.10.9), because $q(v_j - w_j) = q(v_j)$ for each j.

6.11 Two-sided ideals

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers. A linear subspace $\mathcal{I}$ of $\mathcal{A}$ is said to be a *two-sided ideal* in $\mathcal{A}$ if for every $a \in \mathcal{A}$ and $x \in \mathcal{I}$, we have that

$$a\,x,\ x\,a \in \mathcal{I}. \tag{6.11.1}$$

Of course, the quotient $\mathcal{A}/\mathcal{I}$ may be defined initially as a vector space over the real or complex numbers, as appropriate, as in the previous section. Let q be the natural quotient mapping from $\mathcal{A}$ onto $\mathcal{A}/\mathcal{I}$, as before.

It is well known that one can define multiplication on $A/\mathcal{I}$ in such a way that

$$q(a)\,q(b) = q(a\,b) \tag{6.11.2}$$

for all $a, b \in \mathcal{A}$. One can check that the right side depends only on $q(a)$, $q(b)$, and not on the particular choices of a, b, because $\mathcal{I}$ is a two-sided ideal in $\mathcal{A}$. Thus $\mathcal{A}/\mathcal{I}$ may be considered as an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$, as appropriate.

If $\mathcal{A}$ is an associative algebra, then it is easy to see that $\mathcal{A}/\mathcal{I}$ is associative too. Similarly, if $\mathcal{A}$ is commutative, then $\mathcal{A}/\mathcal{I}$ is commutative.

If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then

$$e_{\mathcal{A}/\mathcal{I}} = q(e_{\mathcal{A}}) \tag{6.11.3}$$

is the multiplicative identity element in $\mathcal{A}/\mathcal{I}$. Note that $\mathcal{I} \neq \mathcal{A}$ exactly when $e_{\mathcal{A}}$ is not contained in $\mathcal{I}$, which means that (6.11.3) is nonzero.

Let $\mathcal{I}_0$ be a proper two-sided ideal in $\mathcal{A}$. One can chow that $\mathcal{I}_0$ is contained in a maximal proper two-sided ideal in $\mathcal{A}$, using Zorn's lemma or Hausdorff's maximality principle.

Let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$. If $\mathcal{I}_1$ is a two-sided ideal in $\mathcal{A}$, then it is easy to see that the closure $\overline{\mathcal{I}_1}$ of $\mathcal{I}_1$ in $\mathcal{A}$ with respect to the metric associated to the norm is a two-sided ideal as well.

If $\mathcal{I}$ is a closed two-sided ideal in $\mathcal{A}$, then

$$\|q(a)\|_{\mathcal{A}/\mathcal{I}} = \inf\{\|a - x\|_{\mathcal{A}} : x \in \mathcal{I}\} \tag{6.11.4}$$

defines a norm on $\mathcal{A}/\mathcal{I}$, as in the previous section. If $a, b \in \mathcal{A}$ and $x, y \in \mathcal{I}$, then

$$q(a)\,q(b) = q(a-x)\,q(b-y) = q((a-x)\,(b-y)). \tag{6.11.5}$$

This implies that

$$\begin{aligned} \|q(a)\,q(b)\|_{\mathcal{A}/\mathcal{I}} &= \|q((a-x)\,(b-y))\|_{\mathcal{A}/\mathcal{I}} \\ &\le \|(a-x)\,(b-y)\|_{\mathcal{A}} \le \|a-x\|_{\mathcal{A}}\,\|b-y\|_{\mathcal{A}}. \end{aligned} \tag{6.11.6}$$

One can use this to get that

$$\|q(a)\,q(b)\|_{\mathcal{A}/\mathcal{I}} \le \|q(a)\|_{\mathcal{A}/\mathcal{I}}\,\|q(b)\|_{\mathcal{A}/\mathcal{I}}. \tag{6.11.7}$$

Suppose now that $\mathcal{A}$ is a Banach algebra with respect to a norm $\|\cdot\|_{\mathcal{A}}$, with a multiplicative identity element $e_{\mathcal{A}}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $\mathcal{I}_1$ is a proper two-sided ideal in $\mathcal{A}$, then $\mathcal{I}_1$ does not contain any invertible elements of $\mathcal{A}$. This means that $\mathcal{I}_1$ does not contain any elements of the open ball in $\mathcal{A}$ centered at $e_{\mathcal{A}}$ with radius 1, because the elements of that ball are invertible in $\mathcal{A}$, as in Section 6.5. It follows that $\overline{\mathcal{I}_1}$ is a proper two-sided ideal in $\mathcal{A}$ too. In particular, a maximal proper two-sided ideal in $\mathcal{A}$ is a closed set in $\mathcal{A}$, with respect to the metric associated to the norm.

Suppose that $\mathcal{I}$ is a proper closed two-sided ideal in $\mathcal{A}$, and let us check that

$$\|e_{\mathcal{A}/\mathcal{I}}\|_{\mathcal{A}/\mathcal{I}} = \|q(e_{\mathcal{A}})\|_{\mathcal{A}/\mathcal{I}} = 1. \tag{6.11.8}$$

Clearly

$$\|q(e_{\mathcal{A}})\|_{\mathcal{A}/\mathcal{I}} \le \|e_{\mathcal{A}}\|_{\mathcal{A}} = 1. \tag{6.11.9}$$

In order to get

$$\|q(e_{\mathcal{A}})\|_{\mathcal{A}/\mathcal{I}} \ge 1, \tag{6.11.10}$$

we need to have that

$$\|e_{\mathcal{A}} - x\|_{\mathcal{A}} \ge 1 \tag{6.11.11}$$

for every $x \in \mathcal{I}$. This condition holds because $\mathcal{I}$ is a proper two-sided ideal in $\mathcal{A}$, as in the preceding paragraph.

Note that $\mathcal{A}/\mathcal{I}$ is a Banach space with respect to the quotient norm, as in the previous section. This means that $\mathcal{A}/\mathcal{I}$ is a Banach algebra with respect to the quotient norm.

6.12 Ideals in commutative algebras

Let $\mathcal{A}$ be a commutative associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. A linear subspace $\mathcal{I}$ of $\mathcal{A}$ is said to be an *ideal* in $\mathcal{A}$ if for every $a \in \mathcal{A}$ and $x \in \mathcal{I}$ we have that

$$a\,x \in \mathcal{I}, \tag{6.12.1}$$

which is the same as saying that $\mathcal{I}$ is a two-sided ideal in $\mathcal{A}$ in this case. In this case, $\mathcal{A}/\mathcal{I}$ is a commutative associative algebra over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with a multiplicative identity element, as in the previous section.

If $y \in \mathcal{A}$, then

$$\{a\,y : a \in \mathcal{A}\} \tag{6.12.2}$$

is an ideal in $\mathcal{A}$. This ideal is equal to $\mathcal{A}$ if and only if y is invertible in $\mathcal{A}$.

Let $\mathcal{I}$ be a proper ideal in $\mathcal{A}$. If every nonzero element of $\mathcal{A}/\mathcal{I}$ is invertible in $\mathcal{A}/\mathcal{I}$, then one can check that $\mathcal{I}$ is maximal with respect to inclusion among proper ideals in $\mathcal{A}$. Conversely, if $\mathcal{I}$ is maximal among proper ideals in $\mathcal{A}$, then one can verify that every nonzero element of $\mathcal{A}/\mathcal{I}$ is invertible.

Suppose that $\mathcal{A}$ is a complex Banach algebra, and that $\mathcal{I}$ is a maximal proper ideal in $\mathcal{A}$. This implies that $\mathcal{I}$ is a closed set in $\mathcal{A}$, and that $\mathcal{A}/\mathcal{I}$ is a complex Banach algebra as well, as in the previous section. In fact, $\mathcal{A}/\mathcal{I}$ is isomorphic to $\mathbf{C}$, by the theorem of Gelfand and Mazur, because every nonzero element of $\mathcal{A}/\mathcal{I}$ is invertible, as before.

If $y \in \mathcal{A}$ is not invertible, then (6.12.2) is a proper ideal in $\mathcal{A}$, which is contained in a maximal proper ideal in $\mathcal{A}$, as in the previous section. This leads to a homomorphism h from $\mathcal{A}$ into $\mathbf{C}$ such that $h(e_{\mathcal{A}}) = 1$ and

$$h(y) = 0, \tag{6.12.3}$$

as in the preceding paragrah.

If $x \in \mathcal{A}$ and $\lambda \in \sigma_{\mathcal{A}}(x)$, then $\lambda\,e_{\mathcal{A}} - x$ is not invertible, and there is a homomorphism h from $\mathcal{A}$ into $\mathbf{C}$ such that $h(e_{\mathcal{A}}) = 1$ and

$$h(\lambda\,e_{\mathcal{A}} - x) = 0, \tag{6.12.4}$$

as in the previous paragraph. This implies that

$$h(x) = \lambda. \tag{6.12.5}$$

Let $\mathcal{A}$ be a commutative ring with a multiplicative identity element $e_{\mathcal{A}}$. An element x of $\mathcal{A}$ is said to be *nilpotent* if

$$x^l = 0 \tag{6.12.6}$$

for some positive integer l. The set $\mathcal{N} = \mathcal{N}(\mathcal{A})$ of all nilpotent elements of $\mathcal{A}$ is called the *nilradical* of $\mathcal{A}$, as on p5 of [10]. One can check that $\mathcal{N}$ is an ideal in $\mathcal{A}$, as in Proposition 1.7 in [10]. In fact, it is well known that $\mathcal{N}$ is the same as the intersection of all of the proper prime ideals in $\mathcal{A}$, as in Proposition 1.8 in [10].

The *Jacobson radical* $\mathcal{R} = \mathcal{R}(\mathcal{A})$ of $\mathcal{A}$ is defined to be the intersection of all of the maximal proper ideals in $\mathcal{A}$, as on p5 of [10]. If $\mathcal{A}$ is a commutative Banach algebra over the real or complex numbers, then this may be simply called the *radical* of $\mathcal{A}$, and denoted $\mathrm{rad}(\mathcal{A})$ or $\mathrm{rad}\,\mathcal{A}$, as in Exercise 12 on p401 of [160], and on p268 of [162]. If

$$\mathrm{rad}(\mathcal{A}) = \{0\}, \tag{6.12.7}$$

then $\mathcal{A}$ is said to be *semisimple* as a commutative Banach algebra, as on p268 of [162]. In the complex case, the radical may be described equivalently as the intersection of all of the kernels of the homomorphisms from $\mathcal{A}$ onto $\mathbf{C}$, or as the set of $x \in \mathcal{A}$ such that

$$\sigma_{\mathcal{A}}(x) = \{0\}. \tag{6.12.8}$$

The radical is defined a bit differently in the complex case on p27 of [8], using a result that will be mentioned in Section 6.14.

6.13 More on invertibility

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. Suppose that $x, y \in \mathcal{A}$ commute with each other, so that

$$x\,y = y\,x. \tag{6.13.1}$$

If x is invertible in $\mathcal{A}$, then

$$y\,x^{-1} = x^{-1}\,y. \tag{6.13.2}$$

Of course, (6.13.1) implies that $x\,y$ commutes with x and y. Suppose that $x\,y$ is invertible in $\mathcal{A}$, so that $(x\,y)^{-1}$ commutes with x and y, as before. It is easy to see that x and y are both invertible in $\mathcal{A}$ in this case, with

$$x^{-1} = y\,(x\,y)^{-1},\ y^{-1} = x\,(x\,y)^{-1}. \tag{6.13.3}$$

If

$$e_{\mathcal{A}} - x^l \in G(\mathcal{A}) \tag{6.13.4}$$

for some positive integer l, then

$$e_{\mathcal{A}} - x \in G(\mathcal{A}). \tag{6.13.5}$$

This follows from (6.5.2) and the remarks in the preceding paragraph.

If x is *nilpotent* in $\mathcal{A}$, so that $x^l = 0$ for some $l \geq 1$, then (6.13.4) holds, so that (6.13.5) holds. In fact,

$$(e_{\mathcal{A}} - x)^{-1} = \sum_{j=0}^{l-1} x^j, \tag{6.13.6}$$

by (6.5.2).

Suppose that $\mathcal{A}$ is a Banach algebra with respect to a norm $\|\cdot\|_{\mathcal{A}}$, with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If

$$\|x^l\|_{\mathcal{A}} < 1 \tag{6.13.7}$$

for some $l \geq 1$, then (6.13.4) holds, as in Section 6.5. This implies that (6.13.5) holds, as before.

Let λ be a real or complex number, as appropriate. Suppose for the moment that

$$|\lambda| > \|x^l\|_{\mathcal{A}}^{1/l} \tag{6.13.8}$$

for some $l \geq 1$. This means that $|\lambda^l| = |\lambda|^l > \|x^l\|_{\mathcal{A}}$, so that

$$\|(\lambda^{-1}\, x)^l\|_{\mathcal{A}} < 1. \tag{6.13.9}$$

It follows that

$$e_{\mathcal{A}} - (\lambda^{-1}\, x)^l \in G(\mathcal{A}), \tag{6.13.10}$$

and thus

$$e_{\mathcal{A}} - \lambda^{-1}\, x \in G(\mathcal{A}), \tag{6.13.11}$$

as before. This implies that

$$\lambda\, e_{\mathcal{A}} - x = \lambda\, (e_{\mathcal{A}} - \lambda^{-1}\, x) \in G(\mathcal{A}). \tag{6.13.12}$$

This shows that for each $l \geq 1$,

$$|\lambda| \leq \|x^l\|_{\mathcal{A}}^{1/l} \text{ when } \lambda \in \sigma_{\mathcal{A}}(x). \tag{6.13.13}$$

Put

$$r_{\mathcal{A}}(x) = \inf_{l \geq 1} \|x^l\|_{\mathcal{A}}^{1/l}. \tag{6.13.14}$$

It follows that

$$|\lambda| \leq r_{\mathcal{A}}(x) \text{ when } \lambda \in \sigma_{\mathcal{A}}(x). \tag{6.13.15}$$

In the next section, we shall show that

$$\lim_{j \to \infty} \|x^j\|_{\mathcal{A}}^{1/j} = r_{\mathcal{A}}(x). \tag{6.13.16}$$

6.14 More on $r_{\mathcal{A}}(x)$

Let $\{a_j\}_{j=1}^{\infty}$ be a sequence of nonnegative real numbers that is *submultiplicative*, in the sense that

$$a_{j+l} \leq a_j\, a_l \tag{6.14.1}$$

for all $j, l \geq 1$. Put

$$\alpha = \inf_{n \geq 1} a_n^{1/n}. \tag{6.14.2}$$

We would like to check that

$$a_j^{1/j} \to \alpha \text{ as } j \to \infty. \tag{6.14.3}$$

If $a_l = 0$ for some $l \geq 1$, then $a_j = 0$ for all $j \geq l$, $\alpha = 0$, and (6.14.3) is obvious. Thus we may suppose that $a_j > 0$ for each j.

Let a positive integer m be given, and observe that

$$a_{l\,m} \leq a_m^l \tag{6.14.4}$$

for each $l \geq 1$, so that

$$a_{l\,m}^{1/(l\,m)} \leq a_m^{1/m}. \tag{6.14.5}$$

Similarly,

(6.14.6) $$a_{l\,m+r} \le a_m^l\, a_r$$

for all $l \ge 0$ and $r \ge 1$. If

(6.14.7) $$j = l\,m + r,$$

then we get that

(6.14.8) $$a_j^{1/j} \le a_m^{l/j}\, a_r^{1/j} = (a_m^{1/m})^{(l\,m)/j}\, a_r^{1/j} = (a_m)^{1/m}\,(a_m)^{-r/(j\,m)}\, a_r^{1/j}.$$

Note that every $j \ge 1$ can be expressed as in (6.14.7) with $l, r \ge 0$, $r < m$, and at least one of l, r positive.

One can use this to get that

(6.14.9) $$\limsup_{j \to \infty} a_j^{1/j} \le a_m^{1/m}.$$

This implies that

(6.14.10) $$\limsup_{j \to \infty} a_j^{1/j} \le \alpha,$$

by the definition (6.14.2) of α. It is easy to obtain (6.14.3) from this and the definition of α.

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers, with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $x \in \mathcal{A}$, then

(6.14.11) $$a_j = \|x^j\|_{\mathcal{A}}$$

defines a submultiplicative sequence of nonnegative real numbers. This means that (6.13.16) follows from (6.14.3).

If $\mathcal{A}$ is a complex Banach algebra, then it is well known that

(6.14.12) $$r_{\mathcal{A}}(x) = \max\{|\lambda| : \lambda \in \sigma_{\mathcal{A}}(x)\}.$$

Remember that $\sigma_{\mathcal{A}}(x)$ is a nonempty closed and bounded subset of $\mathbf{C}$ in this case, as in Section 6.8. This implies that $\sigma_{\mathcal{A}}(x)$ is compact with respect to the standard Euclidean metric on $\mathbf{C}$, so that the maximum on the right side of (6.14.12) is attained.

Let us give an outline of the proof of (6.14.12). Let r be the right side of (6.14.12), and note that

(6.14.13) $$r \le r_{\mathcal{A}}(x),$$

by (6.13.15). Thus we need only show that

(6.14.14) $$r_{\mathcal{A}}(x) \le r.$$

If $\lambda \in \mathbf{C}$ and $|\lambda| > r$, then

(6.14.15) $$\lambda \notin \sigma_{\mathcal{A}}(x),$$

so that $\lambda\, e_{\mathcal{A}} - x \in G(\mathcal{A})$.

Let us interpret $1/r$ as being $+\infty$ when $r = 0$. If $\zeta \in \mathbf{C}$ and $|\zeta| < 1/r$, then

$$e_{\mathcal{A}} - \zeta\, x \in G(\mathcal{A}). \tag{6.14.16}$$

This is trivial when $\zeta = 0$, and otherwise it follows from the fact that ζ^{-1} is not an element of $\sigma_{\mathcal{A}}(x)$, because $|\zeta^{-1}| = 1/|\zeta| > r$. Thus

$$(e_{\mathcal{A}} - \zeta\, x)^{-1} \tag{6.14.17}$$

is an $\mathcal{A}$-valued function of ζ defined on

$$\{\zeta \in \mathbf{C} : |\zeta| < 1/r\}. \tag{6.14.18}$$

In fact, (6.14.17) is holomorphic on (6.14.18). Note that (6.14.17) is equal to

$$\sum_{j=0}^{\infty} \zeta^j\, x^j \tag{6.14.19}$$

when ζ is small enough, as in Section 6.5. If

$$0 < t < 1/r, \tag{6.14.20}$$

then one can show that there is a nonnegative real number $C(t)$ such that

$$t^j\, \|x^j\|_{\mathcal{A}} \le C(t) \tag{6.14.21}$$

for each $j \ge 1$, using complex analysis. This implies that

$$\|x^j\|_{\mathcal{A}}^{1/j} \le t^{-1}\, C(t)^{1/j} \tag{6.14.22}$$

for each $j \ge 1$. One can use this to get that

$$r_{\mathcal{A}}(x) \le 1/t, \tag{6.14.23}$$

which implies (6.14.14).

6.15 Identity elements and essential ranges

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers. Let us define $\mathcal{A}_1$ initially as a vector space over the real or complex numbers, as appropriate, by taking the direct sum of $\mathcal{A}$ and $\mathbf{R}$ or $\mathbf{C}$, as appropriate. More precisely, $\mathcal{A}_1$ may be defined as the Cartesian product $\mathcal{A} \times \mathbf{R}$ or $\mathcal{A} \times \mathbf{C}$, as appropriate, using coordinatewise addition and scalar multiplication, as in Section 5.12. We can define multiplication on $\mathcal{A}_1$ by

$$(a, t)\,(b, z) = (a\,b + z\,a + t\,b, t\,z) \tag{6.15.1}$$

for every $a, b \in \mathcal{A}$ and $t, z \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. This is a bilinear mapping from $\mathcal{A}_1 \times \mathcal{A}_1$ into $\mathcal{A}_1$, so that $\mathcal{A}_1$ is an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$, as appropriate.

By construction,

$$a \mapsto (a, 0) \tag{6.15.2}$$

is an injective algebra homomorphism from $\mathcal{A}$ into $\mathcal{A}_1$. We also have that

$$e_{\mathcal{A}_1} = (0, 1) \tag{6.15.3}$$

is the multiplicative identity element in $\mathcal{A}_1$. If $\mathcal{A}$ is associative, then one can check that $\mathcal{A}_1$ is associative. Similarly, if $\mathcal{A}$ is commutative, then $\mathcal{A}_1$ is commutative.

If $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$, then one can check that

$$\|(a, t)\|_{\mathcal{A}_1} = \|a\|_{\mathcal{A}} + |t| \tag{6.15.4}$$

defines a submultiplicative norm on $\mathcal{A}_1$. Of course, $\|(a, 0)\|_{\mathcal{A}_1} = \|a\|_{\mathcal{A}}$ for every $a \in \mathcal{A}$, and $\|e_{\mathcal{A}_1}\|_{\mathcal{A}_1} = 1$. If $\mathcal{A}$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, then $\mathcal{A}_1$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}_1}$, as in Section 5.12. This corresponds to part of Proposition 2.5.4 on p58 of [8], Theorem C.3 on p470 of [91], Exercise 15 on p402 of [160], and some remarks on p228 of [162].

Let $(X, \mathcal{M}, \mu)$ be a nonempty measure space, and note that the corresponding spaces $L^\infty(X, \mathbf{R})$, $L^\infty(X, \mathbf{C})$ are commutative associative Banach algebras over $\mathbf{R}$, $\mathbf{C}$, respectively, with respect to the usual L^∞ norms. The function $\mathbf{1}_X$ equal to 1 at every point in X is the multiplicative identity element of these algebras. Let us suppose that $\mu(X) > 0$, so that $\|\mathbf{1}_X\|_\infty = 1$. If $f \in L^\infty(X, \mathbf{R})$ or $L^\infty(X, \mathbf{C})$, then f is invertible in $L^\infty(X, \mathbf{R})$ or $L^\infty(X, \mathbf{C})$, as appropriate, if and only if there is a positive real number c such that

$$|f(x)| \ge c \tag{6.15.5}$$

for almost every $x \in X$ with respect to μ. In this case, the multiplicative inverse of f is equal to $1/f$ almost everywhere on X with respect to μ.

If $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $f - \lambda \mathbf{1}_X$ is invertible in $L^\infty(X, \mathbf{R})$ or $L^\infty(X, \mathbf{C})$, as appropriate, if and only if

$$|f(x) - \lambda| \ge c(\lambda) \tag{6.15.6}$$

for some $c(\lambda) > 0$ and almost every $x \in X$ with respect to μ. Thus the spectrum of f in $L^\infty(X, \mathbf{R})$ or $L^\infty(X, \mathbf{C})$, as appropriate, consists of the $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that

$$\mu(\{x \in X : |f(x) - \lambda| < r\}) > 0 \tag{6.15.7}$$

for every $r > 0$. This is called the *essential range* of f.

Chapter 7

Algebras, norms, and linear mappings

7.1 Quaternions

A *quaternion* may be expressed as

$$x = x_0 + x_1\, i + x_2\, j + x_3\, k, \tag{7.1.1}$$

where x_0, x_1, x_2, x_3 are real numbers. The space $\mathbf{H}$ of quaternions may be considered initially as a vector space over the real numbers, which may be identified with $\mathbf{R}^4$. We may also identify $x_0 \in \mathbf{R}$ with the quaternion x as in (7.1.1), with $x_1 = x_2 = x_3 = 0$.

Multiplication on $\mathbf{R}$ can be extended to a bilinear mapping from $\mathbf{H} \times \mathbf{H}$ into $\mathbf{H}$, with the following properties. The product of a real number and a quaternion, in either order, is the same as scalar multiplication. In particular, $1 \in \mathbf{R}$ is the multiplicative identity element in $\mathbf{H}$. We put

$$i^2 = j^2 = k^2 = -1 \tag{7.1.2}$$

and

$$i\, j = -j\, i = k,\ \ j\, k = -k\, j = i,\ \ k\, i = -i\, k = j, \tag{7.1.3}$$

as on p256 of [30]. One can check that this makes $\mathbf{H}$ an associative algebra over $\mathbf{R}$.

If $x \in \mathbf{H}$ is as in (7.1.1), then put

$$x^* = x_0 - x_1\, i - x_2\, j - x_3\, k. \tag{7.1.4}$$

Clearly $x \mapsto x^*$ is a linear mapping from $\mathbf{H}$ into itself, as a vector space over $\mathbf{R}$, and $(x^*)^* = x$ for every $x \in \mathbf{H}$. It is easy to see that $(x\, y)^* = y^*\, x^*$ for every $x, y \in \mathbf{H}$, so that $x \mapsto x^*$ is an algebra involution on $\mathbf{H}$.

If $x \in \mathbf{H}$ is as in (7.1.1), then

$$x\,x^* = x^*\,x = x_0^2 + x_1^2 + x_2^2 + x_3^2. \tag{7.1.5}$$

Put

$$\|x\|_{\mathbf{H}} = (x_0^2 + x_1^2 + x_2^2 + x_3^2)^{1/2}, \tag{7.1.6}$$

using the nonnegative square root on the right side. This corresponds exactly to the standard Euclidean norm on $\mathbf{R}^4$, and in particular this defines a norm on $\mathbf{H}$, as a vector space over $\mathbf{R}$. Using this notation, (7.1.5) is the same as saying that

$$x\,x^* = x^*\,x = \|x\|_{\mathbf{H}}^2. \tag{7.1.7}$$

Note that

$$\|x^*\|_{\mathbf{H}} = \|x\|_{\mathbf{H}}. \tag{7.1.8}$$

If $x \in \mathbf{H}$ and $x \neq 0$, then $\|x\|_{\mathbf{H}} > 0$, and (7.1.7) implies that x is invertible in $\mathbf{H}$, with

$$x^{-1} = \|x\|_{\mathbf{H}}^{-2}\,x^*. \tag{7.1.9}$$

If $x, y \in \mathbf{H}$, then

$$\|x\,y\|_{\mathbf{H}}^2 = (x\,y)\,(x\,y)^* = x\,y\,y^*\,x^* = \|y\|_{\mathbf{H}}^2\,x\,x^* = \|x\|_{\mathbf{H}}^2\,\|y\|_{\mathbf{H}}^2. \tag{7.1.10}$$

This implies that

$$\|x\,y\|_{\mathbf{H}} = \|x\|_{\mathbf{H}}\,\|y\|_{\mathbf{H}}. \tag{7.1.11}$$

7.2 Some additional properties of $r_{\mathcal{A}}(x)$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $x \in \mathcal{A}$, then we can define $r_{\mathcal{A}}(x)$ as in (6.13.14), and (6.13.16) holds. More precisely, (6.14.11) defines a submultiplicative sequence of nonnegative real numbers, so that (6.13.16) follows from (6.14.3), as before.

Let $\mathcal{B}$ be another associative algebra over the real or complex numbers, as appropriate, with a submultiplicative norm $\|\cdot\|_{\mathcal{B}}$. Also let ϕ be an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ that is bounded as a linear mapping. If $x \in \mathcal{A}$, then

$$\|\phi(x)^j\|_{\mathcal{B}}^{1/j} = \|\phi(x^j)\|_{\mathcal{B}}^{1/j} \le \|\phi\|_{op,\mathcal{AB}}^{1/j}\,\|x^j\|_{\mathcal{A}}^{1/j} \tag{7.2.1}$$

for each $j \ge 1$. This implies that

$$r_{\mathcal{B}}(\phi(x)) \le r_{\mathcal{A}}(\phi(x)), \tag{7.2.2}$$

where the left side is defined in the same way as before. This uses the fact that $C^{1/j} \to 1$ as $j \to \infty$ for any positive real number C.

Similarly, if ϕ is an isomorphism from $\mathcal{A}$ onto $\mathcal{B}$, and ϕ and its inverse are bounded as linear mappings, then

$$r_{\mathcal{B}}(\phi(x)) = r_{\mathcal{A}}(x). \tag{7.2.3}$$

There are analogous statements for opposite algebra homomorphisms and isomorphisms.

If $x, y \in \mathcal{A}$ commute, then

$$\|(x\,y)^j\|_{\mathcal{A}}^{1/j} = \|x^j\,y^j\|_{\mathcal{A}}^{1/j} \le \|x^j\|_{\mathcal{A}}^{1/j}\,\|y^j\|_{\mathcal{A}}^{1/j} \tag{7.2.4}$$

for each $j \ge 1$. This implies that

$$r_{\mathcal{A}}(x\,y) \le r_{\mathcal{A}}(x)\,r_{\mathcal{A}}(y). \tag{7.2.5}$$

Suppose for the moment that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$. Observe that

$$r_{\mathcal{A}}(e_{\mathcal{A}}) = 1, \tag{7.2.6}$$

even if $\|e_{\mathcal{A}}\|_{\mathcal{A}}$ is not asked to be equal to 1. If x is an invertible element of $\mathcal{A}$, then we get that

$$1 \le r_{\mathcal{A}}(x)\,r_{\mathcal{A}}(x^{-1}), \tag{7.2.7}$$

by taking $y = x^{-1}$ in (7.2.5).

As before, $x \in \mathcal{A}$ is said to be *nilpotent* if $x^l = 0$ for some $l \ge 1$. If

$$r_{\mathcal{A}}(x) = \lim_{j \to \infty} \|x^j\|_{\mathcal{A}}^{1/j} = 0, \tag{7.2.8}$$

then x is said to be *quasinilpotent* in $\mathcal{A}$ with respect to $\|\cdot\|_{\mathcal{A}}$, as in Definition 1.7.4 on p20 of [8].

Suppose that x is an *idempotent* element of $\mathcal{A}$, so that

$$x^2 = x. \tag{7.2.9}$$

This implies that $x^j = x$ for each $j \ge 1$, so that

$$r_{\mathcal{A}}(x) = 1, \tag{7.2.10}$$

unless $x = 0$.

7.3 Invertibility and subalgebras

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e = e_{\mathcal{A}}$, and let $\mathcal{B}$ be a subalgebra of $\mathcal{A}$ that contains e. Observe that any invertible element of $\mathcal{B}$ is also invertible as an element of $\mathcal{A}$, so that

$$G(\mathcal{B}) \subseteq G(\mathcal{A}). \tag{7.3.1}$$

If $x \in \mathcal{B}$, then it follows that

$$\sigma_{\mathcal{A}}(x) \subseteq \sigma_{\mathcal{B}}(x), \tag{7.3.2}$$

where $\sigma_{\mathcal{A}}(x)$, $\sigma_{\mathcal{B}}(x)$ are as in Section 6.8.

Suppose that $\|\cdot\| = \|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$, and that

$$\mathcal{B} \text{ is a closed set in } \mathcal{A}, \tag{7.3.3}$$

with respect to the associated metric. Let $\{x_j\}_{j=1}^\infty$ be a sequence of invertible elements of $\mathcal{B}$ that converges to $x \in \mathcal{B}$ with respect to the metric associated to $\|\cdot\|$, and suppose that x is invertible in $\mathcal{A}$. This implies that

$$\{x_j^{-1}\}_{j=1}^\infty \text{ converges to } x^{-1} \tag{7.3.4}$$

in $\mathcal{A}$, as in Section 6.7. It follows that

$$x^{-1} \in \mathcal{B}, \tag{7.3.5}$$

because of (7.3.3).

Suppose now that $\mathcal{A}$ is a Banach algebra with respect to $\|\cdot\|$. This means that $\mathcal{B}$ is complete with respect to the metric associated to the restriction of $\|\cdot\|$ to $\mathcal{B}$, because of (7.3.3), as in Section 1.6. Thus $\mathcal{B}$ is a Banach algebra too, with respect to the restriction of $\|\cdot\|$ to $\mathcal{B}$. If $x \in \mathcal{B}$, then

$$\partial\sigma_{\mathcal{B}}(x) \subseteq \sigma_A(x). \tag{7.3.6}$$

Here $\partial\sigma_{\mathcal{B}}(x)$ is the boundary of $\sigma_{\mathcal{B}}(x)$ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to the standard Euclidean metric.

To see this, let $\lambda \in \partial\sigma_{\mathcal{B}}(x)$ be given. This implies that there is a sequence $\{\lambda_j\}_{j=1}^\infty$ of real or complex numbers, as appropriate, that converges to λ, with

$$\lambda_j\, e - x \in G(\mathcal{B}) \tag{7.3.7}$$

for each j. If

$$\lambda\, e - x \in G(\mathcal{A}), \tag{7.3.8}$$

then

$$\lambda\, e - x \in G(\mathcal{B}), \tag{7.3.9}$$

as in (7.3.5). Remember that $\sigma_{\mathcal{B}}(x)$ is a closed set in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, because $\mathcal{B}$ is a Banach algebra, as in Section 6.8. This means that $\lambda \in \sigma_{\mathcal{B}}(x)$, which is a contradiction.

This corresponds to Theorem 1.11.3 on p32 of [8], and to part of part (b) of Theorem 10.18 on p238 of [162]. This is known as the *spectral permanence theorem*, as in [8].

7.4 Some Volterra integral operators

Let k be a continuous real or complex-valued function on the closed triangle

$$\Delta = \{(x, y) \in \mathbf{R}^2 : 0 \le y \le x \le 1\} \tag{7.4.1}$$

in the plane. Note that k is uniformly continuous on Δ, with respect to the restriction of the standard Euclidean metric on $\mathbf{R}^2$ to Δ, because Δ is compact.

If f is a continuous real or complex-valued function on the closed unit interval $[0,1]$, then let $T_k(f)$ be the function defined on $[0,1]$ by

$$(T_k(f))(x) = \int_0^x k(x,y)\, f(y)\, dy, \tag{7.4.2}$$

as in the left side of (1.3) on p3 of [8]. In particular, let $T(f)$ be the function defined on $[0,1]$ by

$$(T(f))(x) = \int_0^x f(y)\, dy, \tag{7.4.3}$$

as in Exercise (4) on p5 of [8], which corresponds to taking $k \equiv 1$ on Δ in (7.4.2).

One can check that (7.4.2) is continuous on $[0,1]$, as in Exercise (3) on p5 of [8]. It is easy to see that

$$\|T_k(f)\|_{sup} \le \|k\|_{sup,\Delta}\, \|f\|_{sup}, \tag{7.4.4}$$

where $\|\cdot\|_{sup}$ is the usual supremum norm of a continuous real or complex-valued function on $[0,1]$, and $\|k\|_{sup,\Delta}$ is the supremum norm of k on Δ. This implies that T_k is a bounded linear mapping from $C([0,1], \mathbf{R})$ or $C([0,1], \mathbf{C})$ into itself, as appropriate, with respect to the supremum norm, as in [8].

More precisely,

$$|(T_k(f))(x)| \le (T_{|k|}(|f|))(x) \le \|k\|_{sup,\Delta}\, (T(|f|))(x) \tag{7.4.5}$$

for every $x \in [0,1]$. We can use this repeatedly to get that

$$|(T_k^n(f))(x)| \le \|k\|_{sup,\Delta}^n\, (T^n(|f|))(x) \le \|k\|_{sup,\Delta}^n\, \|f\|_{sup}\, (T^n(1))(x) \tag{7.4.6}$$

for every $n \ge 1$ and $x \in [0,1]$. Of course,

$$(T^n(1))(x) = x^n/n! \tag{7.4.7}$$

for every $n \ge 1$ and $x \in [0,1]$, by calculus. It follows that

$$\|T_k^n(f)\|_{sup} \le \|k\|_{sup,\Delta}^n\, \|f\|_{sup}/n! \tag{7.4.8}$$

for every $n \ge 1$, so that

$$\|T_k^n\|_{op} \le \|k\|_{sup,\Delta}^n/n!, \tag{7.4.9}$$

as in Exercise (3) on p21 of [8]. Note that equality holds in (7.4.9) when k is constant on Δ.

Alternatively, consider the n-dimensional simplex

$$\Delta_n = \{x \in \mathbf{R}^n : 0 \le x_1 \le x_2 \le \cdots \le x_n \le 1\} \tag{7.4.10}$$

for each positive integer n. The n-dimensional volume of Δ_n is the same as $(T^n(1))(1)$, which is equal to $1/n!$, as before. The volume of Δ_n can also be obtained by considering permutations of the coordinates, as in Exercise (2) on p20 of [8].

Let V be $C([0,1],\mathbf{R})$ or $C([0,1],\mathbf{C})$, as appropriate, with the supremum norm, and let $\mathcal{A}$ be $\mathcal{BL}(V)$, with the corresponding operator norm. One can use (7.4.9) to get that

$$r_{\mathcal{A}}(T_k) = 0, \tag{7.4.11}$$

where $r_{\mathcal{A}}$ is as in Section 6.13. This means that T_k is quasinilpotent in $\mathcal{A}$, as in Section 7.2. This is related to Exercise (4) on p21 of [8]. Note that T_k is not invertible on V, because $(T_k(f))(0) = 0$ for every $f \in V$.

7.5 Self-adjointness and involutions

Let W be a vector space over the complex numbers, which may also be considered as a vector space over the real numbers. If $W_0 \subseteq W$ is a linear subspace of W as a vector space over the real numbers, then we may refer to W_0 as a *real-linear subspaces* of W.

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers, and let

$$x \mapsto x^* \tag{7.5.1}$$

be an algebra involution on $\mathcal{A}$, as in Section 6.4. Remember that this means that (7.5.1) is an opposite algebra isomorphism from $\mathcal{A}$ onto itself such that

$$(x^*)^* = x \tag{7.5.2}$$

for every $x \in \mathcal{A}$. In the complex case, we may also be interested in conjugate-linear involutions, as before.

Let us say that $x \in \mathcal{A}$ is *self-adjoint* with respect to (7.5.1) if

$$x^* = x, \tag{7.5.3}$$

and *anti-self-adjoint* if

$$x^* = -x. \tag{7.5.4}$$

If $\mathcal{A}$ is complex and (7.5.1) is conjugate-linear, then $x \in \mathcal{A}$ is anti-self-adjoint if and only if $i\,x$ is self-adjoint. In this case, self-adjoint elements of $\mathcal{A}$ are also said to be *Hermitian.*

The spaces of self-adjoint and anti-self-adjoint elements of $\mathcal{A}$ are linear subspaces of $\mathcal{A}$ in the real case, and when the involution is complex-linear in the complex case. If $\mathcal{A}$ is complex and the involution is conjugate-linear, then the spaces of self-adjoint and anti-self-adjoint elements of $\mathcal{A}$ are real-linear subspaces of $\mathcal{A}$.

If x is any element of $\mathcal{A}$, then

$$(1/2)\,(x + x^*) \tag{7.5.5}$$

is self-adjoint with respect to (7.5.1), and

$$(1/2)\,(x - x^*) \tag{7.5.6}$$

is anti-self-adjoint. Note that x is equal to the sum of (7.5.5) and (7.5.6). We also have that

$$x^* \, x \tag{7.5.7}$$

is self-adjoint.

If

$$x \, x^* = x^* \, x, \tag{7.5.8}$$

then one may say that x is *normal* with respect to (7.5.1). This terminology is perhaps most commonly used when $\mathcal{A}$ is complex, and the involution is conjugate-linear.

Suppose that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$. Note that

$$e_{\mathcal{A}}^* = e_{\mathcal{A}}, \tag{7.5.9}$$

as in Section 6.4. Suppose that $\mathcal{A}$ is associative as well. If $x \in \mathcal{A}$ is invertible, then it is easy to see that x^* is invertible, with

$$(x^*)^{-1} = (x^{-1})^*. \tag{7.5.10}$$

Consider

$$U(\mathcal{A}) = \{x \in \mathcal{A} : x \, x^* = x^* \, x = e_{\mathcal{A}}\}. \tag{7.5.11}$$

Equivalently, $U(\mathcal{A})$ consists of the invertible elemets x of $\mathcal{A}$ such that

$$x^{-1} = x^*. \tag{7.5.12}$$

One can check that

$$U(\mathcal{A}) \text{ is a subgroup of } G(\mathcal{A}). \tag{7.5.13}$$

If $\mathcal{A} = \mathcal{BL}(V)$ for some Hilbert space V, with the Hilbert space adjoint as the involution, then $U(\mathcal{A})$ consists of the unitary transformations from V onto itself. This is related to some remarks in Sections 2.10 and 3.5.

7.6 Some remarks about idempotents

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers. An element a of $\mathcal{A}$ is said to be an *idempotent* if

$$a \, a = a. \tag{7.6.1}$$

If V is a vector space over the real or complex numbers, then an idempotent element of the algebra of linear mappings from V into itself is the same as a projection on V. These will be discussed further in Section 8.2.

Suppose for the moment that $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$. If a is an idempotent element of $\mathcal{A}$, then

$$\|a\|_{\mathcal{A}} = \|a \, a\|_{\mathcal{A}} \le \|a\|_{\mathcal{A}}^2. \tag{7.6.2}$$

If $a \neq 0$, then we get that

$$\|a\|_{\mathcal{A}} \geq 1. \tag{7.6.3}$$

Suppose for the moment again that $x \mapsto x^*$ is an algebra involution on $\mathcal{A}$, which may be conjugate-linear in the complex case. If a is an idempotent element of $\mathcal{A}$, then

$$a^* = (a\,a)^* = a^*\,a^*, \tag{7.6.4}$$

so that a^* is idempotent as well.

Suppose now that $\mathcal{A}$ is an associative algebra with a multiplicative identity element $e_{\mathcal{A}}$. In particular, $e_{\mathcal{A}}$ is an idempotent element of $\mathcal{A}$. If an idempotent element a of $\mathcal{A}$ is invertible in $\mathcal{A}$, then

$$a = e_{\mathcal{A}}. \tag{7.6.5}$$

If a is any idempotent element of $\mathcal{A}$, then

$$a\,(e_{\mathcal{A}} - a) = (e_{\mathcal{A}} - a)\,a = 0. \tag{7.6.6}$$

This implies that

$$(e_{\mathcal{A}} - a)^2 = e_{\mathcal{A}} - a, \tag{7.6.7}$$

so that $e_{\mathcal{A}} - a$ is idempotent in $\mathcal{A}$ too. If $e_{\mathcal{A}} - a$ is invertible in $\mathcal{A}$, then $e_{\mathcal{A}} - a = e_{\mathcal{A}}$, as in the preceding paragraph, so that $a = 0$.

Let $\lambda \in \mathbf{R}$ or $\mathbf{C}$ be given, as appropriate. Observe that

$$\lambda\,e_{\mathcal{A}} - a = \lambda\,(e_{\mathcal{A}} - a) + (\lambda - 1)\,a. \tag{7.6.8}$$

If $\lambda \neq 0, 1$, then

$$\lambda^{-1}\,(e_{\mathcal{A}} - a) + (\lambda - 1)^{-1}\,a \tag{7.6.9}$$

defines an element of $\mathcal{A}$. One can check that this is the multiplicative inverse of $\lambda\,e_{\mathcal{A}} - a$ in this case.

This shows that

$$\sigma_{\mathcal{A}}(a) \subseteq \{0, 1\}, \tag{7.6.10}$$

where $\sigma_{\mathcal{A}}(a)$ is as in Section 6.8. We also get that equality holds unless $a = 0$ or $e_{\mathcal{A}}$.

7.7 The C^* identity

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with an involution $x \mapsto x^*$, which may be conjugate-linear in the complex case. Also let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$. Thus

$$\|x^*\,x\|_{\mathcal{A}} \leq \|x^*\|_{\mathcal{A}}\,\|x\|_{\mathcal{A}} \tag{7.7.1}$$

for every $x \in \mathcal{A}$. If

$$\|x^*\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}, \tag{7.7.2}$$

then (7.7.1) is the same as saying that

$$\|x^* \, x\|_{\mathcal{A}} \le \|x\|_{\mathcal{A}}^2. \tag{7.7.3}$$

Suppose that $\| \cdot \|_{\mathcal{A}}$ satisfies the C^* *identity*

$$\|x^* \, x\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^2 \tag{7.7.4}$$

for every $x \in \mathcal{A}$. This implies that

$$\|x\|_{\mathcal{A}}^2 \le \|x\|_{\mathcal{A}} \, \|x^*\|_{\mathcal{A}} \tag{7.7.5}$$

for every $x \in \mathcal{A}$. It follows that

$$\|x\|_{\mathcal{A}} \le \|x^*\|_{\mathcal{A}} \tag{7.7.6}$$

when $x \ne 0$, and which is trivial when $x = 0$. If we replace x with x^*, then we get that

$$\|x^*\|_{\mathcal{A}} \le \|x\|_{\mathcal{A}}. \tag{7.7.7}$$

This shows that (7.7.2) holds for every $x \in \mathcal{A}$ under these conditions.

More precisely, if

$$\|x\|_{\mathcal{A}}^2 \le \|x^* \, x\|_{\mathcal{A}} \tag{7.7.8}$$

for every $x \in \mathcal{A}$, then (7.7.5) holds for every $x \in \mathcal{A}$. In this case, the argument in the preceding paragraph shows that (7.7.2) holds for every $x \in \mathcal{A}$. Using this and (7.7.1), we get that (7.7.3) holds for every $x \in \mathcal{A}$. Combining this with (7.7.8), we get (7.7.4). This corresponds to Remark 2.9.1 on p75 of [8], and to part (a) of Exercise 5 on p300 of [167].

As another variant, if (7.7.2) holds, then (7.7.4) is the same as saying that

$$\|x^* \, x\|_{\mathcal{A}} = \|x^*\|_{\mathcal{A}} \, \|x\|_{\mathcal{A}} \tag{7.7.9}$$

for every $x \in \mathcal{A}$. This is mentioned on p276 of [162]. Note that $\|x \, x^*\|_{\mathcal{A}}$ is normally considered in [162], rather than $\|x^* \, x\|_{\mathcal{A}}$. Of course, the arguments for the two versions are very similar, and they are related using the involution.

If x is self-adjoint or anti-self-adjoint in $\mathcal{A}$, then (7.7.4) is the same as saying that

$$\|x \, x\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^2. \tag{7.7.10}$$

Note that

$$(x \, x)^* = x^* \, x^* = x \, x \tag{7.7.11}$$

in both cases. If x is a self-adjoint idempotent element of $\mathcal{A}$, then (7.7.10) implies that

$$\|x\|_{\mathcal{A}} = \|x \, x\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^2. \tag{7.7.12}$$

This means that

$$\|x\|_{\mathcal{A}} = 1 \tag{7.7.13}$$

when $x \ne 0$.

Suppose for the moment that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$. Remember that $e_{\mathcal{A}}$ is self-adjoint, as in Sections 6.4 and 7.5. It follows that

$$\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1, \tag{7.7.14}$$

as in (7.7.13), unless $e_{\mathcal{A}} = 0$, which would mean that $\mathcal{A} = \{0\}$. This corresponds to part (b) of Exercise (2) on p49 of [8].

Suppose from now on in this section that $\mathcal{A}$ is associative. If $x \in \mathcal{A}$ is self-adjoint, then

$$\|x^{2^l}\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^{2^l} \tag{7.7.15}$$

for every nonnegative integer l, because of (7.7.10) and (7.7.11). This also works when x is anti-self-adjoint. In both cases, one can use this to get that

$$\|x^j\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^j \tag{7.7.16}$$

for each $j \ge 1$.

If x is any element of $\mathcal{A}$, then $x^*\,x$ is self-adjoint, so that

$$\|(x^*\,x)^j\|_{\mathcal{A}} = \|x^*\,x\|_{\mathcal{A}}^j = \|x\|_{\mathcal{A}}^{2\,j} \tag{7.7.17}$$

for each $j \ge 1$. If x is normal in $\mathcal{A}$, in the sense that x commutes with x^*, then

$$(x^*\,x)^j = (x^*)^j\,x^j = (x^j)^*\,x^j \tag{7.7.18}$$

for each j. This means that

$$\|(x^*\,x)^j\|_{\mathcal{A}} = \|(x^j)^*\,x^j\|_{\mathcal{A}} = \|x^j\|_{\mathcal{A}}^2 \tag{7.7.19}$$

for each j. It follows that (7.7.16) holds in this case as well.

7.8 Continuous vector-valued functions

Let X, Y be nonempty metric or topological spaces, and let $C(X,Y)$ be the space of continuous mappings from X into Y. Of course, if X is equipped with the discrete metric or topology, then every mapping from X into Y is continuous.

Suppose from now on in this section that (Y, d_Y) is a metric space. A mapping f from X into Y is said to be *bounded* if $f(X)$ is a bounded set in Y, which means that it is contained in a ball in Y. Let $\mathcal{B}(X,Y)$ be the set of all bounded mappings from X into Y. One can define the *supremum metric* on $\mathcal{B}(X,Y)$ in a standard way, using d_Y. If Y is complete as a metric space with respect to d_Y, then it is well known that

(7.8.1) $\mathcal{B}(X,Y)$ is complete with respect to the supremum metric.

Let

$$C_b(X,Y) = \mathcal{B}(X,Y) \cap C(X,Y) \tag{7.8.2}$$

be the space of all bounded continuous mappings from X into Y. If X is compact, then every continuous mapping f from X into Y is bounded. More precisely, $f(X)$ is a compact subset of Y, which is bounded in particular.

One can check that

$$C_b(X, Y) \text{ is a closed set in } \mathcal{B}(X, Y), \tag{7.8.3}$$

with respect to the supremum metric. This basically corresponds to the fact that if $\{f_j\}_{j=1}^\infty$ is a sequence of continuous mappings from X into Y that converges uniformly to a mapping f from X into Y, then

$$f \text{ is continuous on } X \tag{7.8.4}$$

too. If Y is complete, then it follows that $C_b(X, Y)$ is complete with respect to the supremum metric, as in Section 1.6.

If X is a metric space, then we let $UC(X, Y)$ be the space of all uniformly continuous mappings from X into Y. If X is compact, then it is well known that every continuous mapping from X into Y is uniformly continuous. If X is equipped with the discrete metric, then every mapping from X into Y is uniformly continuous.

Let

$$UC_b(X, Y) = \mathcal{B}(X, Y) \cap UC(X, Y) = C_b(X, Y) \cap UC(X, Y) \tag{7.8.5}$$

be the space of bounded uniformly continuous mappings from X into Y. One can verify that

$$UC_b(X, Y) \text{ is a closed set in } \mathcal{B}(X, Y), \tag{7.8.6}$$

with respect to the supremum metric. As before, this basically corresponds to the fact that if $\{f_j\}_{j=1}^\infty$ is a sequence of uniformly continuous mappings from X into Y that converges uniformly to a mapping f from X into Y, then

$$f \text{ is uniformly continuous on } X \tag{7.8.7}$$

as well. If Y is complete, then we get that $UC_b(X, Y)$ is complete with respect to the supremum metric, as before.

Suppose now that Y is a vector space over the real or complex numbers with a norm $\|\cdot\|_Y$, using the metric on Y associated to $\|\cdot\|_Y$. If X is any nonempty topological or metric space again, then one can check that $C(X, Y)$ is a linear subspace of the space of all Y-valued functions on X, with respect to pointwise addition and scalar multiplication.

In this case, $\mathcal{B}(X, Y)$ is the same as $\ell^\infty(X, Y)$, as in Section 2.3, on which the supremum metric is the same as the metric associated to the supremum norm, as before. The space

$$C_b(X, Y) = C(X, Y) \cap \ell^\infty(X, Y) \tag{7.8.8}$$

of all bounded continuous mappings from X into Y is a linear subspace of each of $C(X, Y)$ and $\ell^\infty(X, Y)$. The restriction of the supremum norm on $\ell^\infty(X, Y)$

to $C_b(X, Y)$ defines a norm on $C_b(X, Y)$, which may be denoted $\|\cdot\|_\infty$, $\|\cdot\|_{sup}$, or $\|\cdot\|_{C_b(X,Y)}$. If Y is a Banach space, then we get that

$$C_b(X, Y) \text{ is a Banach space} \tag{7.8.9}$$

with respect to the supremum norm as well.

If X is a metric space, then one can verify that $UC(X, Y)$ is a linear subspace of $C(X, Y)$. Similarly, $UC_b(X, Y)$ is a linear subspace of $C_b(X, Y)$ and $UC(X, Y)$. If Y is a Banach space, then $UC_b(X, Y)$ is a Banach space with respect to the supremum norm too.

7.9 Algebra-valued functions

Let X be a nonempty set, and let $\mathcal{A}$ be an algebra over the real or complex numbers in the strict sense. The space of all $\mathcal{A}$-valued functions on X is an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to pointwise multiplication of functions. If $\mathcal{A}$ is associative or commutative, then the algebra of all $\mathcal{A}$-valued functions on X has the same property. If $\mathcal{A}$ has a multiplicative identity element $e_\mathcal{A}$, then the function on X equal to $e_\mathcal{A}$ at every point is the multiplicative identity element in the algebra of all $\mathcal{A}$-valued functions on X. If $\mathcal{A}$ is associative and $\mathcal{A}$ has a multiplicative identity element $e_\mathcal{A}$, then an $\mathcal{A}$-valued function f on X is invertible as an $\mathcal{A}$-valued function on X if and only if

$$f(x) \text{ is invertible in } \mathcal{A} \text{ for every } x \in X. \tag{7.9.1}$$

Suppose that $\|\cdot\|_\mathcal{A}$ is a submultiplicative norm on $\mathcal{A}$, and remember that $\ell^\infty(X, \mathcal{A})$ is the space of all bounded $\mathcal{A}$-valued functions on X, as in Section 2.3. If $f, g \in \ell^\infty(X, \mathcal{A})$, then it is easy to see that their product $f(x)\,g(x)$ is bounded on X too, with

$$\|f\,g\|_\infty \le \|f\|_\infty\,\|g\|_\infty. \tag{7.9.2}$$

Thus $\ell^\infty(X, \mathcal{A})$ is a subalgebra of the algebra of all $\mathcal{A}$-valued functions on X. If $\mathcal{A}$ has a multiplicative identity element $e_\mathcal{A}$, then the function on X equal to $e_\mathcal{A}$ at every point is bounded, with supremum norm equal to $\|e_\mathcal{A}\|_\mathcal{A}$. If $\mathcal{A}$ is also associative, then $f \in \ell^\infty(X, \mathcal{A})$ is invertible in $\ell^\infty(X, \mathcal{A})$ if and only if (7.9.1) holds, and

$$f(x)^{-1} \text{ is bounded on } X. \tag{7.9.3}$$

If f is a bounded real or complex-valued function on X, then one can check that

$$\|f^l\|_\infty = \|f\|_\infty^l \tag{7.9.4}$$

for every positive integer l. Of course, this uses the standard absolute value function on the real or complex numbers.

Suppose now that X is a metric or topological space. If f, g are continuous $\mathcal{A}$-valued functions on X, then one can check that their product $f(x)\,g(x)$ is continuous on X as well. This implies that the space $C(X, \mathcal{A})$ of all continuous $\mathcal{A}$-valued functions on X is a subalgebra of the space of all $\mathcal{A}$-valued functions on X. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then the function on X equal to $e_{\mathcal{A}}$ at every point is continuous, and thus is the multiplicative identity element in $C(X, \mathcal{A})$. If $\mathcal{A}$ is associative as well, then $f \in C(X, \mathcal{A})$ is invertible in $C(X, \mathcal{A})$ if and only if (7.9.1) holds, because of continuity of taking inverses on $G(\mathcal{A})$, as in Section 6.7.

Similarly, the space

$$C_b(X, \mathcal{A}) = C(X, \mathcal{A}) \cap \ell^\infty(X, \mathcal{A}) \tag{7.9.5}$$

of all bounded continuous $\mathcal{A}$-valued functions on X is a subalgebra of each of $C(X, \mathcal{A})$ and $\ell^\infty(X, \mathcal{A})$. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then the function on X equal to $e_{\mathcal{A}}$ at every point is the multiplicative identity element in $C_b(X, \mathcal{A})$. If $\mathcal{A}$ is also associative, then $f \in C_b(X, \mathcal{A})$ is invertible in $C_b(X, \mathcal{A})$ if and only if (7.9.1) and (7.9.3) hold, because of continuity of taking inverses on $G(\mathcal{A})$.

Suppose that X is a metric space, and that f, g are bounded uniformly continuous $\mathcal{A}$-valued functions on X. One can check that

$$f\,g \text{ is uniformly continuous on } X, \tag{7.9.6}$$

using (7.10.7). This implies that $UC_b(X, \mathcal{A})$ is a subalgebra of $C_b(X, \mathcal{A})$.

Suppose that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, so that the function on X equal to $e_{\mathcal{A}}$ at every point is the multiplicative identity element in $UC_b(X, \mathcal{A})$. If $\mathcal{A}$ is associative, and $f \in UC_b(X, \mathcal{A})$ satisfies (7.9.1) and (7.9.3), then one can verify that

$$f(x)^{-1} \text{ is uniformly continuous on } X, \tag{7.9.7}$$

using (7.10.12).

7.10 Bounded Lipschitz functions

Let (X, d_X), (Y, d_Y) be nonempty metric spaces. Also let $\{f_j\}_{j=1}^\infty$ be a sequence of Lipschitz mappings from X into Y that converges pointwise to a mapping f from X into Y. Suppose that there is a nonnegative real number C such that

$$\mathrm{Lip}_{X,Y}(f_j) \le C \tag{7.10.1}$$

for each j, where $\mathrm{Lip}_{X,Y}(f_j)$ is the minimal Lipschitz constant of f_j, as in Section 2.1. Under these conditions, one can check that f is Lipschitz, with

$$\mathrm{Lip}_{X,Y}(f) \le C. \tag{7.10.2}$$

Let W be a vector space over the real or complex numbers with a norm $\|\cdot\|_W$. Consider the space

$$\mathrm{Lip}_b(X, W) = \mathrm{Lip}(X, W) \cap C_b(X, W) \tag{7.10.3}$$

of all bounded Lipschitz mappings from X into W. This is a linear subspace of each of $\mathrm{Lip}(X, W)$ and $C_b(X, W)$.

If f is a Lipschitz mapping from X into W, then we let $\mathrm{Lip}(f)$ be the minimal Lipschitz constant of f, as in Section 2.1 again. This defines a seminorm on $\mathrm{Lip}(X, W)$, as before. If t is a nonnegative real number and $f \in \mathrm{Lip}_b(X, W)$, then put

$$\|f\|_{\mathrm{Lip}_b, t} = \|f\|_{\mathrm{Lip}_b(X,W), t} = \|f\|_{sup} + t\ \mathrm{Lip}(f). \tag{7.10.4}$$

It is easy to see that this defines a norm on $\mathrm{Lip}_b(X, W)$.

If W is a Banach space and $t > 0$, then

$$\mathrm{Lip}_b(X, W) \text{ is a Banach space with respect to } \|\cdot\|_{\mathrm{Lip}_{b,t}}. \tag{7.10.5}$$

To see this, let $\{f_j\}_{j=1}^\infty$ be a sequence of elements of $\mathrm{Lip}_b(X, W)$ that is a Cauchy sequence with respect to the metric associated to $\|\cdot\|_{\mathrm{Lip}_b, t}$. In particular, $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to the supremum metric, so that $\{f_j\}_{j=1}^\infty$ converges uniformly to a bounded continuous W-valued function f on X. One can use the fact that $\{f_j\}_{j=1}^\infty$ is bounded with respect to $\|\cdot\|_{\mathrm{Lip}_b, t}$ to get that f is Lipschitz on X, as in (7.10.2). One can also use the Cauchy condition for $\{f_j\}_{j=1}^\infty$ with respect to the metric associated to $\|\cdot\|_{\mathrm{Lip}_b, t}$ to get that $\{f_j\}_{j=1}^\infty$ converges to f with respect to this metric.

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers, with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If f, g are $\mathcal{A}$-valued functions on X, then

$$f(x)\,g(x) - f(w)\,g(w) = (f(x) - f(w))\,g(x) + f(w)\,(g(x) - g(w)) \tag{7.10.6}$$

for every $x, w \in X$, so that

$$\begin{aligned} \|f(x)\,g(x) - f(w)\,g(w)\|_{\mathcal{A}} &\le \|f(x) - f(w)\|_{\mathcal{A}}\,\|g(x)\|_{\mathcal{A}} \\ &\quad + \|f(w)\|_{\mathcal{A}}\,\|g(x) - g(w)\|_{\mathcal{A}}. \end{aligned} \tag{7.10.7}$$

If $f, g \in \mathrm{Lip}_b(X, \mathcal{A})$, then it follows that $f\,g$ is Lipschitz on $\mathcal{A}$ too, with

$$\mathrm{Lip}(f\,g) \le \mathrm{Lip}(f)\,\|g\|_{sup} + \|f\|_{sup}\ \mathrm{Lip}(g). \tag{7.10.8}$$

This implies that $\mathrm{Lip}_b(X, \mathcal{A})$ is a subalgebra of $C_b(X, \mathcal{A})$. One can also check that

$$\|f\,g\|_{\mathrm{Lip}_b, t} \le \|f\|_{\mathrm{Lip}_b, t}\,\|g\|_{\mathrm{Lip}_b, t} \tag{7.10.9}$$

for each $t \ge 0$.

Suppose that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, so that the function $\mathbf{1}_X\,e_{\mathcal{A}}$ equal to $e_{\mathcal{A}}$ at every point in X is the multiplicative identity element in $\mathrm{Lip}_b(X, \mathcal{A})$. Note that

$$\|\mathbf{1}_X\,e_{\mathcal{A}}\|_{\mathrm{Lip}_b, t} = \|e_{\mathcal{A}}\|_{\mathcal{A}} \tag{7.10.10}$$

for each $t \ge 0$.

Suppose that $\mathcal{A}$ is also associative, and that f is an $\mathcal{A}$-valued function on X such that $f(x)$ is invertible in $\mathcal{A}$ for every $x \in X$. Observe that

$$f(x)^{-1} - f(w)^{-1} = f(x)^{-1}\,(f(w) - f(x))\,f(w)^{-1} \tag{7.10.11}$$

for every $x, w \in X$, so that

$$\|f(x)^{-1} - f(w)^{-1}\|_{\mathcal{A}} \le \|f(x)^{-1}\|_{\mathcal{A}} \, \|f(w) - f(x)\|_{\mathcal{A}} \, \|f(w)^{-1}\|_{\mathcal{A}}. \tag{7.10.12}$$

If $f(x)^{-1}$ is bounded on X, and f is Lipschitz on X, then it follows that $f(x)^{-1}$ is Lipschitz on X, with

$$\mathrm{Lip}(f(\cdot)^{-1}) \le \|f(\cdot)^{-1}\|_{sup}^2 \, \mathrm{Lip}(f). \tag{7.10.13}$$

Suppose now that f is a real or complex-valued function on X that is bounded and Lipschitz. If $l \ge 2$ is an integer, then one can check that

$$\mathrm{Lip}(f^l) \le l \, \|f\|_{sup}^{l-1} \, \mathrm{Lip}(f), \tag{7.10.14}$$

using (7.10.8). This implies that

$$\|f\|_{sup}^l = \|f^l\|_{sup} \le \|f^l\|_{\mathrm{Lip}_b, t} \le \|f\|_{sup}^l + t \, l \, \|f\|_{sup}^{l-1} \, \mathrm{Lip}(f) \tag{7.10.15}$$

for each $t \ge 0$. One can use this to get that

$$\lim_{l \to \infty} \|f^l\|_{\mathrm{Lip}_b, t}^{1/l} = \|f\|_{sup} \tag{7.10.16}$$

for each $t > 0$.

7.11 Bilipschitz conditions and invertibility

Let (X, d_X) and (Y, d_Y) be metric spaces, and let f be a mapping from X into Y. Suppose that there is a positive real number c such that

$$c \, d_X(x, w) \le d_Y(f(x), f(w)) \tag{7.11.1}$$

for every $x, w \in X$. This implies in particular that f is one-to-one on X, so that the inverse mapping f^{-1} may be defined on the image $f(X)$ of X under f in Y. More precisely, (7.11.1) is the same as saying that

$$f^{-1} \text{ is Lipschitz with constant } 1/c \text{ on } f(X), \tag{7.11.2}$$

with respect to the restriction of d_Y to $f(X)$.

If f is also Lipschitz on X, then f is said to be a *bilipschitz mapping* from X into Y. In this case, if X is complete as a metric space, then it is easy to see that

$$f(X) \text{ is complete, with respect to the restriction of } d_Y \text{ to } f(X). \tag{7.11.3}$$

This implies that

$$f(X) \text{ is a closed set in } Y, \tag{7.11.4}$$

as in Section 1.6. If $f(X)$ is dense in Y, then it follows that

$$f(X) = Y. \tag{7.11.5}$$

Suppose now that Y is a vector space over the real or complex numbers with a norm $\|\cdot\|_Y$, and that d_Y is the metric on Y associated to $\|\cdot\|_Y$. Let f be a mapping from X into Y that satisfies (7.11.1) for some $c > 0$ again, and let g be a mapping from X into Y such that

$$f - g \text{ is Lipschitz on } X. \tag{7.11.6}$$

If $x, w \in X$, then

$$\begin{aligned} c\, d_X(x,w) &\le \|f(x) - f(w)\|_Y \\ &\le \|(f(x) - g(x)) - (f(w) - g(w))\|_Y + \|g(x) - g(w)\|_Y \\ &\le \mathrm{Lip}_{X,Y}(f-g)\, d_X(x,w) + \|g(x) - g(w)\|_Y, \end{aligned} \tag{7.11.7}$$

where $\mathrm{Lip}_{X,Y}(f-g)$ is the minimal Lipschitz constant of $f - g$ on X, as in Section 2.1. This implies that

$$(c - \mathrm{Lip}_{X,Y}(f-g))\, d_X(x,w) \le \|g(x) - g(w)\|_Y. \tag{7.11.8}$$

If

$$\mathrm{Lip}_{X,Y}(f-g) < c, \tag{7.11.9}$$

then this is the same type of condition as before.

Suppose that X is also a vector space over the real or complex numbers, as appropriate, with a norm $\|\cdot\|_X$, and that d_X is the metric on X associated to $\|\cdot\|_X$. If f is a linear mapping from X into Y, then (7.11.1) implies that

$$c\,\|x\|_X \le \|f(x)\|_Y \tag{7.11.10}$$

for every $x \in X$. Conversely, this condition implies that

$$c\,\|x - w\|_X \le \|f(x - w)\|_Y = \|f(x) - f(w)\|_Y \tag{7.11.11}$$

for every $x, w \in X$, which is the same as (7.11.1) in this case. Note that $f(X)$ is a linear subspace of Y, and that (7.11.10) is the same as saying that f^{-1} is bounded as a linear mapping from $f(X)$ into X, with respect to the restriction of $\|\cdot\|_Y$ to $f(X)$, and with operator norm less than or equal to $1/c$.

Suppose that g is a linear mapping from X into Y such that

$$f - g \in \mathcal{BL}(X, Y). \tag{7.11.12}$$

If f satisfies (7.11.10) for some $c > 0$, then

$$\begin{aligned} c\,\|x\|_X &\le \|f(x) - g(x)\|_Y + \|g(x)\|_Y \\ &\le \|f - g\|_{op,X,Y}\, \|x\|_X + \|g(x)\|_X \end{aligned} \tag{7.11.13}$$

for every $x \in X$. This implies that

$$(c - \|f - g\|_{op,XY})\, \|x\|_X \le \|g(x)\|_Y \tag{7.11.14}$$

for every $x \in X$, which is the same type of condition as before when

$$\|f - g\|_{op,XY} < c. \tag{7.11.15}$$

Of course, this could also be obtained from the analogous statements for Lipschitz mappings, but the arguments are a bit simpler in this case.

Suppose that f satisfies (7.11.1) for some $c > 0$, and that (7.11.6) holds. If $x \in X$, then we get that

$$\|f(x) - g(x)\|_Y \le \|f - g\|_{op,XY} \, \|x\|_X \le c^{-1} \, \|f - g\|_{op,XY} \, \|f(x)\|_Y. \tag{7.11.16}$$

Suppose that (7.11.15) holds as well, which is the same as saying that

$$c^{-1} \, \|f - g\|_{op,XY} < 1. \tag{7.11.17}$$

If $f(X) = Y$, then it follows that

$$g(X) \text{ is dense in } Y, \tag{7.11.18}$$

as in Section 4.6. This implies that $g(X) = Y$ when X is complete, as before.

Similarly, if $f(X)$ is dense in Y, then (7.11.18) holds. Indeed, if $x \in X$ and $y \in Y$, then

$$\begin{aligned} \|y - g(x)\|_Y &\le \|y - f(x)\|_Y + \|f(x) - g(x)\|_Y \\ &\le \|y - f(x)\|_Y + c^{-1} \, \|f - g\|_{op,XY} \, \|f(x)\|_Y, \end{aligned} \tag{7.11.19}$$

using (7.11.16) in the second step. This implies that

$$\begin{aligned} \|y - g(x)\|_Y \le\ & (1 + c^{-1} \, \|f - g\|_{op,XY}) \, \|y - f(x)\|_Y \\ & + c^{-1} \, \|f - g\|_{op,XY} \, \|y\|_Y. \end{aligned} \tag{7.11.20}$$

One can use this to get that the criterion for density in Section 4.6 holds in this case too.

7.12 The contraction mapping theorem

Let (X, d_X) be a nonempty metric space, and let ϕ be a mapping from X into itself. Suppose that ϕ is a *contraction* on X, which means that ϕ is Lipschitz with

$$\mathrm{Lip}(\phi) < 1. \tag{7.12.1}$$

If X is complete, then the *contraction mapping theorem* states that

$$\phi \text{ has a unique fixed point in } X, \tag{7.12.2}$$

which is to say that there is a unique $x \in X$ such that

$$\phi(x) = x. \tag{7.12.3}$$

The uniqueness of the fixed point can be verified directly, and does not use completeness of X.

To get the existence of the fixed point, let $x_1 \in X$ be given, and let $\{x_j\}_{j=1}^\infty$ be the sequence of elements of X defined recursively by

$$x_{j+1} = \phi(x_j) \tag{7.12.4}$$

for each j. One can show that

$$\{x_j\}_{j=1}^\infty \text{ is a Cauchy sequence in } X, \tag{7.12.5}$$

using (7.12.1). This implies that $\{x_j\}_{j=1}^\infty$ converges in X, because X is complete, and we put

$$x = \lim_{j \to \infty} x_j. \tag{7.12.6}$$

One can check that (7.12.3) holds, using (7.12.4), and the fact that ϕ is continuous on X.

Suppose now that X is a Banach space over the real or complex numbers, with norm $\|\cdot\|_X$, and with d_X equal to the metric associated to the norm. Put

$$\psi(x) = x - \phi(x) \tag{7.12.7}$$

for each $x \in X$, which defines a mapping from X into itself. More precisely, ψ is Lipschitz on X, with

$$\mathrm{Lip}(\psi) \le 1 + \mathrm{Lip}(\phi), \tag{7.12.8}$$

as in Section 2.1. In fact,

$$\psi \text{ is bilipschitz on } X, \tag{7.12.9}$$

as in the previous section. We would like to show that

$$\psi(X) = X, \tag{7.12.10}$$

using the contraction mapping theorem.

Let $y \in X$ be given, and put

$$\phi_y(x) = \phi(x) + y \tag{7.12.11}$$

for each $x \in X$. Note that ϕ_y is a Lipschitz mapping from X into itself, with

$$\mathrm{Lip}(\phi_y) = \mathrm{Lip}(\phi). \tag{7.12.12}$$

Thus

$$\phi_y \text{ is a contraction on } X, \tag{7.12.13}$$

because ϕ is a contraction, by hypothesis. The contraction mapping theorem implies that ϕ_y has a unique fixed point $x(y)$ in X, so that

$$\phi(x(y)) + y = \phi_y(x(y)) = x(y). \tag{7.12.14}$$

This implies that

$$\psi(x(y)) = y, \tag{7.12.15}$$

so that (7.12.10) holds.

If ϕ is a bounded linear mapping on X, then (7.12.1) is the same as saying that

$$\|\phi\|_{op} < 1. \tag{7.12.16}$$

This gives another way to look at the invertibility of ψ as a bounded linear mapping on X under these conditions.

7.13 The open mapping theorem

Let X, Y be metric spaces, or topological spaces. A mapping f from X into Y is said to be *open* at a point $x \in X$ if for every open set $U \subseteq X$ with $x \in U$ there is an open set $V \subseteq Y$ such that

$$f(x) \in V \subseteq f(U). \tag{7.13.1}$$

We simply say that f is an *open mapping* if for every open set $U \subseteq X$,

$$f(U) \text{ is an open set in } Y. \tag{7.13.2}$$

If f is an open mapping, then f is clearly open at every $x \in X$. Conversely, if f is open at every $x \in X$, then it is easy to see that f is an open mapping.

Suppose for the moment that

$$f \text{ is a one-to-one mapping from } X \text{ onto } Y. \tag{7.13.3}$$

In this case, f is an open mapping at $x \in X$ if and only if

$$f^{-1} \text{ is continuous at } f^{-1}(x). \tag{7.13.4}$$

Similarly, f is an open mapping if and only if f^{-1} is continuous.

Suppose for the moment again that

$$X \text{ is compact} \tag{7.13.5}$$

and

$$Y \text{ is Hausdorff,} \tag{7.13.6}$$

and note that (7.13.6) holds when Y is a metric space. If $E \subseteq X$ is a closed set, then it is well known that E is compact, because X is compact. If f is a continuous mapping from X into Y, then it follows that $f(E)$ is compact in Y. This implies that

$$f(E) \text{ is a closed set in } Y, \tag{7.13.7}$$

because Y is Hausdorff. If f is also a one-to-one mapping from X onto Y, then we get that

$$f^{-1} \text{ is continuous} \tag{7.13.8}$$

under these conditions.

Now let V, W be vector spaces, both real or both complex, with norms $\|\cdot\|_V$, $\|\cdot\|_W$, respectively. Also let T be a linear mapping from V into W. If

$$T \text{ is an open mapping at } 0, \tag{7.13.9}$$

then it is easy to see that T is an open mapping at every point in V, so that

$$T \text{ is an open mapping from } V \text{ into } W. \tag{7.13.10}$$

Note that

$$T(V) = W \tag{7.13.11}$$

in this case.

Let us use $B_V(v, r)$, $B_W(w, r)$ for the open balls in V, W centered at $v \in V$, $w \in W$ with radius $r > 0$ with respect to the metrics associated to the norms, respectively. If T is open at 0, then there is an $r_1 > 0$ such that

$$B_W(0, r_1) \subseteq T(B_V(0, 1)). \tag{7.13.12}$$

Conversely, this condition implies that

$$B_W(0, r_1\, r) \subseteq T(B_V(0, r)) \tag{7.13.13}$$

for every $r > 0$, because of linearity. It follows that T is an open mapping at 0.

Suppose that V, W are Banach spaces, and that T is a bounded linear mapping from V onto W. Under these conditions, the *open mapping theorem* states that

$$T \text{ is an open mapping.} \tag{7.13.14}$$

It suffices to show that (7.13.12) holds for some $r_1 > 0$, as in the preceding paragraph.

Observe that

$$\bigcup_{j=1}^{\infty} T(B_V(0, j)) = W, \tag{7.13.15}$$

because T maps $\bigcup_{j=1}^{\infty} B_V(0, j) = V$ onto W, by hypothesis. In particular, this means that

$$\bigcup_{j=1}^{\infty} \overline{T(B_V(0, j))} = W. \tag{7.13.16}$$

Because W is a Banach space, we can use the Baire category theorem to get that

$$\overline{T(B_V(0, j))} \text{ has nonempty interior in } W \tag{7.13.17}$$

for some j.

More precisely, this implies that $\overline{T(B_V(0, j))}$ contains an open set that contains an element of $T(B_V(0, j))$. One can use this to get that

$$B_W(0, r_0) \subseteq \overline{T(B_V(0, 2\, j))} \tag{7.13.18}$$

for some $r_0 > 0$.

Using this and linearity, one can check that there is a positive real number C such that for each $w \in W$ there is a $v \in V$ such that

$$\|v\|_V \le C\,\|w\|_W \tag{7.13.19}$$

and $\|w - T(v)\|_W$ is as small as we like.

We can repeat the process and approximate $w - T(v)$ in the same way. Continuing in this manner, we get an infinite series in V that converges absolutely and thus converges in V, because V is a Banach space. By construction, T sends the sum of this series to w. The norm of the sum of the series is less than or equal to the sum of the norms, which can be estimated by $C\,\|w\|_W$ plus an arbitrarily small positive real number.

7.14 Open mappings and quotient spaces

Let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$. Also let V_0 be a closed linear subspace of V, and let V/V_0 be the corresponding quotient space, with the quotient norm $\|\cdot\|_{V/V_0}$ defined as in Section 6.10. One can check that

(7.14.1) the natural quotient mapping q_0 from V onto V/V_0 is an open mapping,

with respect to the metrics associated to the norms. More precisely, q_0 maps the open ball in V centered at 0 with radius $r > 0$ onto the open ball in V/V_0 centered at 0 with radius r, because of the way that $\|\cdot\|_{V/V_0}$ is defined.

Let W be another vector space over the real or complex numbers, as appropriate, and with a norm $\|\cdot\|_W$. If T_0 is a linear mapping from V/V_0 into W, then

$$T = T_0 \circ q_0 \tag{7.14.2}$$

defines a linear mapping from V into W such that

$$V_0 \subseteq \ker T. \tag{7.14.3}$$

If T_0 is bounded as a linear mapping from V/V_0 into W, then it follows that T is a bounded linear mapping from V into W, with

$$\|T\|_{op,VW} \le \|T_0\|_{op,(V/V_0)W}. \tag{7.14.4}$$

In fact, one can check that

$$\|T_0\|_{op,(V/V_0)W} \le \|T\|_{op,VW} \tag{7.14.5}$$

too. This means that

$$\|T_0\|_{op,(V/V_0)W} = \|T\|_{op,VW}. \tag{7.14.6}$$

Conversely, if T is a linear mapping from V into W that satisfies (7.14.3), then there is a unique linear mapping T_0 from V/V_0 into W such that (7.14.2) holds, by standard arguments. If T is a bounded linear mapping from V into W, then one can verify that

(7.14.7) $$T_0 \text{ is bounded on } V/V_0,$$

and that (7.14.5) holds. We also have that (7.14.4) holds, so that (7.14.6) holds, as before.

If T is any bounded linear mappng from V into W, then the kernel of T is a closed linear subspace of V. Thus we can take

(7.14.8) $$V_0 = \ker T$$

in the previous paragraphs. This leads to a bounded linear mapping T_0 from V/V_0 into W as in (7.14.2). In this case, we also have that

(7.14.9) $$\ker T_0 = \{0\},$$

by construction.

Of course,

(7.14.10) $$T(V) = T_0(V/V_0).$$

Suppose that T maps V onto W, so that T_0 maps V/V_0 onto W. Suppose that V and W are Banach spaces as well, and remember that V/V_0 is a Banach space too, as in Section 6.10. The open mapping theorem implies that

(7.14.11) $$T_0^{-1} \text{ is bounded as a linear mapping from } W \text{ onto } V/V_0.$$

7.15 The closed graph theorem

Let X, Y be metric spaces, or topological spaces, and let f be a mapping from X into Y. The *graph* of f is the subset

(7.15.1) $$\{(x, f(x)) : x \in X\}$$

of the Cartesian product $X \times Y$. If f is continuous, and X, Y are metric spaces, then one can check that

(7.15.2) $$\text{the graph of } f \text{ is a closed set in } X \times Y,$$

with respect to a suitable metric on $X \times Y$ as in Section 5.11. This also works when X, Y are topological spaces, using the product topology on $X \times Y$, if Y is Hausdorff.

Suppose for the moment that

(7.15.3) $$X \text{ and } Y \text{ are compact,}$$

so that $X \times Y$ is compact, by Tychonoff's theorem. If the graph of f is a closed set in $X \times Y$, then it is compact as well. Let p be the restriction of the obvious coordinate projection from $X \times Y$ onto X to the graph of f, so that

$$p((x, f(x))) = x \tag{7.15.4}$$

for every $x \in X$. This is a one-to-one mapping from the graph of f onto X. It is easy to see that

$$p \text{ is continuous,} \tag{7.15.5}$$

with respect to the restriction of a suitable metric on $X \times Y$ to the graph of f when X and Y are metric spaces, or with respect to the topology induced on the graph of f by the product topology on $X \times Y$ when X and Y are topological spaces. It follows that

$$p^{-1} \text{ is continuous} \tag{7.15.6}$$

when X, Y are metric spaces, and when X, Y are topological spaces and X is Hausdorff, as in Section 7.13. This implies that

$$f \text{ is continuous} \tag{7.15.7}$$

under these conditions.

Now let V, W be vector spaces, both real or both complex, and let T be a linear mapping from V into W. As before, the graph of T is the subset

$$\{(v, T(v)) : v \in V\} \tag{7.15.8}$$

of the Cartesian product $V \times W$. Remember that $V \times W$ may be considered as a vector space over the real or complex numbers, as appropriate, with respect to coordinatewise addition and scalar multiplication. It is easy to see that the graph of T is a linear subspace of $V \times W$.

Let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively, which can be used to get a suitable norm on $V \times W$, as in Section 5.12. If T is a bounded linear mapping from V into W, then

$$\text{the graph of } T \text{ is a closed set in } V \times W \tag{7.15.9}$$

with respect to the metric associated to such a norm, as before.

The condition that the graph of T be a closed set in $V \times W$ is equivalent to saying that if

$$\begin{gathered}\{(v_j, T(v_j))\}_{j=1}^\infty \text{ is a sequence of elements of the graph of } T \\ \text{that converges to } (v, w) \in V \times W,\end{gathered} \tag{7.15.10}$$

then (v, w) is an element of the graph of T, so that

$$w = T(v). \tag{7.15.11}$$

Thus the graph of T is a closed set in $V \times W$ if and only if for every sequence $\{v_j\}_{j=1}^\infty$ of elements of V that converges to $v \in V$ with respect to the metric associated to $\|\cdot\|_V$, and for which $\{T(v_j)\}_{j=1}^\infty$ converges to $w \in W$ with respect to the metric associated to $\|\cdot\|_W$, we have that (7.15.11) holds. One can also use linearity of T to reduce to the case where $v = 0$.

Suppose that V and W are Banach spaces, and that the graph of T is a closed set in $V \times W$. Under these econditions, the *closed graph theorem* states that

(7.15.12) T is a bounded linear mapping from V into W.

To see this, note that

(7.15.13) $V \times W$ is a Banach space with respect to a suitable norm,

as in Section 5.12. It follows that

(7.15.14) the graph of T is a Banach space

with respect to the restriction of the norm on $V \times W$ to the graph of T, as in Section 1.6.

Let p be the restriction of the obvious coordinate projection from $V \times W$ onto V to the graph of T, as before. This is a bounded linear mapping from the graph of T onto V with respect to the appropriate norms. More precisely, p is a one-to-one mapping from the graph of T onto V, and the open mapping theorem implies that

(7.15.15) p^{-1} is a bounded linear mapping from V onto the graph of T.

This means that T is a bounded linear mapping from V into W.

Chapter 8

Bilinear forms and linear mappings

8.1 Invertible mappings and product spaces

Let X, Y be metric spaces, or topological spaces, and consider the Cartesian products $X \times Y$, $Y \times X$. If X, Y are metric spaces, then we can get suitable metrics on $X \times Y$, $Y \times X$ as in Section 5.11. We may as well use the same way to get these metrics on $X \times Y$ and $Y \times X$, so that

$$(x, y) \mapsto (y, x) \tag{8.1.1}$$

is an isometry from $X \times Y$ onto $Y \times X$. Otherwise, if X, Y are topological spaces, then we can use the product topologies on $X \times Y$ and $Y \times X$, and (8.1.1) is a homeomorphism.

If f is a one-to-one mapping from X onto Y, then the graph

$$\{(y, f^{-1}(y)) : y \in Y\} \tag{8.1.2}$$

of f^{-1} in $Y \times X$ is the same as

$$\{(f(x), x) : x \in X\}. \tag{8.1.3}$$

This corresponds to the graph (7.15.1) of f under the mapping (8.1.1). In particular, this means that the graph of f is a closed set in $X \times Y$ if and only if

$$\text{the graph of } f^{-1} \text{ is a closed set in } Y \times X. \tag{8.1.4}$$

Let V, W be vector spaces, both real or both complex, so that $V \times W$ and $W \times V$ may be considered as real or complex vector spaces, as appropriate, with respect to coordinatewise addition and scalar multiplication. Of course,

$$(v, w) \mapsto (w, v) \tag{8.1.5}$$

defines an isomorphism from $V \times W$ onto $V \times W$, as real or complex vector spaces, as appropriate. If T is a one-to-one linear mapping from V onto W, then the graph

$$\{(w, T^{-1}(w)) : w \in W\} \tag{8.1.6}$$

of T^{-1} is a linear subspace of $W \times V$. This is the same as

$$\{(T(v), v) : v \in V\}. \tag{8.1.7}$$

This corresponds to the graph (7.15.8) of T under the mapping (8.1.5), as before.

Let $\|\cdot\|_V$. $\|\cdot\|_W$ be norms on V, W, respectively, which can be used to get suitable norms on $V \times W$ and $W \times V$, as in Section 5.12. We may as well use the same way to get norms on each of $V \times W$ and $W \times V$, so that (8.1.5) is an isometric linear mapping from $V \times W$ onto $W \times V$. If T is a bounded linear mapping from V into W, then the graph of T is a closed set in $V \times W$, as in Section 7.15.

If T is a one-to-one bounded linear mapping from V onto W, then it follows that

$$\text{the graph of } T^{-1} \text{ is a closed set in } W \times V. \tag{8.1.8}$$

If V, W are Banach spaces, then the conclusion of the closed graph theorem implies that T^{-1} is a bounded linear mapping from W onto V. Of course, this is the same as the version of the open mapping theorem that was used to get the closed graph theorem.

Let V be a vector space over the real or complex numbers again, and let V_1, V_2 be linear subspaces of V. Thus $V_1 \times V_2$ may be considered as a vector space over the real or complex numbers, as appropriate, as well, and

$$(v_1, v_2) \mapsto v_1 + v_2 \tag{8.1.9}$$

defines a linear mapping from $V_1 \times V_2$ into V. This mapping sends $V_1 \times V_2$ onto

$$V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_2,\ v_2 \in V_2\}, \tag{8.1.10}$$

and the kernel of this linear mapping is equal to

$$\{(v_1, v_2) \in V_1 \times V_2 : v_1 = -v_2\}. \tag{8.1.11}$$

In particular, the kernel of (8.1.9) is equal to $\{0\}$ if and only if

$$V_1 \cap V_2 = \{0\}. \tag{8.1.12}$$

Let $\|\cdot\|_V$ be a norm on V again, whose restrictions to V_1, V_2 define norms on those spaces. We can use these norms to get a suitable norm on $V_1 \times V_2$, as in Section 5.12. It is easy to see that (8.1.9) is a bounded linear mapping from $V_1 \times V_2$ with respect to such a norm. Suppose that (8.1.10) is equal to V, and that (8.1.12) holds, so that (8.1.9) is a one-to-one mapping from $V_1 \times V_2$ onto V. We would like to have conditions under which

$$\text{the inverse of (8.1.9) is bounded as a linear mapping from } V \text{ onto } V_1 \times V_2. \tag{8.1.13}$$

A necessary condition for this to hold is that

$$V_1,\ V_2 \text{ be closed sets in } V, \tag{8.1.14}$$

with respect to the metric associated to $\|\cdot\|_V$, because $V_1 \times \{0\}$, $\{0\} \times V_2$ are closed sets in $V_1 \times V_2$. Suppose that this condition holds, and that V is a Banach space. This implies that V_1, V_2 are Banach spaces with respect to the restrictions of $\|\cdot\|_V$ to those spaces, as in Section 1.6. It follows that $V_1 \times V_2$ is a Banach space as well, as in Section 5.12. In this case, we get (8.1.13) from the open mapping theorem.

8.2 Projections on vector spaces

Let V be a vector space over the real or complex numbers. A linear mapping P from V into itself is said to be a *projection* if

$$P \circ P = P \tag{8.2.1}$$

on V, so that

$$P(P(v)) = P(v) \tag{8.2.2}$$

for every $v \in V$. This implies that

$$(I - P) \circ P = P \circ (I - P) = 0. \tag{8.2.3}$$

where $I = I_V$ is the identity map on V. It follows that

$$(I - P) \circ (I - P) = I - P, \tag{8.2.4}$$

so that $I - P$ is a projection on V as well.

Observe that

$$P(V) \subseteq \ker(I - P) \tag{8.2.5}$$

and

$$(I - P)(V) \subseteq \ker P, \tag{8.2.6}$$

by (8.2.3). The opposite inclusions can be verfiied directly. This means that

$$P(V) = \ker(I - P) \tag{8.2.7}$$

and

$$(I - P)(V) = \ker P. \tag{8.2.8}$$

One can use (8.2.7) and (8.2.8) to get that

$$(\ker P) + P(V) = V \tag{8.2.9}$$

and

$$(\ker P) \cap P(V) = \{0\}. \tag{8.2.10}$$

We may consider

$$(\ker P) \times P(V) \tag{8.2.11}$$

as a vector space over the real or complex numbers, as appropriate, with respect to coordinatewise addition and scalar multiplication, as usual. Using (8.2.9) and (8.2.10), we get that

$$(v, w) \mapsto v + w \tag{8.2.12}$$

defines a one-to-one linear mapping from (8.2.11) onto V, as in the previous section. The inverse mapping is given by

$$u \mapsto ((I - P)(u), P(u)). \tag{8.2.13}$$

Let $\|\cdot\|_V$ be a norm on V, whose restrictions to $\ker P$, $P(V)$ define norms on those spaces. This can be used to get a suitable norm on (8.2.11), as in Section 5.12. Note that (8.2.12) is bounded as a linear mapping from (8.2.11) into V, as in the previous section. The inverse mapping is bounded if and only if

$$P \text{ is bounded on } V, \tag{8.2.14}$$

because it can be given as in (8.2.13). Of course, if P is bounded on V, then $I - P$ is bounded on V, and (8.2.7), (8.2.8) are closed sets in V.

If V_1, V_2 are linear subspaces of V such that $V_1 + V_2 = V$ and $V_1 \cap V_2 = \{0\}$, then (8.2.12) defines a one-to-one linear mapping from $V_1 \times V_2$ onto V, as in the previous section. In this case, it is easy to see that there is a unique projection P on V with

$$\ker P = V_1 \tag{8.2.15}$$

and

$$P(V) = V_2. \tag{8.2.16}$$

If V is a Banach space, and V_1, V_2 are closed sets in V, then the inverse of (8.2.12) as a linear mapping from $V_1 \times V_2$ onto V is bounded, with respect to a suitable norm on $V_1 \times V_2$, as before. This means that (8.2.14) holds, which corresponds to a simplification of part (b) of Theorem 5.16 on p126 of [162]. More precisely, the proof in [162] uses the closed graph theorem a bit more directly.

8.3 Orthogonal projections

Let $(V, \langle\cdot,\cdot\rangle_V)$ be an inner product space over the real or complex numbers, with corresponding norm $\|\cdot\|_V$. Also let W be a linear subspace of V, and suppose that for each $v \in V$ there is an element $P_W(v)$ of W such that

$$\langle v - P_W(v), w\rangle_V = 0 \tag{8.3.1}$$

for every $w \in W$. Remember that $P_W(v)$ is uniquely determined by these properties, as in Section 2.12. It is easy to see that P_W is linear, using uniqueness. If $v \in W$, then we have that

$$P_W(v) = v, \tag{8.3.2}$$

by uniqueness.

It follows that P_W is a projection on V, as in the previous section. This is the *orthogonal projection* of V onto W. If W has finite dimension, then one can use an orthonormal basis for W to get P_W, as in Section 2.12. If V is a Hilbert space, and W is a closed linear subspace of V, then one can get P_W as in Sections 2.14 or 2.15.

Note that P_W is a bounded linear mapping from V into itself with respect to $\|\cdot\|_V$, with operator norm less than or equal to 1, as in Section 2.12. More precisely,

$$\|P_W\|_{op} = 1 \tag{8.3.3}$$

when $W \neq \{0\}$, and otherwise $P_W = 0$. We also have that

$$W = P_W(V) = \ker(I - P_W), \tag{8.3.4}$$

as in the previous section. This implies that

$$W \text{ is a closed set in } V, \tag{8.3.5}$$

with respect to the metric associated to $\|\cdot\|_V$, because $I - P_W$ is a bounded linear mapping from V into itself, as before.

Remember that the orthogonal complement of W in V is the closed linear subspace of V defined by

$$W^\perp = \{u \in V : \langle u, w\rangle_V = 0 \text{ for every } w \in W\}, \tag{8.3.6}$$

as in Section 2.15. Note that

$$W^\perp = \ker P_W, \tag{8.3.7}$$

as before. This implies that

$$(I - P_W)(V) = W^\perp, \tag{8.3.8}$$

as in the previous section. Of course, (8.3.1) is the same as saying that

$$(I - P_W)(V) \subseteq W^\perp, \tag{8.3.9}$$

and the opposite inclusion can be obtained from (8.3.7).

Put

$$P_{W^\perp} = I - P_W. \tag{8.3.10}$$

If $v \in V$, then $P_{W^\perp}(v) \in W^\perp$, and

$$\langle v - P_{W^\perp}(v), u\rangle_V = \langle P_W(v), u\rangle_V = 0 \tag{8.3.11}$$

for every $u \in W^\perp$, because $P_W(v) \in W$. These properties determine $P_{W^\perp}(v)$ uniquely, as in Section 2.12 again. This means that $P_{W^\perp}$ is the orthogonal projection of V onto $W^\perp$, as before.

Let V_1, V_2 be linear subspaces of V that are orthogonal to each other with respect to $\langle \cdot, \cdot \rangle_V$ on V, which is to say that

$$\langle v_1, v_2 \rangle_V = 0 \tag{8.3.12}$$

for every $v_1 \in V_1$ and $v_2 \in V_2$. Equivalently, this means that

$$V_1 \subseteq V_2^\perp, \ V_2 \subseteq V_1^\perp. \tag{8.3.13}$$

In particular, this implies that $V_1 \cap V_2 = \{0\}$. Suppose that $V_1 + V_2 = V$, so that every element of V can be expressed in a unique way as a sum of elements of V_1 and V_2. This leads to a unique linear mapping P from V into itself such that

$$P(v_1 + v_2) = v_2 \tag{8.3.14}$$

for every $v_1 \in V_1$ and $V_2 \in V_2$.

It is easy to see that P is the same as the orthogonal projection P_{V_2} of V onto V_2 under these conditions. Similarly,

$$(I - P)(v_1 + v_2) = v_1 \tag{8.3.15}$$

for every $v_1 \in V_1$ and $v_2 \in V_2$, and $I - P$ is the orthogonal projection P_{V_1} of V onto V_1. We also have that

$$V_1 = V_2^\perp, \ V_2 = V_1^\perp \tag{8.3.16}$$

in this case, as in Section 2.15. In particular, this implies that V_1, V_2 are closed sets in V, as before. Remember that P is bounded as a linear mapping from V into itself, as mentioned earlier.

8.4 Projections and distances

Let V be a vector space over the real or complex numbers, and let V_1, V_2 be linear subspaces of V such that $V_1 \cap V_2 = \{0\}$ and $V_1 + V_2 = V$. This means that every element of V can be expressed in a unique way as $v_1 + v_2$ for some $v_1 \in V_1$ and $v_2 \in V_2$, which leads to a unique linear mapping P from V into itself that sends $v_1 + v_2$ to v_2. This is the unique projection on V with kernel V_1 that maps V onto V_2, as in Section 8.2. Similarly, $I - P$ is the unique projection on V with kernel V_2 that maps V onto V_1, and which sends $v_1 + v_2$ to v_1 for every $v_1 \in V_1$, $v_2 \in V_2$. Any projection on V may be considered in this way, as before.

Let $\| \cdot \|_V$ be a norm on V. Observe that P is a bounded linear mapping from V into itself with respect to $\| \cdot \|_V$ if and only if there is a nonnegative real number C_2 such that

$$\|v_2\|_V \le C_2 \, \|v_1 + v_2\|_V \tag{8.4.1}$$

for all $v_1 \in V_1$ and $v_2 \in V_2$. Similarly, $I - P$ is bounded on V if and only if there is a $C_1 \ge 0$ such that

$$\|v_1\|_V \le C_1 \, \|v_1 + v_2\|_V \tag{8.4.2}$$

for all $v_1 \in V_1$ and $v_2 \in V_2$. Of course, P is bounded on V if and only if $I - P$ is bounded on V.

Note that (8.4.1) is the same as saying that

$$\|v_2\|_V \le C_2 \, \|v_2 - v_1\|_V \tag{8.4.3}$$

for all $v_1 \in V_1$ and $v_2 \in V_2$. Equivalently, this means that

$$\|v_2\|_V \le C_2 \; \mathrm{dist}(v_2, V_1) \tag{8.4.4}$$

for all $v_2 \in V_2$, where $\mathrm{dist}(v_2, V_1)$ is as in Section 2.11, using the metric on V associated to $\| \cdot \|_V$. Similarly, (8.4.2) is the same as saying that

$$\|v_1\|_V \le C_1 \, \|v_1 - v_2\|_V \tag{8.4.5}$$

for all $v_1 \in V_1$ and $v_2 \in V_2$. This means that

$$\|v_1\|_V \le C_1 \; \mathrm{dist}(v_1, V_2) \tag{8.4.6}$$

for all $v_1 \in V_1$.

Of course,

$$\mathrm{dist}(v, V_1) \le \|v\|_V \tag{8.4.7}$$

for every $v \in V$, because $0 \in V_2$. If (8.4.4) holds with $C_2 = 1$, then we get that

$$\|v_2\|_V = \mathrm{dist}(v_2, V_1) \tag{8.4.8}$$

for all $v_2 \in V_2$. Similarly,

$$\mathrm{dist}(v, V_2) \le \|v\|_V \tag{8.4.9}$$

for every $v \in V$. If (8.4.6) holds with $C_1 = 1$, then

$$\|v_1\|_V = \mathrm{dist}(v_1, V_2) \tag{8.4.10}$$

for all $v_1 \in V_1$.

If $v \in V$, then

$$\mathrm{dist}(v, V_1) = \mathrm{dist}(P(v), V_1), \tag{8.4.11}$$

because $v - P(v) \in V_1$, by construction. Thus (8.4.4) implies that

$$\|P(v)\|_V \le C_2 \; \mathrm{dist}(v, V_1) \tag{8.4.12}$$

for every $v \in V$, because $P(v) \in V_2$. Conversely, this implies (8.4.4), because $P(v_2) = v_2$ for every $v_2 \in V_2$. Similarly,

$$\mathrm{dist}(v, V_2) = \mathrm{dist}(v - P(v), V_2) \tag{8.4.13}$$

for every $v \in V$, because $P(v) \in V_2$. It follows that (8.4.6) holds if and only if

$$\|v - P(v)\|_V \le C_1 \; \mathrm{dist}(v, V_2) \tag{8.4.14}$$

for every $v \in V$, because $v - P(v) \in V_1$.

Observe that

$$\operatorname{dist}(v, V_1) = \operatorname{dist}(P(v), V_1) \le \|P(v)\|_V \tag{8.4.15}$$

for every $v \in V$, using (8.4.11) in the first step, and the analogue of (8.4.7) with $P(v)$ in place of v in the second step. If (8.4.12) holds with $C_2 = 1$, then we have that

$$\|P(v)\|_V = \operatorname{dist}(v, V_1) \tag{8.4.16}$$

for every $v \in V$. Similarly,

$$\operatorname{dist}(v, V_2) = \operatorname{dist}(v - P(v), V_2) \le \|v - P(v)\|_V \tag{8.4.17}$$

for every $v \in V$. If (8.4.14) holds with $C_1 = 1$, then

$$\|v - P(v)\|_V = \operatorname{dist}(v, V_2) \tag{8.4.18}$$

for every $v \in V$.

Suppose for the moment that $\langle \cdot, \cdot \rangle_V$ is an inner product on V, and that $\| \cdot \|_V$ is the corresponding norm on V. If V_1, V_2 are orthogonal to each other with respect to $\langle \cdot, \cdot \rangle_V$, then (8.4.1) and (8.4.2) hold with $C_1 = C_2 = 1$. Conversely, if (8.4.1) holds with $C_2 = 1$, or (8.4.2) holds with $C_1 = 1$, then one can check that V_1 and V_2 are orthogonal to each other with respect to $\langle \cdot, \cdot \rangle_V$. This is similar to an argument mentioned in Section 2.15.

Let v_0 be any nonzero element of V, and let λ_0 be a linear functional on V such that

$$\lambda_0(v_0) = 1. \tag{8.4.19}$$

Put

$$P_0(v) = \lambda_0(v)\, v_0 \tag{8.4.20}$$

for each $v \in V$, which defines a linear mapping from V into itself. Note that

$$P_0(v_0) = v_0, \tag{8.4.21}$$

so that P_0 is a projection from V onto the linear subspace spanned by v_0. We also have that

$$\ker P_0 = \ker \lambda_0, \tag{8.4.22}$$

by construction.

Let $\| \cdot \|_V$ be a norm on V again. If λ_0 is a bounded linear functional on V with respect to $\| \cdot \|_V$, then P_0 is bounded as a linear mapping from V into itself, with

$$\|P_0\|_{op} = \|\lambda_0\|_{V'}\, \|v_0\|_V. \tag{8.4.23}$$

The Hahn–Banach theorem implies that there is such a λ_0 for which the right side of (8.4.23) is equal to 1, as in Section 3.10.

8.5 Symmetric bilinear forms

Let V be a vector space over the real or complex numbers, and let b be a bilinear form on V, which is to say a bilinear mapping from $V \times V$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. We say that b is *symmetric* on V if

$$b(v, w) = b(w, v) \tag{8.5.1}$$

for all $v, w \in V$, and that b is *antisymmetric* on V if

$$b(v, w) = -b(w, v) \tag{8.5.2}$$

for all $v, w \in V$. If b is any bilinear form on V, then

$$(1/2)\,(b(v, w) + b(w, v)) \tag{8.5.3}$$

is a symmetric bilinear form on V, and

$$(1/2)\,(b(v, w) - b(w, v)) \tag{8.5.4}$$

is an antisymmetric bilinear form on V. Of course, $b(v, w)$ is the same as the sum of (8.5.3) and (8.5.4).

If b is a bilinear form on V, then

$$b(v + w, v + w) = b(v, v) + b(v, w) + b(w, v) + b(w, w) \tag{8.5.5}$$

for all $v, w \in V$. Equivalently,

$$(1/2)\,(b(v, w) + b(w, v)) = (1/2)\,(b(v + w, v + w) - b(v, v) - b(w, w)), \tag{8.5.6}$$

which is a *polarization identity.* If b is antisymmetric on V, then

$$b(v, v) = 0 \tag{8.5.7}$$

for every $v \in V$. Conversely, this condition implies that b is antisymmetric on V, because of (8.5.5). If b is symmetric on V, then (8.5.6) shows that b is determined by the values of $b(u, u)$, $u \in V$.

Suppose now that V is a complex vector space, and that b is a sesquilinear form on V, which is to say a sesquilinear mapping from $V \times V$ into $\mathbf{C}$. Note that

$$\overline{b(w, v)} \tag{8.5.8}$$

is a sesquilinear form on V as well. If

$$b(v, w) = \overline{b(w, v)} \tag{8.5.9}$$

for every $v, w \in V$, then b is said to be *Hermitian symmetric*, or a *Hermitian form* on V. If

$$b(v, w) = -\overline{b(w, v)} \tag{8.5.10}$$

for every $v, w \in V$, then one might say that b is *Hermitian antisymmetric* on V, which is equivalent to $i\,b$ being Hermitian symmetric on V. If b is any sesquilinear form on V, then

$$(1/2)\,(b(v,w) + \overline{b(w,v)}) \tag{8.5.11}$$

is Hermitian symmetric on V,

$$(1/2)\,(b(v,w) - \overline{b(w,v)}) \tag{8.5.12}$$

is Hermitian antisymmetric on V, and $b(v,w)$ is the same as the sum of (8.5.11) and (8.5.12).

If b is a sesquilinear form on V, then (8.5.5) holds, as before. We also have that

$$b(v + i\,w, v + i\,w) = b(v,v) - i\,b(v,w) + i\,b(w,v) + b(w,w) \tag{8.5.13}$$

for all $v, w \in V$, which is another *polarization identity.* Using (8.5.5) and (8.5.13), we get that b can be obtained from the values of $b(u,u)$, $u \in V$.

If b is Hermitian symmetric on V, then

$$b(v,v) \in \mathbf{R} \tag{8.5.14}$$

for every $v \in V$. Conversely, one can check that this condition implies that b is Hermitian symmetric on V, using (8.5.5) and (8.5.13). Alternatively, observe that if $v = w$, then (8.5.11) and (8.5.12) are equal to the real and imaginary parts of $b(v,v)$, respectively. If (8.5.12) holds for each $v \in V$, then it follows that (8.5.12) is equal to 0, as in the preceding paragraph. As another approach, (8.5.14) implies that $b(v,w)$ and $\overline{b(v,w)}$ are the same when $v = w$. It follows that $b(v,w)$ and $\overline{b(w,v)}$ are equal to each other, because they are both sesquilinear forms on V, as before. This is the argument used in the proof of Theorem 3 on p13 of [88].

8.6 Self-adjoint linear operators

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space over the real or complex numbers, with the corresponding norm $\| \cdot \|$. A linear mapping T from V into itself is said to be *symmetric* with respect to $\langle \cdot, \cdot \rangle$ if

$$\langle T(v), w \rangle = \langle v, T(w) \rangle \tag{8.6.1}$$

for every $v, w \in V$. If

$$\langle T(v), w \rangle = -\langle v, T(w) \rangle \tag{8.6.2}$$

for every $v, w \in V$, then T is said to be *antisymmetric* on V with respect to $\langle \cdot, \cdot \rangle$. In the complex case, this is the same as saying that $i\,T$ is symmetric on V.

Suppose for the moment that V is a Hilbert space, and that T is a bounded linear mapping from V into itself. If $w \in V$, then there is a unique element $T^*(w)$ of V such that

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \tag{8.6.3}$$

for every $v \in V$, as in Section 3.5. This defines a bounded linear mapping from V into itself, as before. Under these conditions, (8.6.1) is the same as saying that

$$T^* = T. \tag{8.6.4}$$

This means that T is *self-adjoint* on V, or equivalently that T is *Hermitian* when V is complex. Similarly, (8.6.2) is the same as saying that

$$T^* = -T, \tag{8.6.5}$$

so that T is *anti-self-adjoint* on V. In the complex case, this happens exactly when $i\,T$ is self-adjoint on V. Of course, the Hilbert space adjoint defines an algebra involution on $\mathcal{BL}(V)$, so that this terminology corresponds to that in Section 7.5.

If V is any inner product space, and T is any linear mapping from V into itself, then

$$b_T(v, w) = \langle T(v), w \rangle \tag{8.6.6}$$

defines a bilinear form on V in the real case, and a sesquilinear form on V in the complex case. Remember that T is bounded as a linear mapping from V into itself if and only if b_T is bounded as a bilinear or sesquilinear form on V, as appropriate, as in Section 5.15.

If V is a Hilbert space, and T is symmetric or antisymmetric on V, then T is bounded on V. This corresponds to a theorem of Hellinger and Toeplitz on p110 of [162]. More precisely, if V is any inner product space, and T is symmetric or antisymmetric on V, then one can check that

$$\text{the graph of } T \text{ is a closed set in } V \times V. \tag{8.6.7}$$

If V is a Hilbert space, then the closed graph theorem implies that T is bounded, as in Section 7.15.

Alternatively, if V is any inner product space again, and T is symmetric or antisymmetric on V, then one can verify that

$$b_T(v, w) \text{ is separately continuous in each of } v \text{ and } w. \tag{8.6.8}$$

If V is a Hilbert space, then it follows that b_T is bounded as a bilinear or sesquilinear form on V, as in Section 5.14. In the complex case, we may consider b_T as a real-bilinear mapping from $V \times V$ into $\mathbf{C}$, as in Section 5.15.

Another approach to the boundedness of T under these conditions will be mentioned in the next section.

If V is a real inner product space, then

$$b_T(w, v) = \langle T(w), v \rangle = \langle v, T(w) \rangle \tag{8.6.9}$$

for every $v, w \in V$. This implies that T is symmetric or antisymmetric on V if and only if b_T is symmetric or antisymmetric as a bilinear form on V, respectively. In particular, T is antisymmetric on V if and only if

$$\langle T(v), v\rangle = b_T(v, v) = 0 \tag{8.6.10}$$

for every $v \in V$, as in the previous section.

Similarly, if V is a complex inner product space, then

$$\overline{b_T(w, v)} = \overline{\langle T(w), v\rangle} = \langle v, T(w)\rangle \tag{8.6.11}$$

for every $v, w \in V$. It follows that T is symmetric or antisymmetric on V if and only if b_T is Hermitian symmetric or antisymmetric as a sesquilinear form on V, respectively. This means that T is symmetric on V if and only if

$$\langle T(v), v\rangle = b_T(v, v) \in \mathbf{R} \tag{8.6.12}$$

for every $v \in V$, as in the previous section. This corresponds to part (a) of Exercise (2) on p45 of [8], and to the first part of Theorem 2 on p41 of [88].

Suppose that P_W is the orthogonal projection of V onto a linear subspace W, as in Section 8.3. If $v_1, v_2 \in V$, then

$$\langle P_W(v_1), v_2\rangle = \langle P_W(v_1), P_W(v_2)\rangle = \langle v_1, P_W(v_2)\rangle, \tag{8.6.13}$$

using (8.3.1) in both steps. This implies that P_W is symmetric with respect to $\langle \cdot, \cdot \rangle_V$ on V. If V is a Hilbert space, then this means that P_W is self-adjoint on V.

Conversely, let P be a projection on V, as in Section 8.2. Suppose that P is symmetric with respect to $\langle \cdot, \cdot \rangle$ on V, so that

$$\langle P(v), w\rangle = \langle v, P(w)\rangle \tag{8.6.14}$$

for every $v, w \in V$. This implies that

$$\langle v, P(w)\rangle = 0 \tag{8.6.15}$$

when $P(v) = 0$, so that $\ker P$ and $P(V)$ are orthogonal to each other in V. It follows that P is the orthogonal projection of V onto $P(V)$, as in Section 8.3.

8.7 Some related continuity arguments

Let V, W be vector spaces, both real or both complex, and with norms $\|\cdot\|_V$, $\|\cdot\|_W$, respectively. Also let T be a linear mapping from V into W, and remember that T is bounded if and only if it is continuous at 0, as in Section 2.2. It is well known that this happens if and only if for every sequence $\{v_j\}_{j=1}^\infty$ of elements of V that converges to 0 with respect to the metric associated to $\|\cdot\|_V$, we have that

$$\lim_{j\to\infty} T(v_j) = 0 \tag{8.7.1}$$

with respect to the metric on W associated to $\|\cdot\|_W$.

Of course, (8.7.1) implies that

$$\{T(v_j)\}_{j=1}^\infty \text{ is bounded in } W, \tag{8.7.2}$$

with respect to the metric associated to $\|\cdot\|_W$. In fact, it is well known that T is continuous at 0 when (8.7.2) holds for all sequences $\{v_j\}_{j=1}^\infty$ of elements of V that converges to 0. This corresponds to a simplification of part of Theorem 1.32 on p23 of [162].

Remember that T is bounded as a linear mapping from V into W when

$$\|T(v)\|_W \tag{8.7.3}$$

is bounded on a ball in V of positive radius centered at 0, as in Section 2.2. If this does not happen, then one can get a sequence $\{v_j\}_{j=1}^\infty$ of elements of V that converges to 0 such that (8.7.2) does not hold.

Alternatively, let $\{u_j\}_{j=1}^\infty$ be a sequence of elements of V that converges to 0. One can find a sequence $\{r_j\}_{j=1}^\infty$ of positive real numbers such that $r_j \to \infty$ as $j \to \infty$ and

$$\lim_{j\to\infty} r_j\, u_j = 0 \tag{8.7.4}$$

in V. This is a simplification of part (b) of Theorem 1.28 on p21 of [162]. Here we can take

$$r_j = \|u_j\|_V^{-1/2} \tag{8.7.5}$$

when $u_j \neq 0$, and $r_j = j$ otherwise. If (8.7.2) holds with $v_j = r_j\, u_j$, then it follows that

$$T(u_j) = r_j^{-1}\, T(r_j\, u_j) \to 0 \text{ as } j \to \infty \tag{8.7.6}$$

in W.

Note that (8.7.2) holds when

$$\{T(v_j)\}_{j=1}^\infty \text{ converges to 0 weakly in } W, \tag{8.7.7}$$

as in Section 4.9. If (8.7.7) holds for every sequence $\{v_j\}_{j=1}^\infty$ of elements of V that converges to 0, then it follows that T is continuous at 0.

Let Z be a subset of the dual space W' of bounded linear functionals on W, and suppose that

$$Z \text{ separates points in } W. \tag{8.7.8}$$

This means that for each nonzero $w \in W$ there is a $\mu \in Z$ such that

$$\mu(w) \neq 0. \tag{8.7.9}$$

Suppose also that for every $\mu \in Z$,

$$\mu \circ T \text{ is a bounded linear functional on } V, \tag{8.7.10}$$

so that $\mu \circ T$ is continuous on V. Under these conditions, one can check that

$$\text{the graph of } T \text{ is a closed set in } V \times W. \tag{8.7.11}$$

If V and W are Banach spaces, then it follows that T is continuous, by the closed graph theorem, as in Section 7.15.

This basically corresponds to a simplification of Theorem 5.1 on p110 of [162], and its proof. That result considers other types of topological vector spaces, and all continuous linear functionals on W. However, the proof works as well for sets of continuous linear functionals on W that separate points.

More precisely, (8.7.10) is the same as saying that for every sequence $\{v_j\}_{j=1}^{\infty}$ of elements of V that converges to 0,

$$\lim_{j \to \infty} \mu(T(v_j)) = 0. \tag{8.7.12}$$

If we take $Z = W'$ here, then (8.7.12) reduces to (8.7.7), and we can use the earlier argument, without asking V or W to be complete.

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space over the real or complex numbers, and let T be a linear mapping from V itself. Suppose that T is symmetric or antisymmetric with respect to $\langle \cdot, \cdot \rangle$. If $\{v_j\}_{j=1}^{\infty}$ is a sequence of elements of V that converges to 0 with respect to the metric associated to the corresponding norm on V, then

$$\langle T(v_j), w \rangle = \pm \langle v_j, T(w) \rangle \to 0 \text{ as } j \to \infty. \tag{8.7.13}$$

This corresponds to (8.7.12), where $W = V$, and Z is the linear subspace of $V' = W'$ consisting of linear functionals of the form $\mu_w(v) = \langle v, w \rangle$, $w \in V$. Note that Z separates points in V.

If V is a Hilbert space, then it follows that T is continuous, as before. More precisely, every continuous linear functional on V is in Z in this case, so that (8.7.13) implies (8.7.7). Otherwise, one can use the closed graph theorem, as before.

8.8 Bounded symmetric bilinear forms

Let V be a vector space over the real or complex numbers, and let b be a symmetric bilinear form on V. Observe that

$$b(v, w) = (1/4)\,(b(v + w, v + w) - b(v - w, v - w)) \tag{8.8.1}$$

for every $v, w \in V$. This is another type of *polarization identity*, which can also be used to show that b is determined by the values of $b(u, u)$, $u \in V$.

Let $\langle \cdot, \cdot \rangle_V$ be an inner product on V, and let $\| \cdot \|_V$ be the corresponding norm on V. Suppose that

$$|b(u, u)| \le C\, \|u\|_V^2 \tag{8.8.2}$$

for some $C \ge 0$ and every $u \in V$. If $v, w \in V$, then we get that

$$|b(v, w)| \le (C/4)\,(\|v + w\|_V^2 + \|v - w\|_V^2) = (C/2)\,(\|v\|_V^2 + \|w\|_V^2), \tag{8.8.3}$$

using (8.8.1) in the first step, and the parallelogram law in the second step. This implies that

$$|b(v, w)| \le C \tag{8.8.4}$$

when $\|v\|_V, \|w\|_V \le 1$. It follows that

$$|b(v, w)| \le C\, \|v\|_V\, \|w\|_V \tag{8.8.5}$$

for every $v, w \in V$, by standard arguments. Of course, (8.8.5) implies (8.8.2). This corresponds to the analogue of Theorem 3 on p33 of [88] in the real case.

Suppose now that V is a complex vector space, and that b is a sesquilinear form on V. One can check that

$$b(v, w) = (1/4) \sum_{l=0}^{3} i^l\, b(v + i^l\, w, v + i^l\, w) \tag{8.8.6}$$

for every $v, w \in V$. This is another *polarization identity*, which corresponds to Exercise (1) on p45 of [8], and to Theorem 1 on p12 of [88]. This can be used to show that b is determined by the values $b(u, u)$, $u \in V$, as in Theorem 2 on p13 of [88].

If b is Hermitian symmetric on V, then

$$b(u, u) \in \mathbf{R} \tag{8.8.7}$$

for every $u \in V$, as in Section 8.5. If (8.8.7) holds for every $u \in V$, then (8.8.6) implies that

$$\operatorname{Re} b(v, w) = (1/4)\, (b(v + w, v + w) - b(v - w, v - w)) \tag{8.8.8}$$

and

$$\operatorname{Im} b(v, w) = (1/4)\, (b(v + i\, w, v + i\, w) - b(v - i\, w, v - i\, w)) \tag{8.8.9}$$

for every $v, w \in V$. This can be used to get that b is Hermitian symmetric on V. Alternatively, if (8.8.7) holds, then $b(v, w)$ and $\overline{b(w, v)}$ are sesquilinear forms on V that agree when $v = w$, which implies that they agree for all $v, w \in V$, as in the preceding paragraph. This corresponds to the proof of Theorem 3 on p13 of [88].

Let $\langle \cdot, \cdot \rangle_V$ be an inner product on V again, with corresponding norm $\| \cdot \|_V$, and suppose that (8.8.2) holds for some $C \ge 0$ and every $u \in V$. If b is Hermitian symmetric on V, then we can use (8.8.8) to get that

$$\begin{aligned} |\operatorname{Re} b(v, w)| &\le (C/4)\, (\|v + w\|_V^2 + \|v - w\|_V^2) \\ &= (C/2)\, (\|v\|_V^2 + \|w\|_V^2) \end{aligned} \tag{8.8.10}$$

for every $v, w \in V$. If $\|v\|_V, \|w\|_V \le 1$, then it follows that

$$|\operatorname{Re} b(v, w)| \le C. \tag{8.8.11}$$

If a is a complex number with $|a| \le 1$, then we get that

$$|\operatorname{Re}(a\, b(v, w))| = |\operatorname{Re} b(a\, v, w)| \le C \tag{8.8.12}$$

when $\|v\|_V, \|w\|_V \le 1$. This implies that (8.8.4) holds. One can use this to get (8.8.5), as before. This corresponds to Theorem 3 on p33 of [88].

If b is not necessarily Hermitian symmetric on V, then we can still use (8.8.6) to get that

$$|b(v, w)| \le (1/4) \sum_{l=0}^{3} |b(v + i^l\, w, v + i^l\, w)| \tag{8.8.13}$$

for every $v, w \in V$. Thus (8.8.2) implies that

$$|b(v, w)| \le (C/4) \sum_{l=0}^{3} \|v + i^l\, w\|_V^2 \tag{8.8.14}$$

for every $v, w \in V$. The sum on the right is the same as

$$(\|v + w\|_V^2 + \|v - w\|_V^2) + (\|v + i\, w\|_V^2 + \|v - i\, w\|_V^2), \tag{8.8.15}$$

which is equal to

$$4\, \|v\|_V^2 + 4\, \|w\|_V^2, \tag{8.8.16}$$

by the parallelogram law. Combining this with (8.8.14), we get that

$$|b(v, w)| \le C\, (\|v\|_V^2 + \|w\|_V^2) \tag{8.8.17}$$

for every $v, w \in V$.

It follows that

$$|b(v, w)| \le 2\, C \tag{8.8.18}$$

when $\|v\|_V, \|w\|_V \le 1$. As before, one can use this and standard arguments to get that

$$|b(v, w)| \le 2\, C\, \|v\|_V\, \|w\|_V \tag{8.8.19}$$

for every $v, w \in V$. This corresponds to Theorem 2 on p33 of [88].

8.9 Bounded symmetric linear mappings

Let $(V, \langle \cdot, \cdot \rangle_V)$ be an inner product space over the real or complex numbers, with the corresponding norm $\| \cdot \|_V$. Also let T be a bounded linear mapping from V into itself with respect to $\| \cdot \|_V$ that is symmetric with respect to $\langle \cdot, \cdot \rangle_V$, as in Section 8.6. Thus

$$b_T(v, w) = \langle T(v), w \rangle_V \tag{8.9.1}$$

is a bounded symmetric bilinear form on V in the real case, and a bounded Hermitian symmetric sesquilinear form on V in the complex case.

Suppose that $V \ne \{0\}$, and consider

$$\sup\{|b_T(u, u)| : u \in V,\ \|u\|_V = 1\}. \tag{8.9.2}$$

This is the same as the smallest nonnegative real number C such that (8.8.2) holds, with $b = b_T$. Similarly,

$$\sup\{|b_T(v, w)| : v, w \in V,\ \|v\|_V = \|w\|_V = 1\} \tag{8.9.3}$$

is the same as the smallest nonnegative real number C such that (8.8.5) holds, with $b = b_T$. Clearly (8.9.2) is less than or equal to (8.9.3), and in fact they are equal to each other, as in the previous section.

We also have that (8.9.3) is equal to the operator norm of T on V. This is very similar to a remark at the beginning of Section 3.5. It follows that (8.9.2) is equal to $\|T\|_{op}$ too. This corresponds to the second part of Theorem 2 on p41 of [88].

Put

$$\alpha(T) = \inf\{\langle T(v), v\rangle_V : v \in V,\ \|v\|_V = 1\} \tag{8.9.4}$$

and

$$\beta(T) = \sup\{\langle T(v), v\rangle_V : v \in V,\ \|v\|_V = 1\}. \tag{8.9.5}$$

These are real numbers with $\alpha(T) \le \beta(T)$ and

$$\max(-\alpha(T), \beta(T)) = \sup\{|\langle T(v), v\rangle_V| : v \in V,\ \|v\|_V = 1\}. \tag{8.9.6}$$

This means that

$$\max(-\alpha(T), \beta(T)) = \|T\|_{op}, \tag{8.9.7}$$

as in the previous paragraph.

Note that

$$\alpha(-T) = -\beta(T),\ \beta(-T) = -\alpha(T). \tag{8.9.8}$$

If $a \in \mathbf{R}$, then

$$\alpha(a\, I + T) = a + \alpha(T) \tag{8.9.9}$$

and

$$\beta(a\, I + T) = a + \beta(T), \tag{8.9.10}$$

where $I = I_V$ is the identity mapping on V.

8.10 Nonnegative bilinear forms and operators

Let V be a vector space over the real numbers, and let b be a symmetric bilinear form on V. If

$$b(v, v) \ge 0 \tag{8.10.1}$$

for every $v \in V$, then b is said to be *nonnegative* on V. Under these conditions, it is well known that

$$|b(v, w)| \le b(v, v)^{1/2}\, b(w, w)^{1/2} \tag{8.10.2}$$

for all $v, w \in V$, which is the analogue of the Cauchy–Schwarz inequality in this case.

Similarly, let V be a vector space over the complex numbers, and let b be a Hermitian symmetric sesquilinear form on V. Thus $b(v, v) \in \mathbf{R}$ for every $v \in V$, as in Section 8.5. If (8.10.1) holds for every $v \in V$, then b is said to be *nonnegative* on V. It is well known that (8.10.2) holds in this case too, which is another version of the Cauchy–Schwarz inequality. This corresponds to Proposition 12.1 on p114 of [191].

To get (8.10.2) in the complex case, we start by observing that

$$0 \le b(v + a\,w, v + a\,w) = b(v,v) + 2\ \mathrm{Re}(\overline{a}\,b(v,w)) + |a|^2\,b(w,w) \tag{8.10.3}$$

for every $v, w \in V$ and $a \in \mathbf{C}$. This implies that

$$2\,t\,|b(v,w)| \le b(v,v) + t^2\,b(w,w) \tag{8.10.4}$$

for all nonnegative real numbers t, using a suitable choice of a with $|a| = t$. This also works in the real case, with some simplifications. One can get (8.10.2) from (8.10.4) in much the same way as when b is an inner product, except that one should be a bit more careful about $b(v,v)$ or $b(w,w)$ being 0 even if v or w is not zero, as appropriate.

Now let $(V, \langle\cdot,\cdot\rangle_V)$ be a real or complex inner product space, with the corresponding norm $\|\cdot\|_V$. Also let T be a linear mapping from V into itself that is symmetric with respect to $\langle\cdot,\cdot\rangle_V$, as in Section 8.6. If

$$\langle T(v), v\rangle_V \ge 0 \tag{8.10.5}$$

for every $v \in V$, then T is said to be *nonnegative* with respect to $\langle\cdot,\cdot\rangle_V$ on V. Of course, this is the same as saying that

$$b_T(v,w) = \langle T(v), w\rangle_V \tag{8.10.6}$$

is nonnegative as a symmetric bilinear form on V in the real case, or as a Hermitian symmetric sesquilinear form on V in the complex case. In both cases, we get that

$$|\langle T(v), w\rangle_V| \le \langle T(v), v\rangle_V^{1/2}\,\langle T(w), w\rangle_V^{1/2} \tag{8.10.7}$$

for every $v, w \in V$, as in (8.10.2).

If T is any bounded linear mapping from V into itself with respect to $\|\cdot\|_V$, then

$$|\langle T(w), w\rangle_V| \le \|T(w)\|_V\,\|w\|_V \le \|T\|_{op}\,\|w\|_V^2 \tag{8.10.8}$$

for every $w \in V$. If T is symmetric and nonnegative on V as well, then

$$|\langle T(v), w\rangle_V| \le \|T\|_{op}^{1/2}\,\langle T(v), v\rangle^{1/2}\,\|w\|_V \tag{8.10.9}$$

for every $v, w \in V$, by (8.10.7). This implies that

$$\|T(v)\|_V \le \|T\|_{op}^{1/2}\,\langle T(v), v\rangle_V^{1/2} \tag{8.10.10}$$

for every $v \in V$.

Suppose that

$$c_0\,\|v\|_V \le \|T(v)\|_V \tag{8.10.11}$$

for some $c_0 > 0$ and every $v \in V$. If T is invertible on V, with bounded inverse, then this holds with $c_0 = 1/\|T^{-1}\|_{op}$. If T is bounded, symmetric, and nonnegative on V too, then we can combine (8.10.10) and (8.10.11) to get that

$$c_0^2 \, \|v\|_V^2 \le \|T\|_{op} \, \langle T(v), v \rangle_V \tag{8.10.12}$$

for every $v \in V$.

Suppose that $V \neq \{0\}$, and note that T is nonnegative on V exactly when

$$\alpha(T) \ge 0, \tag{8.10.13}$$

where the left side is as in (8.9.4). If T is invertible on V, with bounded inverse, then

$$\alpha(T) > 0, \tag{8.10.14}$$

by (8.10.12). This means that T does not have a bounded inverse on V when

$$\alpha(T) = 0. \tag{8.10.15}$$

Suppose that V is a Hilbert space, and let $(W, \langle \cdot, \cdot \rangle_W)$ be another real or complex Hilbert space, as appropriate, with associated norm $\| \cdot \|_W$. Also let R be a bounded linear mapping from V into W, so that the adjoint R^* of R is a bounded linear mapping from W into V, as in Section 3.5. Thus $R^* \circ R$ is a bounded linear mapping from V into itself, which is self-adjoint. If $v \in V$, then

$$\langle (R^* \circ R)(v), v \rangle_V = \langle R(v), R(v) \rangle_W = \|R(v)\|_W^2, \tag{8.10.16}$$

so that $R^* \circ R$ is nonnegative on V. Similarly, if $V = W$ is an inner product space, and R is a linear mapping from V into itself that is symmetric with respect to $\langle \cdot, \cdot \rangle_V$, then $R \circ R$ is symmetric and nonnegative with respect to $\langle \cdot, \cdot \rangle$.

8.11 Some remarks about normal operators

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a Hilbert space over the real or complex numbers, with corresponding norm $\| \cdot \|_V$. If a bounded linear mapping T from V into itself commutes with its adjoint, then one may say that T is *normal*, as in Section 7.5. This terminology is perhaps most commonly used in the complex case, as before.

If T is a bounded linear mapping from V into itself and $v \in V$, then

$$\|T(v)\|_V^2 = \langle T(v), T(v) \rangle_V = \langle T^*(T(v)), v \rangle_V \tag{8.11.1}$$

and

$$\|T^*(v)\|_V^2 = \langle T^*(v), T^*(v) \rangle_V = \langle T(T^*(v)), v \rangle_V. \tag{8.11.2}$$

If T is normal, then it follows that

$$\|T(v)\|_V = \|T^*(v)\|_V. \tag{8.11.3}$$

In particular, this means that

$$\ker T = \ker T^*, \tag{8.11.4}$$

as in part (b) of Theorem 12.12 on p298 of [162]. Thus T is injective if and only if T^* is injective under these conditions.

Conversely, if (8.11.3) holds for every $v \in V$, then

$$\langle T^*(T(v)), v\rangle_V = \langle T(T^*(v)), v\rangle_V \tag{8.11.5}$$

for every $v \in V$. This implies that

$$\langle T^*(T(v)), w\rangle_V = \langle T(T^*(v)), w\rangle_V \tag{8.11.6}$$

for every $v, w \in V$, using polarization identities, as in Sections 8.5 or 8.8. More precisely, in the real case, this also uses the fact that $T^* \circ T$ and $T \circ T^*$ are self-adjoint on V, to get that both sides of (8.11.6) are symmetric bilinear forms on V. It is easy to see that T and T^* commute on V, using (8.11.6). This corresponds to Theorem 1 on p42 of [88], and to part (a) of Theorem 12.12 on p298 of [162].

If T is any bounded linear mapping from V into itself, then one can check that

$$\ker T^* = T(V)^\perp \tag{8.11.7}$$

as in Theorem 12.10 on p298 of [162]. If T is normal, then we get that

$$\ker T = T(V)^\perp, \tag{8.11.8}$$

because of (8.11.4).

Suppose that T is a bounded linear mapping from V into itself such that

$$c\,\|v\|_V \le \|T(v)\|_V \tag{8.11.9}$$

for some $c > 0$ and all $v \in V$. This implies that $T(V)$ is a closed set in V, because V is complete with respect to the metric associated to $\|\cdot\|_V$, as in Section 7.11. If T is normal, then we also have that $T(V)$ is dense in V, because of (8.11.8). This means that T is a one-to-one linear mapping from V onto itself wtih bounded inverse in this case.

Suppose now that V is complex, and that T is a bounded self-adjoint linear mapping from V into itself. Also let $\lambda \in \mathbf{C}$ be given, and observe that

$$(\lambda\, I - T)^* = \overline{\lambda}\, I - T^* = \overline{\lambda}\, I - T, \tag{8.11.10}$$

where $I = I_V$ is the identity mapping on V, as usual. In particular, $\lambda\, I - T$ is normal on V.

If $v \in V$, then

$$\langle(\lambda\, I - T)(v), v\rangle_V = \lambda\,\langle v, v\rangle_V - \langle T(v), v\rangle_V = \lambda\,\|v\|_V^2 - \langle T(v), v\rangle_V. \tag{8.11.11}$$

Note that

$$\langle T(v), v\rangle_V \in \mathbf{R}, \tag{8.11.12}$$

because T is self-adjoint on V. This implies that

$$\operatorname{Im}\langle(\lambda\, I - T)(v), v\rangle_V = (\operatorname{Im}\lambda)\, \|v\|_V^2. \tag{8.11.13}$$

It follows that

$$|\operatorname{Im}\lambda|\, \|v\|_V^2 \le \|(\lambda\, I - T)(v)\|_V\, \|v\|_V, \tag{8.11.14}$$

by the Cauchy–Schwarz inequality. This means that

$$|\operatorname{Im}\lambda|\, \|v\|_V \le \|(\lambda\, I - T)(v)\|_V, \tag{8.11.15}$$

which is trivial when $v = 0$.

If $\operatorname{Im}\lambda \ne 0$, then we get that $\lambda\, I - T$ is a one-to-one linear mapping from V onto itself with bounded inverse, as before. This shows that the spectrum of T is contained in the real line, with respect to the algebra of bounded linear mappings from V into itself. This corresponds to Theorem 1 on p54 of [88].

8.12 Positivity and invertibility

Let $(V, \langle\cdot,\cdot\rangle_V)$ be a Hilbert space over the real or complex numbers, with corresponding norm $\|\cdot\|_V$, and let T be a bounded self-adjoint linear mapping from V into itself. Suppose for the moment that there is a positive real number c such that

$$c\, \|v\|_V^2 \le \langle T(v), v\rangle_V \tag{8.12.1}$$

for every $v \in V$. This implies that

$$c\, \|v\|_V^2 \le \|T(v)\|_V\, \|v\|_V \tag{8.12.2}$$

for every $v \in V$, by the Cauchy–Schwarz inequality. It follows that T is a one-to-one mapping from V onto itself, with bounded inverse, as in the previous section.

Suppose from now on in this section that $V \ne \{0\}$, and let $\alpha(T)$ be as in Section 8.9. If

$$\alpha(T) > 0, \tag{8.12.3}$$

then T has a bounded inverse on V, as in the preceding paragraph.

Suppose for the moment again that T is nonnegative on V, and let a be a positive real number. Thus $a\, I + T$ is self-adjoint on V, and

$$\alpha(a\, I + T) = a + \alpha(T) \ge a > 0. \tag{8.12.4}$$

This implies that $a\, I + T$ is a one-to-one linear mapping from V onto itself with bounded inverse, as before. It follows that the spectrum of T is contained in the set of nonnegative real numbers. This uses the fact that the spectrum of T is contained in the real line in the complex case, as in the previous section.

Let T be any bounded self-adjoint linear mapping from V into itself again. If λ is a real number such that

$$\lambda < \alpha(T), \tag{8.12.5}$$

then

$$\alpha(T - \lambda\, I) = \alpha(T) - \lambda > 0, \tag{8.12.6}$$

so that $T - \lambda\, I$ has a bounded inverse on V. Of course, this is the same as saying that $\lambda\, I - T$ has a bounded inverse on V.

Similarly, suppose that $\lambda \in \mathbf{R}$ satisfies

$$\lambda > \beta(T), \tag{8.12.7}$$

where $\beta(T)$ is as in Section 8.9. This implies that

$$\alpha(\lambda\, I - T) = \lambda - \beta(T) > 0, \tag{8.12.8}$$

so that $\lambda\, I - T$ has a bounded inverse on V.

Let $\sigma(T) = \sigma_{\mathcal{BL}(V)}(T)$ be the spectrum of T with respect to the algebra of bounded linear mappings from V into itself, as in Section 6.8. Using the remarks in the previous two paragraphs, we get that

$$\sigma(T) \subseteq [\alpha(T), \beta(T)]. \tag{8.12.9}$$

This also uses the fact that $\sigma(T) \subseteq \mathbf{R}$ in the complex case, as before.

Observe that

$$\alpha(T - \alpha(T)\, I) = \alpha(T) - \alpha(T) = 0, \tag{8.12.10}$$

so that $T - \alpha(T)\, I$ does not have a bounded inverse on V, as in Section 8.10. Equivalently, this means that $\alpha(T)\, I - T$ does not have a bounded inverse on V, so that

$$\alpha(T) \in \sigma(T). \tag{8.12.11}$$

Similarly,

$$\alpha(\beta(T)\, I - T) = \beta(T) - \beta(T) = 0, \tag{8.12.12}$$

so that $\beta(T)\, I - T$ does not have a bounded inverse on V, and thus

$$\beta(T) \in \sigma(T). \tag{8.12.13}$$

In particular, we get that

$$\max\{|\lambda| : \lambda \in \sigma(T)\} = \max(-\alpha(T), \beta(T)). \tag{8.12.14}$$

This implies that

$$\max\{|\lambda| : \lambda \in \sigma(T)\} = \|T\|_{op}, \tag{8.12.15}$$

as in Section 8.9. This corresponds to Theorem 2 on p55 of [88].

8.13 Polynomials and associative algebras

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. Suppose that

$$p(t) = \sum_{j=0}^{n} c_j\, t^j \tag{8.13.1}$$

is a polynomial in a single variable with real or complex coefficients, as appropriate. If $x \in \mathcal{A}$, then

$$\widetilde{p}(x) = \sum_{j=0}^{n} c_j\, x^j \tag{8.13.2}$$

defines an element of $\mathcal{A}$, where x^0 is interpreted as being equal to $e_{\mathcal{A}}$, as usual. This element is often simply denoted $p(x)$, but it is sometimes helpful to use different notation, as on p243 of [162].

The space of all such polynomials p is a commutative associative algebra over the real or complex numbers, as appropriate, with respect to the usual definition of multiplication. It is easy to see that

$$p \mapsto \widetilde{p}(x) \tag{8.13.3}$$

defines a homomorphism from this algebra into $\mathcal{A}$.

Let us check that

$$p(\sigma_{\mathcal{A}}(x)) \subseteq \sigma_{\mathcal{A}}(\widetilde{p}(x)), \tag{8.13.4}$$

where $\sigma_{\mathcal{A}}(y)$ is the spectrum of $y \in \mathcal{A}$, as in Section 6.8. This is at least part of a version of the *spectral mapping theorem.* Let $\lambda \in \sigma_{\mathcal{A}}(x)$ be given, and let us verify that

$$p(\lambda) \in \sigma_{\mathcal{A}}(\widetilde{p}(x)). \tag{8.13.5}$$

It is well known and not difficult to show that

$$p(t) - p(\lambda) = (t - \lambda)\, q(t) \tag{8.13.6}$$

for some polynomial $q(t)$ with real or complex coefficients, as appropriate. This implies that

$$\widetilde{p}(x) - p(\lambda)\, e_{\mathcal{A}} = (x - \lambda\, e_{\mathcal{A}})\, \widetilde{q}(x). \tag{8.13.7}$$

Note that $x - \lambda\, e_{\mathcal{A}}$ and $\widetilde{q}(x)$ commute with each other. If the left side of (8.13.7) is invertible in $\mathcal{A}$, then each of the factors on the right side is invertible as well, as in Section 6.13. In particular, if $x - \lambda\, e_{\mathcal{A}}$ is not invertible, then $\widetilde{p}(x) - p(\lambda)\, e_{\mathcal{A}}$ is not invertible.

We would like to have that

$$\sigma_{\mathcal{A}}(\widetilde{p}(x)) \subseteq p(\sigma_{\mathcal{A}}(x)) \tag{8.13.8}$$

under suitable conditions, which would imply that

$$p(\sigma_{\mathcal{A}}(x)) = \sigma_{\mathcal{A}}(\widetilde{p}(x)). \tag{8.13.9}$$

If p is a constant, then (8.13.8) holds when $\sigma_{\mathcal{A}}(x) \neq \emptyset$. In particular, this works when $\mathcal{A}$ is a complex Banach algebra, as in Section 6.8.

Suppose now that p is not a constant. Let $\mu \in \mathbf{C}$ be given, and observe that

$$p(t) - \mu = c \prod_{j=1}^{n} (t - \lambda_j) \tag{8.13.10}$$

for some complex numbers c and $\lambda_1, \ldots, \lambda_n$, with $c \neq 0$, by the fundamental theorem of algebra. Suppose that $\mathcal{A}$ is complex, so that

$$\widetilde{p}(x) - \mu\, e_{\mathcal{A}} = c \prod_{j=1}^{n} (x - \lambda_j\, e_{\mathcal{A}}). \tag{8.13.11}$$

If $\mu \in \sigma_{\mathcal{A}}(\widetilde{p}(x))$, then it follows that $\lambda_j \in \sigma_{\mathcal{A}}(x)$ for some j. This means that $\mu = p(\lambda_j)$ is an element of $p(\sigma_{\mathcal{A}}(x))$.

Suppose that $\mathcal{A}$ is real, so that $p(t)$ has real coefficients, and let $\mu \in \mathbf{R}$ be given. In this case,

$$p(t) - \mu = c \prod_{j=1}^{r} (t - \lambda_j) \prod_{l=1}^{m} ((t - a_l)^2 + b_l^2) \tag{8.13.12}$$

for some real numbers c, $\lambda_1, \ldots, \lambda_l$, $a_1, \ldots, a_m$, and $b_1, \ldots, b_m$, with $b_1, \ldots, b_m$ and c not equal to 0. This is the same as (8.13.10), with $n = r + 2\,m$, arranged so that the previous λ_j's in $\mathbf{R}$ are listed first, and the remaining λ_j's are of the form $a_l \pm b_l\, i$. Thus

$$\widetilde{p}(x) - \mu\, e_{\mathcal{A}} = c \prod_{j=1}^{r} (x - \lambda_j\, e_{\mathcal{A}}) \prod_{l=1}^{m} ((x - a_l\, e_{\mathcal{A}})^2 + b_l^2\, e_{\mathcal{A}}). \tag{8.13.13}$$

Suppose that

$$(x - a\, e_{\mathcal{A}})^2 + b^2\, e_{\mathcal{A}} \tag{8.13.14}$$

is invertible in $\mathcal{A}$ for all $a, b \in \mathbf{R}$ with $b \neq 0$. If $\mu \in \sigma_{\mathcal{A}}(\widetilde{p}(x))$, then we get that $\lambda_j \in \sigma_{\mathcal{A}}(x)$ for some j. This implies that $\mu = p(\lambda_j) \in p(\sigma_{\mathcal{A}}(x))$, as before.

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real Hilbert space, with associated norm $\| \cdot \|_V$, and with $V \neq \{0\}$. If $\mathcal{A} = \mathcal{B}(V)$ and $x \in \mathcal{BL}(V)$ is self-adjoint, then $\sigma_{\mathcal{A}}(x) \neq \emptyset$, and (8.13.14) is invertible in $\mathcal{A}$ for all $a, b \in \mathbf{R}$ with $b \neq 0$, as in the previous section. This means that (8.13.9) holds for all polynomials p with real coefficients, as before.

If $\mathcal{A}$ is a complex Banach algebra and $x \in \mathcal{A}$, then one can define $\widetilde{f}(x) \in \mathcal{A}$ for complex-valued holomorphic functions f defined on open subsets of the complex plane that contain $\sigma_{\mathcal{A}}(x)$. This is discussed in Section 1.12 of [8], and beginning on p240 of [162]. Part (b) of Theorem 10.28 on p244 of [162] is the spectral mapping theorem for functions of this type.

8.14 Polynomials of self-adjoint operators

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space, with associated norm $\| \cdot \|_V$, and with $V \neq \{0\}$. Also let T be a bounded self-adjoint linear mapping from V into itself, and let p be a polynomial in a single variable with real coefficients. Thus $\widetilde{p}(T)$ may be defined as a bounded linear mapping from V into itself as in the previous section, and it is easy to see that $\widetilde{p}(T)$ is self-adjoint as well. As in (8.13.9), we have that

$$\sigma(\widetilde{p}(T)) = p(\sigma(T)), \tag{8.14.1}$$

where $\sigma(\cdot) = \sigma_{\mathcal{BL}(V)}(\cdot)$ is the spectrum with respect to the algebra of bounded linear mappings from V into itself.

Remember that

$$\|\widetilde{p}(T)\|_{op} = \max\{|\mu| : \mu \in \sigma(\widetilde{p}(T))\}, \tag{8.14.2}$$

as in (8.12.15). This means that

$$\|\widetilde{p}(T)\|_{op} = \max\{|\mu| : \mu \in p(\sigma(T))\}, \tag{8.14.3}$$

because of (8.14.1). Of course, this is the same as saying that

$$\|\widetilde{p}(T)\|_{op} = \max\{|p(\lambda)| : \lambda \in \sigma(T)\}. \tag{8.14.4}$$

This corresponds to Theorem 3 on p55 of [88].

Let $C(\sigma(T), \mathbf{R})$ be the space of continuous real-valued functions on $\sigma(T)$, with respect to the restriction of the standard Euclidean metric on $\mathbf{R}$ to $\sigma(T)$. If $f \in C(\sigma(T), \mathbf{R})$, then the supremum norm of f may be denoted $\|f\|_{sup}$, as usual, or $\|f\|_{sup,\sigma(T)}$, to indicate the role of $\sigma(T)$. We would like to define a mapping

$$f \mapsto \widetilde{f}(T) \tag{8.14.5}$$

from $C(\sigma(T), \mathbf{R})$ into $\mathcal{BL}(V)$ that agrees with the previous definition of $\widetilde{f}(T)$ when f is the restriction to $\sigma(T)$ of a polynomial with real coefficients, and with other nice properties.

Remember that $\mathcal{BL}(V)$ is an associative algebra over the real or complex numbers, depending on whether V is real or complex. We may consider $\mathcal{BL}(V)$ as an associative algebra over the real numbers in both cases. The mapping (8.14.5) is a homomorphism from $C(\sigma(T), \mathbf{R})$ into $\mathcal{BL}(V)$, as associative algebras over $\mathbf{R}$, and an isometry with respect to the supremum norm on $C(\sigma(T), \mathbf{R})$ and the operator norm on $\mathcal{BL}(V)$.

The continuous real-valued functions on $\sigma(T)$ obtained from restrictions of polynomials with real coefficients to $\sigma(T)$ form a subalgebra of $C(\sigma(T), \mathbf{R})$. This subalgebra is dense in $C(\sigma(T), \mathbf{R})$ with respect to the supremum metric, by the Stone–Weierstrass theorem. More precisely, it is well known that a continuous real-valued function on a closed set in $\mathbf{R}$ can be extended to a continuous function on $\mathbf{R}$. One could use this to reduce to Weierstrass' approximation theorem for continuous functions on a closed interval in $\mathbf{R}$.

This leads to a unique bounded linear mapping from $C(\sigma(T), \mathbf{R})$ into $\mathcal{BL}(V)$, as a Banach space over the real numbers, as in Section 2.2. This extension will be expressed as in (8.14.5), to be compatible with the previous notation for polynomials. One can check that this extension is an algebra homomorphism, because of the analogous property for polynomials. Similarly, this extension is an isometry, because of (8.14.4).

Note that

$$\widetilde{f}(T) \text{ is self-adjoint} \tag{8.14.6}$$

for every $f \in C(\sigma(T), \mathbf{R})$, because of the analogous property of polynomials. If

$$f(\lambda) \neq 0 \tag{8.14.7}$$

for each $\lambda \in \sigma(T)$, then $g = 1/f$ is a continuous real-valued function on $\sigma(T)$ too. This implies that

$$\widetilde{f}(T) \text{ is invertible in } \mathcal{BL}(V), \tag{8.14.8}$$

with inverse equal to $\widetilde{g}(T)$.

8.15 Some functions on associative algebras

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. If $a \in \mathcal{A}$ and p is a polynomial in a single variable with real or complex coefficients, as appropriate, then $\widetilde{p}(a)$ may be defined as an element of $\mathcal{A}$ as in Section 8.13. We have also seen some other situations in which it may be reasonable to define $\widetilde{p}(a)$ for other real or complex-valued functions p defined on or around $\sigma_{\mathcal{A}}(a)$. We would like to consider some other aspects of this in this section.

If $\widetilde{p}(a) \in \mathcal{A}$ can be defined for some suitable functions p, then a related question is whether $\widetilde{p}(a)$ depends only on the restriction of p to $\sigma_{\mathcal{A}}(a)$. To put it another way, if

$$p = 0 \text{ on } \sigma_{\mathcal{A}}(a), \tag{8.15.1}$$

then do we have that

$$\widetilde{p}(a) = 0? \tag{8.15.2}$$

If a is a nilpotent element of $\mathcal{A}$, for instance, then it is easy to see that

$$\sigma_{\mathcal{A}}(a) = \{0\}, \tag{8.15.3}$$

using a remark in Section 6.13. If p is a polynomial with real or complex coefficients, then (8.15.1) is the same as saying that $p(0) = 0$ in this case. If $p(t) = t$, then $\widetilde{p}(a) = a$, so that (8.15.2) would mean that $a = 0$.

Let X be a nonempty set, and suppose for the moment that $\mathcal{A}$ is the algebra of all real or complex-valued functions on X. If $a \in \mathcal{A}$, then

$$\sigma_{\mathcal{A}}(a) = a(X), \tag{8.15.4}$$

as mentioned in Section 6.8. If p is a real or complex-valued function on $a(X)$, as appropriate, then

$$\widetilde{p}(a) = p \circ a \tag{8.15.5}$$

defines another real or complex-valued function on X. Note that this agrees with the definition of $\widetilde{p}(a)$ in Section 8.13 when p is a polynomial.

If $\mathcal{A}$ is the algebra of bounded real or complex-valued functions on X and $a \in \mathcal{A}$, then $\sigma_{\mathcal{A}}(a)$ is the closure $\overline{a(X)}$ of $a(X)$ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 6.8. If p is a bounded real or complex-valued function on $a(X)$, then (8.15.5) defines a bounded real or complex-valued function on X, as appropriate.

If X is a metric space or a topological space, $\mathcal{A}$ is the algebra of continuous real or complex-valued functions on X, and $a \in \mathcal{A}$, then (8.15.4) holds again. If p is a continuous real or complex-valued function on $a(X)$, with respect to the restriction of the standard Euclidean metric on $\mathbf{R}$ or $\mathbf{C}$, then (8.15.5) defines a continuous real or complex-valued function on X, as appropriate.

Similarly, if $\mathcal{A}$ is an algebra of functions on some space with some additional properties, then one may need additional properties of p to get that (8.15.5) is an element of $\mathcal{A}$.

Let V be a vector space over the real or complex numbers, and let T be a linear mapping from V into itself. In some cases, we may be able to define $\widetilde{p}(T)$ for suitable functions p on or around the spectrum of T. Some basic examples of this are given by multiplication operators on various spaces of functions.

If $\mathcal{A}$ is a Banach algebra and p is defined by a power series, then $\widetilde{p}(a)$ can be defined for suitable $a \in \mathcal{A}$, as in Section 9.14.

Chapter 9

Algebras, operators, and power series

9.1 Eigenvalues and eigenvectors

Let V be a vector space over the real or complex numbers, and let T be a linear mapping from V into itself. As usual, $v \in V$ is said to be an *eigenvector* of T with *eigenvalue* $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, when

$$T(v) = \lambda\, v. \tag{9.1.1}$$

However, $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, is normally considered to be an *eigenvalue* of T only when it is the eigenvalue associated to a nonzero eigenvector of T.

The set $\sigma_p(T)$ of all eigenvalues of T in this sense is called the *point spectrum* of T. Equivalently, this is the set of $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that

$$\ker(\lambda\, I - T) \neq \{0\}. \tag{9.1.2}$$

Note that

$$\sigma_p(T) \subseteq \sigma_{\mathcal{L}(V)}(T), \tag{9.1.3}$$

where $\sigma_{\mathcal{L}(V)}(T)$ is the spectrum of T with respect to the algebra $\mathcal{L}(V)$ of all linear mappings from V into itself. If V has finite dimension, then

$$\sigma_p(T) = \sigma_{\mathcal{L}(V)}(T). \tag{9.1.4}$$

Let q be a polynomial in a single variable with real or complex coefficients, as appropriate. If $v \in V$ is an eigenvector of T with eigenvalue λ, then

$$(\widetilde{q}(T))(v) = q(\lambda)\, v. \tag{9.1.5}$$

This implies that

$$q(\lambda) \in \sigma_p(\widetilde{q}(T)) \tag{9.1.6}$$

when $v \neq 0$, so that

$$q(\sigma_p(T)) \subseteq \sigma_p(\widetilde{q}(T)). \tag{9.1.7}$$

If q is not a constant, then we would like to have that

$$\sigma_p(\widetilde{q}(T)) \subseteq q(\sigma_p(T)) \tag{9.1.8}$$

under suitable conditions. In the complex case, this can be obtained from the same type of argument as in Section 8.13.

More precisely, if $\mu \in \mathbf{C}$, then one can use the fundamental theorem of algebra to express $q(t) - \mu$ as the product of a nonzero complex number and finitely many linear factors $t - \lambda_j$. If

$$\mu \in \sigma_p(\widetilde{q}(T)), \tag{9.1.9}$$

then $\widetilde{q}(T) - \mu\, I$ is not injective on V. This would imply that one of the corresponding factors $T - \lambda_j\, I$ of $\widetilde{q}(T) - \mu\, I$ is not injective. This means that λ_j is an eigenvalue of T for some j. It follows that

$$\mu \in q(\sigma_p(T)), \tag{9.1.10}$$

because $q(\lambda_j) = \mu$ for each j, by construction.

In the real case, one can use the same type of argument as before as well. This works when

$$(T - a\, I)^2 + b^2\, I \tag{9.1.11}$$

is injective on V for all $a, b \in \mathbf{R}$ with $b \neq 0$. This condition holds when $(V, \langle \cdot, \cdot \rangle_V)$ is a real inner product space, and T is symmetric with respect to $\langle \cdot, \cdot \rangle_V$.

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a complex inner product space with corresponding norm $\| \cdot \|_V$, and suppose that T is symmetric with respect to $\langle \cdot, \cdot \rangle_V$. If $v \in V$ is an eigenvector of T with eigenvalue $\lambda \in \mathbf{C}$, then

$$\lambda\, \|v\|_V^2 = \langle T(v), v \rangle_V = \langle v, T(v) \rangle_V = \overline{\lambda}\, \|v\|_V^2. \tag{9.1.12}$$

This implies that $\lambda \in \mathbf{R}$ when $v \neq 0$.

9.2 Eigenvalues of bounded linear mappings

Let V be a real or complex vector space with a norm $\| \cdot \|_V$, let T be a bounded linear mapping from V into itself, and let $\sigma_{\mathcal{BL}(V)}(T)$ be the spectrum of T with respect to the algebra $\mathcal{BL}(V)$ of all bounded linear mappings from V into itself. Note that

$$\sigma_{\mathcal{L}(V)}(T) \subseteq \sigma_{\mathcal{BL}(V)}(T), \tag{9.2.1}$$

because $\mathcal{BL}(V) \subseteq \mathcal{L}(V)$, as in Section 7.3. In particular,

$$\sigma_p(T) \subseteq \sigma_{\mathcal{BL}(V)}(T). \tag{9.2.2}$$

It is easy to see that

$$|\lambda| \le \|T\|_{op} \text{ when } \lambda \in \sigma_p(T). \tag{9.2.3}$$

More precisely, if $\lambda \in \sigma_p(T)$, then $\lambda^l \in \sigma_p(T^l)$ for every positive integer l, so that

$$|\lambda|^l = |\lambda^l| \le \|T^l\|_{op}. \tag{9.2.4}$$

Equivalently, $|\lambda| \le \|T^l\|_{op}^{1/l}$, so that

$$|\lambda| \le r_{\mathcal{BL}(V)}(T), \tag{9.2.5}$$

where $r_{\mathcal{BL}(V)}(T)$ is as in Section 6.13.

If V is a Banach space, then

$$\sigma_{\mathcal{BL}(V)}(T) = \sigma_{\mathcal{L}(V)}(T), \tag{9.2.6}$$

by the open mapping theorem. In this case, $\mathcal{BL}(V)$ is a Banach algebra with respect to the operator norm, and $|\lambda| \le r_{\mathcal{BL}(V)}(T)$ for every $\lambda \in \sigma_{\mathcal{BL}(V)}(T)$, as in Section 6.13.

Suppose for the moment that V is a complex Banach space, and that T is bounded on V. If q is a complex-valued holomorphic function defined on an open subset U of $\mathbf{C}$ such that

$$\sigma_{\mathcal{BL}(V)}(T) \subseteq U, \tag{9.2.7}$$

then $\widetilde{q}(T)$ may be defined as a bounded linear mapping from V into itself, as mentioned in Section 8.13. The analogues of (9.1.5) and (9.1.7) in this case are given in parts (a) and (b) of Theorem 10.33 on p247 of [162].

Part (d) of that theorem says that (9.1.8) holds when q is not constant on any connected component of U. More precisely, part (c) of that theorem says that if (9.1.9) holds and q is not identically equal to μ on any connected component of U, then (9.1.10) holds.

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space, and let T be a bounded linear mapping from V into itself which is normal, in the sense that T commutes with T^*. In the real case, we have that

$$\ker(T - \lambda\, I) = \ker(T^* - \lambda\, I) \tag{9.2.8}$$

for every $\lambda \in \mathbf{R}$, as in Section 8.11. This means that T and T^* have the same eigenvalues, and with the same eigenvectors. Similarly, in the complex case, if $\lambda \in \mathbf{C}$, then

$$\ker(T - \lambda\, I) = \ker(T^* - \overline{\lambda}\, I). \tag{9.2.9}$$

It follows that λ is an eigenvalue of T if and only if $\overline{\lambda}$ is an eigenvaue of T^*, with the same eigenvectors.

Suppose now that T is self-adjoint, and let f be a continuous real-valued function on the spectrum $\sigma_{\mathcal{BL}(V)}(T)$ of T with respect to $\mathcal{BL}(V)$. If $v \in V$ is an eigenvector of T with eigenvalue λ, then

$$(\widetilde{f}(T))(v) = f(\lambda)\, v. \tag{9.2.10}$$

This can be obtained from the analogous statement for polynomials, by approximating f by polynomials uniformly on $\sigma_{\mathcal{BL}(V)}(T)$. This implies that

$$f(\sigma_p(T)) \subseteq \sigma_p(\widetilde{f}(T)), \tag{9.2.11}$$

as before.

9.3 Approximate eigenvalues

Let V, W, and Z be vector spaces, all real or all complex, and with norms $\|\cdot\|_V$, $\|\cdot\|_W$, and $\|\cdot\|_Z$, respectively. Also let T_1 be a linear mapping from V into W, and let T_2 be a linear mapping from W into Z. Suppose that there are positive real numbers c_1, c_2 such that

$$\|T_1(v)\|_W \ge c_1\,\|v\|_V \tag{9.3.1}$$

for every $v \in V$, and

$$\|T_2(w)\|_Z \ge c_2\,\|w\|_W \tag{9.3.2}$$

for every $w \in W$. Under these conditions, we get that

$$\|(T_2 \circ T_1)(v)\|_Z = \|T_2(T_1(v))\|_Z \ge c_2\,\|T_1(v)\|_W \ge c_1\,c_2\,\|v\|_V \tag{9.3.3}$$

for every $v \in V$.

Suppose that $V \ne \{0\}$, and let T be a bounded linear mapping from V into itself. Also let λ be a real or complex number, as appropriate. If $T - \lambda\,I$ has a bounded inverse on V, then

$$\|(T - \lambda\,I)(v)\|_V \ge c\,\|v\|_V \tag{9.3.4}$$

for some $c > 0$ and all $v \in V$, as in Section 7.11.

If there is no $c > 0$ such that (9.3.4) holds, then λ is said to be an *approximate eigenvalue* of T on V, as on p51 of [88]. Equivalently, this means that there is a sequence $\{v_j\}_{j=1}^\infty$ of elements of V such that

$$\|v_j\|_V = 1 \text{ for each } j \tag{9.3.5}$$

and

$$\|(T - \lambda\,I)(v_j)\|_V = \|T(v_j) - \lambda\,v_j\|_V \to 0 \text{ as } j \to \infty. \tag{9.3.6}$$

Of course, if λ is an eigenvalue of T, then λ is an approximate eigenvalue of T.

The set of approximate eigenvalues of T is known as the *approximate point spectrum* of T, and may be denoted $\sigma_{ap}(T)$. Thus

$$\sigma_p(T) \subseteq \sigma_{ap}(T) \subseteq \sigma_{\mathcal{BL}(V)}(T). \tag{9.3.7}$$

One can check that

$$|\lambda| \le \|T\|_{op} \text{ when } \lambda \in \sigma_{ap}(T). \tag{9.3.8}$$

Alternatively, if $|\lambda| > \|T\|_{op}$, then (9.3.4) holds with

$$c = |\lambda| - \|T\|_{op}, \tag{9.3.9}$$

as in Section 7.11.

The set of $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that (9.3.4) holds for some $c > 0$ is an open set, as in Section 7.11. Equivalently,

$$\sigma_{ap}(T) \text{ is a closed set} \tag{9.3.10}$$

in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to the standard Euclidean metric.

Suppose that $\lambda \in \sigma_{ap}(T)$, and let $\{v_j\}_{j=1}^{\infty}$ be a sequence of elements of V that satisfies (9.3.5) and (9.3.6). If l is a positive integer, then one can verify that

$$\|T^l(v_j) - \lambda^l\, v_j\|_V \to 0 \text{ as } j \to \infty. \tag{9.3.11}$$

This means that

$$\lambda^l \in \sigma_{ap}(T^l). \tag{9.3.12}$$

It follows that $|\lambda|^l = |\lambda^l| \le \|T^l\|_{op}$, as in (9.3.8). This implies that

$$|\lambda| \le r_{\mathcal{BL}(V)}(T), \tag{9.3.13}$$

where $r_{\mathcal{BL}(V)}(T)$ is as in Section 6.13.

9.4 More on approximate eigenvalues

Let us continue with the same notation and hypotheses as in the previous section. If q is a polynomial in a single variable with real or complex coefficients, as appropriate, then we get that

$$\|(\widetilde{q}(T))(v_j) - q(\lambda)\, v_j\|_V \to 0 \text{ as } j \to \infty. \tag{9.4.1}$$

This implies that

$$q(\lambda) \in \sigma_{ap}(\widetilde{q}(T)), \tag{9.4.2}$$

so that

$$q(\sigma_{ap}(T)) \subseteq \sigma_{ap}(\widetilde{q}(T)). \tag{9.4.3}$$

If q is not constant, then we would like to get that

$$\sigma_{ap}(\widetilde{q}(T)) \subseteq q(\sigma_{ap}(T)) \tag{9.4.4}$$

under suitable conditions. In the complex case, we can use the same type of argument as in Sections 8.13 and 9.1, as follows.

Let $\mu \in \mathbf{C}$ be given, and remember that $q(t) - \mu$ can be factored into the product of a nonzero complex number and finitely many linear factors $t - \lambda_j$, by the fundametal theorem of algebra. If each of the factors $T - \lambda_j\, I$ satisfies a condition like (9.3.4), then $\widetilde{q}(T) - \mu\, I$ satisfies the same type of condition. This uses the remark about compositions at the beginning of the previous section.

If

$$\mu \in \sigma_{ap}(\widetilde{q}(T)), \tag{9.4.5}$$

then at least one of the factors $T - \lambda_j\, I$ does not satisfy a condition like (9.3.4). This implies that λ_j is an approximate eigenvalue of T for some j. This means that

$$\mu \in q(\sigma_{ap}(T)), \tag{9.4.6}$$

because $q(\lambda_j) = \mu$ for each j.

In the real case, suppose that for every $a, b \in \mathbf{R}$ with $b \neq 0$ there is a positive real number $c_{a,b}$ such that

$$\|((T - a\, I)^2 + b^2\, I)(v)\|_V \geq c_{a,b}\, \|v\|_V \tag{9.4.7}$$

for all $v \in V$. This permits one to use the same type of argument as in Sections 8.13 and 9.1 again. If $(V, \langle \cdot, \cdot \rangle_V)$ is a real inner product space, and T is symmetric with respect to $\langle \cdot, \cdot \rangle_V$, then (9.4.7) holds with $c_{a,b} = b^2$.

Suppose for the moment that V is a complex Banach space, and that q is a holomorphic function defined on an open subset U of $\mathbf{C}$ with

$$\sigma_{\mathcal{BL}(V)}(T) \subseteq U. \tag{9.4.8}$$

Thus $\widetilde{q}(T)$ may be defined as a bounded linear mapping from V into itself, as mentioned in Sections 8.13 and 9.2. One can check that (9.4.1) and (9.4.3) hold in this case, using essentially the same argument as in the proofs of parts (a) and (b) of Theorem 10.33 on p247 of [162].

Similarly, if (9.4.5) holds and q is not identically equal to μ on any connected component of U, then one can verify that (9.4.6) holds, using essentially the same argument as in the proof of part (c) of that theorem. If q is not constant on any connected component of U, then it follows that (9.4.4) holds, as in part (d) of that theorem.

9.5 Inner products and $\sigma_{ap}(T)$

Let us continue with the same notation and hypotheses as in the previous two sections. Suppose now that $(V, \langle \cdot, \cdot \rangle_V)$ is a real or complex inner product space with corresponding norm $\| \cdot \|_V$. If $\lambda \in \sigma_{ap}(T)$, and $\{v_j\}_{j=1}^{\infty}$ is a sequence of elements of V that satisfies (9.3.5) and (9.3.4), then

$$\lim_{j \to \infty} \langle T(v_j), v_j \rangle_V = \lambda. \tag{9.5.1}$$

In the complex case, we also have that

$$\lim_{j \to \infty} \langle v_j, T(v_j) \rangle_V = \overline{\lambda}. \tag{9.5.2}$$

If T is symmetric with respect to $\langle \cdot, \cdot \rangle_V$, then it follows that $\lambda \in \mathbf{R}$.

Alternatively, in the complex case, if T is symmetric with respect to $\langle \cdot, \cdot \rangle_V$, then

$$\|(T - \lambda I)(v)\|_V \ge |\operatorname{Im} \lambda| \, \|v\|_V \tag{9.5.3}$$

for every $\lambda \in \mathbf{C}$ and $v \in V$, as in Section 8.11. If $\operatorname{Im} \lambda \ne 0$, then it follows that $\lambda \notin \sigma_{ap}(T)$.

Suppose that V is a real or complex Hilbert space, and that T is normal, in the sense that T commutes with T^*. Of course, this implies that $T - \lambda I$ is normal too. This implies that

$$\|(T - \lambda I)(v)\|_V = \|(T - \lambda I)^*(v)\|_V \tag{9.5.4}$$

for every $v \in V$, as in Section 8.11. Thus (9.3.4) holds for some $c > 0$ if and only if $(T - \lambda I)^*$ has the analogous property. Similarly, (9.3.6) holds for some sequence $\{v_j\}_{j=1}^\infty$ of elements of V if and only if

$$\|(T - \lambda I)^*(v_j)\|_V \to 0 \text{ as } j \to \infty. \tag{9.5.5}$$

In particular, this means that λ is an approximate eigenvalue of T if and only if λ is an approximate eigenvalue of T^* in the real case, and if and only if $\overline{\lambda}$ is an approximate eigenvalue of T^* in the complex case. In both cases, one can use the same sequence $\{v_j\}_{j=1}^\infty$ for T and T^*.

If T is normal, so that $T - \lambda I$ is normal as well, then (9.3.4) implies that $T - \lambda I$ has a bounded inverse on V, as in Section 8.11. This means that

$$\sigma_{ap}(T) = \sigma_{\mathcal{BL}(V)}(T) \tag{9.5.6}$$

in this case, as in Theorem 2 on p51 of [88].

Suppose that T is self-adjoint on V, and let f be a continuous real-valued function on $\sigma_{\mathcal{B}(V)}(T)$. Suppose also that $\lambda \in \sigma_{ap}(T)$, and let $\{v_j\}_{j=1}^\infty$ be a sequence of elements of V that satisfies (9.3.5) and (9.3.6). Under these conditions, one can verify that

$$\|(\widetilde{f}(T))(v_j) - f(\lambda)\, v_j\|_V \to 0 \text{ as } j \to \infty, \tag{9.5.7}$$

by approximating f uniformly on $\sigma_{\mathcal{BL}(V)}(T)$ by polynomials with real coefficients, and using the analogous statement (9.4.1) for such polynomials. This means that $f(\lambda)$ is an approximate eigenvalue of $\widetilde{f}(T)$. It follows that

$$f(\sigma_{\mathcal{BL}(V)}(T)) \subseteq \sigma_{\mathcal{BL}(V)}(\widetilde{f}(T)), \tag{9.5.8}$$

because of (9.5.6).

Remember that $\widetilde{f}(T)$ is invertible in $\mathcal{BL}(V)$ when $f \ne 0$ at every point in $\sigma_{\mathcal{BL}(V)}(T)$, as in Section 8.14. Similarly, $\widetilde{f}(T) - \lambda I$ is invertible in $\mathcal{BL}(V)$ when $f \ne \lambda$ at every point in $\sigma_{\mathcal{BL}(V)}(T)$. This implies that

$$\sigma_{\mathcal{BL}(V)}(\widetilde{f}(T)) \subseteq f(\sigma_{\mathcal{BL}(V)}(T)). \tag{9.5.9}$$

This means that

$$f(\sigma_{\mathcal{BL}(V)}(T)) = \sigma_{\mathcal{BL}(V)}(\widetilde{f}(T)), \tag{9.5.10}$$

because of (9.5.8).

9.6 Invertibility and dual linear mappings

Let V, W be vector spaces, both real or both complex, and remember that V^{alg}, W^{alg} are the algebraic dual spaces of V, W, respectively, as in Section 3.1. If T is a linear mapping from V into W, then T^{alg} is the corresponding dual linear mapping from W^{alg} into V^{alg}, as in Section 3.13. If T is a one-to-one linear mapping from V onto W, then it is easy to see that T^{alg} is a one-to-one linear mapping from W^{alg} onto V^{alg}, with

$$(T^{\mathrm{alg}})^{-1} = (T^{-1})^{\mathrm{alg}}. \tag{9.6.1}$$

More precisely, T^{-1} is a linear mapping from W onto V, and the corresponding dual linear mapping $(T^{-1})^{\mathrm{alg}}$ maps V^{alg} onto W^{alg}.

If T is any linear mapping from V into W, then

$$\ker T^{\mathrm{alg}} = \{\lambda \in W^{\mathrm{alg}} : T(V) \subseteq \ker \lambda\}. \tag{9.6.2}$$

This implies that $\ker T^{\mathrm{alg}} = \{0\}$ if and only if $T(V) = W$. If T^{alg} is a one-to-one mapping from V^{alg} onto W^{alg}, then we would like to check that T is a one-to-one mapping from V onto W. It suffices to verify that T is injective, because the previous statement implies that T is surjective.

Let $(T^{\mathrm{alg}})^{\mathrm{alg}}$ be the dual linear mapping from $(V^{\mathrm{alg}})^{\mathrm{alg}}$ into $(W^{\mathrm{alg}})^{\mathrm{alg}}$ associated to T^{alg}, as in Section 3.15. Remember that there are natural one-to-one linear mappings from V, W into $(V^{\mathrm{alg}})^{\mathrm{alg}}$, $(W^{\mathrm{alg}})^{\mathrm{alg}}$, respectively, as in Section 3.14. We have also seen that $(T^{\mathrm{alg}})^{\mathrm{alg}}$ corresponds to T, using these linear mappings. If $(T^{\mathrm{alg}})^{\mathrm{alg}}$ is injective, then it follows that T is injective. If T^{alg} is invertible, then $(T^{\mathrm{alg}})^{\mathrm{alg}}$ is invertible, as before, which implies in particular that $(T^{\mathrm{alg}})^{\mathrm{alg}}$ is injective.

Suppose now that $\|\cdot\|_V$, $\|\cdot\|_W$ are norms on V, W, respectively, and remember that V', W' are the corresponding dual spaces of bounded linear functionals on V, W, respectively, with their associated dual norms, as in Section 3.1. If T is a bounded linear mapping from V into W, then T' is the corresponding dual linear mapping from W' into V', which is a bounded linear mapping with the same operator norm as T, as in Section 3.13. If T is a one-to-one bounded linear mapping from V onto W with bounded inverse, then T' is a one-to-one bounded linear mapping from W' onto V' with bounded inverse, and

$$(T')^{-1} = (T^{-1})'. \tag{9.6.3}$$

As before, T^{-1} is a bounded linear mapping from W onto V, and $(T^{-1})'$ maps V' onto W'.

If T is any bounded linear mapping from V into W, then

$$\ker T' = \{\lambda \in W' : T(V) \subseteq \ker \lambda\}. \tag{9.6.4}$$

Thus

$$\ker T' = \{0\} \tag{9.6.5}$$

when

$$T(V) \text{ is dense in } W, \tag{9.6.6}$$

with respect to the metric associated to $\|\cdot\|_W$. Conversely, one can check that (9.6.5) implies (9.6.6), using the Hahn–Banach theorem. More precisely, if $T(V)$ is not dense in W, then the closure $\overline{T(V)}$ of $T(V)$ in W is a closed linear subspace of W. Under these conditions, the Hahn–Banach theorem implies that there is a nonzero element λ of W' that is in the kernel of T'.

Remember that V', W' are Banach spaces with respect to the appropriate dual norms, as in Sections 2.2 and 3.1. If T' is a one-to-one linear mapping from W' onto V', then the inverse of T' is a bounded linear mapping from V' onto W', by the open mapping theorem, as in Section 7.13. In this case, we would like to check that

$$c\,\|v\|_V \le \|T(v)\|_W \tag{9.6.7}$$

for some $c > 0$ and all $v \in V$.

Let T'' be the dual linear mapping from V'' into W'' associated to T', as in Section 3.15 again. This is a one-to-one bounded linear mapping from V'' onto W'' with bounded inverse, because of the corresponding properties of T', as before. In particular, there is a positive real number c such that

$$c\,\|L\|_{V''} \le \|T''(L)\|_{W''} \tag{9.6.8}$$

for every $L \in V''$, as in Section 7.11. Remember that the natural mappings from V, W into V'', W'' are isometries, as in Section 3.14. One can use this to get (9.6.7) from (9.6.8), because T'' corresponds to T with respect to these embeddings, as in Section 3.15.

If V is a Banach space, then (9.6.7) implies that $T(V)$ is a closed set in W, as in Section 7.11. If T' is one-to-one on W', then $T(V)$ is dense in W, as before, and we get that $T(V) = W$. If T' is a one-to-one bounded linear mapping from W' onto V' with bounded inverse, then it follows that T is a one-to-one bounded linear mapping from V onto W with bounded inverse, when V is a Banach space.

9.7 Subadditivity of $r_{\mathcal{A}}(x)$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $x \in \mathcal{A}$, then

$$r_{\mathcal{A}}(x) = \inf_{l \ge 1} \|x^l\|_{\mathcal{A}}^{1/l} = \lim_{j \to \infty} \|x^j\|_{\mathcal{A}}^{1/j}, \tag{9.7.1}$$

as in Sections 6.13, 6.14, and 7.2. If $x, y \in \mathcal{A}$ commute with each other, then we would like to show that

$$r_{\mathcal{A}}(x + y) \le r_{\mathcal{A}}(x) + r_{\mathcal{A}}(y). \tag{9.7.2}$$

This corresponds to part of Exercise 12 on p289 of [162] when $\mathcal{A}$ is a complex Banach algebra. One can reduce to the case where $\mathcal{A}$ is commutative, by considering the closed subalgebra generated by x, y. This permits one to use a characterization of $r_{\mathcal{A}}(x)$ in terms of complex homomorphisms on $\mathcal{A}$.

Let us consider a more direct approach here. Let R_x, R_y be positive real numbers such that

$$r_{\mathcal{A}}(x) < R_x, \ r_{\mathcal{A}}(y) < R_y. \tag{9.7.3}$$

It suffices to show that

$$r_{\mathcal{A}}(x+y) \le R_x + R_y. \tag{9.7.4}$$

Let n be a positive integer, and observe that

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j \, y^{n-j}, \tag{9.7.5}$$

by the binomial theorem. Using (9.7.3), we get that there are positive integers L_x, L_y such that

$$\|x^l\|_{\mathcal{A}} < R_x^l \tag{9.7.6}$$

when $l \ge L_x$, and

$$\|y^l\|_{\mathcal{A}} < R_y^l \tag{9.7.7}$$

when $l \ge L_y$. We shall only be concerned with large n, and in particular we may as well take n to be larger than $L_x + L_y$. Put

$$\Sigma_1(n) = \sum_{j=0}^{L_x - 1} \binom{n}{j} x^j \, y^{n-j}, \tag{9.7.8}$$

$$\Sigma_2(n) = \sum_{j=L_x}^{n-L_y} \binom{n}{j} x^j \, y^{n-j}, \tag{9.7.9}$$

$$\Sigma_3(n) = \sum_{j=n-L_y+1}^{n} \binom{n}{j} x^j \, y^{n-j}, \tag{9.7.10}$$

so that

$$(x+y)^n = \Sigma_1(n) + \Sigma_2(n) + \Sigma_3(n). \tag{9.7.11}$$

Observe that

$$\begin{aligned} \|(x+y)^n\|_{\mathcal{A}} &\le \|\Sigma_1(n)\|_{\mathcal{A}} + \|\Sigma_2(n)\|_{\mathcal{A}} + \|\Sigma_3(n)\|_{\mathcal{A}} \\ &\le 3 \max(\|\Sigma_1(n)\|_{\mathcal{A}}, \|\Sigma_2(n)\|_{\mathcal{A}}, \|\Sigma_3(n)\|_{\mathcal{A}}). \end{aligned} \tag{9.7.12}$$

Thus

$$\|(x+y)^n\|_{\mathcal{A}}^{1/n} \le 3^{1/n} \max(\|\Sigma_1(n)\|_{\mathcal{A}}^{1/n}, \|\Sigma_2(n)\|_{\mathcal{A}}^{1/n}, \|\Sigma_3(n)\|_{\mathcal{A}}^{1/n}). \tag{9.7.13}$$

We also have that

$$\begin{aligned} \|\Sigma_2(n)\|_{\mathcal{A}} &\le \sum_{j=L_x}^{n-L_y} \binom{n}{j} \|x^j\|_{\mathcal{A}} \, \|y^{n-j}\|_{\mathcal{A}} \le \sum_{j=L_x}^{n-L_y} \binom{n}{j} R_x^j \, R_y^{n-j} \\ &\le \sum_{j=0}^{n} \binom{n}{j} R_x^j \, R_y^{n-j} = (R_x + R_y)^n, \end{aligned} \tag{9.7.14}$$

using (9.7.6) and (9.7.7) in the second step. This implies that

$$\|\Sigma_2(n)\|_{\mathcal{A}}^{1/n} \le R_x + R_y. \tag{9.7.15}$$

If $n \ge L_x + L_y$ and $j \le L_x - 1$, then $n - j \ge L_y + 1$, and we can use (9.7.7) to get that

$$\|\Sigma_1(n)\|_{\mathcal{A}} \le \sum_{j=0}^{L_x - 1} \binom{n}{j} \|x^j\|_{\mathcal{A}} R_y^{n-j}. \tag{9.7.16}$$

This implies that

$$\|\Sigma_1(n)\|_{\mathcal{A}}^{1/n} \le \Big(\sum_{j=0}^{L_x - 1} \binom{n}{j} \|x^j\|_{\mathcal{A}} R_y^{-j} \Big)^{1/n} R_y. \tag{9.7.17}$$

Let us reexpress $\Sigma_3(n)$ as

$$\Sigma_3(n) = \sum_{l=0}^{L_y - 1} \binom{n}{l} x^{n-l} y^l. \tag{9.7.18}$$

If $n \ge L_x + L_y$ and $l \le L_y - 1$, then $n - l \ge L_x + 1$, and we can use (9.7.6) to get that

$$\|\Sigma_3(n)\|_{\mathcal{A}} \le \sum_{l=0}^{L_y - 1} \binom{n}{l} R_x^{n-l} \|y^l\|_{\mathcal{A}}. \tag{9.7.19}$$

This implies that

$$\|\Sigma_3(n)\|_{\mathcal{A}}^{1/n} \le \Big(\sum_{l=0}^{L_y - 1} \binom{n}{l} R_x^{-l} \|y^l\|_{\mathcal{A}} \Big)^{1/n} R_x. \tag{9.7.20}$$

One can use (9.7.13), (9.7.15), (9.7.17) and (9.7.20) to get that

$$r_{\mathcal{A}}(x + y) = \lim_{n \to \infty} \|(x + y)^n\|_{\mathcal{A}}^{1/n} \le R_x + R_y. \tag{9.7.21}$$

9.8 Sums of nonnegative sums

Let X be a nonempty set, and let f be a nonnegative real-valued function on X. If $E \subseteq X$, then $\sum_{x \in E} f(x)$ may be defined as a nonnegative extended real number as in Section 2.4, and we have that

$$\sum_{x \in E} f(x) \le \sum_{x \in X} f(x). \tag{9.8.1}$$

In particular, if f is summable on X, then f is summable on E.

In some cases, it may be convenient to allow sums to be taken of nonnegative extended real numbers. A sum of this type is considered to be $+\infty$ whenever any of the terms is equal to $+\infty$.

Let I be a nonempty set, and let E_j be a subset of X for each $j \in I$. Put

$$E = \bigcup_{j \in I} E_j, \tag{9.8.2}$$

and let us check that

$$\sum_{x \in E} f(x) \le \sum_{j \in I} \Big(\sum_{x \in E_j} f(x) \Big). \tag{9.8.3}$$

More precisely, the sum on the right may be considered as the sum of a non-negative extended real-valued function on I. It suffices to verify that if A is a finite subset of E, then

$$\sum_{x \in A} f(x) \le \sum_{j \in I} \Big(\sum_{x \in E_j} f(x) \Big). \tag{9.8.4}$$

In fact, if B is a finite subset of I such that

$$A \subseteq \bigcup_{j \in B} E_j, \tag{9.8.5}$$

then

$$\sum_{x \in A} f(x) \le \sum_{j \in B} \Big(\sum_{x \in A \cap E_j} f(x) \Big) \le \sum_{j \in B} \Big(\sum_{x \in E_j} f(x) \Big). \tag{9.8.6}$$

Suppose now that the E_j's are pairwise-disjoint, so that

$$E_j \cap E_l = \emptyset \tag{9.8.7}$$

when $j \ne l$. In order to show that

$$\sum_{x \in E} f(x) = \sum_{j \in I} \Big(\sum_{x \in E_j} f(x) \Big), \tag{9.8.8}$$

it is enough to check that

$$\sum_{j \in I} \Big(\sum_{x \in E_j} f(x) \Big) \le \sum_{x \in E} f(x). \tag{9.8.9}$$

To get this, it suffices to have that

$$\sum_{j \in B} \Big(\sum_{j \in E_j} f(x) \Big) \le \sum_{x \in E} f(x). \tag{9.8.10}$$

This can be obtained from an analogous statement about sums over unions of two disjoint sets mentioned in Section 2.4. Alternatively, one can observe that if A_j is a finite subset of E_j for each $j \in B$, then

$$\sum_{j \in B} \Big(\sum_{x \in A_j} f(x) \Big) = \sum_{x \in \bigcup_{j \in B} A_j} f(x) \le \sum_{x \in E} f(x). \tag{9.8.11}$$

Let Y, Z be nonempty sets, and suppose that

$$X = Y \times Z. \tag{9.8.12}$$

Consider the iterated sums

$$\sum_{y \in Y} \Big(\sum_{z \in Z} f(y, z) \Big) \tag{9.8.13}$$

and

$$\sum_{z \in Z} \Big(\sum_{y \in Y} f(y, z) \Big). \tag{9.8.14}$$

Each of these iterated sums is equal to $\sum_{x \in X} f(x)$, as in (9.8.8). This corresponds to considering X as the union of the pairwise-disjoint sets $\{y\} \times Z$, $y \in Y$, or $Y \times \{z\}$, $z \in Z$. In particular, (9.8.13) and (9.8.14) are equal to each other.

9.9 Sums of generalized sums

Let X be a nonempty set, let W be a vector space over the real or complex numbers with a norm $\|\cdot\|_W$, and let f be a W-valued function on X. Suppose that $\sum_{x \in X} f(x)$ satisfies the generalized Cauchy condition, as in Section 2.7. If E is a nonempty subset of X, then it is easy to see that

$$\sum_{x \in E} f(x) \tag{9.9.1}$$

satisfies the generalized Cauchy condition as well. If W is a Banach space, then it follows that (9.9.1) converges in the generalized sense, as in Section 2.7 again. Similarly, if f has bounded finite sums on X, as in Section 2.9, then the restriction of f to E has bounded finite sums too, with BFS norm less than or equal to the BFS norm of f on X.

Let us suppose from now on in this section that W is a Banach space. Let I be a nonempty set, and let $\{E_j\}_{j \in I}$ be a pairwise-disjoint family of nonempty subsets of X. Thus

$$\sum_{x \in E_j} f(x) \tag{9.9.2}$$

converges in the generalized sense for each $j \in I$, as in the preceding paragraph. Put $E = \bigcup_{j \in I} E_j$. If I has only finitely many elements, then one can check that

$$\sum_{j \in I} \Big(\sum_{x \in E_j} f(x) \Big) = \sum_{x \in E} f(x). \tag{9.9.3}$$

If I has infinitely many elements, then we would like to check that the sum over I in the left side of (9.9.3) converges in the generalized sense, and is equal to the right side of (9.9.3). Of course, if $B \subseteq I$ has only finitely many elements, then

$$\sum_{j \in B} \Big(\sum_{x \in E_j} f(x) \Big) = \sum_{x \in \bigcup_{j \in B} E_j} f(x), \tag{9.9.4}$$

as in the preceding paragraph. Let $\epsilon > 0$ be given, and remember that there is a finite set $A_1(\epsilon) \subseteq X$ such that

$$\Big\| \sum_{x \in C} f(x) \Big\|_W < \epsilon \tag{9.9.5}$$

for every finite set $C \subseteq X \setminus A_1(\epsilon)$, as in Section 2.7. Put

$$B_1(\epsilon) = \{ j \in I : A_1(\epsilon) \cap E_j \neq \emptyset \}, \tag{9.9.6}$$

which is a finite subset of I. If $B_0 \subseteq I \setminus B_1(\epsilon)$ is a finite set, then one can verify that

$$\Big\| \sum_{j \in B_0} \Big(\sum_{x \in E_j} f(x) \Big) \Big\|_W = \Big\| \sum_{x \in \bigcup_{j \in B_0} E_j} f(x) \Big\|_W \le \epsilon, \tag{9.9.7}$$

using (9.9.5).

This shows that the sum over I in the left side of (9.9.3) satisfies the generalized Cauchy condition. Similarly, one can show that for suitable finite subsets B of I, the right side of (9.9.4) can be approximated by finite sums that also approximate the right side of (9.9.3). This is a bit simpler when E_j has only finitely many elements for each j, so that $\bigcup_{j \in B} E_j$ is finite when B is finite.

Suppose for the moment that $\|f(x)\|_W$ is summable as a nonnegative real-valued on X. In this case,

$$\Big\| \sum_{x \in E_j} f(x) \Big\|_W \le \sum_{x \in E_j} \|f(x)\|_W \tag{9.9.8}$$

for each $j \in I$. We also have that

$$\sum_{j \in I} \Big(\sum_{x \in E_j} \|f(x)\|_W \Big) = \sum_{x \in E} \|f(x)\|_W, \tag{9.9.9}$$

as in the previous section. It follows that the left side of (9.9.8) is summable as a nonnegative real-valued function on I. If $W = \mathbf{R}$ or $\mathbf{C}$, with the standard absolute value function as the norm, then (9.9.3) can be obtained from the analogous statement in the previous section, by expressing f as a linear combination of nonnegative real-valued summable functions on X.

9.10 Cauchy products

Let $\sum_{j=0}^{\infty} a_j$, $\sum_{l=0}^{\infty} b_l$ be infinite series of real or complex numbers, and put

$$c_n = \sum_{j=0}^{n} a_j \, b_{n-j} \tag{9.10.1}$$

for each nonnegative integer n. The series $\sum_{n=0}^\infty c_n$ is called the *Cauchy product* of $\sum_{j=0}^\infty a_j$ and $\sum_{l=0}^\infty b_l$, and it is easy to see that

$$\sum_{n=0}^\infty c_n = \Big(\sum_{j=0}^\infty a_j\Big)\Big(\sum_{l=0}^\infty b_l\Big), \tag{9.10.2}$$

at least formally. More precisely, both sides of (9.10.2) are formally equal to

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} a_j\, b_l, \tag{9.10.3}$$

where $(\mathbf{Z}_+\cup\{0\})^2 = (\mathbf{Z}_+\cup\{0\})\times(\mathbf{Z}_+\cup\{0\})$. If $a_j = 0$ for all but finitely many j, and $b_l = 0$ for all but finitely many l, then one can check $c_n = 0$ for all but finitely many n. In this case, each of these infinite sums reduces to a finite sum, and both sides of (9.10.2) are equal to (9.10.3).

Suppose for the moment that a_j, b_l are nonnegative real numbers for each j, l, respectively, so that c_n is a nonnegative real number for each n. Under these conditions, the infinite sums mentioned in the preceding paragraph may be defined as nonnegative extended real numbers as in Section 2.4. It is easy to see that the left side of (9.10.2) is equal to (9.10.3) in this sense, as in Section 9.8. One can also check that (9.10.2) holds, or equivalently that the right side of (9.10.2) is equal to (9.10.3), where the right side of (9.10.2) is interpreted as being equal to 0 when either of the factors is equal to 0. In particular, if $\sum_{j=0}^\infty a_j$ and $\sum_{l=0}^\infty b_l$ converge, then $\sum_{n=0}^\infty c_n$ converges, (9.10.2) holds, and (9.10.3) is finite and equal to both sides of (9.10.2).

If the a_j's and b_l's are any real or complex numbers, then

$$|c_n| \le \sum_{j=0}^n |a_j|\,|b_{n-j}| \tag{9.10.4}$$

for each $n \ge 0$. Note that the right side is the same as the nth term of the Cauchy product of $\sum_{j=0}^\infty |a_j|$ and $\sum_{l=0}^\infty |b_l|$. If $\sum_{j=0}^\infty a_j$ and $\sum_{l=0}^\infty b_l$ converge absolutely, then it follows that $\sum_{n=0}^\infty c_n$ converges absolutely, with

$$\sum_{n=0}^\infty |c_n| \le \Big(\sum_{j=0}^\infty |a_j|\Big)\Big(\sum_{l=0}^\infty |b_l|\Big). \tag{9.10.5}$$

We also have that

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} |a_j|\,|b_l| = \Big(\sum_{j=0}^\infty |a_j|\Big)\Big(\sum_{l=0}^\infty |b_l|\Big), \tag{9.10.6}$$

as before. It follows that

$$f(j,l) = a_j\, b_l \tag{9.10.7}$$

is summable as a real or complex-valued function on $(\mathbf{Z}_+\cup\{0\})^2$, which is to say that it is an element of $\ell^1((\mathbf{Z}_+\cup\{0\})^2, \mathbf{R})$ or $\ell^1((\mathbf{Z}_+\cup\{0\})^2, \mathbf{C})$, as appropriate.

This means that the sum (9.10.3) converges in the generalized sense, as in Section 2.8. Using this, it is easy to see that the sum is equal to the left side of (9.10.2), as in the previous section. One can also check that (9.10.3) is equal to the right side of (9.10.2), by considering the sum over $(\mathbf{Z}_+ \cup \{0\})^2$ as an iterated sum over each variable. Of course, this implies that (9.10.2) holds. Alternatively, one can verify (9.10.2) by approximating $\sum_{j=0}^\infty a_j$ and $\sum_{l=0}^\infty b_l$ by finite subsums, using (9.10.5).

If the a_j's and b_l's are real numbers, then the absolute convergence of $\sum_{j=0}^\infty a_j$ and $\sum_{l=0}^\infty b_l$ implies that these series may be expressed as differences of convergent series of nonnegative real numbers. This permits one to reduce (9.10.2) to the analogous statement in that case. Similarly, if the a_j's and b_l's are complex numbers, then one can reduce to the real case by considering their real and imaginary parts.

9.11 Cauchy products and bilinear mappings

Let V, W, and Z be vector spaces all real or all complex, and let B be a bilinear mapping from $V \times W$ into Z. Also let $\sum_{j=0}^\infty v_j$ and $\sum_{l=0}^\infty w_l$ be infinite series of elements of V and W, respectively. Put

$$z_n = \sum_{j=0}^n B(v_j, w_{n-j}) \tag{9.11.1}$$

for each nonnegative integer n. The series $\sum_{n=0}^\infty z_n$ may be described as the *Cauchy product* of $\sum_{j=0}^\infty v_j$ and $\sum_{l=0}^\infty w_l$ with respect to B. It is easy to see that

$$\sum_{n=0}^\infty z_n = B\Big(\sum_{j=0}^\infty v_j, \sum_{l=0}^\infty w_l\Big), \tag{9.11.2}$$

at least formally.

In fact, both sides of (9.11.2) correspond to formally expressing

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} B(v_j, w_l) \tag{9.11.3}$$

as an iterated sum. If $v_j = 0$ for all but finitely many j, and $w_l = 0$ for all but finitely many l, then $z_n = 0$ for all but finitely many n, each of these infinite sums reduces to a finite sum, and both sides of (9.11.2) are equal to (9.11.3).

Let $\|\cdot\|_V$, $\|\cdot\|_W$, and $\|\cdot\|_Z$ be norms on V, W, and Z, respectively, and suppose that B is bounded with respect to these norms, so that

$$\|B(v, w)\|_Z \le C\,\|v\|_V\,\|w\|_W \tag{9.11.4}$$

for some nonnegative real number C and all $v \in V$, $w \in W$. This implies that

$$\|z_n\|_Z \le C \sum_{j=0}^n \|v_j\|_V\,\|w_{n-j}\|_W \tag{9.11.5}$$

for every $n \geq 0$. The sum on the right is the same as the nth term of the Cauchy product of the series $\sum_{j=0}^{\infty} \|v_j\|_V$ and $\sum_{l=0}^{\infty} \|w_l\|_W$. Suppose from now on in this section that $\sum_{j=0}^{\infty} v_j$ and $\sum_{l=0}^{\infty} w_l$ converge absolutely with respect to $\|\cdot\|_V$ and $\|\cdot\|_W$, respectively. This means that $\sum_{n=0}^{\infty} z_n$ converges absolutely with respect to $\|\cdot\|_Z$, with

$$\sum_{n=0}^{\infty} \|z_n\|_Z \leq C \Big(\sum_{j=0}^{\infty} \|v_j\|_V\Big) \Big(\sum_{l=0}^{\infty} \|w_l\|_W\Big), \tag{9.11.6}$$

as in the previous section.

Similarly,

$$\begin{aligned} \sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} \|B(v_j, w_l)\|_Z &\leq C \sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} \|v_j\|_V \, \|w_l\|_W \\ &= C \Big(\sum_{j=0}^{\infty} \|v_j\|_V\Big) \Big(\sum_{l=0}^{\infty} \|w_l\|_W\Big). \end{aligned} \tag{9.11.7}$$

Using the absolute convergence of $\sum_{j=0}^{\infty} v_j$ and $\sum_{l=0}^{\infty} w_l$ with respect to $\|\cdot\|_V$ and $\|\cdot\|_W$, respectively, we get that

$$g(j,l) = B(v_j, w_l) \tag{9.11.8}$$

is summable as a Z-valued function on $(\mathbf{Z}_+ \cup \{0\})^2$ with respect to $\|\cdot\|_Z$, which is to say that it is an element of $\ell^1((\mathbf{Z}_+ \cup \{0\})^2, Z)$, as in Section 2.6. Suppose that V, W, and Z are complete with respect to the metrics associated to their norms, so that absolutely convergent series in these spaces converge, as in Section 1.7. We also get that (9.11.3) converges in the generalized sense, as in Section 2.8. The sum is equal to the left side of (9.11.2), as in Section 9.9.

Similarly, (9.11.3) is equal to the right side of (9.11.2), because the sum over $(\mathbf{Z}_+ \cup \{0\})^2$ can be expressed as an iterated sum. As before, one can get (9.11.2) more directly by approximating $\sum_{j=0}^{\infty} v_j$ and $\sum_{l=0}^{\infty} w_l$ by finite subsums, using (9.11.6). Note that the completeness of Z is only a convenience here, because the convergence of the sums in Z can be obtained from the convergence of $\sum_{j=0}^{\infty} v_j$ and $\sum_{l=0}^{\infty} w_l$ in V and W, respectively.

9.12 Radius of convergence

Let

$$\sum_{j=0}^{\infty} a_j \, z^j \tag{9.12.1}$$

be a power series with real or complex coefficients. Also let E_0 be the set of nonnegative real numbers r such that

$$\{|a_j| \, r^j\}_{j=0}^{\infty} \text{ is a bounded sequence,} \tag{9.12.2}$$

and let E_1 be the set of $r \geq 0$ such that

$$\sum_{j=0}^{\infty} |a_j|\, r^j \tag{9.12.3}$$

converges. Note that $0 \in E_1$, and that $E_1 \subseteq E_0$. If $0 \leq r_1 < r_0$ and $r_0 \in E_0$, then one can check that $r_1 \in E_1$. It follows that

$$\sup E_0 = \sup E_1, \tag{9.12.4}$$

where the suprema are interpreted as being $+\infty$ when E_0, E_1 are unbounded.

Let ρ be the common value in (9.12.4), which is the *radius of convergence* of the power series (9.12.1). If $z \in \mathbf{C}$ and $|z| < \rho$, then

$$|z| \in E_1, \tag{9.12.5}$$

and (9.12.1) converges absolutely, by the comparison test. If $|z| > \rho$, then

$$|z| \not\in E_0, \tag{9.12.6}$$

and (9.12.1) does not converge. It is easy to see that ρ is uniquely determined by these two properties.

If $r \in E_1$, then (9.12.1) converges absolutely when $|z| \leq r$, by the comparison test. This implies that

$$f(z) = \sum_{j=0}^{\infty} a_j\, z^j \tag{9.12.7}$$

defines a complex-valued function on the closed disk

$$\{z \in \mathbf{C} : |z| \leq r\}. \tag{9.12.8}$$

We also get that the sequence of partial sums

$$\sum_{j=0}^{n} a_j\, z^j \tag{9.12.9}$$

converges to f uniformly on (9.12.8), by a well-known criterion for uniform convergence due to Weierstrass. It follows that f is continuous on (9.12.8), with respect to the standard Euclidean metric on $\mathbf{C}$ and its restriction to (9.12.8), because polynomials define continuous functions on $\mathbf{C}$. More precisely, one can use the same argument to get that f is uniformly continuous on (9.12.8), because polynomials are uniformly continuous on bounded subsets of $\mathbf{C}$.

Suppose that $\rho > 0$, so that (9.12.7) defines a complex-valued function on

$$\{z \in \mathbf{C} : |z| < \rho\}. \tag{9.12.10}$$

It is easy to see that f is continuous on (9.12.10), because the restriction of f to (9.12.8) is continuous when $0 \leq r < \rho$, as in the preceding paragraph. Of course, it is well known that f is holomorphic on (9.12.10). Conversely, it is well known that any holomorphic function on an open disk in $\mathbf{C}$ centered at 0 of positive radius can be represented by a unique power series in this way, where the radius of convergence of the power series is at least the radius of the disk.

Let

$$\sum_{l=0}^{\infty} b_l\, z^l \tag{9.12.11}$$

be another power series with resl or complex coefficients. Put $c_n = \sum_{j=0}^{n} a_j\, b_{n-j}$ for each nonnegative integer n, as in Section 9.10, and observe that

$$c_n\, z^n = \sum_{j=0}^{n} (a_j\, z^j)\, (b_{n-j}\, z^{n-j}) \tag{9.12.12}$$

for each $n \geq 0$. Thus

$$\sum_{n=0}^{\infty} c_n\, z^n \tag{9.12.13}$$

is the Cauchy product of (9.12.1) and (9.12.11). If (9.12.1) and (9.12.11) converge absolutely for some $z \in \mathbf{C}$, then it follows that (9.12.13) converges absolutely too, and that

$$\sum_{n=0}^{\infty} c_n\, z^n = \Big(\sum_{j=0}^{\infty} a_j\, z^j\Big) \Big(\sum_{l=0}^{\infty} b_l\, z^l\Big), \tag{9.12.14}$$

as before.

9.13 Cauchy products and Banach algebras

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers, and let $\sum_{j=0}^{\infty} v_j$, $\sum_{l=0}^{\infty} w_l$ be infinite series of elements of $\mathcal{A}$. Put

$$z_n = \sum_{j=0}^{n} v_j\, w_{n-j} \tag{9.13.1}$$

for each nonnegative integer n, so that $\sum_{n=0}^{\infty} z_n$ is the *Cauchy product* of $\sum_{j=0}^{\infty} v_j$ and $\sum_{l=0}^{\infty} w_l$ with respect to multiplication on $\mathcal{A}$, as in Section 9.11. As before,

$$\sum_{n=0}^{\infty} z_n = \Big(\sum_{j=0}^{\infty} v_j\Big) \Big(\sum_{l=0}^{\infty} w_l\Big), \tag{9.13.2}$$

at least formally. In particular, if $v_j = 0$ for all but finitely many j, and $w_l = 0$ for all but finitely many l, then $z_n = 0$ for all but finitely many n. In this case, the infinite series reduce to finite sums that satisfy (9.13.2).

Let $\|\cdot\|$ be a submultiplicative norm on $\mathcal{A}$, and note that

$$\|z_n\|_{\mathcal{A}} \le \sum_{j=0}^{n} \|v_j\|_{\mathcal{A}}\, \|w_{n-j}\|_{\mathcal{A}} \tag{9.13.3}$$

for each $n \ge 0$. If $\sum_{j=0}^{\infty} v_j$, $\sum_{l=0}^{\infty} w_l$ converge absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, then $\sum_{n=0}^{\infty} z_n$ converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, with

$$\sum_{n=0}^{\infty} \|z_n\|_{\mathcal{A}} \le \Big(\sum_{j=0}^{\infty} \|v_j\|_{\mathcal{A}}\Big)\Big(\sum_{l=0}^{\infty} \|w_l\|_{\mathcal{A}}\Big). \tag{9.13.4}$$

If $\mathcal{A}$ is complete with respect to metric associated to $\|\cdot\|_{\mathcal{A}}$, then we have seen that (9.13.2) holds under these conditions.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$. Let $\sum_{j=0}^{\infty} a_j\, z^j$ be a power series with real or complex coefficients, as appropriate. If $x \in \mathcal{A}$ and

$$\sum_{j=0}^{\infty} |a_j|\, \|x^j\|_{\mathcal{A}} \tag{9.13.5}$$

converges as an infinite series of nonnegative real numbers, then

$$\sum_{j=0}^{\infty} a_j\, x^j \tag{9.13.6}$$

converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, and thus converges in $\mathcal{A}$, as in Section 1.7. In particular, if

$$\sum_{j=0}^{\infty} |a_j|\, \|x\|_{\mathcal{A}}^j \tag{9.13.7}$$

converges, then (9.13.5) converges, by the comparison test.

Let $\sum_{l=0}^{\infty} b_l\, z^l$ be another power series with real or complex coefficients, as appropriate, and let $\sum_{n=0}^{\infty} c_n\, z^n$ be the Cauchy product of $\sum_{j=0}^{\infty} a_j\, z^j$ and $\sum_{l=0}^{\infty} b_l\, z^l$. Suppose that

$$\sum_{l=0}^{\infty} |b_l|\, \|x^l\|_{\mathcal{A}} \tag{9.13.8}$$

converges as an infinite series of nonnegative real numbers too, so that

$$\sum_{l=0}^{\infty} b_l\, x^l \tag{9.13.9}$$

converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$. Note that

$$c_n\, x^n = \sum_{j=0}^{n} (a_j\, x^j)\, (b_{n-j}\, x^{n-j}) \tag{9.13.10}$$

for each $n \geq 0$, so that

$$\sum_{n=0}^{\infty} c_n \, x^n \tag{9.13.11}$$

is the Cauchy product of (9.13.6) and (9.13.9) with respect to multiplication on $\mathcal{A}$. Using (9.13.10), we get that

$$|c_n| \, \|x^n\|_{\mathcal{A}} \leq \sum_{j=0}^{n} (|a_j| \, \|x^j\|_{\mathcal{A}}) \, (|b_{n-j}| \, \|x^{n-j}\|_{\mathcal{A}}) \tag{9.13.12}$$

for each $n \geq 0$.

The absolute convergence of (9.13.6) and (9.13.9) with respect to $\| \cdot \|_{\mathcal{A}}$ implies that (9.13.11) converges absolutely with respect to $\| \cdot \|_{\mathcal{A}}$ too, with

$$\sum_{n=0}^{\infty} |c_n| \, \|x^n\|_{\mathcal{A}} \leq \Big(\sum_{j=0}^{\infty} |a_j| \, \|x^j\|_{\mathcal{A}} \Big) \Big(\sum_{l=0}^{\infty} |b_l| \, \|x^l\|_{\mathcal{A}} \Big). \tag{9.13.13}$$

We also have that

$$\sum_{n=0}^{\infty} c_n \, x^n = \Big(\sum_{j=0}^{\infty} a_j \, x^j \Big) \Big(\sum_{l=0}^{\infty} b_l \, x^l \Big) \tag{9.13.14}$$

under these conditions. Of course, if x is nilpotent in $\mathcal{A}$, then (9.13.6), (9.13.9), and (9.13.11) reduce to finite sums, and (9.13.14) holds automatically. More precisely, one does not need a norm on $\mathcal{A}$ in this case.

9.14 Power series and Banach algebras

Let $(\mathcal{A}, \| \cdot \|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers, with a multiplicative identity element $e_{\mathcal{A}}$, and let $\sum_{j=0}^{\infty} a_j \, z^j$ be a power series with real or complex coefficients, as appropriate. Also let r be a nonnegative real number in the set E_1 defined in Section 9.12, so that $\sum_{j=0}^{\infty} |a_j| \, r^j$ converges. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \leq r$, then (9.13.7) converges, by the comparison test. This implies that (9.13.6) converges absolutely with respect to $\| \cdot \|_{\mathcal{A}}$, so that

$$\widetilde{f}(x) = \sum_{j=0}^{\infty} a_j \, x^j \tag{9.14.1}$$

defines an $\mathcal{A}$-valued function on

$$\{x \in \mathcal{A} : \|x\|_{\mathcal{A}} \leq r\}. \tag{9.14.2}$$

Note that (9.14.1) is bounded on (9.14.2), with

$$\|\widetilde{f}(x)\|_{\mathcal{A}} \leq \sum_{j=0}^{\infty} |a_j| \, r^j \tag{9.14.3}$$

when $\|x\|_{\mathcal{A}} \le r$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$.

As in Section 9.12, the sequence of partial sums

$$\sum_{j=0}^{n} a_j\, x^j \tag{9.14.4}$$

converges uniformly to (9.14.1) on (9.14.2), essentially as in Weierstrass' criterion for uniform convergence. These partial sums are continuous as $\mathcal{A}$-valued functions of x on $\mathcal{A}$, with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. In fact, these partial sums are uniformly continuous on bounded subsets of $\mathcal{A}$, and on (9.14.2) in particular. It follows that (9.14.1) is uniformly continuous on (9.14.2).

Let ρ be the radius of convergence of $\sum_{j=0}^{\infty} a_j\, z^j$, as in Section 9.12, and suppose that $\rho > 0$. If $x \in \mathcal{A}$ and

$$\|x\|_{\mathcal{A}} < \rho, \tag{9.14.5}$$

then $\|x\|_{\mathcal{A}} \in E_1$, which is to say that (9.13.7) converges. This implies that (9.13.6) converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, as before, so that (9.14.1) defines an $\mathcal{A}$-valued function on

$$\{x \in \mathcal{A} : \|x\|_{\mathcal{A}} < \rho\}. \tag{9.14.6}$$

We also have that (9.14.1) is continuous on (9.14.6), with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, and its restriction to (9.14.6). This follows from the fact that the restriction of (9.14.1) to (9.14.2) is continuous when $0 \le r < \rho$, as in the preceding paragraph.

Suppose now that $x \in \mathcal{A}$ satisfies

$$r_{\mathcal{A}}(x) = \lim_{j\to\infty} \|x^j\|_{\mathcal{A}}^{1/j} < \rho. \tag{9.14.7}$$

This implies that there is an $r \in \mathbf{R}$ such that

$$\lim_{j\to\infty} \|x^j\|_{\mathcal{A}}^{1/j} < r < \rho. \tag{9.14.8}$$

It follows that

$$\|x^j\|_{\mathcal{A}}^{1/j} < r \tag{9.14.9}$$

for all but finitely many j, so that

$$\|x^j\|_{\mathcal{A}} < r^j \tag{9.14.10}$$

for all but finitely many j. Note that $\sum_{j=0}^{\infty} |a_j|\, r^j$ converges, because $r < \rho$, as in Section 9.12. This implies that (9.13.5) converges, by the comparison test, so that (9.13.6) converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, and we can define $\widetilde{f}(x) \in \mathcal{A}$ as in (9.14.1).

In the complex case, if f is a holomorphic function on an open subset U of $\mathbf{C}$ that contains the spectrum of $x \in \mathcal{A}$. then $\widetilde{f}(x) \in \mathcal{A}$ can be defined in a natural way, as mentioned in Section 8.13. If U is an open disk centered at 0 and $r_{\mathcal{A}}(x)$ is less than the radius of U, then this corresponds to the remarks in the preceding paragraph.

9.15 More on $\widetilde{f}(x)$

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let $\sum_{j=0}^{\infty} a_j z^j$ be a power series with real or complex coefficients, as appropriate, again. Suppose that $\sum_{j=0}^{\infty} |a_j|\, r^j$ converges for some $r \geq 0$, and put $f(z) = \sum_{j=0}^{\infty} a_j\, z^j$ when $z \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|z| \leq r$. Similarly, if $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \leq r$, then $\widetilde{f}(x) \in \mathcal{A}$ may be defined as in (9.14.1).

Put

$$p_n(z) = \sum_{j=0}^{n} a_j\, z^n \tag{9.15.1}$$

for each nonnegative integer n, so that $p_n(z) \to f(z)$ as $n \to \infty$ uniformly for $z \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|z| \leq r$, as in Section 9.12. If $x \in \mathcal{A}$, then

$$\widetilde{p}_n(x) = \sum_{j=0}^{n} a_j\, x^j \tag{9.15.2}$$

defines an element of $\mathcal{A}$ for each n, as in Section 8.13. If $\|x\|_{\mathcal{A}} \leq r$, then

$$\widetilde{p}_n(x) \to \widetilde{f}(x) \text{ as } n \to \infty, \tag{9.15.3}$$

with respect to the metric on $\mathcal{A}$ associated to $\|\cdot\|_{\mathcal{A}}$. Similarly, if $w \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|w| \leq r$, then

$$\widetilde{p}_n(x) - p_n(w)\, e_{\mathcal{A}} \to \widetilde{f}(x) - f(w)\, e_{\mathcal{A}} \text{ as } n \to \infty. \tag{9.15.4}$$

If $w \in \sigma_{\mathcal{A}}(x)$, then $\widetilde{p}_n(x) - p_n(w)\, e_{\mathcal{A}}$ is not invertible in $\mathcal{A}$ for any n, as in Section 8.13. This implies that $\widetilde{f}(x) - f(w)\, e_{\mathcal{A}}$ is not invertible, because of (9.15.4), and the fact that the invertible elements of $\mathcal{A}$ form an open set. This means that

$$f(w) \in \sigma_{\mathcal{A}}(\widetilde{f}(x)). \tag{9.15.5}$$

It follows that

$$f(\sigma_{\mathcal{A}}(x)) \subseteq \sigma_{\mathcal{A}}(\widetilde{f}(x)). \tag{9.15.6}$$

More precisely, if $w \in \sigma_{\mathcal{A}}(x)$, then $|w| \leq r$, as in Section 6.8, so that (9.15.5) holds.

Similarly, if $\sum_{j=0}^{\infty} a_j\, z^j$ has radius of convergence $\rho > 0$, and $x \in \mathcal{A}$ satisfies (9.14.7), then $\widetilde{f}(x) \in \mathcal{A}$ may be defined as in (9.14.1), and satisfies (9.15.3). If $w \in \mathbf{R}$ or $\mathbf{C}$ and $|w| < \rho$, then $f(w)$ can be defined in the usual way, and

$$p_n(w) \to f(w) \text{ as } n \to \infty, \tag{9.15.7}$$

so that (9.15.4) holds. If $w \in \sigma_{\mathcal{A}}(x)$, then

$$|w| \leq r_{\mathcal{A}}(x) < \rho, \tag{9.15.8}$$

where the first step is as in Section 6.13. In this case, (9.15.5) holds, for the same reasons as before. This implies that (9.15.6) holds, as before.

Suppose now that $(V, \|\cdot\|_V)$ is a Banach space over the real or complex numbers with $V \neq \{0\}$, and that $\mathcal{A} = \mathcal{BL}(V)$, with the corresponding operator norm $\|\cdot\|_{op}$. Suppose again that $\sum_{j=0}^{\infty} |a_j|\, r^j$ converges, so that $\widetilde{f}(T) \in \mathcal{BL}(V)$ may be defined as before when $T \in \mathcal{BL}(V)$ and

$$\|T\|_{op} \le r. \tag{9.15.9}$$

Let $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, be an eigenvalue of T with eigenvector v, and remember that $|\lambda| \le \|T\|_{op}$, as in Section 9.2. Observe that

$$(\widetilde{f}(T))(v) = f(\lambda)\, v, \tag{9.15.10}$$

because of the analogous statement for polynomials. This means that

$$f(\lambda) \in \sigma_p(\widetilde{f}(T)), \tag{9.15.11}$$

so that

$$f(\sigma_p(T)) \subseteq \sigma_p(\widetilde{f}(T)). \tag{9.15.12}$$

Now let $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, be an approximate eigenvalue of T, and let $\{v_j\}_{j=1}^{\infty}$ be a corresponding sequence of unit vectors in V, as in Section 9.3. Thus $|\lambda| \le \|T\|_{op}$, as before. One can check that

$$\|(\widetilde{f}(T))(v_j) - f(\lambda)\, v_j\|_V \to 0 \text{ as } j \to \infty, \tag{9.15.13}$$

using the analogous statement for T^l for each $l \ge 1$. This means that $f(\lambda)$ is an approximate eigenvalue of $\widetilde{f}(T)$, so that

$$f(\lambda) \in \sigma_{ap}(\widetilde{f}(T)). \tag{9.15.14}$$

It follows that

$$f(\sigma_{ap}(T)) \subseteq \sigma_{ap}(\widetilde{f}(T)). \tag{9.15.15}$$

If $\sum_{j=0}^{\infty} a_j\, z^j$ has radius of convergence $\rho > 0$ and $T \in \mathcal{BL}(V)$ satisfies

$$r_{\mathcal{BL}(V)}(T) < \rho, \tag{9.15.16}$$

then $\widetilde{f}(T) \in \mathcal{BL}(V)$ may be defined as before. Suppose that $\lambda \in \mathbf{R}$ or $\mathbf{C}$ is an eigenvalue of T with eigenvector v, so that

$$|\lambda| \le r_{\mathcal{BL}(V)}(T) < \rho, \tag{9.15.17}$$

where the first step is as in Section 9.2. In this case, (9.15.10), (9.15.11), and (9.15.12) hold, for essentially the same reasons as before. Similarly, if λ is an approximate eigenvalue of T, then the first inequality in (9.15.17) holds, as in Section 9.3. If $\{v_j\}_{j=1}^{\infty}$ is a corresponding sequence of unit vectors in V, then (9.15.13), (9.15.14), and (9.15.15) hold, for essentially the same reasons as before.

Chapter 10

Some more algebras, power series

10.1 Power series and Lipschitz conditions

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers. If $x, y \in \mathcal{A}$ and n is a positive integer, then

$$x^n - y^n = \sum_{j=1}^{n} (x^j\, y^{n-j} - x^{j-1}\, y^{n-j+1}) = \sum_{j=1}^{n} x^{j-1}\, (x - y)\, y^{n-j}. \tag{10.1.1}$$

More precisely, one can simply drop the factors of x^{j-1} when $j = 1$, as well as the factors of y^{n-j} when $j = n$. If $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$, then we get that

$$\begin{aligned} \|x^n - y^n\|_{\mathcal{A}} &\le \sum_{j=1}^{n} \|x^{j-1}\|_{\mathcal{A}}\, \|x - y\|_{\mathcal{A}}\, \|y^{n-j}\|_{\mathcal{A}} \\ &\le n \max(\|x\|_{\mathcal{A}}, \|y\|_{\mathcal{A}})^{n-1}\, \|x - y\|_{\mathcal{A}}. \end{aligned} \tag{10.1.2}$$

If r is a nonnegative real number and $\|x\|_{\mathcal{A}}, \|y\|_{\mathcal{A}} \le r$, then we get that

$$\|x^n - y^n\|_{\mathcal{A}} \le n\, r^{n-1}\, \|x - y\|_{\mathcal{A}}. \tag{10.1.3}$$

This means that

$$x \mapsto x^n \tag{10.1.4}$$

is Lipschitz with constant $n\, r^{n-1}$, as an $\mathcal{A}$-valued function on

$$\{x \in \mathcal{A} : \|x\|_{\mathcal{A}} \le r\}. \tag{10.1.5}$$

Of course, this uses the metric on $\mathcal{A}$ associated to $\|\cdot\|_{\mathcal{A}}$, and its restriction to (10.1.5). This Lipschitz condition could also be obtained from some of the remarks in Section 7.10.

Suppose now that $\mathcal{A}$ is a Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$, and let $\sum_{j=0}^\infty a_j\, z^j$ be a power series with real or complex coefficients, as appropriate. Suppose also that

$$\sum_{j=1}^\infty j\, |a_j|\, r^{j-1} \tag{10.1.6}$$

converges, which implies in particular that $\sum_{j=0}^\infty |a_j|\, r^j$ converges. Thus

$$\widetilde{f}(x) = \sum_{j=0}^\infty a_j\, x^j \tag{10.1.7}$$

defines an $\mathcal{A}$-valued function on (10.1.5), as in Section 9.14. If $x, y \in \mathcal{A}$ and $\|x\|_{\mathcal{A}}, \|y\|_{\mathcal{A}} \le r$, then

$$\widetilde{f}(x) - \widetilde{f}(y) = \sum_{j=1}^\infty a_j\, (x^j - y^j), \tag{10.1.8}$$

because the $j = 0$ terms cancel. This implies that

$$\|\widetilde{f}(x) - \widetilde{f}(y)\|_{\mathcal{A}} \le \sum_{j=1}^\infty |a_j|\, \|x^j - y^j\|_{\mathcal{A}}. \tag{10.1.9}$$

It follows that

$$\|\widetilde{f}(x) - \widetilde{f}(y)\|_{\mathcal{A}} \le \sum_{j=1}^\infty j\, |a_j|\, r^{j-1}\, \|x - y\|_{\mathcal{A}}, \tag{10.1.10}$$

because of (10.1.3). This means that (10.1.7) is Lipschitz on (10.1.5), with constant (10.1.6). Of course, if $a_j = 0$ for all but finitely many j, then the completeness of $\mathcal{A}$ is not needed here. If $\sum_{j=0}^\infty a_j\, z^j$ has radius of convergence ρ, as in Section 9.12, and if $0 \le r < \rho$, then it is well known that (10.1.6) converges.

Note that

$$\sum_{j=1}^\infty j\, a_j\, z^{j-1} = \sum_{j=0}^\infty (j+1)\, a_{j+1}\, z^j \tag{10.1.11}$$

is the power series obtained by differentiating $\sum_{j=0}^\infty a_j\, z^j$ formally term-by-term. If (10.1.6) converges, then (10.1.11) converges absolutely for every $z \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|z| \le r$, by the comparison test. If $r > 0$, then it is well known that (10.1.11) is in fact the derivative of $\sum_{j=0}^\infty a_j\, z^j$ when $|z| \le r$, considered as a one-sided derivative in the real case when $|z| = r$. In the complex case, one should use an appropriate complex derivative here, and a suitable interpretation for $|z| = r$. In both cases, one can use the same type of Lipschitz conditions as before to show that the derivative of the sum is equal to the sum of the derivatives.

10.2 Another property of power series

Let $\sum_{j=0}^\infty a_j\, z^j$ be a power series with real or complex coefficients, and suppose that (10.1.6) converges for some $r > 0$, so that $\sum_{j=0}^\infty |a_j|\, r^j$ converges, as before. If $w, z \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|z|, |w| \le r$, then $\sum_{j=0}^\infty a_j\, z^j$ and $\sum_{j=0}^\infty a_j\, w^j$ converge absolutely, and

$$\begin{aligned}(10.2.1)\qquad \sum_{j=0}^\infty a_j\, z^j - \sum_{j=0}^\infty a_j\, w^j &= \sum_{j=1}^\infty a_j\, (z^j - w^j)\\ &= \sum_{j=1}^\infty a_j \Big(\sum_{l=0}^{j-1} z^l\, (z-w)\, w^{j-l-1}\Big),\end{aligned}$$

where the second step is as in (10.1.1). We would like to reexpress this as

$$(10.2.2)\qquad \sum_{j=0}^\infty a_j\, z^j - \sum_{j=0}^\infty a_j\, w^j = (z-w) \sum_{j=1}^\infty a_j \Big(\sum_{l=0}^{j-1} z^l\, w^{j-l-1}\Big).$$

If $z \ne w$, then this follows from (10.2.1), including the convergence of the series on the right. If $z = w$, then the convergence of the series on the right follows from the convergence of (10.1.6).

Put

$$(10.2.3)\qquad b_l(w) = \sum_{j=l+1}^\infty a_j\, w^{j-l-1}$$

for each $l \ge 0$, and note that the series on the right converges absolutely. It is easy to see that

$$(10.2.4)\qquad \sum_{l=0}^\infty b_l(w)\, z^l = \sum_{j=1}^\infty a_j \Big(\sum_{l=0}^{j-1} z^l\, w^{j-l-1}\Big),$$

at least formally, by interchanging the order of summation. This means that

$$(10.2.5)\qquad \sum_{j=0}^\infty a_j\, z^j - \sum_{j=0}^\infty a_j\, w^j = (z-w) \sum_{l=0}^\infty b_l(w)\, z^l,$$

at least formally, by (10.2.2).

Let $c_{j,l}$ be defined for $j, l \ge 0$ by

$$\begin{aligned}(10.2.6)\qquad c_{j,l} &= a_j \quad \text{when } 0 \le l \le j-1\\ &= 0 \quad \text{otherwise.}\end{aligned}$$

Observe that

$$(10.2.7)\qquad \sum_{j=0}^\infty \Big(\sum_{l=0}^\infty |c_{j,l}|\, r^{j-1}\Big) = \sum_{j=1}^\infty \Big(\sum_{l=0}^{j-1} |a_j|\, r^{j-1}\Big) = \sum_{j=1}^\infty j\, |a_j|\, r^{j-1}.$$

This means that

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} |c_{j,l}|\, r^{j-1} = \sum_{j=1}^{\infty} j\, |a_j|\, r^{j-1}. \tag{10.2.8}$$

More precisely, the left side of (10.2.8), as defined in Section 2.4, is equal to the left side of (10.2.7), as in Section 9.8.

Similarly,

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} |c_{j,l}|\, r^{j-1} = \sum_{l=0}^{\infty}\Big(\sum_{j=0}^{\infty} |c_{j,l}|\, r^{j-1}\Big) = \sum_{l=0}^{\infty}\Big(\sum_{j=l+1}^{\infty} |a_j|\, r^{j-1}\Big). \tag{10.2.9}$$

Thus

$$\sum_{l=0}^{\infty}\Big(\sum_{j=l+1}^{\infty} |a_j|\, r^{j-1}\Big) = \sum_{j=1}^{\infty} j\, |a_j|\, r^{j-1}, \tag{10.2.10}$$

by (10.2.8). Note that

$$|b_l(w)| \le \sum_{j=l+1}^{\infty} |a_j|\, r^{j-l-1} \tag{10.2.11}$$

for each $l \ge 0$, so that

$$\sum_{l=0}^{\infty} |b_l(w)|\, r^l \le \sum_{j=1}^{\infty} j\, |a_j|\, r^{j-1}. \tag{10.2.12}$$

Clearly

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} |c_{j,l}|\, |z|^l\, |w|^{j-l-1} \le \sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} |c_{j,l}|\, r^{j-1}, \tag{10.2.13}$$

where the summand on the left is interpreted as being equal to 0 unless $l \le j-1$, even when $w = 0$. Of course, the right side is finite, by (10.2.8). It follows that

$$\sum_{(j,l)\in(\mathbf{Z}_+\cup\{0\})^2} c_{j,l}\, z^l\, w^{j-l-1} \tag{10.2.14}$$

converges in the generalized sense, as in Section 2.8, where again the summand is interpreted as being equal to 0 unless $l \le j - 1$. Both sides of (10.2.4) are equal to (10.2.14), by summing in one index at a time, as in Section 9.9. Thus (10.2.4) holds, which implies that (10.2.5) holds.

10.3 Taking w to be fixed

Let us continue with the same notation and hypotheses as in the previous section, and let us now take w to be fixed. Put $f(z) = \sum_{j=0}^{\infty} a_j\, z^j$ and

$$g(z) = \sum_{l=0}^{\infty} b_l(w)\, z^l, \tag{10.3.1}$$

so that

$$f(z) - f(w) = (z - w)\, g(z), \tag{10.3.2}$$

by (10.2.5).

If $\sum_{j=0}^\infty a_j\, z^j$ has radius of convergence $\rho > 0$, then it is well known that (10.1.6) converges when $0 < r < \rho$. This implies that $\sum_{l=0}^\infty b_l(w)\, z^l$ has radius of convergence at least ρ, because of (10.2.12). More precisely, this works for any $w \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that $|w| < \rho$.

Note that the radius of convergence of $\sum_{j=0}^\infty a_j\, z^j$ is less than or equal to the radius of convergence of $\sum_{l=0}^\infty b_l(w)\, z^l$, because of (10.2.5). Thus the radius of convergence of $\sum_{l=0}^\infty b_l(w)\, z^l$ is equal to ρ too,

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{A}}$. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \le r$, then $\widetilde{f}(x) = \sum_{j=0}^\infty a_j\, x^j$ defines an element of $\mathcal{A}$, as in Section 9.14. Similarly,

$$\widetilde{g}(x) = \sum_{l=0}^\infty b_l(w)\, x^l \tag{10.3.3}$$

defines an element of $\mathcal{A}$ under these conditions. We also have that

$$\widetilde{f}(x) - f(w)\, e_{\mathcal{A}} = (x - w\, e_{\mathcal{A}})\, \widetilde{g}(x). \tag{10.3.4}$$

This could be shown in the same way as before, or obtained from (10.3.2).

If $\sum_{j=0}^\infty a_j\, z^j$ has radius of convergence ρ, then $\widetilde{f}(x)$ and $\widetilde{g}(x)$ can be defined in the same way when $\|x\|_{\mathcal{A}} < \rho$, and (10.3.4) holds. Similarly, if $r_{\mathcal{A}}(x) < \rho$, then $\widetilde{f}(x)$ and $\widetilde{g}(x)$ can be defined in the same way, as in Section 9.14. We also have that (10.3.4) holds in this case. As before, we can take w to be any real or complex number with $|w| < \rho$ here.

Of course, $\widetilde{g}(x)$ and $x - w\, e_{\mathcal{A}}$ commute with each other. If $\widetilde{f}(x) - f(w)\, e_{\mathcal{A}}$ is invertible in $\mathcal{A}$, then it follows that $x - w\, e_{\mathcal{A}}$ is invertible in $\mathcal{A}$, as in Section 6.13. This implies that

$$f(w) \in \sigma_{\mathcal{A}}(\widetilde{f}(x)) \tag{10.3.5}$$

when $w \in \sigma_{\mathcal{A}}(x)$.

It follows that

$$f(\sigma_{\mathcal{A}}(x)) \subseteq \sigma_{\mathcal{A}}(\widetilde{f}(x)). \tag{10.3.6}$$

More precisely, if $\|x\|_{\mathcal{A}} \le r$ and $w \in \sigma_{\mathcal{A}}(x)$, then $|w| \le r$, so that (10.3.5) holds. Remember that the same conclusions were obtained in Section 9.15, under the hypothesis that $\sum_{j=0}^\infty |a_j|\, r^j$ converges, without asking that (10.1.6) converge.

Similarly, if $\sum_{j=0}^\infty a_j\, z^j$ has radius of convergence ρ, $r_{\mathcal{A}}(x) < \rho$, and w is an element of $\sigma_{\mathcal{A}}(x)$, then

$$|w| \le r_{\mathcal{A}}(x) < \rho, \tag{10.3.7}$$

as in Section 6.13. In this case, (10.3.5) holds, for essentially the same reasons as before. This implies that (10.3.6) holds, as before. This is another approach to the analogous statements in Section 9.15.

Let $(V, \|\cdot\|_V)$ be a Banach space over the real or complex numbers, as appropriate, with $V \neq \{0\}$, and suppose now that $\mathcal{A} = \mathcal{BL}(V)$, with the corresponding operator norm. If $T \in \mathcal{BL}(V)$ and $\|T\|_{op} \leq r$, and $\lambda \in \mathbf{R}$ or $\mathbf{C}$ is an eigenvalue or approximate eigenvalue of T, then we saw in Section 9.15 that $f(\lambda)$ is an eigenvalue or approximate eigenvector of $\widetilde{f}(T)$, as appropriate. These statements could also be obtained from (10.3.2) when (10.1.6) converges. Similarly, if $\sum_{j=0}^{\infty} a_j\, z^j$ has radius of convergence $\rho > 0$ and $r_{\mathcal{BL}(V)}(T) < \rho$, then the analogous statements could be obtained from (10.3.4).

10.4 The exponential function

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $x \in \mathcal{A}$, then the *exponential* of x in $\mathcal{A}$ may be defined by the usual power series,

$$\exp x = \exp_{\mathcal{A}} x = \sum_{j=0}^{\infty} (1/j!)\, x^j. \tag{10.4.1}$$

More precisely, the series on the right converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, because

$$\sum_{j=0}^{\infty} (1/j!)\, \|x^j\|_{\mathcal{A}} \leq \sum_{j=0}^{\infty} (1/j!) \|x\|_{\mathcal{A}}^j = \exp \|x\|_{\mathcal{A}}, \tag{10.4.2}$$

using the ordinary exponential function on the right side. This also shows that

$$\|\exp x\|_{\mathcal{A}} \leq \exp \|x\|_{\mathcal{A}}. \tag{10.4.3}$$

Note that $\exp_{\mathcal{A}} 0 = e_{\mathcal{A}}$.

If $x, y \in \mathcal{A}$ commute with each other, then

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j\, y^{n-j} \tag{10.4.4}$$

for each nonnegative integer n, with suitable interpretations when $n = 0$, by the binomial theorem. This implies that

$$\begin{aligned} \exp(x+y) &= \sum_{n=0}^{\infty} (1/n!)\, (x+y)^n \\ &= \sum_{n=0}^{\infty} \Big(\sum_{j=0}^{n} (1/j!)\, (1/(n-j)!)\, x^j\, y^{n-j} \Big). \end{aligned} \tag{10.4.5}$$

The right side corresponds to the Cauchy product of the series defining $\exp x$ and $\exp y$, as in Section 9.13. It follows that

$$\exp(x+y) = (\exp x)\, (\exp y) \tag{10.4.6}$$

under these conditions, because these series converge absolutely, as before.

In particular,

$$(\exp x)\,(\exp(-x)) = (\exp(-x))\,(\exp x) = e_{\mathcal{A}} \tag{10.4.7}$$

for every $x \in \mathcal{A}$. This means that $\exp x$ is invertible in $\mathcal{A}$, with

$$(\exp x)^{-1} = \exp(-x). \tag{10.4.8}$$

Let $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ be another Banach algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{B}}$. Also let ϕ be an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ that is bounded as a linear mapping and satisfies $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. If $x \in \mathcal{A}$, then it is easy to see that

$$\phi(\exp_{\mathcal{A}} x) = \exp_{\mathcal{B}} \phi(x). \tag{10.4.9}$$

This works as well when ϕ is an opposite algebra homomorphism, as in Section 6.4. In the complex case, one may wish to consider conjugate-linear mappings too, which may be considered as real-linear in particular.

Let $x \mapsto x^*$ be an algebra involution on $\mathcal{A}$, as in Section 6.4, which may be conjugate-linear in the complex case. Suppose that this involution is bounded as a linear or conjugate-linear mapping from $\mathcal{A}$ into itself, as appropriate, so that

$$\|x^*\|_{\mathcal{A}} \le C\,\|x\|_{\mathcal{A}} \tag{10.4.10}$$

for some nonnegative real number C and every $x \in \mathcal{A}$. Thus

$$(\exp x)^* = \exp(x^*) \tag{10.4.11}$$

for every $x \in \mathcal{A}$, as in (10.4.9). If x is anti-self-adjoint with respect to this involution, in the sense that $x^* = -x$, then

$$(\exp x)^* = \exp(x^*) = \exp(-x) = (\exp x)^{-1}. \tag{10.4.12}$$

This means that $\exp x$ is an element of the subgroup $U(\mathcal{A})$ of the group $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$ defined in Section 7.5, using this involution.

10.5 Algebra derivations

Let $\mathcal{A}$ be an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$. A linear mapping δ from $\mathcal{A}$ into itself is said to be a *derivation* if

$$\delta(x\,y) = \delta(x)\,y + x\,\delta(y) \tag{10.5.1}$$

for every $x, y \in \mathcal{A}$, as on p4 of [97], p7 of [172], and p2 of [173]. The derivations of $\mathcal{A}$ form a linear subspace of the space $\mathcal{L}(\mathcal{A})$ of all linear mappings from $\mathcal{A}$ into itself. If δ_1, δ_2 are derivations on $\mathcal{A}$, then one can check that

$$\delta_1 \circ \delta_2 - \delta_2 \circ \delta_1 \tag{10.5.2}$$

is a derivation on $\mathcal{A}$ as well, as in [97, 172, 173].

Let δ be a derivation on $\mathcal{A}$, and let n be a nonnegative integer. If $x, y \in \mathcal{A}$, then it is well known that

$$\delta^n(x\,y) = \sum_{j=0}^{n} \binom{n}{j} \delta^j(x)\, \delta^{n-j}(y). \tag{10.5.3}$$

This is a version of the *Leibniz rule.*

Let $\|\cdot\|_{\mathcal{A}}$ be a norm on $\mathcal{A}$, and suppose that multiplication on $\mathcal{A}$ is bounded as a bilinear mapping from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$ with respect to $\|\cdot\|_{\mathcal{A}}$, so that

$$\|x\,y\|_{\mathcal{A}} \le C\, \|x\|_{\mathcal{A}}\, \|y\|_{\mathcal{A}} \tag{10.5.4}$$

for some nonnegative real number C and all $x, y \in \mathcal{A}$. Suppose also that $\mathcal{A}$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, so that the space $\mathcal{BL}(\mathcal{A})$ of bounded linear mappings from $\mathcal{A}$ into itself is a Banach algebra with respect to the corresponding operator norm $\|\cdot\|_{op}$. If δ is bounded as a linear mapping from $\mathcal{A}$ into itself, then $\exp \delta$ may be defined as a bounded linear mapping from $\mathcal{A}$ into itself too, as in the previous section.

If $x, y \in \mathcal{A}$, then

$$\begin{aligned} (\exp\delta)(x\,y) &= \sum_{n=0}^{\infty} (1/n!)\, \delta^n(x\,y) \\ &= \sum_{n=0}^{\infty} \Big(\sum_{j=0}^{n} (1/j!)\, (1/(n-j)!)\, \delta^j(x)\, \delta^{n-j}(y) \Big), \end{aligned} \tag{10.5.5}$$

using (10.5.3) in the second step. The right side corresponds to the Cauchy product of

$$(\exp\delta)(x) = \sum_{j=0}^{\infty} (1/j!)\, \delta^j(x) \text{ and } (\exp\delta)(y) = \sum_{l=0}^{\infty} (1/l!)\, \delta^l(y), \tag{10.5.6}$$

as in Section 9.11. Note that these series converge absolutely with respect to $\|\cdot\|_{\mathcal{A}}$, because δ is bounded on $\mathcal{A}$. It follows that

$$(\exp\delta)(x\,y) = ((\exp\delta)(x))\, ((\exp\delta)(y)), \tag{10.5.7}$$

as before. This shows that $\exp \delta$ is an automorphism of $\mathcal{A}$, as an algebra in the strict sense.

10.6 Commutators and conjugations

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers. If $a \in \mathcal{A}$, then put

$$\delta_a(x) = a\,x - x\,a \tag{10.6.1}$$

for every $x \in \mathcal{A}$. This defines a linear mapping from $\mathcal{A}$ into itself. One can check that

$$\delta_a \text{ is a derivation on } \mathcal{A}. \tag{10.6.2}$$

This type of derivation in this case is called an *inner derivation* of $\mathcal{A}$.

Remember that L_a, R_a are the left and right multiplication operators on $\mathcal{A}$ associated to a as in Sections 6.3 and 6.4. Note that

$$\delta_a = L_a - R_a. \tag{10.6.3}$$

We also have that

$$L_a \circ R_a = R_a \circ L_a, \tag{10.6.4}$$

because $\mathcal{A}$ is associative.

Suppose that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Remember that L_a and R_a are bounded linear mappings on $\mathcal{A}$, with

$$\|L_a\|_{op} = \|R_a\|_{op} = \|a\|_{\mathcal{A}}, \tag{10.6.5}$$

as before. This implies that δ_a is bounded as a linear mapping from $\mathcal{A}$ into itself, with

$$\|\delta_a\|_{op} \le 2\,\|a\|_{\mathcal{A}}. \tag{10.6.6}$$

Of course, the space $\mathcal{BL}(\mathcal{A})$ of bounded linear mappings from $\mathcal{A}$ into itself is a Banach algebra with respect to the operator norm, because $\mathcal{A}$ is complete, by hypothesis. Thus the exponentials of L_a, R_a, and δ_a may be defined as elements of $\mathcal{BL}(\mathcal{A})$ as in Section 10.4. In fact, we have that

$$\begin{aligned} \exp \delta_a = \exp(L_a - R_a) &= (\exp L_a) \circ (\exp(-R_a)) \\ &= (\exp L_a) \circ (\exp R_a)^{-1}, \end{aligned} \tag{10.6.7}$$

because of (10.6.4).

Remember that $a \mapsto L_a$ is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{BL}(\mathcal{A})$, with $L_{e_{\mathcal{A}}}$ equal to the identity mapping $I_{\mathcal{A}}$ on $\mathcal{A}$, as in Section 6.3. Similarly, $a \mapsto R_a$ is an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{BL}(\mathcal{A})$, with $R_{e_{\mathcal{A}}} = I_{\mathcal{A}}$, as in Section 6.4. This implies that

$$\exp L_a = \exp_{\mathcal{BL}(\mathcal{A})} L_a = L_{\exp_{\mathcal{A}} a} \tag{10.6.8}$$

and

$$\exp R_a = \exp_{\mathcal{BL}(\mathcal{A})} R_a = R_{\exp_{\mathcal{A}} a}, \tag{10.6.9}$$

as in Section 10.4. If $x \in \mathcal{A}$, then we get that

$$(\exp \delta_a)(x) = (\exp_{\mathcal{BL}(\mathcal{A})} \delta_a)(x) = (\exp_{\mathcal{A}} a)\, x\, (\exp_{\mathcal{A}} a)^{-1}, \tag{10.6.10}$$

by (10.6.7). This corresponds to Exercise 15 on p260 of [162].

10.7 A criterion for local invertibility

Let $(V, \|\cdot\|_V)$ be a Banach space over the real or complex numbers, let $v_0 \in V$ and $r_0 > 0$ be given, and remember that $\overline{B}(v_0, r_0)$ is the closed ball in V centered at v_0 with radius r_0 with respect to the metric associated to $\|\cdot\|_V$, as in Section 1.2. Also let g be a mapping from $\overline{B}(v_0, r_0)$ into V, and put

$$\phi(v) = v - g(v) \tag{10.7.1}$$

on $\overline{B}(v_0, r_0)$. Suppose that

$$\phi \text{ is Lipschitz with constant } c_0 \geq 0 \text{ on } \overline{B}(v_0, r_0), \tag{10.7.2}$$

with respect to the metric on V associated to $\|\cdot\|_V$ and its restriction to $\overline{B}(v_0, r_0)$. This implies that

$$g \text{ is Lipschitz with constant } 1 + c_0 \text{ on } \overline{B}(v_0, r_0), \tag{10.7.3}$$

as in Section 2.1. Suppose too that $c_0 < 1$, and note that

$$(1 - c_0)\,\|v - w\|_V \leq \|g(v) - g(w)\|_V \tag{10.7.4}$$

for every $v, w \in \overline{B}(v_0, r_0)$, as in Section 7.11.

If $y \in V$, then put

$$\phi_y(v) = \phi(v) + y = v - g(v) + y \tag{10.7.5}$$

on $\overline{B}(v_0, r_0)$. Clearly

$$\phi_y \text{ is Lipschitz with constant } c_0 \text{ on } \overline{B}(v_0, r_0), \tag{10.7.6}$$

because of (10.7.2). Observe that

$$\begin{aligned} \phi_y(v) - v_0 = \phi(v) + y - v_0 &= \phi(v) - \phi(v_0) + \phi(v_0) + y - v_0 \\ &= \phi(v) - \phi(v_0) + y - g(v_0) \end{aligned} \tag{10.7.7}$$

on $\overline{B}(v_0, r_0)$. This implies that

$$\begin{aligned} \|\phi_y(v) - v_0\|_V &\leq \|\phi(v) - \phi(v_0)\|_V + \|y - g(v_0)\|_V \\ &\leq c_0\,\|v - v_0\|_V + \|y - g(v_0)\|_V \leq c_0\,r_0 + \|y - g(v_0)\|_V \end{aligned} \tag{10.7.8}$$

on $\overline{B}(v_0, r_0)$.

Suppose that

$$\|y - g(v_0)\|_V \leq (1 - c_0)\,r_0, \tag{10.7.9}$$

so that

$$\|\phi_y(v) - v_0\|_V \leq r_0 \tag{10.7.10}$$

on $\overline{B}(v_0, r_0)$, by (10.7.8). Equivalently, this means that

$$\phi_y(\overline{B}(v_0, r_0)) \subseteq \overline{B}(v_0, r_0). \tag{10.7.11}$$

Note that $\overline{B}(v_0, r_0)$ is complete as a metric space with respect to the restriction of the metric on V associated to $\|\cdot\|_V$, because V is complete, by hypothesis, and $\overline{B}(v_0, r_0)$ is a closed set in V, as in Section 1.6. Thus the contraction mapping theorem implies that there is a unique point $v_y \in \overline{B}(v_0, r_0)$ such that

$$\phi_y(v_y) = v_y, \tag{10.7.12}$$

as in Section 7.12.

Of course, (10.7.12) is the same as saying that

$$g(v_y) = y, \tag{10.7.13}$$

by the definition (10.7.5) of ϕ_y. This shows that

$$\overline{B}(v_0, (1 - c_0)\, r_0) \subseteq g(\overline{B}(v_0, r_0)) \tag{10.7.14}$$

under these conditions. We also have that

$$(1 - c_0)\, \|v_y - v_0\|_V \leq \|g(v_y) - g(v_0)\|_V = \|y - g(v_0)\|_V, \tag{10.7.15}$$

by (10.7.4). This implies that

$$\overline{B}(v_0, (1 - c_0)\, r) \subseteq g(\overline{B}(v_0, r)) \tag{10.7.16}$$

when $0 < r \leq r_0$, which could be obtained from the same argument as before too. Similarly,

$$B(v_0, (1 - c_0)\, r) \subseteq g(B(v_0, r)) \tag{10.7.17}$$

when $0 < r \leq r_0$, where $B(v_0, r)$ is the open ball in V centered at v_0 with radius r with respect to the metric associated to $\|\cdot\|_V$.

This basically corresponds to part of the inverse function theorem for functions defined on subsets of Banach spaces, as in Theorem 10.39 on p252 of [162]. More precisely, let T be a bounded linear mapping from V into itself with bounded inverse. Suppose that f is a mapping from $\overline{B}(v_0, r_0)$ into V such that

$$T - f \text{ is Lipschitz with constant } c_1 \text{ on } \overline{B}(v_0, r_0), \tag{10.7.18}$$

for some $c_1 \geq 0$. This implies that

$$I - T^{-1} \circ f = T^{-1} \circ (T - f) \tag{10.7.19}$$

is Lipschitz with constant $c_1\, \|T^{-1}\|_{op}$ on $\overline{B}(v_0, r_0)$. If

$$c_1\, \|T^{-1}\|_{op} < 1, \tag{10.7.20}$$

then one can apply the previous remarks to $T^{-1} \circ f$, to get analogous properties of f.

10.8 More on local invertibility

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let $r_0 > 0$ be given. We would like to consider some of the remarks in the previous section with $V = \mathcal{A}$, $v_0 = 0$, and where g is given by a power series

$$g(x) = \sum_{j=0}^{\infty} a_j\, x^j, \tag{10.8.1}$$

with real or complex coefficients, as appropriate. Suppose that

$$\sum_{j=0}^{\infty} |a_j|\, r_0^j \tag{10.8.2}$$

converges as an infinite series of nonnegative real numbers, so that the right side of (10.8.1) converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$ when $x \in \mathcal{A}$ satisfies $\|x\|_{\mathcal{A}} \le r_0$, as in Section 9.14. More precisely, this defines g as an $\mathcal{A}$-valued function on the closed ball $\overline{B}(0, r_0)$ in $\mathcal{A}$ centered at 0 with radius r_0 with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. We have also seen that g is bounded and uniformly continuous on $\overline{B}(0, r_0)$, with respect to the metric on $\mathcal{A}$ associated to $\|\cdot\|_{\mathcal{A}}$ and its restriction to $\overline{B}(0, r_0)$.

In fact, we shall ask that

$$\sum_{j=1}^{\infty} j\, |a_j|\, r_0^{j-1} \tag{10.8.3}$$

converges, which implies that (10.8.1) converges. In this case, g is Lipschitz on $\overline{B}(0, r_0)$, with constant (10.8.3), as in Section 10.1. Put

$$\phi(x) = x - g(x) = -a_0\, e_{\mathcal{A}} + (1 - a_1)\, x - \sum_{j=2}^{\infty} a_j\, x^j \tag{10.8.4}$$

on $\overline{B}(0, r_0)$, as in (10.7.1). Observe that ϕ is Lipschitz with constant

$$|a_1 - 1| + \sum_{j=2}^{\infty} j\, |a_j|\, r_0^{j-1} \tag{10.8.5}$$

on $\overline{B}(0, r_0)$, as in Section 10.1. Thus we would like to have that

$$|a_1 - 1| + \sum_{j=2}^{\infty} j\, |a_j|\, r_0^{j-1} < 1, \tag{10.8.6}$$

as before.

If $y \in \mathcal{A}$, then we put

$$\phi_y(x) = \phi(x) + y = x - g(x) + y \tag{10.8.7}$$

on $\overline{B}(0, r_0)$, as in (10.7.5). Note that ϕ_y is Lipschitz with constant (10.8.5) on $\overline{B}(0, r_0)$ too, as before. Equivalently,

$$\phi_y(x) = y - a_0\, e_{\mathcal{A}} + (1 - a_1)\, x - \sum_{j=2}^{\infty} a_j\, x^j \tag{10.8.8}$$

on $\overline{B}(0, r_0)$, so that

$$\|\phi_y(x)\|_{\mathcal{A}} \le \|y - a_0\, e_{\mathcal{A}}\|_{\mathcal{A}} + |1 - a_1|\, r_0 + \sum_{j=2}^{\infty} |a_j|\, r_0^j \tag{10.8.9}$$

on $\overline{B}(0, r_0)$. If

$$\|y - a_0\, e_{\mathcal{A}}\|_{\mathcal{A}} + |1 - a_1|\, r_0 + \sum_{j=2}^{\infty} |a_j|\, r_0^j \le r_0, \tag{10.8.10}$$

then we get that

$$\phi_y(\overline{B}(0, r_0)) \subseteq \overline{B}(0, r_0). \tag{10.8.11}$$

If (10.8.6) holds as well, then the contraction mapping theorem implies that ϕ_y has a unique fixed point in $\overline{B}(0, r_0)$, as in Section 7.12.

We may often be interested in having

$$a_1 = 1, \tag{10.8.12}$$

so that (10.8.6) reduces to

$$\sum_{j=2}^{\infty} j\, |a_j|\, r_0^{j-1} < 1. \tag{10.8.13}$$

Similarly, (10.8.10) reduces to

$$\|y - a_0\, e_{\mathcal{A}}\|_{\mathcal{A}} + \sum_{j=2}^{\infty} |a_j|\, r_0^j \le r_0 \tag{10.8.14}$$

in this case. If

$$\sum_{j=2}^{\infty} |a_j|\, r_0^{j-1} < 1, \tag{10.8.15}$$

then (10.8.14) holds when

$$\|y - a_0\, e_{\mathcal{A}}\|_{\mathcal{R}} \le \Big(1 - \sum_{j=2}^{\infty} |a_j|\, r_0^{j-1}\Big)\, r_0. \tag{10.8.16}$$

Note that (10.8.15) holds when (10.8.13) holds. Under these conditions, we can get (10.8.13) by using a smaller positive real number in place of r_0, if necessary.

10.9 More on the exponential function

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and let r be a positive real number. The restriction of the exponential function to the closed ball $\overline{B}(0, r)$ in $\mathcal{A}$ centered at 0 with radius r is Lipschitz with constant

$$\sum_{j=1}^{\infty} j\,(1/j!)\,r^{j-1}, \tag{10.9.1}$$

as in Section 10.1. Note that (10.9.1) is equal to

$$\sum_{j=1}^{\infty}(1/(j-1)!)\,r^{j-1} = \sum_{j=0}^{\infty}(1/j!)\,r^j = \exp r. \tag{10.9.2}$$

Similarly, the restriction of

$$(\exp x) - x \tag{10.9.3}$$

to $\overline{B}(0, r)$ is Lipschitz with constant

$$\sum_{j=2}^{\infty} j\,(1/j!)\,r^{j-1} = \sum_{j=2}^{\infty}(1/(j-1)!)\,r^{j-1} = \sum_{j=1}^{\infty}(1/j!)\,r^j = (\exp r) - 1. \tag{10.9.4}$$

Of course, this tends to 0 as $r \to 0$. It follows that

$$\exp x \text{ is open at } 0, \tag{10.9.5}$$

as a mapping from $\mathcal{A}$ into itself, as in Section 10.7.

Remember that $G(\mathcal{A})$ is the group of invertible elements of $\mathcal{A}$, as in Section 6.6. Thus

$$\exp \mathcal{A} = \{\exp x : x \in \mathcal{A}\} \tag{10.9.6}$$

is a subset of $G(\mathcal{A})$. Using (10.9.5), we get that

$$e_{\mathcal{A}} \text{ is an element of the interior of } \exp \mathcal{A}. \tag{10.9.7}$$

Let $\Gamma(\mathcal{A})$ be the subgroup of $G(\mathcal{A})$ generated by $\exp \mathcal{A}$, as in the proof of Theorem 10.44 on p258 of [162]. Clearly $e_{\mathcal{A}}$ is an element of the interior of $\Gamma(\mathcal{A})$, because $\exp \mathcal{A} \subseteq \Gamma(\mathcal{A})$. This implies that

$$\Gamma(\mathcal{A}) \text{ is an open set in } \mathcal{A}, \tag{10.9.8}$$

as in [162].

If $\mathcal{A}$ is commutative, then it is easy to see that

$$\exp x \text{ is open at every point in } \mathcal{A}, \tag{10.9.9}$$

using (10.9.5). This means that

$$\exp x \text{ is an open mapping on } \mathcal{A}, \tag{10.9.10}$$

as in Section 7.13. This is related to a remark on p257 of [162]. In this case, $\exp \mathcal{A}$ is a subgroup of $G(\mathcal{A})$, so that

$$\Gamma(\mathcal{A}) = \exp \mathcal{A}. \tag{10.9.11}$$

10.10 The subgroup $G_1(\mathcal{A})$ of $G(\mathcal{A})$

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Remember that the group $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$ is an open set in $\mathcal{A}$, as in Section 6.6. Let $G_1(\mathcal{A})$ be the connected component of $G(\mathcal{A})$ that contains $e_{\mathcal{A}}$, as on p257 of [162]. Equivalently, $G_1(\mathcal{A})$ is the union of all of the connected subsets of $G(\mathcal{A})$ that contain $e_{\mathcal{A}}$.

Remember that open balls in $\mathcal{A}$ are convex, as in Section 1.2. This means that open balls in $\mathcal{A}$ are path connected. It follows that $G(\mathcal{A})$ is locally path connected, because $G(\mathcal{A})$ is an open set in $\mathcal{A}$. This implies that the connected components of $G(\mathcal{A})$ are the same as the path-connected components of $G(\mathcal{A})$, by standard arguments. We also get that the connected components of $G(\mathcal{A})$ are open sets.

Thus $G_1(\mathcal{A})$ is the same as the path-connected component of $G(\mathcal{A})$ that contains $e_{\mathcal{A}}$. This is the set of invertible elements of $\mathcal{A}$ that can be connected to $e_{\mathcal{A}}$ by a continuous path in $G(\mathcal{A})$. Note that

$$G_1(\mathcal{A}) \text{ is an open set in } \mathcal{A}, \tag{10.10.1}$$

as in part (a) of Theorem 10.44 on p258 of [162].

One can check that

$$G_1(\mathcal{A}) \text{ is a subgroup of } G(\mathcal{A}), \tag{10.10.2}$$

using the characterization of $G_1(\mathcal{A})$ in terms of continuous paths in $G(\mathcal{A})$ starting at $e_{\mathcal{A}}$. This is another part of part (a) of Theorem 10.44 on p258 of [162].

In fact,

$$G_1(\mathcal{A}) \text{ is a normal subgroup of } G(\mathcal{A}), \tag{10.10.3}$$

which is the third part of part (a) of Theorem 10.44 on p258 of [162]. This can also be obtained from the characterization of $G_1(\mathcal{A})$ in terms of continuous paths in $G(\mathcal{A})$ starting at $e_{\mathcal{A}}$.

If $x \in \mathcal{A}$, then

$$\exp(t\,x),\ 0 \le t \le 1, \tag{10.10.4}$$

defines a continuous path from $e_{\mathcal{A}}$ to $\exp x$ in $G(\mathcal{A})$. This implies that $\exp x$ is an element of $G_1(\mathcal{A})$, so that

$$\exp \mathcal{A} \subseteq G_1(\mathcal{A}), \tag{10.10.5}$$

as in part (b) of Theorem 10.44 on p258 of [162].

Let $\Gamma(\mathcal{A})$ be the subgroup of $G(\mathcal{A})$ generated by $\exp \mathcal{A}$, as in the previous section. Observe that

$$\Gamma(\mathcal{A}) \subseteq G_1(\mathcal{A}), \tag{10.10.6}$$

because of (10.10.2) and (10.10.5). Part (b) of Theorem 10.44 on p258 of [162] states that

$$\Gamma(\mathcal{A}) = G_1(\mathcal{A}). \tag{10.10.7}$$

Remember that $\Gamma(\mathcal{A})$ is an open set in $\mathcal{A}$, as in (10.9.8). This implies that all of the cosets of $\Gamma(\mathcal{A})$ in $G(\mathcal{A})$ are open sets. It follows that $G(\mathcal{A}) \setminus \Gamma(\mathcal{A})$ is an open set, because it is the union of all of the cosets of $\Gamma(\mathcal{A})$ in $G(\mathcal{A})$ other than $\Gamma(\mathcal{A})$. One can use this to get that

$$G_1(\mathcal{A}) \subseteq \Gamma(\mathcal{A}), \tag{10.10.8}$$

because $G_1(\mathcal{A})$ is a connected subset of $G(\mathcal{A})$ that contains $e_{\mathcal{A}}$.

If $\mathcal{A}$ is commutative, then

$$\exp \mathcal{A} = G_1(\mathcal{A}), \tag{10.10.9}$$

by (10.9.11) and (10.10.7). This is part (c) of Theorem 10.44 on p258 of [162].

10.11 Some induced homomorphisms

Let $\mathcal{A}$, $\mathcal{B}$ be associative algebras, both real or both complex, with multiplicative identity elements $e_{\mathcal{A}}$, $e_{\mathcal{B}}$, respectively. Also let ϕ be an algebra homomrohism from $\mathcal{A}$ into $\mathcal{B}$ such that $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. If $x \in \mathcal{A}$ is invertible, then $\phi(x)$ is invertible in $\mathcal{B}$, with

$$\phi(x^{-1}) = \phi(x)^{-1}. \tag{10.11.1}$$

Thus

$$\phi(G(\mathcal{A})) \subseteq G(\mathcal{B}), \tag{10.11.2}$$

and the restriction of ϕ to $G(\mathcal{A})$ defines a group homomorphism into $G(\mathcal{B})$.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$, $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ are Banach algebras, and that ϕ is bounded as a linear mapping from $\mathcal{A}$ into $\mathcal{B}$. This means that ϕ is continuous, and one can use that to get that

$$\phi(G_1(\mathcal{A})) \subseteq G_1(\mathcal{B}). \tag{10.11.3}$$

Remember that $G_1(\mathcal{A})$, $G_1(\mathcal{B})$ are normal subgroups of $G(\mathcal{A})$, $G(\mathcal{B})$, as in the previous section. Thus the corresponding quotient groups $G(\mathcal{A})/G_1(\mathcal{A})$ and $G(\mathcal{B})/G_1(\mathcal{B})$ may be defined in the usual way. It follows from (10.11.3) that

(10.11.4) ϕ induces a group homomormphism from $G(\mathcal{A})/G_1(\mathcal{A})$ into $G(\mathcal{B})/G_1(\mathcal{B})$.

Let V be a vector space over the real or complex numbers, and let X, Y be nonempty sets. If ψ is a mapping from X into Y, then

$$\Psi(f) = f \circ \psi \tag{10.11.5}$$

defines a mapping from the space of all V-valued functions on Y into the space of all V-valued functions on X. More precisely, the spaces of V-valued functions on X and Y are vector spaces over the real or complex numbers, as appropriate, with respect to pointwise addition and scalar multiplication, and Ψ is a linear mapping. It is easy to see that

(10.11.6) $$\Psi \text{ is one-to-one when } \psi(X) = Y.$$

Let $\|\cdot\|_V$ be a norm on V. If f is a bounded V-valued function on Y, then (10.11.5) is a bounded V-valued function on X, so that the restriction of Ψ to $\ell^\infty(Y, V)$ defines a linear mapping into $\ell^\infty(X, V)$. Observe that

(10.11.7) $$\|\Psi(f)\|_{\ell^\infty(X,V)} \le \|f\|_{\ell^\infty(Y,V)}$$

for every $f \in \ell^\infty(Y, V)$. More precisely, the operator norm of the restriction of Ψ to $\ell^\infty(Y, V)$ is equal to 1, with respect to the corresponding supremum norms. If $\psi(X) = Y$, then

(10.11.8) $$\|\Psi(f)\|_{\ell^\infty(X,V)} = \|\Psi(f)\|_{\ell^\infty(Y,V)}$$

for every $f \in \ell^\infty(Y, V)$.

Suppose now that X, Y are metric or topological spaces, and that ψ is continuous. If f is a continuous mapping from Y into V, then (10.11.5) is a continuous mapping from X into V. Thus the restriction of Ψ to $C(Y, V)$ defines a linear mapping into $C(X, V)$. Note that

(10.11.9) $$\Psi \text{ is one-to-one on } C(Y, V) \text{ when } \psi(X) \text{ is dense in } Y.$$

Similarly, if f is bounded and continuous on Y, then (10.11.5) is bounded and continuous on X, so that the restriction of Ψ to $C_b(Y, V)$ is a bounded linear mapping into $C_b(X, V)$ with respect to the supremum norm. If $\psi(X)$ is dense in Y, then (10.11.8) holds for every $f \in C_b(Y, V)$.

Let $\mathcal{A}_0$ be an algebra in the strict sense over the real or complex numbers, so that the spaces of $\mathcal{A}_0$-valued functions on X and Y are algebras in the strict sense over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to pointwise multiplication of functions, as in Section 7.9. In this case, (10.11.5) defines an algebra homomorphism from the space of $\mathcal{A}_0$-valued functions on Y into the space of $\mathcal{A}_0$-valued functions on X. Suppose that $\mathcal{A}_0$ has a multiplicative identity element $e_{\mathcal{A}_0}$, so that the functions on X and Y equal to $e_{\mathcal{A}_0}$ at every point are the multiplicative identity elements in the algebras of $\mathcal{A}_0$-valued functions on X and Y, respectively, as before. Of course, Ψ sends the function on Y equal to $e_{\mathcal{A}_0}$ at every point to the function on X equal to $e_{\mathcal{A}_0}$ at every point.

10.12 Some continuity conditions

Let X, Y be nonempty metric or topological spaces, and let (Z, d_Z) be a metric space. Also let F be a mapping from $X \times Y$ into Z. If $y \in Y$, then

(10.12.1) $$F_y(x) = F(x, y)$$

defines a mapping from X into Z. Thus

$$y \mapsto F_y \tag{10.12.2}$$

defines a mapping from Y into the space of mappings from X into Z. Conversely, every mapping from Y into the space of mappings from X into Z corresponds to a mapping from $X \times Y$ into Z in this way.

Similarly, if $x \in X$, then $F(x, y)$ may be considered as a mapping from Y into Z. Consider the collection

$$\mathcal{E} = \{F(x, \cdot) : x \in X\} \tag{10.12.3}$$

of all of the mappings from Y into Z that correspond to some $x \in X$ in this way. This collection is said to be *equicontinuous* at a point $y_0 \in Y$ if for every $\epsilon > 0$ there is an open set $V \subseteq Y$ such that $y_0 \in V$ and

$$d_Z(F(x, y), F(x, y_0)) < \epsilon \tag{10.12.4}$$

for every $x \in X$ and $y \in V$. This condition clearly implies that $F(x, y)$ is continuous as a function of y on Y at y_0 for every $x \in X$. If X has only finitely many elements, and if $F(x, y)$ is continuous in y at y_0 for every $x \in X$, then $\mathcal{E}$ is equicontinuous at y_0.

Suppose for the moment that

$$F_y \text{ is bounded as a mapping from } X \text{ into } Z \tag{10.12.5}$$

for each $y \in Y$. Remember that the space $\mathcal{B}(X, Z)$ of all mappings from X into Z is a metric space with respect to the supremum metric associated to d_Z. One can check that

$$\mathcal{E} \text{ is equicontinuous at } y_0 \tag{10.12.6}$$

if and only if (10.12.2) is continuous at y_0 as a mapping from Y into $\mathcal{B}(X, Z)$, with respect to the supremum metric on $\mathcal{B}(X, Z)$.

Suppose for the moment again that (10.12.6) holds. Let $x_0 \in X$ be given, and suppose also that

$$F_{y_0}(x) = F(x, y_0) \text{ is continuous in } x \text{ at } x_0, \tag{10.12.7}$$

as a mapping from X into Z. Under these conditions, one can verify that

$$F(x, y) \text{ is continuous at } (x_0, y_0), \tag{10.12.8}$$

as a mapping from $X \times Y$ into Z, with respect to the product topology on $X \times Y$. If X and Y are both metric spaces, then one can use a suitable product metric on $X \times Y$, as in Section 5.11.

Let $\epsilon > 0$ be given. If (10.12.8) holds, then there are open sets $U(x_0) \subseteq X$ and $V(x_0) \in Y$ such that $x_0 \in U(x_0)$, $y_0 \in V(x_0)$, and

$$d_Z(F(x, y), F(x_0, y_0)) < \epsilon/2 \tag{10.12.9}$$

for every $x \in U(x_0)$ and $y \in V(x_0)$. This implies that

$$\begin{aligned} d_Z(F(x,y), F(x,y_0)) &\le d_Z(F(x,y), F(x_0,y_0)) + d_Z(F(x_0,y_0), F(x,y_0)) \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned} \tag{10.12.10}$$

for every $x \in U(x_0)$ and $y \in V(x_0)$.

Suppose that (10.12.8) holds for every $x_0 \in X$. If X is compact, then X can be covered by finitely many open sets $U(x_0)$ as in the preceding paragraph. If we take V to be the intersection of the finitely many corresponding open sets $V(x_0)$ in Y, then we get that (10.12.6) holds.

If X and Y are both metric spaces, then one can use a suitable product metric on $X \times Y$, as before. If X and Y are both compact, then $X \times Y$ is compact, by Tychonoff's theorem. Under these conditions, the continuity of F on $X \times Y$ implies that F is uniformly continuous. In particular, this implies that $\mathcal{E}$ is equicontinuous at every $y_0 \in Y$.

10.13 Some remarks about homotopies

Let X and Z be nonempty metric or topological spaces, and let f, g be continuous mappings from X into Z. A *homotopy* between f and g is a continuous mapping from $X \times [0,1]$ into Z that corresponds to f on $X \times \{0\}$ and to g on $X \times \{1\}$. This uses the product topology on $X \times [0,1]$ associated to the topology induced on $[0,1]$ by the standard topology on the real line. If X is a metric space, then one can use a suitable product metric on $X \times [0,1]$, as in Section 5.11.

In particular, a homotopy between f and g can be used to define a mapping from $[0,1]$ into the space $C(X,Z)$ of continuous mappings from X into Z that is equal to f at 0 and to g at 1. The continuity of the homotopy as a mapping from $X \times [0,1]$ into Z may be considered as a continuity condition on this mapping from $[0,1]$ into $C(X,Z)$.

Suppose that (Z, d_Z) is a metric space, so that the space $C_b(X,Z)$ of bounded continuous mappings from X into Z is a metric space with respect to the supremum metric. A mapping from $[0,1]$ into $C_b(X,Z)$ determines a mapping from $X \times [0,1]$ into Z, as before. The continuity of this mapping from $[0,1]$ into $C_b(X,Y)$ with respect to the supremum metric on $C_b(X,Z)$ is equivalent to the equicontinuity condition for the corresponding mapping from $X \times [0,1]$ into Z at every point in $Y = [0,1]$ discussed in the previous section. It follows that a continuous mapping from $[0,1]$ into $C_b(X,Y)$ determines a continuous mapping from $X \times [0,1]$ into Z, with respect to the usual product topology on $X \times [0,1]$, as before.

Suppose that X is compact, so that every continuous mapping from X into Z is bounded. In this case, a continuous mapping from $X \times [0,1]$ into Z corresponds to a continuous mapping from $[0,1]$ into $C_b(X,Z)$. More precisely, continuous mappings from $X \times [0,1]$ into Z satisfy the equicontinuity condition discussed in the previous section at every point in $[0,1]$, because X is compact. This implies that the corresponding mapping from $[0,1]$ into $C_b(X,Z)$ is continuous, as before.

Let $(\mathcal{A}_0, \|\cdot\|_{\mathcal{A}_0})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}_0}$. If X is any metric or topological space, then $C(X,\mathcal{A}_0)$ is an associative algebra over the real or complex numbers, as appropriate, as in Section 7.9. Remember that $f \in C(X,\mathcal{A}_0)$ is invertible if and only if $f(x)$ is invertible in $\mathcal{A}_0$ for every $x \in X$. Thus

$$G(C(X,\mathcal{A}_0)) = C(X, G(\mathcal{A}_0)), \tag{10.13.1}$$

where $G(\mathcal{A}_0)$ is considered as a metric space, using the resptriction of the metric on $\mathcal{A}_0$ associated to $\|\cdot\|_{\mathcal{A}_0}$ to $G(\mathcal{A}_0)$.

The space $C_b(X,\mathcal{A}_0)$ of all continuous $\mathcal{A}_0$-valued functions on X is a Banach algebra with respect to the supremum norm, as in Sections 7.8 and 7.9. Suppose that X is compact again, so that this is the same as $C(X,\mathcal{A}_0)$. In this case, a continuous path in (10.13.1) corresponds to a homotopy between continuous mappings from X into $G(\mathcal{A}_0)$, as before.

Thus the path connected components of (10.13.1) are the same as the homotopy classes of continuous mappings from X into $G(\mathcal{A}_0)$. In particular,

$$G_1(C(X,\mathcal{A}_0)) \tag{10.13.2}$$

consists of continuous mappings from X into $G(\mathcal{A}_0)$ that are homotopic to the constant function on X equal to $e_{\mathcal{A}_0}$ at every point, as continuous mappings from X into $G(\mathcal{A}_0)$. The quotient group

$$G(C(X,\mathcal{A}_0))/G_1(C(X,\mathcal{A}_0)) \tag{10.13.3}$$

corresponds to the set of homotopy classes of continuous mappings from X into $\mathcal{A}_0$. This is related to Exercise 16 on p261 of [162].

10.14 More on $G_1(\mathcal{A})$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $x \in \mathcal{A}$, then $r_{\mathcal{A}}(x)$ can be defined in the usual way, as in Sections 6.13 and 7.2. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then one can check that

$$r_{\mathcal{A}}(t\,x) = |t|\, r_{\mathcal{A}}(x). \tag{10.14.1}$$

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Suppose also that $x \in \mathcal{A}$ is quasinilpotent, so that $r_{\mathcal{A}}(x) = 0$, as in Section 7.2. This implies that $e_{\mathcal{A}} - x$ is invertible in $\mathcal{A}$, as in Section 6.13.

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $r_{\mathcal{A}}(t\,x) = 0$, by (10.14.1), so that $t\,x$ is quasinilpotent as well. This implies that

$$e_{\mathcal{A}} - t\,x \in G(\mathcal{A}), \tag{10.14.2}$$

as in the preceding paragraph. It follows that

$$e_{\mathcal{A}} - t\,x \in G_1(\mathcal{A}), \tag{10.14.3}$$

by the definition of $G_1(\mathcal{A})$, as in Section 10.10.

Now let $x \in G(\mathcal{A})$ be given, and put

$$f(\lambda) = \lambda\,x - (\lambda - 1)\,e_{\mathcal{A}} = \lambda\,x + (1 - \lambda)\,e_{\mathcal{A}}. \tag{10.14.4}$$

for every real or complex number λ, as appropriate. Note that

$$f(0) = e_{\mathcal{A}},\ \ f(1) = x. \tag{10.14.5}$$

Let E be the set of $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that

$$f(\lambda) \in G(\mathcal{A}). \tag{10.14.6}$$

Thus $0, 1 \in E$, by (10.14.5). We also have that E is an open set in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to the standard Euclidean metric, because $G(\mathcal{A})$ is an open set in $\mathcal{A}$ with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, as in Section 6.6.

If $\lambda \neq 0$, then (10.14.6) is the same as saying that

$$x - \lambda^{-1}\,(\lambda - 1)\,e_{\mathcal{A}} \in G(\mathcal{A}). \tag{10.14.7}$$

Equivalently, this means that

$$\lambda^{-1}\,(\lambda - 1) = 1 - \lambda^{-1} \tag{10.14.8}$$

is not in the spectrum $\sigma_{\mathcal{A}}(x)$ of x with respect to $\mathcal{A}$. If

(10.14.9) 0, 1 are in the same connected component of E,

then it follows that

$$x \in G_1(\mathcal{A}). \tag{10.14.10}$$

This corresponds to part of the proof of part (d) of Theorem 10.44 on p258 of [162]. This is also related to Exercise 26 on p262 of [162].

Of course, (10.14.9) holds when

$$\{\lambda \in \mathbf{R} : 0 \le \lambda \le 1\} \subseteq E. \tag{10.14.11}$$

One can verify that this happens exactly when

$$\sigma_{\mathcal{A}}(x) \cap \{\lambda \in \mathbf{R} : \lambda \le 0\} = \emptyset, \tag{10.14.12}$$

because $0 \in E$ automatically. In the real case, (10.14.9) is equivalent to (10.14.11), and (10.14.12) is the same as saying that

$$\sigma_{\mathcal{A}}(x) \subseteq (0, +\infty). \tag{10.14.13}$$

In the complex case, there is exactly one unbounded connected component of $\mathbf{C} \setminus \sigma_{\mathcal{A}}(x)$, because $\sigma_{\mathcal{A}}(x)$ is bounded, as in Section 6.8. In this case, one can check that (10.14.9) holds if and only if

(10.14.14) 0 is in the unbounded connected component of $\mathbf{C} \setminus \sigma_{\mathcal{A}}(x)$.

10.15 Some additional properties of $G_1(\mathcal{A})$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. Suppose that $x \in \mathcal{A}$ satisifes

$$x^n = e_{\mathcal{A}} \tag{10.15.1}$$

for some positive integer n. Consider the polynomial $p(t) = t^n$. Under these conditions, we have that

$$p(\sigma_{\mathcal{A}}(x)) \subseteq \sigma_{\mathcal{A}}(\widetilde{p}(x)) = \{1\}, \tag{10.15.2}$$

where the first step is as in Section 8.13.

In the real case, (10.15.2) implies that

$$\sigma_{\mathcal{A}}(x) \subseteq \{1\} \tag{10.15.3}$$

when n is odd. In the complex case, (10.15.2) implies that

$$\mathbf{C} \setminus \sigma_{\mathcal{A}}(x) \text{ is connected} \tag{10.15.4}$$

for all $n \geq 1$, because $\sigma_{\mathcal{A}}(x)$ has only finitely many elements.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra, with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. In the real case, we get that (10.14.10) holds when n is odd, because (10.15.3) implies (10.14.13). In the complex case, (10.14.10) holds for every $n \geq 1$, because (10.15.4) implies (10.14.14). This corresponds to part of the proof of part (d) of Theorem 10.44 on p258 of [162], and to Exercise 26 on p262 of [162].

Suppose from now on in this section that $\mathcal{A}$ is commutative. Suppose also that $w \in G(\mathcal{A})$ satisfies

$$w^n \in G_1(\mathcal{A}) \tag{10.15.5}$$

for some positive integer n. Remember that $G_1(\mathcal{A}) = \exp \mathcal{A}$ when $\mathcal{A}$ is commutative, as in Section 10.10. Thus

$$w^n = \exp a \tag{10.15.6}$$

for some $a \in \mathcal{A}$.

Put

$$y = \exp(n^{-1}\, a) \tag{10.15.7}$$

and

$$z = w\, y^{-1}. \tag{10.15.8}$$

Observe that

$$z^n = w^n\, y^{-n} = (\exp a)\,(\exp(-a)) = e_{\mathcal{A}}. \tag{10.15.9}$$

In the real case, this implies that

$$z \in G_1(\mathcal{A}) \tag{10.15.10}$$

when n is odd, as before. In the complex case, we get that (10.15.10) holds for any $n \geq 1$.

Remember that $y \in G_1(\mathcal{A})$, as in Section 10.10. If (10.15.10) holds, then it follows that

$$w = z\, y \in G_1(\mathcal{A}), \tag{10.15.11}$$

because $G_1(\mathcal{A})$ is a subgroup of $G(\mathcal{A})$.

Let $[w]$ be the image of w under the natural quotient homomorphism from $G(\mathcal{A})$ onto the quotient group $G(\mathcal{A})/G_1(\mathcal{A})$. Thus (10.15.5) is the same as saying that

$$[w]^n = [w^n] \text{ is the identity element in } G(\mathcal{A})/G_1(\mathcal{A}). \tag{10.15.12}$$

Similarly, (10.15.11) is the same as saying that

$$[w] \text{ is the identity element in } G(\mathcal{A})/G_1(\mathcal{A}). \tag{10.15.13}$$

This corresponds to part (d) of Theorem 10.44 on p258 of [162].

Part III

Algebras, norms, and operators, 2

Chapter 11

Algebras, norms, and square roots

11.1 Exponentials of nilpotent elements

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. Suppose that $x \in \mathcal{A}$ is nilpotent, so that $x^l = 0$ for some positive integer l. Under these conditions, the *exponential* of x in $\mathcal{A}$ may be defined by

$$\exp x = \exp_{\mathcal{A}} x = \sum_{j=0}^{\infty} (1/j!)\, x^j = \sum_{j=0}^{l-1} (1/j!)\, x^j. \tag{11.1.1}$$

Similarly, if $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $(t\,x)^l = t^l\, x^l = 0$, and

$$\exp(t\,x) = \sum_{j=0}^{l-1} (1/j!)\, t^j\, x^j. \tag{11.1.2}$$

Let y be another nilpotent element of $\mathcal{A}$, so that $y^m = 0$ for some $m \ge 1$. Suppose that x and y commute, and let us check that $x + y$ is nilpotent as well. More precisely,

$$(x + y)^{l+m-1} = 0. \tag{11.1.3}$$

Indeed, $(x + y)^{l+m-1}$ can be expanded into a sum of products of powers of x and y, where the sum of the powers of x and y is equal to $l+m-1$. This implies that in each term, the power of x is at least l, or the power of y is at least m.

This implies that the exponential of $x+y$ may be defined as before. We also have that

$$\exp(x + y) = (\exp x)\,(\exp y), \tag{11.1.4}$$

as in Section 10.4. In particular, we can take $y = -x$, to get that $\exp x$ is invertible in $\mathcal{A}$, with

$$(\exp x)^{-1} = \exp(-x). \tag{11.1.5}$$

Let $\mathcal{B}$ be another associative algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{B}}$, and let ϕ be an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ with $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. If $x \in \mathcal{A}$ is nilpotent, then $\phi(x)$ is nilpotent in $\mathcal{B}$, and

$$\phi(\exp_{\mathcal{A}} x) = \exp_{\mathcal{B}} \phi(x). \tag{11.1.6}$$

This also works when ϕ is an oppositve algebra homomorphism, which may by conjugate-linear in the complex case.

Suppose for the moment that $\mathcal{A}$ is an algebra in the strict sense over $\mathbf{R}$ or $\mathbf{C}$, and let δ be a derivation on $\mathcal{A}$, as in Section 10.5. If δ is nilpotent as a linear mapping from $\mathcal{A}$ into itself, then $\exp \delta$ may be defined as a linear mapping from $\mathcal{A}$ into itself, as before. In fact,

$$\exp \delta \text{ is an automorphism of } \mathcal{A}, \tag{11.1.7}$$

as an algebra in the strict sense, as in Section 10.5. This corresponds to some remarks on p8f of [97] when $\mathcal{A}$ is a Lie algebra.

Let $\mathcal{A}$ be an associative algebra over $\mathbf{R}$ or $\mathbf{C}$ with a multiplicative identity element $e_{\mathcal{A}}$ again, and let L_a, R_a be the left and right multiplication operators on $\mathcal{A}$ associated to $a \in \mathcal{A}$. If a is nilpotent, then it is easy to see that L_a and R_a are nilpotent as elements of the algebra $\mathcal{L}(\mathcal{A})$ of linear mappings from $\mathcal{A}$ into itself, so that their exponentials are defined as elements of $\mathcal{L}(V)$ too. In this case, we have that

$$\exp L_a = \exp_{\mathcal{L}(\mathcal{A})} L_a = L_{\exp_{\mathcal{A}} a} \tag{11.1.8}$$

and

$$\exp R_a = \exp_{\mathcal{L}(\mathcal{A})} R_a = R_{\exp_{\mathcal{A}} a}, \tag{11.1.9}$$

as in (11.1.6).

Remember that $\delta_a = L_a - R_a$ is a derivation on $\mathcal{A}$, and that L_a and R_a commute on $\mathcal{A}$, as in Section 10.6. If a is nilpotent, then it follows that δ_a is nilpotent in $\mathcal{L}(\mathcal{A})$, as in (11.1.3). We also get that

$$\exp \delta_a = (\exp L_a) \circ (\exp R_a)^{-1}, \tag{11.1.10}$$

as in (11.1.4). This means that

$$(\exp \delta_a)(x) = (\exp_{\mathcal{L}(\mathcal{A})} \delta_a)(x) = (\exp_{\mathcal{A}} a)\, x\, (\exp_{\mathcal{A}} a)^{-1} \tag{11.1.11}$$

for every $x \in \mathcal{A}$, because of (11.1.8) and (11.1.9). This is related to some remarks on p9 of [97].

11.2 One-sided inverses

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let x be an element of $\mathcal{A}$. We say that $w \in \mathcal{A}$ is a *left inverse* of x if

$$w\, x = e_{\mathcal{A}}. \tag{11.2.1}$$

Similarly, $z \in \mathcal{A}$ is a *right inverse* of x if

$$x\,z = e_{\mathcal{A}}. \tag{11.2.2}$$

If w and z are left and right inverses of x in $\mathcal{A}$, respectively, then

$$w = w\,(x\,z) = (w\,x)\,z = z. \tag{11.2.3}$$

This means that x is invertible in $\mathcal{A}$, with

$$x^{-1} = w = z. \tag{11.2.4}$$

Let V, W be vector spaces, both real or both complex, and let T be a linear mapping from V into W. A linear mapping L from W into V is said to be a *left inverse* of T if

$$L \circ T = I_V, \tag{11.2.5}$$

the identity mapping on V. This means that

$$L(T(v)) = v \tag{11.2.6}$$

for every $v \in V$, which implies that T is one-to-one on V. If $V = W$, then this is the same as a left inverse of T in the algebra $\mathcal{L}(V)$ of linear mappings from V into itself.

Similarly, a linear mapping R from W into V is said to be a *right inverse* of T if

$$T \circ R = I_W. \tag{11.2.7}$$

This is the same as saying that

$$T(R(w)) = w \tag{11.2.8}$$

for every $w \in W$, which implies that T maps V onto W. If $V = W$, then this is the same as a right inverse of T in $\mathcal{L}(V)$.

If V and W have the same finite dimension, then it is well known that T is one-to-one on V if and only if $T(V) = W$. In this case, we get that T is invertbile as a linear mapping from V into W when T has a left or right inverse.

Let $\|\cdot\|_V$, $\|\cdot\|_W$ be norms on V, W, respectively, and suppose that T is a bounded linear mapping from V into W, with respect to these norms. A bounded linear mapping L from W into V is a *left inverse* of T as a bounded linear mapping from V into W if (11.2.5) holds. This implies that

$$\|v\|_V \le \|L\|_{op,WV}\,\|T(v)\|_W \tag{11.2.9}$$

for every $v \in V$. If $V = W$, and the two norms are the same, then L is the same as a left inverse of T in the algebra $\mathcal{BL}(V)$ of bounded linear mappings from V into itself.

Similarly, a bounded linear mapping R from W into V is said to be a *right inverse* of T as a bounded linear mapping from V into W if (11.2.7) holds. One can check that this implies that

(11.2.10) $\qquad T$ is an open mapping from V onto W.

If $V = W$, and the two norms are the same, then R is the same as a right inverse of T in $\mathcal{BL}(V)$.

Let $\mathcal{A}$ be any associative algebra over $\mathbf{R}$ or $\mathbf{C}$ with a multiplicative identity element $e_{\mathcal{A}}$ again, and let $x, y \in \mathcal{A}$ be given. If $x\,y$ is invertible in $\mathcal{A}$, then

$$x\,(y\,(x\,y)^{-1}) = (x\,y)\,(x\,y)^{-1} = e_{\mathcal{A}} \tag{11.2.11}$$

and

$$((x\,y)^{-1}\,x)\,y = (x\,y)^{-1}\,(x\,y) = e_{\mathcal{A}}, \tag{11.2.12}$$

so that $y\,(x\,y)^{-1}$ is a right inverse of x in $\mathcal{A}$, and $(x\,y)^{-1}\,x$ is a left inverse of y. Similarly, if $y\,x$ is invertible in $\mathcal{A}$, then

$$y\,(x\,(y\,x)^{-1}) = (y\,x)\,(y\,x)^{-1} = e_{\mathcal{A}} \tag{11.2.13}$$

and

$$((y\,x)^{-1}\,y)\,x = (y\,x)^{-1}\,(y\,x) = e_{\mathcal{A}}, \tag{11.2.14}$$

so that $x\,(y\,x)^{-1}$ is a right inverse of y in $\mathcal{A}$, and $(y\,x)^{-1}\,y$ is a left inverse of x in $\mathcal{A}$. If $x\,y$ and $y\,x$ are both invertible in $\mathcal{A}$, then it follows that x and y are invertible in $\mathcal{A}$, as in part (b) of Exercise 1 on p259 of [162]. If x and y commute, then this corresponds to a remark in Section 6.13.

11.3 Topological divisors of zero

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a sub-multiplicative norm $\|\cdot\|_{\mathcal{A}}$, and let x be an element of $\mathcal{A}$. Let us say that x is a *left topological divisor of zero* in $\mathcal{A}$ if there is a sequence $\{z_j\}_{j=1}^{\infty}$ of elements of $\mathcal{A}$ such that

$$\|z_j\|_{\mathcal{A}} = 1 \tag{11.3.1}$$

for each j and

$$\lim_{j\to\infty} x\,z_j = 0, \tag{11.3.2}$$

as in Exercise 4 on p300 of [167]. Similarly, let us say that x is a *right topological divisor of zero* in $\mathcal{A}$ if there is a sequence $\{w_j\}_{j=1}^{\infty}$ of elements of $\mathcal{A}$ such that

$$\|w_j\|_{\mathcal{A}} = 1 \tag{11.3.3}$$

for each j and

$$\lim_{j\to\infty} w_j\,x = 0, \tag{11.3.4}$$

as in [167]. If there is a sequence $\{y_j\}_{j=1}^{\infty}$ of elements of $\mathcal{A}$ such that

$$\|y_j\|_{\mathcal{A}} = 1 \tag{11.3.5}$$

for each j and

$$\lim_{j\to\infty} x\,y_j = \lim_{j\to\infty} y_j\,x = 0, \tag{11.3.6}$$

then we may say that x is a *topological divisor of zero* in $\mathcal{A}$, as in Exercise 5 on p259 of [162]. In this case, x is both a left and right topological divisor of zero, with

$$w_j = z_j = y_j \tag{11.3.7}$$

for each j.

Suppose that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$. If x is a left topological divisor of zero in $\mathcal{A}$, then it is easy to see that

$$x \text{ does not have a left inverse in } \mathcal{A}. \tag{11.3.8}$$

Similarly, if x is a right topological divisor of zero in $\mathcal{A}$, then

$$x \text{ does not have a right inverse in } \mathcal{A}. \tag{11.3.9}$$

Let $\mathcal{B}$ be another associative algebra over the real or complex numbers, as appropriate, with a submultiplicative norm $\|\cdot\|_{\mathcal{B}}$. Suppose that ϕ is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ that is bounded as a linear mapping. If $x \in \mathcal{A}$ is a left or right topological divisor of zero in $\mathcal{A}$, or simply a topological divisor of zero in $\mathcal{A}$, then $\phi(x)$ has the analogous property in $\mathcal{B}$.

Suppose now that $\mathcal{A}$ is a Banach algebra, with a multiplicative identity element $e_{\mathcal{A}}$. Let $x \in \mathcal{A}$ be an element of the boundary of the set $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$. This implies that x is not invertible in $\mathcal{A}$, because $G(\mathcal{A})$ is an open set in $\mathcal{A}$, as in Section 6.6, and that there is a sequence $\{u_j\}_{j=1}^\infty$ of elements of $G(\mathcal{A})$ that converges to x with respect to the metric on $\mathcal{A}$ associated to $\|\cdot\|_{\mathcal{A}}$.

We also have that

$$\|u_j^{-1}\|_{\mathcal{A}} \to \infty \text{ as } j \to \infty \tag{11.3.10}$$

under these conditions. More precisely, the norms of the u_j^{-1}'s are not bounded, because x is not invertible, as in Section 6.7. This means that (11.3.10) holds after passing to a subsequence, which would suffice for the remarks in the next paragraph. The same argument shows that the norms of the u_j^{-1}'s are not bounded along any subsequence, which implies (11.3.10). This corresponds to Lemma 10.17 on p238 of [162].

Put

$$y_j = u_j^{-1}/\|u_j^{-1}\|_{\mathcal{A}} \tag{11.3.11}$$

for each j, which automatically satisfies (11.3.5). One can check that (11.3.6) holds under these conditions, so that

$$x \text{ is a topological divisor of zero in } \mathcal{A}. \tag{11.3.12}$$

This corresponds to part (a) of Exercise 5 on p259 of [162].

Let $a \in \mathcal{A}$ be given, and remember that $\sigma_{\mathcal{A}}(a)$ is the spectrum of a with respect to $\mathcal{A}$, as in Section 6.8. Suppose that $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, is an element of the boundary of $\sigma_{\mathcal{A}}(a)$, with respect to the standard Euclidean metric on **R** or **C**, as appropriate. This implies that $\lambda\, e_{\mathcal{A}} - a$ is an element of the boundary of $G(\mathcal{A})$, so that

$$\lambda\, e_{\mathcal{A}} - a \text{ is a topological divisor of zero in } \mathcal{A}, \tag{11.3.13}$$

as in the preceding paragraph. This corresponds to part (a) of Exercise 4 on p300 of [167]. Note that this could be used as another way to obtain some of the remarks in Section 7.3.

11.4 Another condition on a norm

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. Consider the condition that there be a nonnegative real number C such that

$$\|x\|_{\mathcal{A}}\, \|y\|_{\mathcal{A}} \le C\, \|x\, y\|_{\mathcal{A}} \tag{11.4.1}$$

for every $x, y \in \mathcal{A}$. It is easy to see that this implies that

$$0 \text{ is the only left or right topological divisor of zero in } \mathcal{A}. \tag{11.4.2}$$

Of course, (11.4.2) implies that

$$0 \text{ is the only topological divisor of zero in } \mathcal{A}. \tag{11.4.3}$$

Suppose that $\mathcal{A}$ is a Banach algebra, with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Note that (11.4.2) holds when every nonzero element of $\mathcal{A}$ is invertible in $\mathcal{A}$, because of (11.3.8) and (11.3.9). Let $a \in \mathcal{A}$ be given, and suppose that $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, is an element of the boundary of the spectrum $\sigma_{\mathcal{A}}(a)$ of a with respect to $\mathcal{A}$. If (11.4.3) holds, then we get that

$$a = \lambda\, e_{\mathcal{A}}, \tag{11.4.4}$$

because of (11.3.13).

Remember that $\sigma_{\mathcal{A}}(a)$ is a bounded subset of **R** or **C**, as appropriate, as in Section 6.8. If

$$\sigma_{\mathcal{A}}(a) \neq \emptyset, \tag{11.4.5}$$

then it follows that

$$\partial\sigma_{\mathcal{A}}(a) \neq \emptyset, \tag{11.4.6}$$

where $\partial\sigma_{\mathcal{A}}(a)$ is the boundary of $\sigma_{\mathcal{A}}$ in **R** or **C**, as appropriate. This uses the well-known fact that **R**, **C** are connected with respect to their standard Euclidean metrics. If (11.4.3) holds, then we get that (11.4.4) holds for some $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, as in the preceding paragraph.

In the complex case, (11.4.5) holds for every $a \in \mathcal{A}$, as in Section 6.8. This means that

$$\mathcal{A} = \{\lambda\, e_{\mathcal{A}} : \lambda \in \mathbf{C}\} \tag{11.4.7}$$

when (11.4.3) holds. This corresponds to Theorem 10.19 on p239 of [162] when (11.4.1) holds. The analogous statement when (11.4.3) holds corresponds to part (b) of Exercise 5 on p259 of [162].

11.5 Some properties of the spectrum

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. Also let ϕ be an algebra *automorphism* of $\mathcal{A}$, which is to say an algebra isomorphism of $\mathcal{A}$ onto itself, or an opposite algebra isomorphism from $\mathcal{A}$ onto itself. As usual, we allow ϕ to be conjugate-linear in the complex case. Remember that

$$\phi(e_{\mathcal{A}}) = e_{\mathcal{A}}, \tag{11.5.1}$$

as in Sections 6.3 and 6.4.

If $a \in \mathcal{A}$ and $\lambda \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then

$$\phi(\lambda\, e_{\mathcal{A}} - a) = \lambda\, e_{\mathcal{A}} - \phi(a), \tag{11.5.2}$$

except in the complex case when ϕ is conjugate-linear, for which we have

$$\phi(\lambda\, e_{\mathcal{A}} - a) = \overline{\lambda}\, e_{\mathcal{A}} - \phi(a). \tag{11.5.3}$$

It follows that

$$\sigma_{\mathcal{A}}(\phi(a)) = \sigma_{\mathcal{A}}(a) \tag{11.5.4}$$

in the first case, and

$$\sigma_{\mathcal{A}}(\phi(a)) = \{\overline{\lambda} : \lambda \in \sigma_{\mathcal{A}}(a)\} \tag{11.5.5}$$

in the second case. The right side may also be denoted $\overline{\sigma_{\mathcal{A}}(a)}$ sometimes, although this notation may be used for the closure of $\sigma_{\mathcal{A}}(a)$ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as well.

In particular, if x is an invertible element of $\mathcal{A}$, then conjugation by x defines an algebra automorphism of $\mathcal{A}$, and we get that

$$\sigma_{\mathcal{A}}(x\, a\, x^{-1}) = \sigma_{\mathcal{A}}(a), \tag{11.5.6}$$

as in (11.5.4). If $y \in \mathcal{A}$, then we can take $a = y\, x$ to get that

$$\sigma_{\mathcal{A}}(x\, y) = \sigma_{\mathcal{A}}(y\, x). \tag{11.5.7}$$

This corresponds to part (c) of Exercise 2 on p259 of [162].

Let $x, y \in \mathcal{A}$ be given, and suppose that

$$e_{\mathcal{A}} - x\, y \in G(\mathcal{A}). \tag{11.5.8}$$

Under these conditions, it is well known that

$$e_{\mathcal{A}} - y\, x \in G(\mathcal{A}). \tag{11.5.9}$$

This corresponds to part (a) of Exercise 2 on p259 of [162], and to the hint in part (b) of Exercise 1 on p299 of [167]. This also corresponds to Exercise (3) on p7 of [8] in the case of linear mappings. More precisely, one can check that

$$(e_{\mathcal{A}} - y\, x)^{-1} = e + y\, (e_{\mathcal{A}} - x\, y)^{-1}\, x, \tag{11.5.10}$$

as in [162, 167]. In [8], the hint is given to look at how the formal Neumann series for $(e_{\mathcal{A}} - x\,y)^{-1}$ and $(e_{\mathcal{A}} - y\,x)^{-1}$ are related, and to use that to give a rigorous proof of (11.5.9). Indeed, these formal Neumann series are related as in (11.5.10).

One can use this to get that

$$\sigma_{\mathcal{A}}(x\,y) \setminus \{0\} = \sigma_{\mathcal{A}}(y\,x) \setminus \{0\}. \tag{11.5.11}$$

This corresponds to Exercise (4) on p7 of [8] in the case of linear operators, the first part of part (b) of Exercise 2 on p259 of [162], and part (b) of Exercise 1 on p299 of [167].

Now let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. Thus $r_{\mathcal{A}}(x)$ may be defined for $x \in \mathcal{A}$ as in Sections 6.13 and 7.2. If $x, y \in \mathcal{A}$, then

$$r_{\mathcal{A}}(x\,y) = r_{\mathcal{A}}(y\,x), \tag{11.5.12}$$

as in Exercise 24 on p262 of [162]. If $\mathcal{A}$ is a complex Banach algebra with a multiplicative identity element, then this can be obtained from (11.5.11), using the relationship between $r_{\mathcal{A}}(\cdot)$ and the spectrum mentioned in Section 6.14. This can also be shown more directly using the hint mentioned in [162], as follows.

Observe that

$$(x\,y)^j = x\,(y\,x)^{j-1}\,y \tag{11.5.13}$$

for every positive integer j. This implies that

$$\|(x\,y)^j\|_{\mathcal{A}} \le \|x\|_{\mathcal{A}}\,\|(y\,x)^{j-1}\|_{\mathcal{A}}\,\|y\|_{\mathcal{A}} \tag{11.5.14}$$

for each j. Let r be a real number such that

$$r_{\mathcal{A}}(y\,x) < r, \tag{11.5.15}$$

so that

$$\|(y\,x)^j\|_{\mathcal{A}} < r^j \tag{11.5.16}$$

when j is sufficiently large. It follows that

$$\|(x\,y)^j\|_{\mathcal{A}} \le \|x\|_{A}\,\|y\|_{\mathcal{A}}\,r^{j-1} \tag{11.5.17}$$

when j is sufficiently large, and thus

$$\|(x\,y)^j\|_{\mathcal{A}}^{1/j} \le \|x\|_{\mathcal{A}}^{1/j}\,\|y\|_{\mathcal{A}}^{1/j}\,r^{1-(1/j)}. \tag{11.5.18}$$

Using this, we get that

$$r_{\mathcal{A}}(x\,y) \le r. \tag{11.5.19}$$

This implies that

$$r_{\mathcal{A}}(x\,y) \le r_{\mathcal{A}}(y\,x). \tag{11.5.20}$$

Of course, the opposite inequality can be obtained in the same way, to get (11.5.12).

11.6 Commutativity and norms

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Suppose that there is a nonnegative real number C such that

$$\|x\,y\|_{\mathcal{A}} \le C\,\|y\,x\|_{\mathcal{A}} \tag{11.6.1}$$

for all $x, y \in \mathcal{A}$. We would like to show that $\mathcal{A}$ is commutative, as in part (a) of Exercise 25 on p262 of [162].

If $z \in \mathcal{A}$ is invertible, then one can check that

$$\|z\,y\,z^{-1}\|_{\mathcal{A}} \le C\|y\|_{\mathcal{A}} \tag{11.6.2}$$

for every $y \in \mathcal{A}$, using (11.6.1). This implies that

$$\|(\exp(\lambda\,x))\,y\,(\exp(-\lambda\,x))\|_{\mathcal{A}} \le C\|y\|_{\mathcal{A}} \tag{11.6.3}$$

for all $x, y \in \mathcal{A}$ and $\lambda \in \mathbf{C}$.

Let $x, y \in \mathcal{A}$, and consider

$$f(\lambda) = f_{x,y}(\lambda) = (\exp(\lambda\,x))\,y\,(\exp(-\lambda\,x)) \tag{11.6.4}$$

as a function of $\lambda \in \mathbf{C}$ with values in $\mathcal{A}$. More precisely, we would like to consider this as a holomorphic function of $\lambda \in \mathbf{C}$ with values in $\mathcal{A}$. The right side may be expressed as a power series in λ with coefficients in $\mathcal{A}$ that converges absolutely with respect to $\|\cdot\|_{\mathcal{A}}$ for every $\lambda \in \mathbf{C}$, using Cauchy products, for instance.

In particular, if μ is a bounded linear functional on $\mathcal{A}$, then

$$\mu(f(\lambda)) = \mu((\exp(\lambda\,x))\,y\,(\exp(-\lambda\,x))) \tag{11.6.5}$$

may be expressed as a power series in λ with complex coefficients that converges absolutely for every $\lambda \in \mathbf{C}$. Thus (11.6.5) is holomorphic as a complex-valued function of $\lambda \in \mathbf{C}$ in the usual sense. This function is also bounded on $\mathbf{C}$, because of (11.6.3). Liouville's theorem implies that this function is constant on $\mathbf{C}$, so that

$$\mu(f(\lambda)) = \mu(f(0)) \tag{11.6.6}$$

for every $\lambda \in \mathbf{C}$. One can use this and the Hahn–Banach theorem to get that

$$f(\lambda) = f(0) \tag{11.6.7}$$

for every $\lambda \in \mathbf{C}$.

It follows that

$$(\exp(\lambda\,x))\,y = y\,(\exp(\lambda\,x)) \tag{11.6.8}$$

for all $x, y \in \mathcal{A}$ and $\lambda \in \mathbf{C}$. One can use this to get that $\mathcal{A}$ is commutative.

11.7 More on commutativity and norms

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. Of course,

$$\|x^2\|_{\mathcal{A}} \le \|x\|_{\mathcal{A}}^2 \tag{11.7.1}$$

for every $x \in \mathcal{A}$. Suppose that there is a nonnegative real number C such that

$$\|x\|_{\mathcal{A}}^2 \le C\,\|x^2\|_{\mathcal{A}} \tag{11.7.2}$$

for every $x \in \mathcal{A}$. If $x \in \mathcal{A}$, then one can check that

$$\|x\|_{\mathcal{A}}^{2^l} \le C^{2^l-1}\,\|x^{2^l}\|_{\mathcal{A}} \tag{11.7.3}$$

for every positive integer l. This corresponds to (3) in the proof of Lemma 11.11 on p270 of [162].

Thus

$$\|x\|_{\mathcal{A}} \le C^{1-2^{-l}}\,\|x^{2^l}\|_{\mathcal{A}}^{2^{-1}} \tag{11.7.4}$$

for every $l \ge 1$. This implies that

$$\|x\|_{\mathcal{A}} \le C\,r_{\mathcal{A}}(x), \tag{11.7.5}$$

where $r_{\mathcal{A}}(x)$ is as in Sections 6.13 and 7.2, as in (4) on p270 of [162]. If $x, y \in \mathcal{A}$, then it follows that

$$\|x\,y\|_{\mathcal{A}} \le C\,r_{\mathcal{A}}(x\,y) = C\,r_{\mathcal{A}}(y\,x) \le C\,\|y\,x\|_{\mathcal{A}}, \tag{11.7.6}$$

using (11.5.12) in the second step.

Suppose for the moment that $\mathcal{A}$ is a complex Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. In this case, (11.7.6) implies that $\mathcal{A}$ is commutative, as in the previous section. This shows that $\mathcal{A}$ is commutative when (11.7.2) holds. This corresponds to part (b) of Exercise 25 on p262 of [162] when $C = 1$.

Let $\mathcal{A}$ be any associative algebra over $\mathbf{R}$ or $\mathbf{C}$ with a submultiplicative norm again. Also let $m \ge 2$ be an integer, and suppose that there is a nonnegative real number $C(m)$ such that

$$\|x\|_{\mathcal{A}}^m \le C(m)\,\|x^m\|_{\mathcal{A}} \tag{11.7.7}$$

for every $x \in \mathcal{A}$. If $n \ge 2$ is an integer, then there is a nonnegative real number $C_n(m)$ such that

$$\|x\|_{\mathcal{A}}^n \le C_n(m)\,\|x^n\|_{\mathcal{A}} \tag{11.7.8}$$

for every $x \in \mathcal{A}$. This corresponds to Exercise 16 on p291 of [162]. In particular, we can take $n = 2$, to get a condition like (11.7.2). One could use the same type of arguments as before in this case as well. If $\mathcal{A}$ is a complex Banach algebra with a multiplicative identity element, then it follows that $\mathcal{A}$ is commutative, as in the preceding paragraph.

If $x \in \mathcal{A}$, then

$$r_{\mathcal{A}}(x) = \lim_{j \to \infty} \|x^j\|_{\mathcal{A}}^{1/j} = \lim_{l \to \infty} \|x^{l\,m}\|_{\mathcal{A}}^{1/(l\,m)} = r_{\mathcal{A}}(x^m)^{1/m}. \tag{11.7.9}$$

If (11.7.5) holds for some $C \geq 0$ and all $x \in \mathcal{A}$, then

$$\|x\|_{\mathcal{A}}^m \leq C^m\, r_{\mathcal{A}}(x)^m = C^m\, r_{\mathcal{A}}(x^m) \leq C^m\, \|x^m\|_{\mathcal{A}} \tag{11.7.10}$$

for every $x \in \mathcal{A}$. This corresponds to another part of Lemma 11.11 on p270 of [162] when $m = 2$.

11.8 A continuity property

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space, with associated norm $\| \cdot \|_V$, and $V \neq \{0\}$. Also let f be a continuous real-valued function on the real line. If T is a bounded self-adjoint linear mapping from V into itself, then $\widetilde{f}(T)$ may be defined as a bounded self-adjoint linear mapping from V into itself as in Section 8.14. More precisely, $\widetilde{f}(T)$ only depends on the restriction of f to the spectrum $\sigma(T) = \sigma_{\mathcal{BL}(V)}(T)$ of T with respect to the algebra $\mathcal{BL}(V)$ of bounded linear mappings from V into itself. In fact,

$$\|\widetilde{f}(T)\|_{op} = \max\{|f(\lambda)| : \lambda \in \sigma(T)\}, \tag{11.8.1}$$

as before.

The space of bounded self-adjoint linear mappings from V into itself is a real-linear subspace of $\mathcal{BL}(V)$. Let us check that

$$T \mapsto \widetilde{f}(T) \tag{11.8.2}$$

is continuous on this space, with respect to the metric associated to the operator norm. This is a version of Exercise (3) on p51 of [8]. More precisely, if r is a positive real number, then (11.8.2) is uniformly continuous on

$$\{T \in \mathcal{BL}(V) : T \text{ is self-adjoint, and } \|T\|_{op} \leq r\}. \tag{11.8.3}$$

If p is a polynomial with real coefficients, then

$$T \mapsto \widetilde{p}(T) \tag{11.8.4}$$

is uniformly continuous on

$$\{T \in \mathcal{BL}(V) : \|T\|_{op} \leq r\} \tag{11.8.5}$$

with respect to the metric associated to the operator norm. This corresponds to a remark in Section 9.14. We also have that (11.8.4) is Lipschitz on (11.8.5) with respect to the metric associated to the operator norm, where the constant can be estimated as in Section 10.1. In particular, (11.8.4) has these properties on (11.8.3).

Otherwise, one can approximate f by polynomials uniformly on $[-r, r]$, by Weierstrass' approximation theorem. Of course, if T is a bounded self-adjoint linear mapping from V into itself with $\|T\|_{op} \le r$, then

$$\sigma(T) \subseteq [-r, r]. \tag{11.8.6}$$

Indeed, if $\lambda \in \sigma(T)$, then

$$|\lambda| \le \|T\|_{op} \le r, \tag{11.8.7}$$

where the first step is as in Section 6.8. In the complex case, we are also using the fact that the spectrum of T is contained in the real line when T is self-adjoint, as in Section 8.11.

If T is a bounded self-adjoint linear mapping from V into itself and p is a polynomial with real coefficients, then

$$\|\widetilde{f}(T) - \widetilde{p}(T)\|_{op} = \max\{|f(\lambda) - p(\lambda)) : \lambda \in \sigma(T)\}, \tag{11.8.8}$$

as in (11.8.1). This implies that

$$\|\widetilde{f}(T) - \widetilde{p}(T)\|_{op} \le \max\{|f(\lambda) - p(\lambda)| : \lambda \in [-r, r]\} \tag{11.8.9}$$

when $\|T\|_{op} \le r$, by (11.8.6). Weierstrass' theorem implies that for each $\epsilon > 0$ there is a polynomial p with real coefficients such that

$$\max\{|f(\lambda) - p(\lambda)| : \lambda \in [-r, r]\} < \epsilon. \tag{11.8.10}$$

In this case, we get that

$$\|\widetilde{f}(T) - \widetilde{p}(T)\|_{op} < \epsilon \tag{11.8.11}$$

when $\|T\|_{op} \le r$, as in (11.8.9). One can use this to get the uniform continuity of (11.8.2) on (11.8.3), because of the analogous statement for polynomials.

11.9 More on self-adjoint operators

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space again, with associated norm $\| \cdot \|_V$, and $V \ne \{0\}$. Also let T be a bounded self-adjoint linear mapping from V into itself, with spectrum $\sigma(T) = \sigma_{\mathcal{BL}(V)}$ with respect to $\mathcal{BL}(V)$. Remember that T is said to be nonnegative on V with respect to $\langle \cdot, \cdot \rangle_V$ when $\langle T(v), v \rangle_V \ge 0$ for every $v \in V$, as in Section 8.10. This may be expressed by saying that

$$T \ge 0 \tag{11.9.1}$$

on V. This is equivalent to the condition that $\alpha(T) \ge 0$, in the notation of Section 8.9.

Remember that $\sigma(T) \subseteq \mathbf{R}$ in the complex case, because T is self-adjoint, as in Section 8.11. It is well known that (11.9.1) holds if and only if

$$\sigma(T) \subseteq \{\lambda \in \mathbf{R} : \lambda \ge 0\}. \tag{11.9.2}$$

This follows from the characterization of $\alpha(T)$ in Section 8.12.

Let f be a continuous real-valued function on $\sigma(T)$, and let $\widetilde{f}(T)$ be the corresponding bounded self-adjoint linear mapping from V into itself, as in Section 8.14. Remember that

$$f(\sigma(T)) = \sigma(\widetilde{f}(T)), \tag{11.9.3}$$

as in Section 9.5. It follows that

$$\widetilde{f}(T) \geq 0 \tag{11.9.4}$$

on V if and only if

$$f \geq 0 \tag{11.9.5}$$

on $\sigma(T)$, as in the preceding paragraph.

Remember that we may consider $\mathcal{BL}(V)$ as an associative algebra over the real numbers, even when V is complex, as in Section 8.14. Let $\mathcal{A}_0(T)$ be the subalgebra of $\mathcal{BL}(V)$, as an algebra over the real numbers, generated by T and the identity mapping $I = I_V$ on V. Equivalently,

$$\mathcal{A}_0(T) = \{\widetilde{p}(T) : p \text{ is a polynomial with real coefficients}\}. \tag{11.9.6}$$

Note that $\mathcal{A}_0(T)$ is commutative, and that the elements of $\mathcal{A}_0(T)$ are self-adjoint.

Consider the closure

$$\mathcal{A}(T) = \overline{\mathcal{A}_0(T)} \tag{11.9.7}$$

of $\mathcal{A}_0(T)$ in $\mathcal{BL}(V)$ with respect to the metric associated to the operator norm. This is also a commutative subalgebra of $\mathcal{BL}(V)$, as an algebra over the real numbers. The elements of $\mathcal{A}(T)$ are self-adjoint, as before.

Remember that

$$f \mapsto \widetilde{f}(T) \tag{11.9.8}$$

is a homomorphism from $C(\sigma(T), \mathbf{R})$ into $\mathcal{BL}(V)$, as algebras over the real numbers, as in Section 8.14. This mapping is also an isometry with respect to the supremum norm on $C(\sigma(T), \mathbf{R})$ and the operator norm on $\mathcal{BL}(V)$, as in (11.8.1). It is easy to see that (11.9.8) maps $C(\sigma(T), \mathbf{R})$ into $\mathcal{A}(T)$, because the continuous functions on $\sigma(T)$ defined by polynomials with real coefficients are dense in $C(\sigma(T), \mathbf{R})$ with respect to the supremum metric, by the Stone–Weierstrass theorem.

Of course, $C(\sigma(T), \mathbf{R})$ is complete as a metric space, with respect to the supremum metric. This implies that the image

$$\{\widetilde{f}(T) : f \in C(\sigma(T), \mathbf{R})\} \tag{11.9.9}$$

of $C(\sigma(T), \mathbf{R})$ under (11.9.8) is complete with respect to the metric associated to the operator norm, because (11.9.8) is an isometry, as before. It follows that (11.9.9) is a closed set in $\mathcal{BL}(V)$, with respect to the metric associated to the operator norm, as in Section 1.6. Clearly (11.9.9) contains $\mathcal{A}_0(T)$, and (11.9.9) is contained in $\mathcal{A}(T)$, as in the preceding paragraph. This means that (11.9.9) is equal to $\mathcal{A}(T)$.

11.10 Nonnegative square roots

Let us continue with the same notation and hypotheses as in the previous section. If f is a continuous nonnegative real-valued function on $\sigma(T)$, then f has a unique nonnegative square root in $C(\sigma(T), \mathbf{R})$. This implies that every nonnegative element of $\mathcal{A}(T)$ has a unique nonnegative square root in $\mathcal{A}(T)$.

Suppose now that T is nonnegative on V. It follows that there is a unique $R \in \mathcal{A}(T)$ such that $R \geq 0$ on V and

$$R \circ R = T. \tag{11.10.1}$$

We would like to show that R is the unique square root of T among all nonnegative self-adjoint bounded linear mappings from V into itself, as in Theorem 12.33 on p3.14 of [162]. Let B be a nonnegative self-adjoint bounded linear mapping from V into itself such that

$$B \circ B = T. \tag{11.10.2}$$

The same remarks as in the previous section and at the beginning of this section can be used for B in place of T. Thus we let $\mathcal{A}_0(B)$ be the subalgebra of $\mathcal{BL}(V)$, as an algebra over the real numbers, generated by B and I, and we let $\mathcal{A}(B)$ be its closure in $\mathcal{BL}(V)$.

Of course, $T \in \mathcal{A}_0(B)$, by (11.10.2). Note that B is the unique nonnegative square root of T in $\mathcal{A}(B)$, by the remarks at the beginning of the section. We also have that

$$\mathcal{A}_0(T) \subseteq \mathcal{A}_0(B), \tag{11.10.3}$$

by (11.10.2). This implies that

$$\mathcal{A}(T) \subseteq \mathcal{A}(B). \tag{11.10.4}$$

In particular, $R \in \mathcal{A}(B)$, so that $B = R$.

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. If $x, y \in \mathcal{A}$ satisfy

$$x = y^2, \tag{11.10.5}$$

and if x is invertible in $\mathcal{A}$, then y is invertible in $\mathcal{A}$ too, as in Section 6.13.

11.11 Polar decompositions

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space, with associated norm $\| \cdot \|_V$, and $V \neq \{0\}$. If T is any bounded linear mapping from V into itself, then $T^* \circ T$ is a bounded self-ajoint linear mapping from V into itself that is also nonnegative, as in Section 8.10. This implies that there is a unique bounded self-adjoint linear mapping P from V into itself that is nonnegative on V and satisfies

$$P \circ P = T^* \circ T, \tag{11.11.1}$$

as in the previous section.

Of course, (11.11.1) is equivalent to saying that

$$\langle (P \circ P)(v), w\rangle_V = \langle (T^* \circ T)(v), w\rangle_V \tag{11.11.2}$$

for all $v, w \in V$. This implies that

$$\langle (P \circ P)(v), v\rangle_V = \langle (T^* \circ T)(v), v\rangle_V \tag{11.11.3}$$

for every $v \in V$. Conversely, (11.11.3) implies (11.11.2), using suitable polarization identities, as in Section 8.5. This also uses the fact that $P \circ P$ and $T^* \circ T$ are self-adjoint in the real case, to get that both sides of (11.11.2) are symmetric bilinear forms on V.

Observe that (11.11.3) is the same as saying that

$$\langle P(v), P(v)\rangle_V = \langle T(v), T(v)\rangle_V \tag{11.11.4}$$

for every $v \in V$. Equivalently, this means that

$$\|P(v)\|_V = \|T(v)\|_V \tag{11.11.5}$$

for every $v \in V$. It follows that P is the unique bounded self-adjoint linear mapping from V into itself that is nonnegative and satisfies (11.11.5). This corresponds to Theorem 12.34 on p314 of [162].

Suppose now that T is invertible on V, so that T^* is invertible, and thus $T^* \circ T$ is invertible too. This implies that P is invertible on V as well, as in the previous section. Put

$$U = T \circ P^{-1}, \tag{11.11.6}$$

so that

$$T = U \circ P. \tag{11.11.7}$$

Note that U is invertible on V.

Observe that $U^* = P^{-1} \circ T^*$, so that

$$U^* \circ U = P^{-1} \circ T^* \circ T \circ P^{-1} = I. \tag{11.11.8}$$

This means that

$$U^{-1} = U^*, \tag{11.11.9}$$

because U is invertible. Thus U is a unitary mapping from V onto itself, or equivalently an orthogonal transformation on V in the real case. One could also use (11.11.4) to get that $\langle \cdot, \cdot \rangle_V$ is invariant under U.

Conversely, suppose that (11.11.7) holds for some unitary mapping U on V and bounded self-adjoint linear mapping P from V into itself that is nonnegative on V. One can check directly that (11.11.1) holds, or equivalently that (11.11.5) holds. This implies that P is uniquely determined by T, as before. It follows that U is uniquely determined by T as well, because P is invertible. This corresponds to part (a) of Theorem 12.35 on p315 of [162].

The expression (11.11.7) for T is called the *polar decomposition* of T.

Suppose that T is normal on V, in the sense that T commutes with T^*, so that $T^* \circ T$ commutes with T and T^*. If P is the nonnegative self-adjoint square root of $T^* \circ T$, as before, then P is in the closure of the subalgebra of $\mathcal{BL}(V)$, as an algebra over the real numbers, generated by $T^* \circ T$ and I, as in the previous section. This means that P commutes with T and T^* too, so that the corresponding unitary operator U commutes with T and T^* as well. In particular, U commutes with $T^* \circ T$, and thus also with P. This corresponds to part of part (b) of Theorem 12.35 on p315 of [162], which will be discussed further in the next section.

11.12 More on polar decompositions

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space again, with associated norm $\| \cdot \|_V$, and $V \neq \{0\}$, and let T be a bounded linear mapping from V into itself. An expression for T as in (11.11.7), where P is a bounded nonnegative self-adjoint linear mapping from V into itself, and U is a unitary mapping from V onto itself, is called a *polar decomposition* of T, as on p315 of [162]. Of course, (11.11.7) is the same as saying that

$$T(v) = U(P(v)) \tag{11.12.1}$$

for every $v \in V$.

In this case, P satisfies (11.11.5), or equivalently

$$\langle P(v), P(w) \rangle_V = \langle T(v), T(w) \rangle_V \tag{11.12.2}$$

for every $v, w \in V$, because U is unitary. This means that (11.11.2) holds, so that (11.11.1) holds. It follows that P is uniquely determined by T, as in Section 11.10, and as mentioned in [162].

This implies that the restriction of U to $P(V)$ is uniquely determined by T as well. More precisely, the restriction of U to the closure $\overline{P(V)}$ of $P(V)$ in V is uniquely determined by T.

Note that

$$\ker P = \ker T, \tag{11.12.3}$$

by (11.11.5). We also have that

$$\ker P = P(V)^{\perp}, \tag{11.12.4}$$

as in Section 8.11, because P is self-adjoint, and thus normal. Thus $P(V)$ is dense in V exactly when $\ker P = \{0\}$, which is the same as saying that

$$\ker T = \{0\}, \tag{11.12.5}$$

because of (11.12.3). In this case, it follows that U is uniquely determined by P. This was mentioned in the previous section when T is invertible on V.

In order to try to get a polar decomposition for T, we can start with the unique bounded nonnegative self-adjoint linear mapping P from V into itself that satisfies (11.11.1). We would like to define U initially as a mapping from $P(V)$ onto $T(V)$ by (11.12.1). More precisely, U is well-defined on $P(V)$ in this way because of (11.12.3). Remember that (11.11.1) is equivalent to (11.11.2), which is equivalent to (11.12.2). This implies that U is unitary as a linear mapping from $P(V)$ onto $T(V)$, with respect to the restrictions of the inner product to these linear subspaces.

Equivalently, U is an isometric linear mapping from $P(V)$ onto $T(V)$ with respect to the restriction of the norm to these linear subspaces. This implies that U has a unique extension to an isometric linear mapping from $\overline{P(V)}$ onto $\overline{T(V)}$, as in Section 2.2. This uses the fact that V is complete with respect to the metric associated to $\| \cdot \|_V$, so that $\overline{P(V)}$ and $\overline{T(V)}$ are complete with respect to the restrictions of the metric to these linear subspaces, as in Section 1.6. Let us also use U to denote this extension, which is unitary with respect to the restrictions of the inner product on V to $\overline{P(V)}$ and $\overline{T(V)}$.

Thus a unitary mapping corresponds exactly to an extension of U to a unitary mapping from V onto itself. Such an extension of U should map $P(V)^\perp$ onto $T(V)^\perp$, and is uniquely determined by this mapping. Remember that

$$\ker T^* = T(V)^\perp, \tag{11.12.6}$$

as in Section 8.11. This means that extensions of U to unitary mappings from V onto itself correspond exactly to unitary mappings from $\ker T$ onto $\ker T^*$, if there are any, because of (11.12.3) and (11.12.4).

Suppose for the moment that T is normal, so that

$$\ker T = \ker T^*, \tag{11.12.7}$$

as in Section 8.11. In this case, extensions of U to unitary mappings from V onto itself correspond exactly to unitary mappings from $\ker T$ onto itself, as in the preceding paragraph. Let us simply take the extension corresponding to the identity mapping on $\ker T$, which we shall denote U as well. This corresponds to part (b) of Theorem 12.35 on p315 of [162].

Remember that P commutes with T and T^* when T is normal, as mentioend at the end of the previous section. One can check directly that U commutes with P and T under these conditions, which is another part of part (b) of Theorem 12.35 on p315 of [162]. This implies that U commutes with T^* too, because $U^* = U^{-1}$.

If V has finite dimension, then $P(V)$ and $T(V)$ have the same codimension in V, because they have the same dimension. This implies that U can be extended to a unitary mapping from V onto itself, as mentioned on p316 of [162].

Of course, any extension of U to a bounded linear mapping from V into itself satisfies (11.11.7), because that only involves the restriction of U to $P(V)$. Such an extension corresponds exactly to a bounded linear mapping from $P(V)^\perp$ into V. The extension defined by taking $U = 0$ on $P(V)^\perp$ is called a *partial isometry* on V, as on p316 of [162].

11.13 Commuting with adjoints

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with an involution $x \mapsto x^*$ that may be conjugate-linear in the complex case. If $x, y \in \mathcal{A}$ commute, then

$$x^* \, y^* = y^* \, x^*. \tag{11.13.1}$$

In some case one might also like to have that

$$x \, y^* = y^* \, x, \tag{11.13.2}$$

which is the same as saying that

$$x^* \, y = y \, x^*. \tag{11.13.3}$$

Of course, x automatically commutes with itself, but it may not commute with x^*.

Suppose that x commutes with y again. If x is self-adjoint or anti-self-adjoint, then (11.13.3) holds, so that (11.13.2) holds. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$ and x is invertible in $\mathcal{A}$, then x^{-1} commutes with y. If $x^* = x^{-1}$, then it follows that (11.13.3) holds, so that (11.13.2) holds again.

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a complex Hilbert space with associated norm $\| \cdot \|_V$ and $V \neq \{0\}$, and let A, B, and T be bounded linear mappings from V into itself. Suppose that A, B are normal, and that

$$A \circ T = T \circ B. \tag{11.13.4}$$

Under these conditions,

$$A^* \circ T = T \circ B^*, \tag{11.13.5}$$

as in Theorem 12.16 on p300 of [162]. This was initially shown by Fuglede when $A = B$, and extended to this formulation by Putnam. The proof that follows was found by Rosenblum, as in [162].

Using (11.13.4), we get that

$$A^l \circ T = T \circ B^l \tag{11.13.6}$$

for all positive integer l. This implies that

$$(\exp(\overline{\lambda} \, A)) \circ T = T \circ (\exp(\overline{\lambda} \, B)) \tag{11.13.7}$$

for all $\lambda \in \mathbf{C}$. Equivalently, this means that

$$T = (\exp(-\overline{\lambda} \, A)) \circ T \circ (\exp(\overline{\lambda} \, B)) \tag{11.13.8}$$

for every $\lambda \in \mathbf{C}$.

Put

$$f(\lambda) = (\exp(\lambda \, A^*)) \circ T \circ (\exp(-\lambda \, B^*)) \tag{11.13.9}$$

for every $\lambda \in \mathbf{C}$. Observe that

$$f(\lambda) = (\exp(\lambda\, A^* - \overline{\lambda}\, A)) \circ T \circ (\exp(-\lambda\, B^* + \overline{\lambda}\, B)) \tag{11.13.10}$$

for every $\lambda \in \mathbf{C}$, because of (11.13.8), and the normality of A, B. We also have that

$$\exp(\lambda\, A^* - \overline{\lambda}\, A),\ \exp(-\lambda\, B^* + \overline{\lambda}\, B) \tag{11.13.11}$$

are unitary operators on V for every $\lambda \in \mathbf{C}$, because

$$\lambda\, A^* - \overline{\lambda}\, A,\ -\lambda\, B^* + \overline{\lambda}\, B \tag{11.13.12}$$

are anti-self-adjoint on V. This implies that

$$\|f(\lambda)\|_{op} = \|T\|_{op} \tag{11.13.13}$$

for every $\lambda \in \mathbf{C}$.

We may consider $f(\lambda)$ as a holomorphic function of $\lambda \in \mathbf{C}$ with values in the algebra $\mathcal{BL}(V)$ of bounded linear mappings from V into itself. In fact, $f(\lambda)$ may be expressed as an absolutely convergent power series in λ with coefficients in $\mathcal{BL}(V)$, using Cauchy products. In particular, if $v, w \in V$, then

$$f_{v,w}(\lambda) = \langle (f(\lambda))(v), w\rangle_V \tag{11.13.14}$$

is a holomorphic complex-valued function of $\lambda \in \mathbf{C}$. Note that

$$|f_{v,w}(\lambda)| \le \|T\|_{op}\, \|v\|_V\, \|w\|_V \tag{11.13.15}$$

for every $\lambda \in \mathbf{C}$, because of (11.13.13).

Thus Liouville's theorem implies that $f_{v,w}(\lambda)$ is constant as a function of $\lambda \in \mathbf{C}$, so that

$$f_{v,w}(\lambda) = f_{v,w}(0) \tag{11.13.16}$$

for every $\lambda \in V$. This means that

$$(\exp(\lambda\, A^*)) \circ T = T \circ (\exp(\lambda\, B^*)) \tag{11.13.17}$$

for every $\lambda \in \mathbf{C}$. One can get (11.13.5) by considering the derivative of both sides of (11.13.17) in λ at 0.

The analogous statement for real Hilbert spaces can be reduced to the complex case using complexification. This will be discussed further in the next chapter.

11.14 Similar normal operators

Let $(V, \langle \cdot, \cdot \rangle_V)$ be a complex Hilbert space with associated norm $\| \cdot \|_V$ and $V \neq \{0\}$ again. Also let A, B, and T be bounded linear mappings from V into itself such that A and B are normal, T is invertible, and

$$A = T \circ B \circ T^{-1}. \tag{11.14.1}$$

If $T = U \circ P$ is the polar decomposition of T, as in Section 11.11, then

$$A = U \circ B \circ U^{-1}. \tag{11.14.2}$$

This is Theorem 12.36 on p316 of [162], which is due to Putnam.

Of course, (11.14.1) is the same as saying that $A \circ T = T \circ B$. This implies that $A^* \circ T = T \circ B^*$, as in the previous section. It follows that

$$T^* \circ A = (A^* \circ T)^* = (T \circ B^*)^* = B \circ T^*. \tag{11.14.3}$$

Remember that $P \circ P = T^* \circ T$, so that

$$B \circ P \circ P = B \circ T^* \circ T = T^* \circ A \circ T = T^* \circ T \circ B = P \circ P \circ B. \tag{11.14.4}$$

We also have that P is in the closure of the subalgebra of $\mathcal{BL}(V)$, as an algebra over the real numbers, generated by $T^* \circ T$ and I, as in Section 11.11. This means that P is in the closure of the subalgebra of $\mathcal{BL}(V)$, as an algebra over the real numbers, generated by $P \circ P$ and I. It follows that

$$B \circ P = P \circ B, \tag{11.14.5}$$

because of (11.14.4). It is easy to obtain (11.14.2) from (11.14.1) and (11.14.5).

The arguments in this section also work for real Hilbert spaces, as long as we have (11.13.5). That will be discussed in the next chapter, as mentioned in the previous section.

11.15 More on square roots

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. Put

$$f(x) = (e_{\mathcal{A}} + x)^2 = e_{\mathcal{A}} + 2\,x + x^2 \tag{11.15.1}$$

for each $x \in \mathcal{A}$. Thus

$$\begin{aligned} f(x) - f(y) &= 2\,(x-y) + x^2 - y^2 \\ &= 2\,(x-y) + x\,(x-y) + (x-y)\,y \end{aligned} \tag{11.15.2}$$

for every $x, y \in \mathcal{A}$.

It follows that

$$2\,\|x-y\|_{\mathcal{A}} \le \|f(x) - f(y)\|_{\mathcal{A}} + (\|x\|_{\mathcal{A}} + \|y\|_{\mathcal{A}})\,\|x-y\|_{\mathcal{A}} \tag{11.15.3}$$

for every $x, y \in \mathcal{A}$. Equivalently, this means that

$$(2 - \|x\|_{\mathcal{A}} - \|y\|_{\mathcal{A}})\,\|x-y\|_{\mathcal{A}} \le \|f(x) - f(y)\|_{\mathcal{A}} \tag{11.15.4}$$

for every $x, y \in \mathcal{A}$. In particular, if

$$\|x\|_{\mathcal{A}} + \|y\|_{\mathcal{A}} < 2 \tag{11.15.5}$$

and

$$f(x) = f(y), \tag{11.15.6}$$

then we get that

$$x = y. \tag{11.15.7}$$

If $x \in \mathcal{A}$, then

$$2\,\|x\|_{\mathcal{A}} \le \|f(x) - e_{\mathcal{A}}\|_{\mathcal{A}} + \|x\|_{\mathcal{A}}^2, \tag{11.15.8}$$

by (11.15.3), with $y = 0$. In particular,

$$\|x\|_{\mathcal{A}} \le \|f(x) - e_{\mathcal{A}}\|_{\mathcal{A}} \tag{11.15.9}$$

when $\|x\|_{\mathcal{A}} \le 1$. More precisely, (11.15.8) implies that

$$1 - \|f(x) - e_{\mathcal{A}}\|_{\mathcal{A}} \le 1 - 2\,\|x\|_{\mathcal{A}} + \|x\|_{\mathcal{A}}^2 = (1 - \|x\|_{\mathcal{A}})^2. \tag{11.15.10}$$

If $\|x\|_{\mathcal{A}}, \|f(x) - e_{\mathcal{A}}\|_{\mathcal{A}} \le 1$, then it follows that

$$(1 - \|f(x) - e_{\mathcal{A}}\|_{\mathcal{A}})^{1/2} \le 1 - \|x\|_{\mathcal{A}}. \tag{11.15.11}$$

This means that

$$\|x\|_{\mathcal{A}} \le 1 - (1 - \|f(x) - e_{\mathcal{A}}\|_{\mathcal{A}})^{1/2}. \tag{11.15.12}$$

Let $a \in \mathcal{A}$ be given, and suppose that we want to find $x \in \mathcal{A}$ such that

$$f(x) = e_{\mathcal{A}} + a. \tag{11.15.13}$$

Put

$$g(x) = x + (1/2)\,x^2 \tag{11.15.14}$$

for each $x \in \mathcal{A}$, so that (11.15.13) is the same as saying that

$$g(x) = (1/2)\,a. \tag{11.15.15}$$

Similarly, put

$$\phi_a(x) = (1/2)\,a - (1/2)\,x^2 \tag{11.15.16}$$

for each $x \in \mathcal{A}$, so that (11.15.15) is equivalent to

$$\phi_a(x) = x. \tag{11.15.17}$$

If $x, y \in \mathcal{A}$, then

$$\begin{aligned} \phi_a(x) - \phi_a(y) &= (1/2)\,(y^2 - x^2) \\ &= (1/2)\,(y - x)\,y + (1/2)\,x\,(y - x). \end{aligned} \tag{11.15.18}$$

This implies that

$$\|\phi_a(x) - \phi_a(y)\|_{\mathcal{A}} \le (1/2)\,(\|x\|_{\mathcal{A}} + \|y\|_{\mathcal{A}})\,\|x - y\|_{\mathcal{A}}. \tag{11.15.19}$$

Note that

$$\|\phi_a(x)\|_{\mathcal{A}} \le (1/2)\,\|a\|_{\mathcal{A}} + (1/2)\,\|x\|_{\mathcal{A}}^2 \tag{11.15.20}$$

for every $x \in \mathcal{A}$.

Let $\overline{B}(0, r)$ be the closed ball in $\mathcal{A}$ centered at 0 with radius $r \geq 0$ with respect to the metric associated to $\| \cdot \|_{\mathcal{A}}$. If $x, y \in \overline{B}(0, r)$, then

$$\|\phi_a(x) - \phi_a(y)\|_{\mathcal{A}} \leq r \, \|x - y\|_{\mathcal{A}}, \tag{11.15.21}$$

by (11.15.19). Similarly, if $a, x \in \overline{B}(0, r)$, then

$$\|\phi_a(x)\|_{\mathcal{A}} \leq (1/2) \, r + (1/2) \, r^2, \tag{11.15.22}$$

by (11.15.20). If $r \leq 1$, then (11.15.22) implies that

$$\|\phi_a(x)\| \leq r. \tag{11.15.23}$$

This means that

$$\phi_a(\overline{B}(0, r)) \subseteq \overline{B}(0, r) \tag{11.15.24}$$

when $\|a\|_{\mathcal{A}} \leq r$ and $r \leq 1$.

Suppose that $\mathcal{A}$ is complete with respect to the metric associated to $\| \cdot \|_{\mathcal{A}}$, so that $\overline{B}(0, r)$ with respect to the restriction of this metric to $\overline{B}(0, r)$, as in Section 1.6. If $\|a\|_{\mathcal{A}} < 1$, then we take take $r = \|a\|_{\mathcal{A}}$ in (11.15.21) and (11.15.24) to get that ϕ_a has a unique fixed point x_a in $\overline{B}(0, \|a\|_{\mathcal{A}})$, by the contraction mapping theorem. This means that x_a is the unique element of $\overline{B}(0, \|a\|_{\mathcal{A}})$ that satisfies (11.15.15), or equivalently (11.15.13).

Chapter 12

Complexifications, nets, and C^* algebras

12.1 Complexifying real vector spaces

Let V be a vector space over the real numbers. We would like to define the *complexification* of V, which is a vector space $V_{\mathbf{C}}$ over the complex numbers. We start by defining $V_{\mathbf{C}}$ as a vector space over the real numbers to be the Cartesian product $V \times V$ with itself, with respect to coordinatewise addition and scalar multiplication. This is the same as the direct sum of V with itself, as a vector space over $\mathbf{R}$, as in Section 5.12.

Let us define multiplication by i on $V_{\mathbf{C}}$ by putting

$$i\,(v_1, v_2) = (-v_2, v_1) \tag{12.1.1}$$

for every $v_1, v_2 \in V$. Note that

$$i\,(i\,(v_1, v_2)) = i\,(-v_2, v_1)) = (-v_1, -v_2) \tag{12.1.2}$$

for every $v_1, v_2 \in V$. One can use (12.1.1) to define multiplication by any complex number on $V_{\mathbf{C}}$ in an obvious way, using also scalar multiplication by real numbers. One can check that this makes $V_{\mathbf{C}}$ into a vector space over the complex numbers.

We may identify $v \in V$ with $(v, 0) \in V_{\mathbf{C}}$, so that V corresponds to $V \times \{0\}$, as a real-linear subspace of $V_{\mathbf{C}}$. Thus $(v_1, v_2) \in V_{\mathbf{C}}$ may be expressed as $v_1 + i\,v_2$. Observe that

$$C_V(v_1 + i\,v_2) = v_1 - i\,v_2 \tag{12.1.3}$$

defines a one-to-one conjugate-linear mapping from $V\mathbf{C}$ onto itself.

Suppose for the moment that V is a linear subspace of the space of real-valued functions on a nonempty set X. In this case, $V_{\mathbf{C}}$ may be identified with the linear subspace of the space of all complex-valued functions on X whose real and imaginary parts are elements of V. Using this identification, (12.1.3) corresponds to taking the complex conjugate of a complex-valued function on X.

Let W be a vector space over the complex numbers, which may also be considered as a vector space over the real numbers. If T is a real-linear mapping from V into W, then

$$T_{\mathbf{C}}(v_1 + i\, v_2) = T(v_1) + i\, T(v_2) \tag{12.1.4}$$

define s complex-linear mapping from $V_{\mathbf{C}}$ into W. This is the unique complex-linear mapping from $V_{\mathbf{C}}$ into W that agrees with T on V.

Let Y be another vector space over the real numbers, and let $Y_{\mathbf{C}}$ be its complexification, as before. Suppose that R is a linear mapping from V into Y, which may be considered as a real-linear maping from V into $Y_{\mathbf{C}}$, by identifying Y with a real-linear subspace of $Y_{\mathbf{C}}$, as before. This leads to a complex-linear mapping $R_{\mathbf{C}}$ from $V_{\mathbf{C}}$ into $Y_{\mathbf{C}}$, as in the preceding paragraph. This may be called the *complexification* of R. Observe that

$$R_{\mathbf{C}} \circ C_V = C_Y \circ R_{\mathbf{C}}, \tag{12.1.5}$$

where C_Y is defined on $Y_{\mathbf{C}}$ as in (12.1.3).

Conversely, let A be a complex-linear mapping from $V_{\mathbf{C}}$ into $Y_{\mathbf{C}}$ such that

$$A \circ C_V = C_Y \circ A. \tag{12.1.6}$$

This implies that A maps V into Y, as real-linear subspaces of $V_{\mathbf{C}}$ and $Y_{\mathbf{C}}$, respectively. Let A_0 be the restriction of A to V, considered as a real-linear mapping into Y. It is easy to see that

$$A = (A_0)_{\mathbf{C}}. \tag{12.1.7}$$

Let Z be another vector space over the real numbers, with complexification $Z_{\mathbf{C}}$, and let B be a linear mapping from Y into Z. This leads to a complex-linear mapping $B_{\mathbf{C}}$ from $Y_{\mathbf{C}}$ into $Z_{\mathbf{C}}$, as before. Similarly, $B \circ R$ is a linear mapping from V into Z, which leads to a complex-linear mapping $(B \circ R)_{\mathbf{C}}$ from $V_{\mathbf{C}}$ into $Z_{\mathbf{C}}$. One can check that

$$(B \circ R)_{\mathbf{C}} = B_{\mathbf{C}} \circ R_{\mathbf{C}}. \tag{12.1.8}$$

12.2 Complexifying inner products

Let V be a vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_V$, and let $V_{\mathbf{C}}$ be the complexification of V, as in the previous sections. The *complexification* of $\langle \cdot, \cdot \rangle_V$ may be defined on $V_{\mathbf{C}}$ by

$$\begin{aligned} \langle v_1 + i\, v_2, w_1 + i\, w_2 \rangle_{V_{\mathbf{C}}} &= \langle v_1, v_2 \rangle_V + i\, \langle v_2, w_1 \rangle_V \\ &\quad - i\, \langle v_1, w_2 \rangle_V + \langle v_2, w_2 \rangle_V . \end{aligned} \tag{12.2.1}$$

It is easy to see that this defines a sesquilinear form on $V_{\mathbf{C}}$. In fact, it is an inner product on $V_{\mathbf{C}}$, because

$$\langle v_1 + i\, v_2, v_1 + i\, v_2\rangle_{V_{\mathbf{C}}} = \langle v_1, v_1\rangle_V + \langle v_2, v_2\rangle_V \tag{12.2.2}$$

for every $v_1, v_2 \in V$.

Note that

$$\operatorname{Re}\langle v_1 + i\, v_2, w_1 + i\, w_2\rangle_{V_{\mathbf{C}}} = \langle v_1, w_1\rangle_V + \langle v_2, w_2\rangle_V \tag{12.2.3}$$

for all $v_1, v_2, w_1, w_2 \in V$. This is the same as the inner product on $\mathbf{C}$, as a vector space over the real numbers, and considered as the direct sum of two copies of V, obtained from $\langle \cdot, \cdot\rangle_V$ on both copies of V as in Section 5.12. If $\|\cdot\|_V$, $\|\cdot\|_{V_{\mathbf{C}}}$ are the norms on V, $V_{\mathbf{C}}$, respectively, associated to their inner products, then

$$\|v_1 + i\, v_2\|_{V_{\mathbf{C}}}^2 = \|v_1\|_V^2 + \|v_2\|_V^2 \tag{12.2.4}$$

for all $v_1, v_2 \in V$. If V is complete with respect to the metric associated to $\|\cdot\|_V$, then $V_{\mathbf{C}}$ is complete with respect to the metric associated to $\|\cdot\|_{V_{\mathbf{C}}}$, as before.

Let $(W, \langle \cdot, \cdot\rangle_W)$ be another real inner product space, with complexification $(W_{\mathbf{C}}, \langle \cdot, \cdot\rangle_{W_{\mathbf{C}}})$, and associated norms $\|\cdot\|_W$, $\|\cdot\|_{W_{\mathbf{C}}}$, respectively. Also let T be a linear mapping from V into W, and let $T_{\mathbf{C}}$ be the corresponding complex-linear mapping from $V_{\mathbf{C}}$ into $W_{\mathbf{C}}$, as in the previous section. If $v_1, v_2 \in V$, then

$$\|T(v_1 + i\, v_2)\|_{W_{\mathbf{C}}}^2 = \|T(v_1) + i\, T(v_2)\|_{W_{\mathbf{C}}}^2 = \|T(v_1)\|_W^2 + \|T(v_2)\|_W^2, \tag{12.2.5}$$

where the second step is as in (12.2.4). If T is a bounded linear mapping from V into W, then one can use this to check that $T_{\mathbf{C}}$ is bounded as a linear mapping from $V_{\mathbf{C}}$ into $W_{\mathbf{C}}$, with

$$\|T_{\mathbf{C}}\|_{op, V_{\mathbf{C}} W_{\mathbf{C}}} = \|T\|_{op, VW}. \tag{12.2.6}$$

Suppose that V, W are real Hilbert spaces, so that $V_{\mathbf{C}}$, $W_{\mathbf{C}}$ are complex Hilbert spaces, as before. If T is a bounded linear mapping from V into W, then the adjoint T^* of T is a bounded linear mapping from W into V, as in Section 3.5. This leads to a bounded linear mapping $(T^*)_{\mathbf{C}}$ from $W_{\mathbf{C}}$ into $V_{\mathbf{C}}$, as in the preceding paragraph. One can verify that

$$(T^*)_{\mathbf{C}} = (T_{\mathbf{C}})^*, \tag{12.2.7}$$

the adjoint of $T_{\mathbf{C}}$. Thus one may denote this operator simply as $T_{\mathbf{C}}^*$.

One can use this to get the analogue of the statement in Section 11.13 for real Hilbert spaces from the previous version for complex Hilbert spaces, as mentioned earlier. This permits one to use the same arguments as in Section 11.14 for real Hilbert spaces, as before.

12.3 Convergence of nets

A binary relation $\preceq$ on a set A is said to be a *partial ordering* if it is reflexive, transitive, and satisfies

$$a = b \text{ when } a \preceq b \text{ and } b \preceq a \tag{12.3.1}$$

for all $a, b \in A$. Sometimes (12.3.1) is not included in the definition of a partial ordering, which may not lead to any additional complications for the present purposes. The term *pre-order* may also be used for binary relations that are reflexive and transitive, but may not satisfy (12.3.1). If a partial ordering $\preceq$ on A has the additional property that for every $a, b \in A$,

$$a \preceq b \text{ or } b \preceq a, \tag{12.3.2}$$

then $\preceq$ is said to be a *linear ordering* or *total ordering* on A.

A partially-ordered set $(A, \preceq)$ is said to be a *directed system* if

$$\text{for every } a, b \in A \text{ there is a } c \in A \text{ such that } a, b \preceq c. \tag{12.3.3}$$

Note that linearly-ordered sets are directed systems. Similarly, let us say that a pre-ordered set $(A, \preceq)$ is a *pre-directed system* if it satisfies (12.3.3). As before, the term directed system is sometimes used for this, which may not lead to additional complications for the present purposes.

Let $(A, \preceq)$ be a nonempty directed system or pre-directed system, and let Z be a set. A *net* of elements of Z indexed by A is a family $\{z_a\}_{a \in A}$ that associates to each $a \in A$ an element z_a of Z. This is the same as a function on A with values in Z, but we shall normally use this notation and terminology in this situation. If A is the set of positive integers with the standard ordering, then a net of elements of Z indexed by A is the same as a sequence of elements of Z.

Suppose now that Z is a metric or topological space. A net $\{z_a\}_{a \in A}$ of elements of Z indexed by A is said to *converge* to a point $z \in Z$ if for every open set $V \subseteq Z$ with $z \in V$ there is an $a \in A$ such that

$$z_b \in V \tag{12.3.4}$$

for every $b \in A$ with $a \preceq b$. If (Z, d_Z) is a metric space, then this is equivalent to asking that for each $\epsilon > 0$ there be an $a \in A$ such that

$$d_Z(z, z_b) < \epsilon \tag{12.3.5}$$

for every $b \in A$ with $a \preceq b$. If A is the set of positive integers with the standard ordering, then convergence of a net of elements of Z indexed by A is the same as convergence of a sequence of elements of Z.

Let Y be a topological space, and let $\mathcal{B}(y)$ be a local base for the topology of Y at a point $y \in Y$. If $U, V \in \mathcal{B}(y)$, then put

$$U \preceq V \text{ when } V \subseteq U. \tag{12.3.6}$$

One can check that $\mathcal{B}(y)$ is a directed system with respect to $\preceq$. Let $\{y_U\}_{U \in \mathcal{B}(y)}$ be a net of elements of Y indexed by $\mathcal{B}(y)$. If

$$y_U \in U \tag{12.3.7}$$

for every $U \in \mathcal{B}(y)$, then it is easy to see that

$$\{y_U\}_{U \in \mathcal{B}(y)} \text{ converges to } y \tag{12.3.8}$$

as a net of elements of Y.

One can check that the limit of a convergent net in a metric space is unique. Similarly, the limit of a convergent net in a Hausdorff topological space is unique.

Let $(A, \preceq)$ be a nonempty directed system or pre-directed system again, and let (Z, d_Z) be a metric space. A net $\{z_a\}_{a \in A}$ of elements of Z is said to be a *Cauchy net* if for every $\epsilon > 0$ there is an $a \in A$ such that

$$d_Z(z_b, z_c) < \epsilon \tag{12.3.9}$$

for every $b, c \in A$ with $a \preceq b, c$. If A is the set of positive integrs with the standard ordering, then a Cauchy net of elements of Z indexed by A is the same as a Cauchy sequence in Z. It is easy to see that a convergence net of elements of Z is a Cauchy net.

If Z is complete with respect to d_Z, and $\{z_a\}_{a \in A}$ is a Cauchy net of elements of Z indexed by A, then it is well known that $\{z_a\}_{a \in A}$ converges to an element z of Z. Indeed, for each positive integer j there is an $a_j \in A$ such that

$$d_Z(z_b, z_c) < 1/j \tag{12.3.10}$$

for every $b, c \in A$ with $a_j \preceq b_j, c_j$. We can also choose $\alpha_j \in A$ recursively for each $j \geq 1$ in such a way that

$$a_j \preceq \alpha_j \tag{12.3.11}$$

and

$$\alpha_{j-1} \preceq \alpha_j \tag{12.3.12}$$

when $j \geq 2$, because A is a (pre-)directed system. It is easy to see that $\{z_{\alpha_j}\}_{j=1}^{\infty}$ is a Cauchy sequence in Z under these conditions. If Z is complete, then $\{z_{\alpha_j}\}_{j=1}^{\infty}$ converges to an element z of Z, and one can verify that $\{z_a\}_{a \in A}$ converges to z as well.

12.4 More on convergence of nets

Let $(A, \preceq)$ be a nonempty directed or pre-directed system again, and let Y, Z be metric or topological spaces. Suppose that $\{y_a\}_{a \in A}$ is a net of elements of Y indexed by A that converges to $y \in Y$. If a mapping ϕ from Y into Z is continuous at y, then it is easy to see that

$$\{\phi(y_a)\}_{a \in A} \text{ converges to } \phi(y), \tag{12.4.1}$$

as a net of elements of Z.

Conversely, suppose that ϕ is not continuous at y. This means that there is an open set $W \subseteq Z$ such that $\phi(y) \in W$ and

$$\phi(U) \not\subseteq W \tag{12.4.2}$$

for any open set $U \subseteq Y$ with $y \in U$. Let $\mathcal{B}(y)$ be a local base for the topology of Y at y, which may be considered as a directed system with as in the previous section. If $U \in \mathcal{B}(y)$, then let y_U be an element of U such that

$$\phi(y_U) \notin W. \tag{12.4.3}$$

Under these conditions, $\{y_U\}_{U \in \mathcal{B}(x)}$ converges to y in Y, as in (12.3.8), but $\{\phi(y_U)\}_{U \in \mathcal{B}(x)}$ does not converge to $\phi(y)$ in Z, because of (12.4.3).

Now le Y be a nonempty set, let $\{f_a\}_{a \in A}$ be a net of functions on Y with values in Z indexed by A, and let f be another Z-valued function on Y. We say that $\{f_a\}_{a \in A}$ converges *pointwise* to f on Y if

$$\{f_a(y)\}_{a \in A} \text{ converges to } f(y) \tag{12.4.4}$$

in Z for every $y \in Y$. Similarly, if (Z, d_Z) is a metric space, then $\{f_a\}_{a \in A}$ converges to f *uniformly* on Y if for every $\epsilon > 0$ there is an $a \in A$ such that

$$d_Z(f(y), f_b(y)) < \epsilon \tag{12.4.5}$$

for every $b \in A$ with $a \preceq b$ and $y \in Y$. These definitions reduce to the usual ones for sequences when A is the set of positive integers with the standard ordering. Of course, uniform convergence automatically implies pointwise convergence.

Suppose that (Z, d_Z) is a metric space, and let $\mathcal{B}(Y, Z)$ be the space of all bounded mappings from Y into Z, as in Section 7.8. As in the case of sequences, convergence of a net in $\mathcal{B}(Y, Z)$ with respect to the supremum metric is equivalent to uniform convergence. Let $\{f_a\}_{a \in A}$ be any net of Z-valued function on Y indexed by A again, and let f be another Z-valued function on Y. If Y is a metric or topological space, $y \in Y$, f_a is continuous at y for each $a \in A$, and $\{f_a\}_{a \in A}$ converges to f uniformly on Y, then

$$f \text{ is continuous at } y \tag{12.4.6}$$

as well. If Y is a metric space, f_a is uniformly continuous on Y for each $a \in A$, and $\{f_a\}_{a \in A}$ converges to f uniformly on Y, then

$$f \text{ is uniformly continuous on } Y \tag{12.4.7}$$

too.

Suppose now that Z is a vector space over the real or complex numbers with a norm $\|\cdot\|_Z$, let $\{z_a\}_{a \in A}$ be a net of elements of Z indexed by A, and let z be another element of Z. Suppose for the moment that $\{z_a\}_{a \in A}$ converges to z with respect to the metric on Z associated to $\| \cdot \|_Z$. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then one can check that

$$\{t\, z_a\}_{a \in A} \text{ converges to } t\, z \text{ in } Z. \tag{12.4.8}$$

Similarly, if $\{w_a\}_{a \in A}$ is a net of elements of Z that converges to $w \in Z$, then

$$\{w_a + z_a\}_{a \in A} \text{ converges to } w + z \text{ in } Z. \tag{12.4.9}$$

We say that $\{z_a\}_{a \in A}$ *converges weakly* to $z \in Z$ if for every bounded linear functional λ on Z,

$$\{\lambda(z_a)\}_{a \in A} \text{ converges to } \lambda(z) \tag{12.4.10}$$

in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. If A is the set of positive integers with the standard ordering, then this is the same weak convergence of sequences, as in Section 4.9. Otherwise, this is the same as the convergence of the net with respect to the weak topology on Z. Weak convergence is implied by convergence with respect to the metric associated to the norm, as before. The weak limit of a net in Z is unique when it exists, because the dual space Z' of bounded linear functionals on Z separates points in Z, by the Hahn–Banach theorem. Note that weak convergence of nets satisfies the same type of properties for sums and scalar multiples of convergent nets as in the preceding paragraph. This follows from the analogous statements for convergent nets in $\mathbf{R}$ and $\mathbf{C}$, which may be considered as particular cases of the previous statements.

Suppose that Y is another vector space over the real or complex numbers, as appropriate, with a norm $\| \cdot \|_Y$. Also let $\{T_a\}_{a \in A}$ be a net of bounded linear mappings from Y into Z indexed by A, and let T be another bounded linear mapping from Y into Z. Under these conditions, $\{T_a\}_{a \in A}$ converges to T with respect to the strong operator topology on the space $\mathcal{BL}(Y, Z)$ of bounded linear mappings from Y into Z if and only if $\{T_a\}_{a \in A}$ converges to T pointwise on Y. If A is the set of positive integers with the standard ordering, then this corresponds to the type of convergence discussed in Section 4.4. If $\{T_a\}_{a \in A}$ converges to T with respect to the metric on $\mathcal{BL}(Y, Z)$ associated to the operator norm, then it is easy to see that $\{T_a\}_{a \in A}$ converges to T pointwise on Y.

In particular, we can take $Z = \mathbf{R}$ or $\mathbf{C}$, with the standard absolute value function as the norm. Let $\{\mu_a\}_{a \in A}$ be a net of bounded linear functionals on Y indexed by A, and let μ be another bounded linear functional on Y. Pointwise convergence of $\{\mu_a\}_{a \in A}$ to μ on Y is equivalent to the convergence of $\{\mu_a\}_{a \in A}$ to μ with respect to the weak* topology on the dual space Y' of bounded linear functionals on Y with respect to $\| \cdot \|_Y$. If A is the set of positive integers with the standard ordering, then this corresponds to the type of convergence mentioned in Section 4.5.

12.5 Nets and approximate eigenvalues

Let V be a vector space over the real or complex numbers, and let v be a nonzero element of V. Consider the collection $\mathcal{A}_v$ of linear mappings T from V into itself such that v is an eigenvector of T with eigenvalue $\lambda_v(T)$. Note that $\lambda_v(T)$ is unique in this case, because $v \neq 0$. It is easy to see that $\mathcal{A}_v$ is a subalgebra of the algebra $\mathcal{L}(V)$ of all linear mappings from V into itself, and of course $\mathcal{A}_v$ contains the identity mapping $I = I_V$ on V. We also have that λ_v is an algebra homomorphism from $\mathcal{A}_v$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with $\lambda_v(I) = 1$.

Suppose now that $\|\cdot\|_V$ is a norm on V, and that $(A, \preceq)$ is a nonempty directed or pre-directed system. Let $\{v_a\}_{a \in A}$ be a net of elements of V such that

$$\|v_a\|_V = 1 \tag{12.5.1}$$

for each $a \in A$. Consider the collection $\mathcal{A}$ of bounded linear mappings T from V into itself for which there is a real or complex number $\lambda(T)$, as appropriate, such that

$$\{\|T(v_a) - \lambda(T)\, v_a\|_V\}_{a \in A} \text{ converges to } 0, \tag{12.5.2}$$

as a net of real numbers. One can check that $\lambda(T)$ is uniquely determined by (12.5.2), and satisfies

$$|\lambda(T)| \le \|T\|_{op}. \tag{12.5.3}$$

We also have that $\lambda(T)$ is an approximate eigenvalue of T, as in Section 9.3.

It is easy to see that $\mathcal{A}$ is a linear subspace of the space $\mathcal{BL}(V)$ of all bounded linear mappings from V into itself, and that λ defines a linear functional on $\mathcal{A}$. In fact, one can verify that $\mathcal{A}$ is a subalgebra of $\mathcal{BL}(V)$, and that λ is an algebra homomorphism from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Clearly $I \in \mathcal{A}$, with $\lambda(I) = 1$.

Let us check that $\mathcal{A}$ is a closed set in $\mathcal{BL}(V)$, with respect to the metric associated to the operator norm. Let $\overline{\mathcal{A}}$ be the closure of $\mathcal{A}$ in $\mathcal{BL}(V)$ with respect to this metric. There is a unique extension of λ to a bounded linear functional on $\overline{\mathcal{A}}$ with respect to the operator norm, as in Section 2.2, and we shall also use λ to denote this extension. If $T \in \overline{\mathcal{A}}$, then one can verify that (12.5.2) holds, by approximating T by elements of $\mathcal{A}$. This implies that $T \in \mathcal{A}$, so that $\overline{\mathcal{A}} = \mathcal{A}$.

Now let $(V, \langle \cdot, \cdot \rangle_V)$ be a real or complex Hilbert space, and let T_0 be a bounded linear mapping from V into itself that is normal, in the sense that it commutes with its adjoint. Also let μ_0 be a real or complex number, as appropriate, in the spectrum of T_0. This implies that μ_0 is an approximate eigenvalue of T_0, as in Section 9.5.

It follows that there is a sequence $\{v_j\}_{j=1}^\infty$ of unit vectors in V such that

$$\lim_{j \to \infty} \|T_0(v_j) - \mu_0\, v_j\|_V = 0, \tag{12.5.4}$$

as in Section 9.3. We also get that

$$\lim_{j \to \infty} \|T_0^*(v_j) - \mu_0\, v_j\|_V = 0 \tag{12.5.5}$$

in the real case, and

$$\lim_{j \to \infty} \|T_0^*(v_j) - \overline{\mu_0}\, v_j\|_V = 0 \tag{12.5.6}$$

in the complex case, as in Section 9.5.

Consider the collection $\mathcal{A}_0$ of bounded linear mappings T from V into itself for which there is a real or complex number $\lambda_0(T)$, as appropriate, such that

$$\lim_{j \to \infty} \|T(v_j) - \lambda_0(T)\, v_j\|_V = 0. \tag{12.5.7}$$

This is a closed subalgebra of $\mathcal{BL}(V)$ that contains I, as before, and which contains T_0, T_0^* in this case.

12.6 C^* Algebras

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the complex numbers with a conjugate-linear algebra involution $x \mapsto x^*$. Suppose that $\|\cdot\|_{\mathcal{A}}$ satisfies the C^* identity

$$\|x^*\, x\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^2 \tag{12.6.1}$$

for every $x \in \mathcal{A}$, as in Section 7.7. Under these conditions, $\mathcal{A}$ is called a C^* *algebra* with respect to the norm and involution. Remember that (12.6.1) implies that

$$\|x^*\|_{\mathcal{A}} = \|x\|_{\mathcal{A}} \tag{12.6.2}$$

for every $x \in \mathcal{A}$, as in Section 7.7.

This corresponds to Definition 2.2.1 on p46 of [8], and is mentioned on p260 of [167]. If $\mathcal{A}$ has a nonzero multiplicative identity element $e_{\mathcal{A}}$, then we have that $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, as in Section 7.7. This may sometimes be included in the definition of a C^* algebra.

Sometimes the term B^* *algebra* is used for what is called a C^* algebra here, as in Definition 11.17 on p276 of [162]. However, this term is also sometimes used for a complex Banach algebra with a norm-preserving conjugate-linear involution, as on p260 of [167]. The latter may be called a *Banach $*$-algebra* as well, as in Definition 2.5.1 on p57 of [8]. A conjugate-linear algebra involution on a complex associative algebra may be denoted $x \mapsto x^{\sim}$, as in Definition 21.6 on p313 of [91]. Similarly, a complex Banach algebra with a conjugate-linear involution that preserves the norm may be caller a *Banach $^{\sim}$-algebra*, as in [91].

If X is a nonempty set, then the space $\ell^\infty(X, \mathbf{C})$ of bounded complex-valued functions on X is a commutative complex Banach algebra with respect to the supreum norm. It is easy to see that

$$f \mapsto \overline{f} \tag{12.6.3}$$

is a conjugate-linear algebra involution on $\ell^\infty(X, \mathbf{C})$. If $f \in \ell^\infty(X, \mathbf{C})$, then

$$\|\overline{f}\, f\|_\infty = \||f|^2\|_\infty = \|f\|_\infty^2, \tag{12.6.4}$$

so that $\ell^\infty(X, \mathbf{C})$ is a C^* algebra with respect to this involution.

Let $(V, \langle\cdot,\cdot\rangle_V)$ be a complex Hilbert space, with the associated norm $\|\cdot\|_V$, and remember that the algebra $\mathcal{BL}(V)$ of bounded linear mappings from V into itself is a complex Banach algebra with respect to the corresponding operator norm. We have seen that the Hilbert space adjoint defines a conjugate-linear involution on $\mathcal{BL}(V)$ that satisfies the C^* identity, as in Section 3.5, so that $\mathcal{BL}(V)$ is a C^* algebra.

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a complex Banach algebra again. If $\mathcal{A}_1$ is a subalgebra of $\mathcal{A}$ that is also a closed set with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, then $\mathcal{A}_1$ may be considered as a complex Banach algebra with respect to the restriction of $\|\cdot\|_{\mathcal{A}}$ to $\mathcal{A}_1$. If $x \mapsto x^*$ is a conjugate-linear involution on $\mathcal{A}$, and if $x^* \in \mathcal{A}_1$ for each $x \in \mathcal{A}_1$, then the restriction of $x \mapsto x^*$ to $x \in \mathcal{A}_1$ defines an involution on $\mathcal{A}_1$. If $\mathcal{A}$ is a C^* algebra, and $\mathcal{A}_1$ has all of the properties just mentioned, then $\mathcal{A}_1$ is a C^* algebra with respect to the restrictions of the norm and involution to $\mathcal{A}_1$.

Let V be a complex Hilbert space again, and let $\mathcal{A}_1$ be a closed subalgebra of $\mathcal{BL}(V)$ that contains the adjoints of all of its elements. Thus $\mathcal{A}_1$ is a C^* algebra with respect to the restrictions of the operator norm and the involution defined by taking the adjoint to $\mathcal{A}_1$, as in the preceding paragraph. In this case, $\mathcal{A}_1$ may be called a C^* *algebra of operators*, as in Definition 2.1.2 on p42 of [8].

Let X be a nonempty metric or topological space. The algebra $C_b(X, \mathbf{C})$ of bounded continuous complex-valued functions on X is a closed subalgebra of $\ell^\infty(X, \mathbf{C})$ that is invariant under complex conjugation, and thus a C^* algebra with respect to the supremum norm and (12.6.3).

12.7 Some remarks about involutions

Let $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{A}_3$ be algebras in the strict sense, all real or all complex. Also let ϕ be a mapping from $\mathcal{A}_1$ into $\mathcal{A}_2$, and let ψ be a mapping from $\mathcal{A}_2$ into $\mathcal{A}_3$. Of course, if ϕ and ψ are linear mappings, then their composition $\psi \circ \phi$ is a linear mapping from $\mathcal{A}_1$ into $\mathcal{A}_3$. In the complex case, we may be concerned with situations where one of ϕ and ψ is complex-linear and the other is conjugate-linear, which implies that

$$\psi \circ \phi \text{ is conjugate-linear} \tag{12.7.1}$$

as well. If ϕ and ψ are both conjugate-linear, then

$$\psi \circ \phi \text{ is complex-linear.} \tag{12.7.2}$$

If ϕ and ψ are both algebra homomorphisms, then $\psi \circ \phi$ is an algebra homomorphism too, as in Section 6.3. If one of ϕ and ψ is an algebra homomorphism and the other is an opposite algebra homomorphism, then

$$\psi \circ \phi \text{ is an opposite algebra homomorphism.} \tag{12.7.3}$$

If ϕ and ψ are both opposite algebra homomorphisms, then

$$\psi \circ \phi \text{ is an algebra homomorphism.} \tag{12.7.4}$$

Let $a_1 \mapsto a_1^{*_1}$ and $a_2 \mapsto a_2^{*_2}$ be involutions on $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively. In the complex case, these involutions should be both complex-linear or both conjugate-linear. An algebra homomorphism ϕ from $\mathcal{A}_1$ into $\mathcal{A}_2$ is said to be a $*$-*homomorphism* if

$$\phi(a_1^{*_1}) = \phi(a_1)^{*_2} \tag{12.7.5}$$

for every $a_1 \in \mathcal{A}_1$. One may also consider ϕ to be a homomorphism from $\mathcal{A}_1$ into $\mathcal{A}_2$ as algebras in the strict sense with involutions in this case. Similarly, we may refer to $*$-*isomorphisms* and $*$-*automorphisms* for algebra isomorphisms and automorphisms that are $*$-homomorphisms, which may be considered as isomorphisms or automorphisms of algebras in the strict sense with involutions, as appropriate, as well.

Now let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$, and let $x \mapsto x^*$ be an algebra involution on $\mathcal{A}$, which may be conjugate-linear in the complex case. Suppose that a is an invertible element of $\mathcal{A}$, so that

$$\phi_a(x) = a\, x\, a^{-1} \tag{12.7.6}$$

defines an algebra automorphism of $\mathcal{A}$. Thus

$$x^{\#} = a\, x^*\, a^{-1} \tag{12.7.7}$$

defines an opposite algebra isomorphism from $\mathcal{A}$ onto itself, as before. In the complex case, if $x \mapsto x^*$ is conjugate-linear, then (12.7.7) is conjugate-linear in x too.

If $x \in \mathcal{A}$, then

$$(x^{\#})^{\#} = a\,(x^{\#})^*\, a^{-1} = a\,(a\, x^*\, a^{-1})^*\, a^{-1} = a\,(a^{-1})^*\, x\, a^*\, a^{-1}. \tag{12.7.8}$$

Suppose for the moment that

$$a^* = t\, a \tag{12.7.9}$$

for some nonzero real or complex number t, as appropriate. This implies that

$$(a^{-1})^* = (a^*)^{-1} = t^{-1}\, a^{-1}. \tag{12.7.10}$$

It follows that

$$(x^{\#})^{\#} = x. \tag{12.7.11}$$

This means that $x \mapsto x^{\#}$ defines another algebra involution on $\mathcal{A}$ when (12.7.9) holds, as in Exercise 27 on p326 of [162].

Note that

$$\phi_a(x)^* = (a^{-1})^*\, x^*\, a^* \tag{12.7.12}$$

for every $x \in \mathcal{A}$. Suppose for the moment again that

$$a^* = \tau a^{-1} \tag{12.7.13}$$

for some nonzero real or complex number τ, as appropriate. This means that

$$(a^{-1})^* = (a^*)^{-1} = \tau^{-1}\, a. \tag{12.7.14}$$

Using this, we get that

$$\phi_a(x)^* = \phi_a(x^*) \tag{12.7.15}$$

for every $x \in \mathcal{A}$, so that ϕ_a is a $*$-automorphism of $\mathcal{A}$. This is related to Exercises (7) and (8) on p50 of [8].

Let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and note that $1 \le \|a\|_{\mathcal{A}}\,\|a^{-1}\|_{\mathcal{A}}$. Suppose that

$$\|a\|_{\mathcal{A}}\,\|a^{-1}\|_{\mathcal{A}} = 1. \tag{12.7.16}$$

If $y \in \mathcal{A}$, then

$$\|y\|_{\mathcal{A}} \le \|a^{-1}\|_{\mathcal{A}}\,\|a\,y\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}^{-1}\,\|a\,y\|_{\mathcal{A}}. \tag{12.7.17}$$

This implies that

$$\|a\,y\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}\,\|y\|_{\mathcal{A}}, \tag{12.7.18}$$

because of submultiplicativity of $\|\cdot\|_{\mathcal{A}}$. Similarly,

$$\|y\,a\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}\,\|y\|_{\mathcal{A}} \tag{12.7.19}$$

and

$$\|a^{-1}\,y\|_{\mathcal{A}} = \|y\,a^{-1}\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}^{-1}\,\|y\|_{\mathcal{A}}. \tag{12.7.20}$$

It follows that

$$\|\phi_a(y)\|_{\mathcal{A}} = \|a\,y\,a^{-1}\|_{\mathcal{A}} = \|y\|_{\mathcal{A}} \tag{12.7.21}$$

under these conditions. This implies that

$$\|x^{\#}\|_{\mathcal{A}} = \|x^*\|_{\mathcal{A}} \tag{12.7.22}$$

for every $x \in \mathcal{A}$. In particular, if $x \mapsto x^*$ preserves the norm, then $x \mapsto x^{\#}$ has the same property.

12.8 Self-adjoint elements

Let $\mathcal{A}$ be an associative algebra over the complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$ and a conjugate-linear algebra involution $x \mapsto x^*$. Suppose that $a \in \mathcal{A}$ is self-adjoint with respect to this involution, so that $a^* = a$. It would be nice if

$$\sigma_{\mathcal{A}}(a) \subseteq \mathbf{R}. \tag{12.8.1}$$

Remember that this holds when $\mathcal{A}$ is the algebra of bounded linear mappings from a complex Hilbert space into itself, with the involution defined by the adjoint, as in Section 8.11.

Similarly, let h be a nonzero algebra homomorphism from $\mathcal{A}$ into $\mathbf{C}$, so that $h(e_{\mathcal{A}}) = 1$. It would be nice if

$$h(a) \in \mathbf{R}. \tag{12.8.2}$$

Remember that $h(a) \in \sigma_{\mathcal{A}}(a)$, as in Section 6.9, so that (12.8.1) implies (12.8.2). If $\mathcal{A}$ is a commutative Banach algebra, then every element of $\sigma_{\mathcal{A}}(a)$ is of the form $h(a)$ for some nonzero homomorphism h from $\mathcal{A}$ into $\mathbf{C}$, as in Section 6.12. This means that (12.8.1) would follow from (12.8.2) in this case.

Let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$ that satisfies the C^* identity $\|x^*\,x\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^2$ for every $x \in \mathcal{A}$. Let us show that (12.8.2) holds in this case, using the argument on p277 of [162]. Let α, β be the real and imaginary parts of $h(a)$, so that we would like to get that $\beta = 0$. Also let $t \in \mathbf{R}$ be given, and put

$$z = a + i\,t\,e_{\mathcal{A}}. \tag{12.8.3}$$

Thus

$$h(z) = h(a) + i\,t = \alpha + i\,(\beta + t) \tag{12.8.4}$$

and

$$z^*\,z = a^*\,a + t^2\,e_{\mathcal{A}} = a^2 + t^2\,e_{\mathcal{A}}. \tag{12.8.5}$$

Observe that

$$\alpha^2 + (\beta + t)^2 = |h(z)|^2 \le \|z\|_{\mathcal{A}}^2, \tag{12.8.6}$$

where the second step is as in Section 6.9. Using the C^* identity, we get that

$$\|z\|_{\mathcal{A}}^2 = \|z^*\,z\|_{\mathcal{A}} \le \|a^*\,a\|_{\mathcal{A}} + t^2 = \|a\|_{\mathcal{A}}^2 + t^2. \tag{12.8.7}$$

Combining (12.8.6) and (12.8.7), we obtain that

$$\alpha^2 + \beta^2 + 2\,\beta\,t + t^2 \le \|a\|_{\mathcal{A}}^2 + t^2, \tag{12.8.8}$$

so that

$$\alpha^2 + 2\,\beta\,t \le \|a\|_{\mathcal{A}}^2. \tag{12.8.9}$$

This implies that $\beta = 0$, because $t \in \mathbf{R}$ is arbitrary.

Alternatively, suppose that $\mathcal{A}$ is a C^* algebra, and let $t \in \mathbf{R}$ be given again. Remember that $\exp_{\mathcal{A}}(i\,t\,a)$ may be defined as in Section 10.4, and satisfies

$$(\exp_{\mathcal{A}}(i\,t\,a))^* = \exp_{\mathcal{A}}(-i\,t\,a^*) = \exp_{\mathcal{A}}(-i\,t\,a) = (\exp_{\mathcal{A}}(i\,t\,a))^{-1}. \tag{12.8.10}$$

This implies that

$$\|\exp_{\mathcal{A}}(i\,t\,a)\|_{\mathcal{A}} = 1, \tag{12.8.11}$$

by the C^* identity. We also have that

$$h(\exp_{\mathcal{A}}(i\,t\,a)) = \exp(i\,t\,h(a)), \tag{12.8.12}$$

where the right side uses the usual complex exponential function. This is because h is a bounded linear functional on $\mathcal{A}$, as well as an algebra homomorphism, as in Section 6.9.

Using the boundedness of h as a linear functional on $\mathcal{A}$ again, we get that

$$|\exp(i\,t\,h(a))| = |h(\exp_{\mathcal{A}}(i\,t\,a))| \le 1. \tag{12.8.13}$$

This implies (12.8.2), because $t \in \mathbf{R}$ is arbitrary. This is the argument used on p48 of [8].

If $\mathcal{A}$ is a commutative C^* algebra, then it follows that (12.8.1) holds. If $\mathcal{A}$ is a C^* algebra that is not necessarily commutative, then observe that the subalgebra $\mathcal{A}_0$ of $\mathcal{A}$ generated by a and $e_{\mathcal{A}}$ is commutative and invariant under the involution. This means that the closure $\overline{\mathcal{A}_0}$ of $\mathcal{A}_0$ in $\mathcal{A}$ with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$ is a commutative C^* algebra. It follows that

$$\sigma_{\overline{\mathcal{A}_0}}(a) \subseteq \mathbf{R}, \tag{12.8.14}$$

as before. Remember that $\sigma_{\mathcal{A}}(a)$ is contained in $\sigma_{\overline{\mathcal{A}_0}}(a)$, because $\overline{\mathcal{A}_0}$ is a subalgebra of $\mathcal{A}$, as in Section 7.3. Thus (12.8.1) follows from (12.8.14). This corresponds to Corollary 1 on p48 of [8], and to part (a) of Theorem 11.28 on p282 of [162].

As another approach, let $t \in \mathbf{R}$ be given again, and put $f_t(z) = \exp(i\,t\,z)$ on $\mathbf{C}$, which can be defined by a convergent power series using the power series for the exponential function on $\mathbf{C}$. Note that

$$f_t(\sigma_{\mathcal{A}}(a)) \subseteq \sigma_{\mathcal{A}}(\exp_{\mathcal{A}}(i\,t\,a)), \tag{12.8.15}$$

as in Section 10.3. The right side is contained in the closed unit disk in $\mathbf{C}$, because of (12.8.11), as in Section 6.8. This implies (12.8.1), because $t \in \mathbf{R}$ is arbitrary. This is a variant of the proof of the first part of part (iv) of 2.1 on p262 of [167]. In [167], one takes $t = 1$, and observes more precisely that the spectrum of $\exp_{\mathcal{A}}(i\,a)$ is contained in the unit circle. This permits one to obtain (12.8.1) from (12.8.15) with $t = 1$.

12.9 Some maximal ideals

Let X be a nonempty metric or topological space, and note that the spaces $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ of bounded continuous real and complex-valued functions on X, respectively, are Banach algebras with respect to the supremum norm. If $x \in X$, then

$$h_x(f) = f(x) \tag{12.9.1}$$

defines an algebra homomorphism from $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ onto $\mathbf{R}$, $\mathbf{C}$, respectively, as in Section 6.9. Thus the kernel of h_x is a maximal proper ideal in each of $C_b(X, \mathbf{R})$ and $C_b(X, \mathbf{C})$, as in Section 6.12.

Remember that X is said to be a Urysohn space if continuous real-valued functions on X separate points in X, as in Section 5.5. More precisely, this implies that bounded continuous real-valued functions on X separate points in X, as before. This means that if $x, y \in X$ and $x \neq y$, then

$$h_x(f) \neq h_y(f) \tag{12.9.2}$$

for some $f \in C_b(X, \mathbf{R})$. Of course, metric spaces are Urysohn spaces, as before.

Suppose that X is compact, so that $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ are the same as the algebras $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ of continuous real and complex-valued functions on

X, respectively. If $\mathcal{I}$ is a maximal proper ideal in $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, then it is well known that

$$\mathcal{I} = \ker h_x \tag{12.9.3}$$

for some $x \in X$. To see this, it suffices to show that

$$\mathcal{I} \subseteq \ker h_x \tag{12.9.4}$$

for some $x \in X$, because $\mathcal{I}$ is supposed to be a maximal proper ideal.

Suppose for the sake of a contradiction that for each $x \in X$, (12.9.4) does not hold. This means that for every $x \in X$ there is an $f_x \in \mathcal{I}$ such that

$$f_x(x) \neq 0. \tag{12.9.5}$$

We may as well suppose that f_x is a nonnegative real-valued continuous function on X for each $x \in X$, by replacing f_x with f_x^2 in the real case, and with $|f_x|^2 = f_x \, \overline{f_x}$ in the complex case, if necessary.

If $x \in X$, then $f_x^{-1}((0, +\infty))$ is an open subset of X that contains x. If X is compact, then there are finitely many elements $x_1, \ldots, x_n$ of X such that

$$X \subseteq \bigcup_{j=1}^{n} f_{x_j}^{-1}((0, +\infty)). \tag{12.9.6}$$

This means that

$$f = \sum_{j=1}^{n} f_{x_j} \tag{12.9.7}$$

is strictly positive at every point in X. Note that $f \in \mathcal{I}$, because $f_{x_j} \in \mathcal{I}$ for each j. This is a contradiction, because $\mathcal{I}$ is supposed to be a proper ideal in $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, and $1/f$ is continuous on X.

If h is any nonzero algebra homomorphism from $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, then the kernel of h is a maximal proper ideal in $C_b(X, \mathbf{R})$ or $C(X, \mathbf{C})$, as appropriate, as in Section 6.12. This implies that

$$\ker h = \ker h_x \tag{12.9.8}$$

for some $x \in X$, as before. One can use this to get that

$$h = h_x, \tag{12.9.9}$$

because $h(\mathbf{1}_X) = 1$, as in Section 6.9.

This is essentially the same argument as used in Example 11.13 (a) on p271 of [162]. Another argument is used in the proof of Theorem 1.10.4 on p28 of [8].

If X is a compact Hausdorff topological space, then it is well known that X is normal. In particular, this means that X is a Urysohn space, by Urysohn's lemma. This implies that there is only one $x \in X$ such that (12.9.9) holds in this case.

12.10 Commutative Banach algebras

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a commutative Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. The set of all nonzero algebra homomorphisms from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, may be denoted $\mathrm{Sp}(\mathcal{A})$, or $\mathrm{Sp}_{\mathbf{R}}(\mathcal{A})$ or $\mathrm{Sp}_{\mathbf{C}}(\mathcal{A})$, to indicate whether $\mathcal{A}$ is considered as an algebra over $\mathbf{R}$ or $\mathbf{C}$. This may be called the *Gelfand spectrum* of $\mathcal{A}$, at least in the complex case, as on p25 of [8]. This is also known as the *maximal ideal space* of $\mathcal{A}$ in the complex case, because of the correspondence with maximal proper ideals in $\mathcal{A}$, as in Section 6.12.

If $h \in \mathrm{Sp}(\mathcal{A})$, then h is a bounded linear functional on $\mathcal{A}$, with dual norm less than or equal to 1, as in Section 6.9. More precisely,

$$\|h\|_{\mathcal{A}'} = 1, \tag{12.10.1}$$

because $h(e_{\mathcal{A}}) = 1$, as before.

If $a \in \mathcal{A}$, then let $\widehat{a}$ be the real or complex-valued function, as appropriate, defined on $\mathrm{Sp}(\mathcal{A})$ by

$$\widehat{a}(h) = h(a) \tag{12.10.2}$$

for every $h \in \mathrm{Sp}(\mathcal{A})$. This is called the *Gelfand transform* of a, at least in the complex case, as on p26 of [8], and p268 of [162]. Remember that

$$\widehat{a}(\mathrm{Sp}(\mathcal{A})) \subseteq \sigma_{\mathcal{A}}(a), \tag{12.10.3}$$

as in Section 6.9. In the complex case, we have that

$$\widehat{a}(\mathrm{Sp}(\mathcal{A})) = \sigma_{\mathcal{A}}(a), \tag{12.10.4}$$

as in Section 6.12. This corresponds to Theorem 1.9.5 on p26 of [8], and to the first part of part (c) of Theorem 11.9 on p268 of [162].

As mentioned in Section 6.9,

$$\mathrm{Sp}(\mathcal{A}) \text{ is compact} \tag{12.10.5}$$

as a subset of the dual space $\mathcal{A}'$ of bounded linear functionals on $\mathcal{A}$ with respect to the weak* topology. More precisely, $\mathrm{Sp}(\mathcal{A})$ is contained in the closed unit ball $\overline{B}_{\mathcal{A}'}(0,1)$ in $\mathcal{A}'$, and $\overline{B}_{\mathcal{A}'}(0,1)$ is compact with respect to the weak* topology on $\mathcal{A}'$, by the Banach–Alaoglu theorem. To get (12.10.5), it suffices to check that $\mathrm{Sp}(\mathcal{A})$ is a closed set in $\mathcal{A}'$ with respect to the weak* topology, as before.

Let us consider $\mathrm{Sp}(\mathcal{A})$ as a topological space, using the topology induced by the weak* topology on the dual space $\mathcal{A}'$ of bounded linear functionals on $\mathcal{A}$. Equivalently, this is the weakest topology on $\mathrm{Sp}(\mathcal{A})$ such that $\widehat{a}$ is continuous for every $a \in \mathcal{A}$. It is easy to see that $\mathrm{Sp}(\mathcal{A})$ is Hausdorff with respect to this topology, and it is also compact, by (12.10.5). This corresponds to Proposition 1.9.3 on p25 of [8], and to part (a) of Theorem 11.9 on p268 of [162].

Note that

$$\mathrm{Sp}(\mathcal{A}) \neq \emptyset \tag{12.10.6}$$

in the complex case, because of (12.10.4). This corresponds to Exercise (1) on p27 of [8]. If $\mathcal{A}$ is a Banach algebra over the real numbers, then let us suppose that (12.10.6) holds for the rest of the section. In both cases,

$$a \mapsto \widehat{a} \tag{12.10.7}$$

defines an algebra homomorphism from $\mathcal{A}$ into $C(\mathrm{Sp}(\mathcal{A}), \mathbf{R})$ or $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as appropriate. This may be called the *Gelfand map*, at least in the complex case, as on p26 of [8], although the term "Gelfand transform" may also be used for this mapping, as on p268 of [162].

The kernel of the Gelfand map (12.10.7) is equal to

$$\bigcap_{h \in \mathrm{Sp}(\mathcal{A})} \ker h. \tag{12.10.8}$$

In the complex case, this is the same as the intersection of all of the maximal proper ideals in $\mathcal{A}$, which is the Jacobson radical $\mathrm{rad}(\mathcal{A})$ of $\mathcal{A}$, as in Section 6.12. This corresponds to part of part (b) of Theorem 11.9 on p268 of [162]. Thus, in the complex case, the Gelfand map is one-to-one if and only if $\mathcal{A}$ is semisimple in the sense that $\mathrm{rad}(\mathcal{A}) = \{0\}$.

Observe that

$$\|\widehat{a}\|_{sup} = \|\widehat{a}\|_{sup, \mathrm{Sp}(\mathcal{A})} \le \|a\|_{\mathcal{A}} \tag{12.10.9}$$

for every $a \in \mathcal{A}$, where the left side is the supremum norm of $\widehat{a}$ on $\mathrm{Sp}(\mathcal{A})$. This means that the Gelfand map is bounded as a linear mapping from $\mathcal{A}$ into $C(\mathrm{Sp}(\mathcal{A}), \mathbf{R})$ or $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as appropriate, with respect to the supremum norm. More precisely, the corresponding operator norm of the Gelfand map is equal to 1, because it sends $e_{\mathcal{A}}$ to the constant function equal to 1 on $\mathrm{Sp}(\mathcal{A})$. This corresponds to part of Remark 1.9.4 on p26 of [8].

In fact, we have that

$$\|\widehat{a}\|_{sup} \le r_{\mathcal{A}}(a) \tag{12.10.10}$$

for every $a \in \mathcal{A}$, where the right side is as in Section 6.13. This follows from (12.10.3), because

$$|\lambda| \le r_{\mathcal{A}}(a) \tag{12.10.11}$$

for every $\lambda \in \sigma_{\mathcal{A}}(a)$, as before. In the complex case, we have that

$$\|\widehat{a}\|_{sup} = r_{\mathcal{A}}(a) \tag{12.10.12}$$

for every $a \in \mathcal{A}$, as in Section 6.14. This corresponds to part of part (c) of Theorem 11.9 on p268 of [162].

12.11 More on the Gelfand map

Let us continue with the same notation and hypotheses as in the previous section. Note that the image

$$\widehat{\mathcal{A}} = \{\widehat{a} : a \in \mathcal{A}\} \tag{12.11.1}$$

of $\mathcal{A}$ under the Gelfand map is a subalgebra of $C(\mathrm{Sp}(\mathcal{A}), \mathbf{R})$ or $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as appropriate, as in Remark 1.9.4 on p26 of [8], and part (b) of Theorem 11.9 on p268 of [162]. We also have that $\widehat{\mathcal{A}}$ separates points in $\mathrm{Sp}(\mathcal{A})$, by construction, and that $\widehat{\mathcal{A}}$ contains the constant functions on $\mathrm{Sp}(\mathcal{A})$.

Consider the condition that there be a positive real number c_0 such that

$$c_0 \, \|a\|_{\mathcal{A}} \le \|\widehat{a}\|_{sup} \tag{12.11.2}$$

for every $a \in \mathcal{A}$. Of course, this condition implies that the Gelfand map is injective. This condition also implies that $\widehat{A}$ is a closed set in $C(\mathrm{Sp}(\mathcal{A}), \mathbf{R})$ or $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as appropriate, with respect to the supremum metric, because $\mathcal{A}$ is complete, as in Section 7.11. Conversely, if $\widehat{A}$ is a closed set in $C(\mathrm{Sp}(\mathcal{A}), \mathbf{R})$ of $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as appropriate, then $\widehat{\mathcal{A}}$ is complete with respect to the supremum metric, as in Section 1.6. If the Gelfand map is injective as well, then the open mapping theorem implies that (12.11.2) holds for some $c_0 > 0$.

If $a \in \mathcal{A}$, then

$$\|\widehat{(a^2)}\|_{sup} = \|(\widehat{a})^2\|_{sup} = \|\widehat{a}\|_{sup}^2. \tag{12.11.3}$$

If (12.11.2) holds for some $c_0 > 0$, then we get that

$$c_0^2 \, \|a\|_{\mathcal{A}}^2 \le \|\widehat{a}\|_{sup}^2 = \|\widehat{(a^2)}\|_{sup} \le \|a^2\|_{\mathcal{A}}, \tag{12.11.4}$$

using (12.10.9) in the third step. This corresponds to part of Lemma 11.11 on p270 of [162]. This is also related to part of part (b) of Theorem 11.12 on p270 of [162].

If (12.11.2) holds with $c_0 = 1$, then

$$\|a\|_{\mathcal{A}} = \|\widehat{a}\|_{sup} \tag{12.11.5}$$

for every $a \in \mathcal{A}$, because of (12.10.9). This implies that

$$\|a^2\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}^2 \tag{12.11.6}$$

for every $a \in \mathcal{A}$. This corresponds to part of Exercise (3) on p27 of [8], and to part of part (a) of Theorem 11.12 on p270 of [162].

Suppose for the moment that

$$\|a\|_{\mathcal{A}}^2 \le C \, \|a^2\|_{\mathcal{A}} \tag{12.11.7}$$

for some positive real number C and every $a \in \mathcal{A}$. This implies that

$$\|a\|_{\mathcal{A}} \le C \, r_{\mathcal{A}}(a) \tag{12.11.8}$$

for every $a \in \mathcal{A}$, as in Section 11.7. This means that

$$\|a\|_{\mathcal{A}} \le C \, \|\widehat{a}\|_{sup} \tag{12.11.9}$$

in the complex case, by (12.10.12). Of course, this is the same as saying that (12.11.2) holds with $c_0 = C^{-1}$. This corresponds to parts of Lemma 11.11 and part (b) of Theorem 11.12 on p270 of [162].

In particular, (12.11.6) implies (12.11.5) in the complex case. This corresponds to parts of Exercise (3) on p27 of [8] and part (a) of Theorem 11.12 on p270 of [162].

Suppose from now on in this section that $\mathcal{A}$ is a commutative C^* algebra. This implies that

$$r_{\mathcal{A}}(a) = \|a\|_{\mathcal{A}} \tag{12.11.10}$$

for every $a \in \mathcal{A}$, as in Section 7.7. It follows that (12.11.5) holds for every $a \in \mathcal{A}$, because of (12.10.12). This corresponds to parts of Theorem 2.2.4 on p47 of [8], and Theorem 11.18 on p276 of [162].

If $a \in \mathcal{A}$ is self-adjoint, then $\sigma_{\mathcal{A}}(a) \subseteq \mathbf{R}$, as in Section 12.8. This means that $\widehat{a}$ is a real-valued function on $\mathrm{Sp}(\mathcal{A})$, because of (12.10.4). One can use this to get that

$$\widehat{(a^*)} = \overline{(\widehat{a})} \tag{12.11.11}$$

for every $a \in \mathcal{A}$, by expressing a as $a_1 + i a_2$, where $a_1, a_2 \in \mathcal{A}$ are self-adjoint. We also obtain that

$$\widehat{\mathcal{A}} = C(\mathrm{Sp}(\mathcal{A}), \mathbf{C}) \tag{12.11.12}$$

under these conditions, by the Stone–Weierstrass theorem. These statements correspond to additional parts of Theorem 2.2.4 on p47 of [8], and Theorem 11.18 on p276 of [162].

12.12 Bounded continuous functions

Let X be a nonempty metric or topological space, and let us take $\mathcal{A} = C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, equipped with the supremum norm. Thus $\mathcal{A}$ is a commutative Banach algebra over the real or complex numbers with a multiplicative identity element, so that $\mathrm{Sp}(\mathcal{A})$ may be defined as in Section 12.10. If $x \in X$, then $h_x(f) = f(x)$ defines an element of $\mathrm{Sp}(\mathcal{A})$, as in Section 12.9. This means that

$$x \mapsto h_x \tag{12.12.1}$$

defines a mapping from X into $\mathrm{Sp}(\mathcal{A})$. This mapping is injective exactly when X is a Urysohn space, as before.

If $f \in \mathcal{A}$, then $\widehat{f}(h) = h(f)$ defines a continuous real or complex-valued function on $\mathrm{Sp}(\mathcal{A})$, as appropriate, as in Section 12.10. Note that

$$\widehat{f}(h_x) = h_x(f) = f(x) \tag{12.12.2}$$

is continuous as a real or complex-valued function of $x \in X$. One can use this to check that (12.12.1) is continuous as a mapping from X into $\mathrm{Sp}(\mathcal{A})$, with respect to the topology defined on $\mathrm{Sp}(\mathcal{A})$ previously.

If X is completely regular in the strict sense, then one can check that the topology on X is the weakest topology with respect to which the elements of $\mathcal{A}$ are all continuous. This means that the topology on X is the weakest topology with respect to which (12.12.1) is continuous as a mapping into $\mathrm{Sp}(\mathcal{A})$. If X is completely regular in the strong sense, then (12.12.1) is a homeomorphism from X onto its image in $\mathrm{Sp}(\mathcal{A})$, with respect to the topology induced on the image by the topology defined on $\mathrm{Sp}(\mathcal{A})$ in Section 12.10.

If $f \in \mathcal{A}$, then

$$\sigma_{\mathcal{A}}(f) = \overline{f(X)}, \tag{12.12.3}$$

where the right side is the closure of $f(X)$ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. This is analogous to the corresponding statement for arbitrary bounded functions on X, mentioned in Section 6.8. If $h \in \mathrm{Sp}(\mathcal{A})$, then it follows that

$$h(f) \in \overline{f(X)}, \tag{12.12.4}$$

because $h(f) \in \sigma_{\mathcal{A}}(f)$, as in Section 6.9.

Suppose for the moment that $\mathcal{A} = C_b(X, \mathbf{C})$. If $f \in C_b(X, \mathbf{R})$, then

$$\sigma_{\mathcal{A}}(f) \subseteq \mathbf{R}, \tag{12.12.5}$$

by (12.12.3). In particular, if $h \in \mathrm{Sp}(\mathcal{A})$, then

$$h(f) \in \mathbf{R}, \tag{12.12.6}$$

by (12.12.4). This implies that

$$h(\overline{a}) = \overline{h(a)} \tag{12.12.7}$$

for every $a \in C_b(X, \mathbf{C})$.

In both the real and complex cases, we would like to check that the image

$$\{h_x : x \in X\} \tag{12.12.8}$$

of X in $\mathrm{Sp}(\mathcal{A})$ under the mapping (12.12.1) is dense in $\mathrm{Sp}(\mathcal{A})$. Suppose for the sake of a contradiction that there is an element h of $\mathrm{Sp}(\mathcal{A})$ that is not in the closure of (12.12.8) in $\mathrm{Sp}(\mathcal{A})$. This means that there is an open set $U \subseteq \mathrm{Sp}(\mathcal{A})$ such that $h \in U$ and

$$h_x \not\in U \tag{12.12.9}$$

for every $x \in X$.

Because of the way that the topology on $\mathrm{Sp}(\mathcal{A})$ is defined, there are finitely many elements $f_1, \ldots, f_n$ of $\mathcal{A}$ and positive real numbers $r_1, \ldots, r_n$ such that

$$\{\phi \in \mathrm{Sp}(\mathcal{A}) : |\phi(f_j) - h(f_j)| < r_j \text{ for each } j = 1, \ldots, n\} \subseteq U. \tag{12.12.10}$$

We can reduce to the case where

$$h(f_j) = 0 \tag{12.12.11}$$

for each $j = 1, \ldots, n$, by subtracting the constant function on X equal to $h(f_j)$ from f_j if necessary. Using this, (12.12.10) reduces to

$$\{\phi \in \mathrm{Sp}(\mathcal{A}) : |\phi(f_j)| < r_j \text{ for each } j = 1, \ldots, n\} \subseteq U. \tag{12.12.12}$$

If $x \in X$, then (12.12.9) implies that

$$|f_j(x)| = |h_x(f_j)| \geq r_j \tag{12.12.13}$$

for some j.

Put

$$g = \sum_{j=1}^{n} f_j^2 \tag{12.12.14}$$

in the real case, and

$$g = \sum_{j=1}^{n} |f_j|^2 = \sum_{j=1}^{n} f_j \, \overline{f_j} \tag{12.12.15}$$

in the complex case. In both cases, we have that

$$h(g) = 0, \tag{12.12.16}$$

because of (12.12.11). We also have that

$$g(x) = \sum_{j=1}^{n} |f_j(x)|^2 \geq \min(r_1^2, \ldots, r_n^2) \tag{12.12.17}$$

for every $x \in X$, because of (12.12.13). This implies that g is invertible in $\mathcal{A}$, contradicting (12.12.16).

12.13 Some remarks about weak topologies

Let X, I be nonempty sets, and suppose that f_j is a mapping from X into a metric or topological space Y_j for each $j \in I$. Under these conditions, it is well known there is a weakest topology on X such that

$$f_j \text{ is continuous for each } j \in I. \tag{12.13.1}$$

This is known as the *weak topology* on X associated to the family of f_j's, $j \in I$.

Let

$$Y = \prod_{j \in I} Y_j \tag{12.13.2}$$

be the Cartesian product of the Y_j's, $j \in I$, equipped with the corresponding product topology. Using the f_j's, we get a mapping F from X into Y, whose jth coordinate is equal to f_j for each $j \in I$. The weak topology on X associated to the f_j's as in the preceding paragraph is the same as the weakest topology on X such that

$$F \text{ is continuous.} \tag{12.13.3}$$

Suppose for the moment that the family of f_j's, $j \in I$, separates points in X, so that

$$F \text{ is one-to-one.} \tag{12.13.4}$$

In this case,

$$F \text{ is a homeomorphism from } X \text{ onto its image in } Y, \tag{12.13.5}$$

with respect to the corresponding weak topology on X, and the topology induced on $F(X)$ by the product topology on Y.

Suppose for the moment again that

$$I \text{ has only finitely or countably many elements,} \tag{12.13.6}$$

and that

$$\text{the topology on } Y_j \text{ is determined by a metric for each } j \in I. \tag{12.13.7}$$

It is well known that

$$\text{the product topology on } Y \text{ is determined by a metric} \tag{12.13.8}$$

as well under these conditions. If the family of f_j's, $j \in I$, separates points in X, then it follows that

$$\text{the corresponding weak topology on } X \text{ is determined by a metric} \tag{12.13.9}$$

too.

Suppose for the moment that for each $j \in I$, there is a base $\mathcal{B}_j$ for the topology of Y_j with only finitely or countably many elements. If (12.13.6) holds, then it is well known that there is a base $\mathcal{B}_Y$ for the product topology on Y with only finitely or countably many elements. Similarly, there is a base $\mathcal{B}_X$ for the weak topology on X associated the family of mappings f_j, $j \in I$ with only finitely or countably many elements under these conditions.

Let X_0 be a subset of X. One can check that the topology induced on X_0 by the weak topology on X associated to the family of mappings f_j, $j \in I$, is the same as the weak topology on X_0 associated to the restrictions of the f_j's to X_0, $j \in I$.

Sometimes we may be concerned with the weak topology on X determined by a vector space Z of real or complex-valued functions on X. The same topology is determined by any subset of Z whose linear span is Z. Similarly, if Z is an algebra with respect to pointwise multiplication of functions, then the same topology on X is determined by any subset of Z that generates Z as an algebra over the real or complex numbers, as appropriate.

Let V be a vector space over the real or complex numbers, with a norm $\|\cdot\|_V$. Remember that the dual space V' of bounded linear functionals on V with respect to $\|\cdot\|_V$ is a vector space over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with a corresponding dual norm $\|\cdot\|_{V'}$, as in Section 3.1. The weak* topology on V' is the weak topology corresponding to the linear functions

$$\lambda \mapsto \lambda(v), \tag{12.13.10}$$

$v \in V$, as in Section 4.5. One can get the same topology on V' using the linear functions (12.13.10) associated to a subset of V whose linear span is equal to V, as in the preceding paragraph.

Let E be a subset of V'. The topology induced on E by the weak* topology on V' is the same as the weak topology on E associated to the restrictions of the functions (12.13.10) to E, as before. If E is bounded with respect to $\|\cdot\|_{V'}$, then one can check that this is the same as the weak topology associated to the restrictions of the functions (12.13.10) to E, with v in a dense subset of V. More precisely, it suffices to use a set of v's whose linear span is dense in V.

Suppose that V is separable with respect to the metric associated to $\|\cdot\|_V$. If E is bounded with respect to $\|\cdot\|_{V'}$, then it follows that the topology induced on E by the weak* topology on V' is the same as the weak topology associated to the restrictions of the functions (12.13.10) to E corresponding to finitely or countably many $v \in V$. This implies that there is a metric on E that determines the same topology, as before.

12.14 Some remarks about $\mathrm{Sp}(\mathcal{A})$

Let $\mathcal{A}$ be a commutative associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. As in Section 12.10, we let $\mathrm{Sp}(\mathcal{A})$ be the set of all nonzero algebra homomorphisms α from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, and we may use $\mathrm{Sp}_{\mathbf{R}}(\mathcal{A})$ or $\mathrm{Sp}_{\mathbf{C}}(\mathcal{A})$ to indicate whether $\mathcal{A}$ is considered as an elgebra over $\mathbf{R}$ or $\mathbf{C}$. Remember that

$$\alpha(e_{\mathcal{A}}) = 1, \tag{12.14.1}$$

because $\alpha \neq 0$, as in Section 6.9.

If $a \in \mathcal{A}$, then we let $\widehat{a}$ be the real or complex-valued function defined on $\mathrm{Sp}(\mathcal{A})$ by

$$\widehat{a}(\alpha) = \alpha(a), \tag{12.14.2}$$

as in Section 12.10. Remember that $\widehat{a}$ maps $\mathrm{Sp}(\mathcal{A})$ into $\sigma_{\mathcal{A}}(a)$, as in Section 6.9.

We may consider $\mathrm{Sp}(\mathcal{A})$ as a topological space, using the weakest topology with respect to which $\widehat{a}$ is continuous for each $a \in \mathcal{A}$, as before. Note that the collection $\widehat{\mathcal{A}}$ of the functions $\widehat{a}$, $a \in \mathcal{A}$, on $\mathrm{Sp}(\mathcal{A})$ separates points in $\mathrm{Sp}(\mathcal{A})$, by construction. This implies that $\mathrm{Sp}(\mathcal{A})$ is Hausdorff with respect to this weak topology. If $\mathrm{Sp}(\mathcal{A}) \neq \emptyset$, then $a \mapsto \widehat{a}$ defines an algebra homomorphism from $\mathcal{A}$ into $C(\mathrm{Sp}(\mathcal{A}), \mathbf{R})$ or $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as before.

Let $\mathcal{B}$ be another commutative associative algebra over the real or complex numbers, as appropriate, with a nonzero multiplicative identity element $e_{\mathcal{B}}$. Also let ϕ be an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ with $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. If $\beta \in \mathrm{Sp}(\mathcal{B})$, then it is easy to see that

$$\widehat{\phi}(\beta) = \beta \circ \phi \tag{12.14.3}$$

is an element of $\mathrm{Sp}(\mathcal{A})$. In particular, note that

$$(\widehat{\phi}(\beta))(e_{\mathcal{A}}) = \beta(\phi(e_{\mathcal{A}})) = \beta(e_{\mathcal{B}}) = 1. \tag{12.14.4}$$

Let $a \in \mathcal{A}$ be given, so that $\phi(a) \in \mathcal{B}$, and $(\widehat{\phi(a)})$ is a real or complex-valued function on $\mathrm{Sp}(\mathcal{B})$, as appropriate. If $\beta \in \mathrm{Sp}(\mathcal{B})$, then

$$(\widehat{\phi(a)})(\beta) = \beta(\phi(a)) = (\widehat{\phi}(\beta))(a) = \widehat{a}(\widehat{\phi}(\beta)). \tag{12.14.5}$$

This means that

$$(\widehat{\phi(a)}) = \widehat{a} \circ \widehat{\phi}. \tag{12.14.6}$$

One can use this to get that

$$\widehat{\phi} \text{ is continuous as a mapping from } \mathrm{Sp}(\mathcal{B}) \text{ into } \mathrm{Sp}(\mathcal{A}). \tag{12.14.7}$$

This uses the fact that $(\widehat{\phi(a)})$ is continuous on $\mathrm{Sp}(\mathcal{B})$, by construction.

Let $\mathcal{C}$ be a third commutative associative algebra over the real or complex numbers, as appropriate, with a nonzero multiplicative identity element $e_{\mathcal{C}}$, and let ψ be an algebra homomorphism from $\mathcal{B}$ into $\mathcal{C}$ with $\psi(e_{\mathcal{B}}) = e_{\mathcal{C}}$. Thus $\psi \circ \phi$ is an algebra homomorphism from $\mathcal{A}$ into $\mathbf{C}$ that sends $e_{\mathcal{A}}$ to $e_{\mathcal{C}}$. It is easy to see that

$$(\widehat{\psi \circ \phi}) = \widehat{\phi} \circ \widehat{\psi}, \tag{12.14.8}$$

as mappings from $\mathrm{Sp}(\mathcal{C})$ into $\mathrm{Sp}(\mathcal{A})$.

Of course, the identity mapping on $\mathcal{A}$ is an algebra homomorphism, for which the induced mapping on $\mathrm{Sp}(\mathcal{A})$ is the identity mapping. If ϕ is an algebra isomorphism from $\mathcal{A}$ onto $\mathcal{B}$, then ϕ^{-1} is an algebra homomorphism from $\mathcal{B}$ into $\mathcal{A}$, which induces a continuous mapping from $\mathrm{Sp}(\mathcal{A})$ into $\mathrm{Sp}(\mathcal{B})$, as before. Under these conditions,

$$\widehat{\phi} \text{ is a homeomorphism from } \mathrm{Sp}(\mathcal{B}) \text{ onto } \mathrm{Sp}(\mathcal{A}), \tag{12.14.9}$$

with

$$\widehat{\phi}^{-1} = (\widehat{\phi^{-1}}). \tag{12.14.10}$$

This uses the remarks in the preceding paragraph.

12.15 Some more remarks about $\mathrm{Sp}(\mathcal{A})$

Let us continue with the same notation and hypotheses as in the previous section. Note that

$$\ker \phi \subseteq \ker \beta \circ \phi = \ker \widehat{\phi}(\beta) \tag{12.15.1}$$

for every $\beta \in \mathrm{Sp}(\mathcal{B})$.

If

$$\phi(\mathcal{A}) = \mathcal{B}, \tag{12.15.2}$$

then

(12.15.3) $\quad \widehat{\phi}$ is one-to-one as a mapping from $\mathrm{Sp}(\mathcal{B})$ into $\mathrm{Sp}(\mathcal{A})$.

More precisely, if $\beta_1, \beta_2 \in \mathrm{Sp}(\mathcal{B})$ and

$$\widehat{\phi}(\beta_1) = \widehat{\phi}(\beta_2), \tag{12.15.4}$$

then $\beta_1 = \beta_2$ on $\phi(\mathcal{A})$, and thus on $\mathcal{B}$. In this case,

$$\widehat{\phi}(\mathrm{Sp}(\mathcal{B})) = \{\alpha \in \mathrm{Sp}(\mathcal{A}) : \ker \phi \subseteq \ker \alpha\}. \tag{12.15.5}$$

In fact,

(12.15.6) $\quad \widehat{\phi}$ is a homeomorphism from $\mathrm{Sp}(\mathcal{B})$ onto its image in $\mathrm{Sp}(\mathcal{A})$,

with respect to the induced topology on $\widehat{\phi}(\mathrm{Sp}(\mathcal{B}))$ when (12.15.2) holds. This is because the induced topology on $\widehat{\phi}(\mathrm{Sp}(\mathcal{B}))$ is the same as the weak topology associated to the restrictions of the functions $\widehat{a}$, $a \in A$, to $\widehat{\phi}(\mathrm{Sp}(\mathcal{B}))$, as in Section 12.13. The restrictions of these functions to $\widehat{\phi}(\mathrm{Sp}(\mathcal{B}))$ correspond exactly to the functions on $\mathrm{Sp}(\mathcal{B})$ of the form (12.14.6), with $a \in \mathcal{A}$. These are the same as the functions $\widehat{b}$, $b \in \mathcal{B}$, used to define the usual topology on $\mathrm{Sp}(\mathcal{B})$ when (12.15.2) holds.

Suppose now that $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a Banach algebra, so that every $h \in \mathrm{Sp}(\mathcal{B})$ is a bounded linear functional on $\mathcal{B}$, as in Section 6.9. Suppose also that

$$\phi(\mathcal{A}) \text{ is dense in } \mathcal{B}, \tag{12.15.7}$$

with respect to the metric associated to $\|\cdot\|_{\mathcal{B}}$. One can verify that (12.15.3) holds in this case too. Indeed, if $\beta_1, \beta_2 \in \mathrm{Sp}(\mathcal{B})$ satisfy (12.15.4), then $\beta_1 = \beta_2$ on $\phi(\mathcal{A})$, which implies that they are equal to $\mathcal{B}$, because h_1, h_2 are continuous on $\mathcal{B}$.

One can check that (12.15.7) implies (12.15.6) in this case as well. One can start in the same way as before, to get that $\widehat{\phi}$ is a homeomorphism from $\mathrm{Sp}(\mathcal{B})$ onto its image in $\mathrm{Sp}(\mathcal{A})$, when $\mathrm{Sp}(\mathcal{B})$ is equipped with the weak topology associated to the family of functions of the form $(\widehat{\phi(a)})$, $a \in \mathcal{A}$. Thus one would like to verify that this is the same as the usual topology on $\mathrm{Sp}(\mathcal{B})$ under these conditions. This corresponds to a remark in Section 12.13, because $\mathrm{Sp}(\mathcal{B})$ is bounded with respect to the dual norm on $\mathcal{B}'$, as in Section 6.9.

If $\alpha = \widehat{\phi}(\beta) = \beta \circ \phi$ for some $\beta \in \mathrm{Sp}(\mathcal{B})$, then

$$|\alpha(a)| \le \|\phi(a)\|_{\mathcal{B}} \tag{12.15.8}$$

for every $a \in \mathcal{A}$, because the dual norm of β on $\mathcal{B}$ is less than or equal to 1, as in Section 6.9. Note that this includes the condition that

$$\ker \phi \subseteq \ker \alpha, \tag{12.15.9}$$

as in (12.15.1). Of course, (12.15.8) does not use (12.15.7).

If (12.15.7) holds, then

$$\widehat{\phi}(\mathrm{Sp}(\mathcal{B})) = \{\alpha \in \mathrm{Sp}(\mathcal{A}) : (12.15.8) \text{ holds}\}. \tag{12.15.10}$$

To see this, suppose that $\alpha \in \mathrm{Sp}(\mathcal{A})$ satisfies (12.15.8), and let us define β initially on $\phi(\mathcal{A})$ by

$$\beta(\phi(a)) = \alpha(a) \tag{12.15.11}$$

for each $a \in \mathcal{A}$. Clearly β is well defined on $\phi(\mathcal{A})$, because of (12.15.9). There is a unique extension of β to a bounded linear functional on $\mathcal{B}$, because of (12.15.7) and (12.15.8), as in Section 2.2. One can check that this extension defines an element of $\mathrm{Sp}(\mathcal{B})$ under these conditions.

Chapter 13

Algebras, polynomials, and $\mathrm{Sp}(\mathcal{A})$

13.1 Some remarks about continuous functions

Let X be a nonempty metric or topological space, so that $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ are commutative associative algebras over the real and complex numbers, respectively, with a nonzero multiplicative identity element. If $x \in X$, then $h_x(f) = f(x)$ defines an algebra homomorphism from each of $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ onto $\mathbf{R}$, $\mathbf{C}$, respectively, as before. Thus

$$x \mapsto h_x \tag{13.1.1}$$

defines a mapping from X into each of $\mathrm{Sp}(C(X, \mathbf{R}))$, $\mathrm{Sp}(C(X, \mathbf{C}))$. This mapping is injective exactly when X is a Urysohn space, as in Sections 12.9 and 12.12.

If $f \in C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, then $\widehat{f}(h) = h(f)$ defines a continuous real or complex-valued function on $\mathrm{Sp}(C(X, \mathbf{R}))$ or $\mathrm{Sp}(C(X, \mathbf{C}))$, as appropriate, with respect to the topology mentioned in Section 12.14. Of course,

$$\widehat{f}(h_x) = h_x(f) = f(x) \tag{13.1.2}$$

is continuous as a real or complex-valued function of $x \in X$, as appropriate. One can use this to get that (13.1.1) is continuous as a mapping from X into each of $\mathrm{Sp}(C(X, \mathbf{R}))$, $\mathrm{Sp}(C(X, \mathbf{C}))$, as in Section 12.12.

If X is completely regular in the strict sense, then the topology on X is the weakest topology with respect to which the elements of $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$ are continuous, as in Section 12.12. This implies that the topology on X is the weakest with respect to which (13.1.1) is continuous as a mapping from X into $\mathrm{Sp}(C(X, \mathbf{R}))$ or $\mathrm{Sp}(C(X, \mathbf{C}))$, as before. If X is regular in the strong sense, then (13.1.1) is a homeomorphism from X onto its image in each of $\mathrm{Sp}(C(X, \mathbf{R}))$, $\mathrm{Sp}(C(X, \mathbf{C}))$, with respect to the induced topology on the image, as before.

If $\mathcal{A} = C(X, \mathbf{R})$ or $C(X, \mathbf{C})$ and $f \in \mathcal{A}$, then

$$\sigma_{\mathcal{A}}(f) = f(X), \tag{13.1.3}$$

as in Section 8.15. If $h \in \mathrm{Sp}(\mathcal{A})$, then we get that

$$h(f) \in f(X), \tag{13.1.4}$$

because $h(f) \in \sigma_{\mathcal{A}}(f)$, as in Section 6.9. If $\mathcal{A} = C(X, \mathbf{C})$ and $f \in C(X, \mathbf{R})$, then it follows that

$$h(f) \in \mathbf{R}. \tag{13.1.5}$$

One can use this to get that

$$h(\overline{a}) = \overline{h(a)} \tag{13.1.6}$$

for every $a \in C(X, \mathbf{C})$.

Suppose that h, h' are both elements of $\mathrm{Sp}(C(X, \mathbf{R}))$ or of $\mathrm{Sp}(C(X, \mathbf{C}))$, and that

$$h(f) = h'(f) \tag{13.1.7}$$

for every $f \in C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate. Let $f \in C(X, \mathbf{R})$ or $C(X, \mathbf{C})$ be given, as appropriate, and let us check that (13.1.7) also holds in this case. If ϵ is any positive real number, then

$$f_\epsilon = \frac{f}{1 + \epsilon\,|f|^2} \tag{13.1.8}$$

is bounded and continuous on X, so that

$$h(f_\epsilon) = h'(f_\epsilon), \tag{13.1.9}$$

by hypothesis. It is easy to see that

$$h(f_\epsilon) = \frac{h(f)}{1 + \epsilon\,|h(f)|^2}, \tag{13.1.10}$$

and similarly for h', using (13.1.6) in the complex case. One can get (13.1.7) by taking the limit as $\epsilon \to 0$ of both sides of (13.1.9).

Let ϕ be the obvious inclusion mapping from $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$ into $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, as appropriate. In both cases, ϕ is an algebra homomorphism which is compatible with the multiplicative identity elements. This leads to a continuous mapping $\widehat{\phi}$ from $\mathrm{Sp}(C(X, \mathbf{R}))$ or $\mathrm{Sp}(C(X, \mathbf{C}))$ into $\mathrm{Sp}(C_b(X, \mathbf{R}))$ or $\mathrm{Sp}(C_b(X, \mathbf{C}))$, as appropriate, as in Section 12.14. Equivalently, $\widehat{\phi}$ sends an element h of $\mathrm{Sp}(C(X, \mathbf{R}))$ or $\mathrm{Sp}(C(X, \mathbf{C}))$ to its restriction to $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate. Note that $\widehat{\phi}$ is one-to-one in both cases, as in the preceding paragraph.

Let Y be another nonempty metric or topological space, and let θ be a continuous mapping from X into Y. Observe that

$$\Theta(f) = f \circ \theta \tag{13.1.11}$$

defines an algebra homomorphism from each of $C(Y, \mathbf{R})$, $C(Y, \mathbf{C})$ into $C(X, \mathbf{R})$, $C(X, \mathbf{C})$, respectively. Of course, Θ sends constant functions on Y to constant functions on X, with the same constant value. Using Θ, we get a continuous mapping $\widehat{\Theta}$ from each of $\mathrm{Sp}(C(X, \mathbf{R}))$, $\mathrm{Sp}(C(X, \mathbf{C}))$ into $\mathrm{Sp}(C(Y, \mathbf{R}))$, $\mathrm{Sp}(C(Y, \mathbf{C}))$, respectively, as in Section 12.14. If $x \in X$, then $\widehat{\Theta}$ sends the homomorphism associated to evaluation at x to the analogous homomorphism associated to evaluation to $\theta(x) \in Y$.

If $f \in C_b(Y, \mathbf{R})$ or $C_b(Y, \mathbf{C})$, then $\Theta(f) \in C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate, with

$$\|\Theta(f)\|_{sup,X} \le \|f\|_{sup,Y}. \tag{13.1.12}$$

If $\theta(X)$ is dense in Y, then

$$\|\Theta(f)\|_{sup,X} = \|f\|_{sup,Y}. \tag{13.1.13}$$

Let Θ_b be the restriction of Θ to $C_b(Y, \mathbf{R})$ or $C_b(Y, \mathbf{C})$, considered as an algebra homomorphism into $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate. This leads to a continuous mapping $\widehat{\Theta_b}$ from $\mathrm{Sp}(C_b(X, \mathbf{R}))$ or $\mathrm{Sp}(C_b(X, \mathbf{C}))$ into $\mathrm{Sp}(C_b(Y, \mathbf{R}))$ or $\mathrm{Sp}(C_b(Y, \mathbf{C}))$, as appropriate.

Let ψ be the obvious inclusion mapping from $C_b(Y, \mathbf{R})$ or $C_b(Y, \mathbf{C})$ into $C(Y, \mathbf{R})$ or $C(Y, \mathbf{C})$, as appropriate, and let ϕ be the analogous inclusion mapping for X, as before. Thus

$$\Theta \circ \phi = \psi \circ \Theta_b, \tag{13.1.14}$$

by construction. This implies that

$$\widehat{\phi} \circ \widehat{\Theta} = \widehat{\Theta_b} \circ \widehat{\psi}, \tag{13.1.15}$$

as in Section 12.14. Here $\widehat{\phi}$ is as before, and $\widehat{\psi}$ is the analogue for ψ.

13.2 Some conditions related to compactness

Let X be a topological space. A point $p \in X$ is said to be a *limit point* of a set $E \subseteq X$ if for every open set $U \subseteq X$ with $p \in U$, there is a $q \in E \cap U$ such that $q \neq p$. Let us say that p is a *strong limit point* of E in X if for every open set $U \subseteq X$ with $p \in U$, we have that $E \cap U$ has infinitely many elements. Thus strong limit points of E in X are automatically limit points of E in X. If X satisfies the first separation condition, then one can check that limit points of E in X are strong limit points of E in X.

We say that E has the *limit point property* if every infinite subset of E has a limit point in E. Let us say that E has the *strong limit point property* if every infinite subset of E has a strong limit point in E. This implies that E has the limit point property, and the converse holds when X satisfies the first separation condition. If E is compact, then one can show that E has the strong limit point property, using a standard argument. If E has the limit point property, and the topology on X is determined by a metric, then it is well known that E is compact.

We say that E is *countably compact* if every covering of E by countably many open subsets of X can be reduced to a finite subcovering. Similarly, we say that E has the *Lindelöf propery* if every open covering of E in X can be reduced to a subcovering with only finitely or countably many elements. Thus E is compact if and only if E is countably compact and E has the Lindelöf property. It is well known that E is countably compact if and only if E has the strong limit point property.

A collection $\mathcal{B}$ of open subsets of X is said to be a *base for the topology of* X if every open set in X can be expressed as a union of elements of $\mathcal{B}$. If there is a base $\mathcal{B}$ for the topology of X with only finitely or countably many elements, then *Lindelöf's theorem* implies that every subset of X has the Lindelöf property.

We say that X is *separable* if there is a dense set in X with only finitely or countably many elements. If there is a base for the topology of X with only finitely or countably many elements, then one can check that X is separable. It is well known that the converse holds when the topology on X is determined by a metric.

Let f be a continuous mapping from X into another topological space Y. If $K \subseteq X$ is compact, then it is well known that $f(K)$ is compact in Y. One can verify that the analogous statements for countable compactness and the Lindelöf property hold as well, using essentially the same argument.

We say that X is *pseudocompact* if every continuous real-valued function on X is bounded on X. Of course, if X is compact, then X is pseudocompact. More precisely, one can check that X is pseudocompact when X is countably compact. The converse holds when X is normal in the strong sense, as on p20 of [178].

13.3 Polynomials in n variables

Let n be a positive integer. A *multi-index* is an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers. In this case, it is convenient to put

$$|\alpha| = \sum_{j=1}^{n} \alpha_j. \tag{13.3.1}$$

Note that the set of all multi-indices is the same as the set $(\mathbf{Z}_+ \cup \{0\})^n$ of all n-tuples of elements of $\mathbf{Z}_+ \cup \{0\}$.

Let $T_1, \ldots, T_n$ be commuting indeterminates. We shall normally try to use upper-case letters for indeterminates, and lower-case letters for elements of $\mathbf{R}$, $\mathbf{C}$, or other associative algebras, as in [44, 80]. A *formal polynomial* in $T_1, \ldots, T_n$ with real or complex coefficients may be expressed as

$$p(T) = p(T_1, \ldots, T_n) = \sum_{|\alpha| \le N} c_\alpha \, T^\alpha, \tag{13.3.2}$$

where N is a nonnegative integer, and the sum is taken over all multi-indices α with $|\alpha| \le N$. Of course, c_α should be an element of $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for each such α. If α is any multi-index, then

$$T^\alpha = T_1^{\alpha_1} \cdots T_n^{\alpha_n} \tag{13.3.3}$$

is the corresponding formal monomial in $T_1, \ldots, T_n$.

The spaces of all formal polynomials in $T_1, \ldots, T_n$ with real or complex coefficients may be denoted

$$\mathbf{R}[T_1, \ldots, T_n],\ \mathbf{C}[T_1, \ldots, T_n], \tag{13.3.4}$$

respectively. The coefficients c_α of a formal polynomial as in (13.3.2) should be considered as being defined for all multi-indices α, with $c_\alpha = 0$ when $|\alpha|$ is strictly larger then N. This means that

$$\alpha \mapsto c_\alpha \tag{13.3.5}$$

defines a real or complex-valued function, as appropriate, on $(\mathbf{Z}_+ \cup \{0\})^n$, with finite support. Thus the spaces (13.3.4) of formal polynomials may be defined more precisely as the spaces

$$c_{00}((\mathbf{Z}_+ \cup \{0\})^n, \mathbf{R}),\ c_{00}((\mathbf{Z}_+ \cup \{0\})^n, \mathbf{C}) \tag{13.3.6}$$

of all real or complex-valued functions on $(\mathbf{Z}_+ \cup \{0\})^n$, as appropriate, with finite support. These are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions on $(\mathbf{Z}_+ \cup \{0\})^n$, which corresponds to termwise addition and scalar multiplication of formal polynomials as in (13.3.2).

If α, β are multi-indices, then $\alpha + \beta$ is the multi-index defined by coordinatewise addition, as usual. Similarly, multiplication of the corresponding formal monomials is defined by

$$T^\alpha \, T^\beta = T^{\alpha+\beta}. \tag{13.3.7}$$

This can be extended to a bilinear operation of multiplication of formal polynomials, which is commutative and associative. This means that the spaces (13.3.4) of formal polynomials are commutative associative algebras over the real and complex numbers, as appropriate, with respect to this definition of multiplication. The "constant" polynomial for which the coefficient of T^α is equal to 1 when $\alpha = 0$ and to 0 otherwise is the multiplicative identity element in each of these algebras.

Of course, formal polynomials like these determine polynomial functions on $\mathbf{R^n}$ or $\mathbf{C}^n$, as appropriate, in the usual way, and an extension of this will be discussed in the next section. Note that the coefficients of such a polynomial are determined by the derivatives of the corresponding polynomial function at 0.

13.4 Polynomials and homomorphisms

Let us continue with the same notation and hypotheses as in the previous section. Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. Also let $a = (a_1, \ldots, a_n)$ be an n-tuple of commuting elements of $\mathcal{A}$, so that

$$a_j \, a_l = a_l \, a_j \tag{13.4.1}$$

for all $j, l = 1, \ldots, n$. If α is a multi-index, then put

$$a^\alpha = a_1^{\alpha_1} \cdots a_n^{\alpha_n}, \tag{13.4.2}$$

where $a_j^{\alpha_j}$ is interpreted as being equal to $e_{\mathcal{A}}$ when $\alpha_j = 0$, as usual. Note that

$$a^{\alpha+\beta} = a^\alpha \, a^\beta \tag{13.4.3}$$

for all multi-indices α, β.

Let $p(T)$ be a formal polynomial in $T_1, \ldots, T_n$ with real or complex coefficients, as appropriate, as in (13.3.2). Put

$$p(a) = p(a_1, \ldots, a_n) = \sum_{|\alpha| \le N} c_\alpha \, a^\alpha, \tag{13.4.4}$$

where the sum is taken over all multi-indices α with $|\alpha| \le N$, as before. This defines $p(a)$ as an element of $\mathcal{A}$. This is essentially the same as in Section 8.13 when $n = 1$. We may also use $p_{\mathcal{A}}(a) = p_{\mathcal{A}}(a_1, \ldots, a_n)$ for (13.4.4), to indicate the role of $\mathcal{A}$.

It is easy to see that

$$p(T) \mapsto p(a) \tag{13.4.5}$$

defines an algebra homomorphism from $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate, into $\mathcal{A}$. Note that (13.4.5) sends the multiplicative identity element in $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate, to $e_{\mathcal{A}}$. Similarly, (13.4.5) sends T_j to a_j for each $j = 1, \ldots, n$. We also have that (13.4.5) is uniquely determined by these properties.

Let $\mathcal{B}$ be another associative algebra over the real or complex numbers, as appropriate, and with a multiplicative identity element $e_{\mathcal{B}}$, and let ϕ be an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ with $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. Put $b_j = \phi(a_j)$ for each $j = 1, \ldots, n$, so that $b = (b_1, \ldots, b_n)$ is a commuting n-tuple of elements of $\mathcal{B}$. Observe that

$$\phi(a^\alpha) = b^\alpha \tag{13.4.6}$$

for every multi-index α. This implies that

$$\phi(p_{\mathcal{A}}(a_1, \ldots, a_n)) = p_{\mathcal{B}}(\phi(a_1), \ldots, \phi(a_n)). \tag{13.4.7}$$

Let X be a nonempty set, and suppose that $\mathcal{A}$ is a subalgebra of the algebra of all real or complex-valued functions on X, as appropriate. It is easy to see that

$$(p_{\mathcal{A}}(a_1, \ldots, a_n))(x) = p(a_1(x), \ldots, a_n(x)) \tag{13.4.8}$$

for every $x \in X$.

13.5 More on polynomials, homomorphisms

Let n be a positive integer, and let $T_1, \ldots . T_n$ be n commuting indeterminates again. If $w \in \mathbf{R}^n$ or $\mathbf{C}^n$, then put

$$h_w(p(T)) = p(w) \tag{13.5.1}$$

for every $p(T) \in \mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate. This defines an algebra homomorphism from $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in the previous section. Note that

$$h_w(T_j) = w_j \tag{13.5.2}$$

for each $j = 1, \ldots, n$, as before.

Using this, we get a mapping

$$w \mapsto h_w \tag{13.5.3}$$

from each of $\mathbf{R}^n$, $\mathbf{C}^n$ into $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$, $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$, respectively. It is easy to see that

(13.5.4) (13.5.3) is one-to-one,

because of (13.5.2). We also have that

(13.5.5) (13.5.3) maps $\mathbf{R}^n$, $\mathbf{C}^n$ onto $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$, $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$,

respectively, because any element of $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$ or $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$ is uniquely determined by its values at $T_1, \ldots, T_n$. More precisely, the inverse mapping is given by

$$h \mapsto (h(T_1), \ldots, h(T_n)). \tag{13.5.6}$$

If $p(T) \in \mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, then we get a real or complex-valued function $(\widehat{p(T)})$ on $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$ or $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$, as appropriate, defined by

$$(\widehat{p(T)})(h) = h(p(T)), \tag{13.5.7}$$

as in Section 12.14. We take $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$ and $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$ to be equipped with the weakest topologies with respect to which these functions are continuous, as before. Of course,

$$(\widehat{p(T)})(h_w) = h_w(p(T)) = p(w) \tag{13.5.8}$$

for every $w \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. Note that the standard topologies on $\mathbf{R}^n$, $\mathbf{C}^n$ are the weakest topologies with respect to which all polynomial functions are continuous. One can use this to get that

(13.5.9) (13.5.3) is a homeomorphism

from each of $\mathbf{R}^n$, $\mathbf{C}^n$ onto $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$, $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$, respectively.

13.6 Polynomials, homomorphisms, and $\mathrm{Sp}(\mathcal{A})$

Let n be a positive integer, and let $T_1, \ldots, T_n$ be n commuting indetermiinates. Also let $\mathcal{A}$ be a commutative associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$, and let $a = (a_1, \ldots, a_n)$ be an n-tuple of elements of $\mathcal{A}$. Thus

$$\psi(p(T)) = p_{\mathcal{A}}(a) \tag{13.6.1}$$

defines an algebra homomorphism from $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$ into $\mathcal{A}$, as appropriate, as in Section 13.4. This leads to an induced mapping

$$\widehat{\psi}(\alpha) = \alpha \circ \psi \tag{13.6.2}$$

from $\mathrm{Sp}(\mathcal{A})$ into $\mathrm{Sp}(\mathbf{R}[T_1, \ldots, T_n])$ or $\mathrm{Sp}(\mathbf{C}[T_1, \ldots, T_n])$, as appropriate, as in Section 12.14. Remember that

$$\widehat{\psi} \text{ is continuous} \tag{13.6.3}$$

with respect to the usual topology on $\mathrm{Sp}(\mathcal{A})$, as before.

Put

$$\phi(\alpha) = (\alpha(\psi(T_1)), \ldots \alpha(\psi(T_n))) = (\alpha(a_1), \ldots, \alpha(a_n)) \tag{13.6.4}$$

for each $\alpha \in \mathrm{Sp}(\mathcal{A})$, where the right side is an element of $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. This is the same as the composition of $\widehat{\psi}$ with (13.5.6). Equivalently,

$$\widehat{\psi}(\alpha) = h_{\phi(\alpha)} \tag{13.6.5}$$

for every $\alpha \in \mathrm{Sp}(\mathcal{A})$, This means that

$$(\widehat{\psi}(\alpha))(p(T)) = \alpha(\psi(p(T))) = \alpha(p_{\mathcal{A}}(a)) = p(\phi(\alpha)) \tag{13.6.6}$$

for every $\alpha \in \mathrm{Sp}(\mathcal{A})$ and $p(T) \in \mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate.

We may also express (13.6.4) as

$$\phi(\alpha) = (\widehat{a_1}(\alpha), \ldots, \widehat{a_n}(\alpha)), \tag{13.6.7}$$

where $\widehat{a}(\alpha) = \alpha(a)$ for each $a \in \mathcal{A}$, as in Section 12.14. This implies that

$$\phi \text{ is continuous as a mapping from } \mathrm{Sp}(\mathcal{A}) \text{ into } \mathbf{R}^n \text{ or } \mathbf{C}^n, \tag{13.6.8}$$

as appropriate, sith respect to the usual topology on $\mathrm{Sp}(\mathcal{A})$. This could be obtained from (13.5.9) and (13.6.3) as well.

Clearly

$$\ker \psi \subseteq \ker \alpha \circ \psi = \ker \widehat{\psi}(\alpha) \tag{13.6.9}$$

for every $\alpha \in \mathrm{Sp}(\mathcal{A})$. The kernel of ψ consists of the $p(T) \in \mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate, such that

$$p_{\mathcal{A}}(a) = 0. \tag{13.6.10}$$

In this case, we have that

$$p(\phi(a)) = 0, \tag{13.6.11}$$

as in (13.6.6).

Suppose now that

$$\psi \text{ maps } \mathbf{R}[T_1, \ldots, T_n] \text{ or } \mathbf{C}[T_1, \ldots, T_n] \text{ onto } \mathcal{A}, \tag{13.6.12}$$

as appropriate. This implies that

$$\widehat{\psi} \text{ is one-to-one on } \operatorname{Sp}(\mathcal{A}), \tag{13.6.13}$$

as in Section 12.15. Under these conditions, $\widehat{\psi}(\operatorname{Sp}(\mathcal{A}))$ consists exactly of the $h \in \operatorname{Sp}(\mathbf{R}[T_1, \ldots, T_n])$ or $\operatorname{Sp}(\mathbf{C}[T_1, \ldots, T_n])$, as appropriate, such that

$$\ker \psi \subseteq \ker h, \tag{13.6.14}$$

as before. This means that $\phi(\operatorname{Sp}(\mathcal{A}))$ consists exactly of the $w \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, such that

$$\ker \psi \subseteq \ker h_w. \tag{13.6.15}$$

This is the same as saying that $\phi(\operatorname{Sp}(\mathcal{A}))$ consists exactly of the $w \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, such that

$$p(w) = 0 \tag{13.6.16}$$

for every $p(T) \in \mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate, that satisfies (13.6.10).

13.7 A class of Banach algebras

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a commutative Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Also let $a_1, \ldots, a_n$ be finitely many elements of $\mathcal{A}$, and let $\mathcal{A}_0$ be the subalgebra of $\mathcal{A}$ generated by $e_{\mathcal{A}}$ and $a_1, \ldots, a_n$. Suppose that

$$\mathcal{A}_0 \text{ is dense in } \mathcal{A}, \tag{13.7.1}$$

with respect to the metric associated to the norm. This corresponds to Example 11.13 (d) on p271 of [162], at least in the complex case.

Let ψ be the algebra homomorphism from $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate, into $\mathcal{A}$ as in (13.6.1), with $a = (a_1, \ldots, a_n)$. Observe that $\mathcal{A}_0$ is the same as the image of $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$, as appropriate, under ψ.

If $\alpha \in \operatorname{Sp}(\mathcal{A})$, then put

$$\phi(\alpha) = (\widehat{a_1}(\alpha), \ldots, \widehat{a_n}(\alpha)) = (\alpha(a_1), \ldots, \alpha(a_n)), \tag{13.7.2}$$

as in (13.6.4) and (13.6.7). This defines a mapping from $\operatorname{Sp}(\mathcal{A})$ into $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. This mapping is continuous with respect to the topology defined on $\mathrm{Sp}(\mathcal{A})$ as in Section 12.10 and the standard metric on $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, because $\widehat{a}$ is continuous on $\mathrm{Sp}(\mathcal{A})$ for every $a \in \mathcal{A}$, as before. Thus

$$K = \phi(\mathrm{Sp}(\mathcal{A})) \tag{13.7.3}$$

is a compact subset of $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, because $\mathrm{Sp}(\mathcal{A})$ is compact, as in Section 12.10.

If $\alpha \in \mathrm{Sp}(\mathcal{A})$, then α is uniquely determined by its restriction to $\mathcal{A}_0$, because of (13.7.1), and the fact that α is a bounded linear functional on $\mathcal{A}$, as in Section 6.9. It is easy to see that the restriction of α to $\mathcal{A}_1$ is uniquely determined by $\alpha(a_1), \ldots, \alpha(a_n)$. This means that

$$\phi \text{ is one-to-one on } \mathrm{Sp}(\mathcal{A}). \tag{13.7.4}$$

Alternatively, $\widehat{\psi}$ is one-to-one on $\mathrm{Sp}(\mathcal{A})$ under these conditions, as in Section 12.15. One can use this to get (13.7.4) from (13.6.5).

It follows that

$$\phi \text{ is a homeomorphism from } \mathrm{Sp}(\mathcal{A}) \text{ onto } K, \tag{13.7.5}$$

with respect to the restriction to K of the standard metric on $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. This uses the compactness of $\mathrm{Sp}(\mathcal{A})$ and a well-known result in topology. This could also be obtained more directly here, because of the way that the topology on $\mathrm{Sp}(\mathcal{A})$ is defined, and because the dual norm of every $h \in \mathrm{Sp}(\mathcal{A})$ is equal to 1. This corresponds to some remarks in Section 12.15, using (13.5.9) and (13.6.5).

Remember that $\mathrm{Sp}(\mathcal{A}) \neq \emptyset$ in the complex case. If $\mathcal{A}$ is a Banach algebra over the real numbers, then let us suppose that $\mathrm{Sp}(\mathcal{A}) \neq \emptyset$ for the rest of the section. Of course, this means that

$$K \neq \emptyset. \tag{13.7.6}$$

Note that ϕ^{-1} is a homeomorphism from K onto $\mathrm{Sp}(\mathcal{A})$. If $b \in \mathcal{A}$, then $\widehat{b}$ is a continuous function on $\mathrm{Sp}(\mathcal{A})$, so that

$$\eta(b) = \widehat{b} \circ \phi^{-1} \tag{13.7.7}$$

is a continuous real or complex-valued function on K, as appropriate. This defines an algebra homomorphism from $\mathcal{A}$ into $C(K, \mathbf{R})$ or $C(K, \mathbf{C})$, as appropriate.

Remember that $\widehat{a_j}(\alpha)$ is the jth coordinate of $\phi(\alpha)$ for $\alpha \in \mathrm{Sp}(\mathcal{A})$ and $j = 1, \ldots, n$, as in (13.7.2). This implies that $\eta(a_j)$ is the same as the restriction to K of the projection onto the jth coordinate for each $j = 1, \ldots, n$. It follows that

$$\eta(\mathcal{A}_0) \tag{13.7.8}$$

is the subalgebra of $C(K, \mathbf{R})$ or $C(K, \mathbf{C})$, as appropriate, consisting of the restrictions to K of polynomials on $\mathbf{R}^n$ or $\mathbf{C}^n$ with real or complex coefficients, as appropriate. In fact,

$$\eta \circ \psi \tag{13.7.9}$$

is the homomorphism from $\mathbf{R}[T_1, \ldots, T_n]$ or $\mathbf{C}[T_1, \ldots, T_n]$ into $C(K, \mathbf{R})$ or $C(K, \mathbf{C})$, as appropriate, that sends a formal polynomial to the restriction of the corresponding polynomial function to K.

If $b \in \mathcal{A}$, then

$$\|\eta(b)\|_{sup,K} = \|\widehat{b}\|_{sup,\mathrm{Sp}(\mathcal{A})} \le \|b\|_{\mathcal{A}}, \tag{13.7.10}$$

where $\|\eta(b)\|_{sup,K}$ is the supremum norm of $\eta(b)$ on K, and the second step is as in Section 12.10. It follows that

$$\eta(\mathcal{A}) \tag{13.7.11}$$

is contained in the closure of the subalgebra of $C(K, \mathbf{R})$ or $C(K, \mathbf{C})$, as appropriate, consisting of restrictions to K of polynomials with real or complex coefficients, as appropriate, with respect to the supremum metric. In the real case, the restrictions to K of polynomials are dense in $C(K, \mathbf{R})$ with respect to the supremum metric, by the Stone–Weierstrass theorem.

13.8 Polynomial convexity

Let us continue with the same notation and hypotheses as in the previous section, except that now we suppose that $\mathcal{A}$ is a complex Banach algebra. Let $w \in \mathbf{C}^n$ be given, and suppose that

$$|p(w)| \le \sup_{z \in K} |p(z)| \tag{13.8.1}$$

for every polynomial p on $\mathbf{C}^n$ with complex coefficients. We would like to show that

$$w \in K \tag{13.8.2}$$

under these conditions. This means that K is *polynomially convex* in $\mathbf{C}^n$, as on p272 of [162].

If p is a polynomial on $\mathbf{C}^n$ with complex coefficients, then p corresponds to a formal polynomial in n commuting indeterminates with complex coefficients, as in Section 13.3. Thus we can define $p_{\mathcal{A}}(a_1, \ldots, a_n)$ as an element of $\mathcal{A}$ as before. We would like to put

$$\alpha_w(p_{\mathcal{A}}(a_1, \ldots, a_n)) = p(w). \tag{13.8.3}$$

This basically corresponds to an argument mentioned in Section 12.15.

Observe that

$$\eta(p_{\mathcal{A}}(a_1, \ldots, a_n)) \tag{13.8.4}$$

is the same as the restriction of p to K, as an element of $C(K, \mathbf{C})$. It follows that

$$\sup_{z \in K} |p(z)| \le \|p_{\mathcal{A}}(a_1, \ldots, a_n)\|_{\mathcal{A}}, \tag{13.8.5}$$

because of (13.7.10). This means that

$$|p(w)| \le \|p_{\mathcal{A}}(a_1, \ldots, a_n)\|_{\mathcal{A}}, \tag{13.8.6}$$

by (13.8.1). This corresponds to the condition that was used in Section 12.15.

Every element of $\mathcal{A}_0$ is of the form $p_{\mathcal{A}}(a_1, \ldots, a_n)$ for some polynomial p, by definition of $\mathcal{A}_0$. Using (13.8.6), we get that α_w is well-defined as a linear functional on $\mathcal{A}_0$. In fact, α_w is a bounded linear functional on $\mathcal{A}_0$, with respect to the restriction of $\| \cdot \|_{\mathcal{A}}$ to $\mathcal{A}_0$. This implies that α_w has a unique extension to a bounded linear functional on $\mathcal{A}$, because of (13.7.1), as in Section 2.2.

It is easy to see that α_w is an algebra homomorphism from $\mathcal{A}_0$ into $\mathbf{C}$. It follows that the extension of α_w to $\mathcal{A}$ mentioned in the preceding paragraph is an algebra homomorphism into $\mathbf{C}$ as well. Clearly

$$\alpha_w(e_{\mathcal{A}}) = 1, \tag{13.8.7}$$

by (13.8.3). Let us also use α_w to denote this extension to $\mathcal{A}$, which is an element of Sp($\mathcal{A}$).

Similarly,

$$\alpha_w(a_j) = w_j \tag{13.8.8}$$

for each $j = 1, \ldots, n$. This means that

$$\phi(\alpha_w) = (\alpha_w(a_1), \ldots, \alpha_w(a_n)) = w. \tag{13.8.9}$$

Thus (13.8.2) holds, by the definition (13.7.3) of K.

Suppose now that $n = 1$, so that

$$\phi(\alpha) = \widehat{a_1}(\alpha) = \alpha(a_1) \tag{13.8.10}$$

for every $\alpha \in \mathrm{Sp}(\mathcal{A})$, as in (13.7.2). We also have that

$$K = \phi(\mathrm{Sp}(\mathcal{A})) = \sigma_{\mathcal{A}}(a_1), \tag{13.8.11}$$

where the second step is as in Sections 6.9 and 6.12. We would like to show that

$$\mathbf{C} \setminus \sigma_{\mathcal{A}}(a_1) \text{ is connected,} \tag{13.8.12}$$

as mentioned on p272 of [162]. This corresponds to Exercise (2) on p33 of [8] and Exercise 7 on p401 of [160] as well.

Remember that $\sigma_{\mathcal{A}}(a_1)$ is a closed and bounded set in $\mathbf{C}$, as in Section 6.8. This implies that $\mathbf{C} \setminus \sigma_{\mathcal{A}}(a_1)$ is an open set with exactly one unbounded connected component. If U is a bounded component of $\mathbf{C} \setminus \sigma_{\mathcal{A}}(a_1)$, then the boundary of U is contained in $\sigma_{\mathcal{A}}(a_1)$. This means every $w \in U$ satisfies (13.8.1), by the maximum principle. Thus (13.8.2) implies that U is contained in $\sigma_{\mathcal{A}}(a_1)$, which is a contradiction.

13.9 A class of C^* algebras

Let $(\mathcal{A}, \| \cdot \|_{\mathcal{A}})$ be a C^* algebra with nonzero multiplicative identity element $e_{\mathcal{A}}$ and involution $b \mapsto b^*$. Suppose that a_1 is a normal element of $\mathcal{A}$, so that a_1 commutes with a_1^*, and let $\mathcal{A}_1$ be the subalgebra of $\mathcal{A}$ generated by $e_{\mathcal{A}}$, a_1, and

a_1^*. Note that $\mathcal{A}_1$ is a commutative algebra, and that if $b \in \mathcal{A}_1$, then $b^* \in \mathcal{A}_1$ too. Suppose also that

$$\mathcal{A}_1 \text{ is dense in } \mathcal{A}, \tag{13.9.1}$$

with respect to the metric associated to the norm. Of course, this implies that $\mathcal{A}$ is commutative as well.

Let $\mathrm{Sp}(\mathcal{A})$ be the space of nonzero algebra homomorphisms from $\mathcal{A}$ into $\mathbf{C}$, with its usual topology, as in Section 12.10. If $b \in \mathcal{A}$, then $\widehat{b}(\alpha) = \alpha(b)$ defines a continuous mapping from $\mathrm{Sp}(\mathcal{A})$ onto $\sigma_{\mathcal{A}}(a)$, as before. Remember that $\alpha(b) \in \mathbf{R}$ for every $\alpha \in \mathrm{Sp}(\mathcal{A})$ when $b \in \mathcal{A}$ is self-adjoint, as in Section 12.8. One can use this to get that

$$\alpha(b^*) = \overline{\alpha(b)} \tag{13.9.2}$$

for every $b \in \mathcal{A}$, which was mentioned in slightly different notation in Section 12.11. In particular,

$$\alpha(a_1^*) = \overline{\alpha(a_1)}. \tag{13.9.3}$$

If $\alpha \in \mathrm{Sp}(\mathcal{A})$, then the restriction of α to $\mathcal{A}_1$ is uniquely determined by $\alpha(a_1)$, because of (13.9.3). This implies that α is uniquely determined on $\mathcal{A}$ by $\alpha(a_1)$, because of (13.9.1), and because α is a bounded linear functional on $\mathcal{A}$, as in Section 6.9. It follows that

$$\widehat{a_1} \text{ is one-to-one on } \mathrm{Sp}(\mathcal{A}). \tag{13.9.4}$$

In fact,

$$\widehat{a_1} \text{ is a homeomorphism from } \mathrm{Sp}(\mathcal{A}) \text{ onto } \sigma_{\mathcal{A}}(a_1), \tag{13.9.5}$$

with respect to the restriction of the standard metric on $\mathbf{C}$ to $\sigma_{\mathcal{A}}(a_1)$. This can be obtained from the compactness of $\mathrm{Sp}(\mathcal{A})$ and a well-known result in topology, or using more direct arguments in this case.

Remember that the Gelfand map $b \mapsto \widehat{b}$ is an isometric algebra isomorphism from $\mathcal{A}$ onto $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as in Section 12.11. We also have that the involution on $\mathcal{A}$ corresponds to complex-conjugation on $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, as before. It follows that

$$b \mapsto \widehat{b} \circ \widehat{a_1}^{-1} \tag{13.9.6}$$

is an isometric algebra isomorphism from $\mathcal{A}$ onto $C(\sigma_{\mathcal{A}}(a_1), \mathbf{C})$, with respect to which the involution on $\mathcal{A}$ corresponds to complex-conjugation on $C(\sigma_{\mathcal{A}}(a_1), \mathbf{C})$.

If $f \in C(\sigma_{\mathcal{A}}(a_1), \mathbf{C})$, then let $f(a_1) = f_{\mathcal{A}}(a_1)$ be the unique element of $\mathcal{A}$ such that

$$(\widehat{f_{\mathcal{A}}(a_1)}) \circ \widehat{a_1}^{-1} = f. \tag{13.9.7}$$

Of course, this is the same as saying that

$$(\widehat{f_{\mathcal{A}}(a_1)}) = f \circ \widehat{a_1} \tag{13.9.8}$$

on $\mathrm{Sp}(\mathcal{A})$. Thus

$$f \mapsto f_{\mathcal{A}}(a_1) \tag{13.9.9}$$

defines an isometric algebra isomorphism from $C(\sigma_{\mathcal{A}}(a_1), \mathbf{C})$ onto $\mathcal{A}$, which is the inverse of (13.9.6). Note that

$$(\overline{f})_{\mathcal{A}}(a_1) = f_{\mathcal{A}}(a_1)^*, \tag{13.9.10}$$

as before.

If $f \equiv 1$ on $\sigma_{\mathcal{A}}(a_1)$, then $f_{\mathcal{A}}(a_1) = e_{\mathcal{A}}$. If $f(z) = z$ on $\sigma_{\mathcal{A}}(a_1)$, then $f_{\mathcal{A}}(a_1) = a_1$. If $f(z) = \overline{z}$ on $\sigma_{\mathcal{A}}(a_1)$, then it follows that $f_{\mathcal{A}}(a_1) = a_1^*$. If f is given by a polynomial in z and $\overline{z}$ on $\sigma_{\mathcal{A}}(a_1)$, with complex coefficients, then we get that $f_{\mathcal{A}}(a_1)$ is given by the same polynomial in a_1 and a_1^*. It is well known that every continuous complex-valued function on $\sigma_{\mathcal{A}}(a_1)$ can be uniformly approxiimated by polynomials in z and $\overline{z}$ with complex coefficients, by the Stone–Weierstrass theorem.

These remarks basically correspond to Theorem 11.19 on p277 of [162]. This basically corresponds to Theorem 2.3.1 on p51 of [8], which is stated for bounded normal operators on complex Hilbert spaces, and uses an additional fact that will be discussed in the next section. This is also related to Corollary 1 on p263 of [167].

13.10 Spectral permanence and C^* algebras

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra with a nonzero multiplicative identity element $e_{\mathcal{A}}$ and involution $a \mapsto a^*$. Also let $\mathcal{B}$ be a subalgebra of $\mathcal{A}$ that is a closed set with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, that contains $e_{\mathcal{A}}$, and for which $b^* \in \mathcal{B}$ for every $b \in \mathcal{B}$. If $b \in \mathcal{B}$, then we would like to show that

$$\sigma_{\mathcal{A}}(b) = \sigma_{\mathcal{B}}(b). \tag{13.10.1}$$

This corresponds to Corollary 2 on p49 of [8], and Theorem 11.29 on p283 of [162]. A version of this is also mentioned in the proof of Corollary 1 on p263 of [167].

Of course, any invertible element of $\mathcal{B}$ is invertible as an element of $\mathcal{A}$. It suffices to show that if $x \in \mathcal{B}$ and x has an inverse x^{-1} in $\mathcal{A}$, then

$$x^{-1} \in \mathcal{B}. \tag{13.10.2}$$

Note that $x^*\, x$ is a self-adjoint element of $\mathcal{B}$. This implies that

$$\sigma_{\mathcal{B}}(x^*\, x) \subseteq \mathbf{R}, \tag{13.10.3}$$

because $\mathcal{B}$ is a C^* algebra, as in Section 12.8. This means that

$$\partial\sigma_{\mathcal{B}}(x^*\, x) = \sigma_{\mathcal{B}}(x), \tag{13.10.4}$$

where the left side is the boundary of $\sigma_{\mathcal{B}}(x^*\, x)$ in the complex plane. Remember that

$$\partial\sigma_{\mathcal{B}}(x^*\, x) \subseteq \sigma_{\mathcal{A}}(x^*\, x), \tag{13.10.5}$$

as in Section 7.3. It follows that

$$\sigma_{\mathcal{B}}(x^*\,x) \subseteq \sigma_{\mathcal{A}}(x^*\,x). \tag{13.10.6}$$

If x is invertible in $\mathcal{A}$, then x^* is invertible in $\mathcal{A}$ too. This implies that $x^*\,x$ is invertible in $\mathcal{A}$, so that $0 \not\in \sigma_{\mathcal{A}}(x^*\,x)$. Thus $0 \not\in \sigma_{\mathcal{B}}(x^*\,x)$, by (13.10.6), which means that

$$(x^*\,x)^{-1} \in \mathcal{B}. \tag{13.10.7}$$

We also have that $(x^*\,x)^{-1}\,x^*\,x = e_{\mathcal{A}}$, so that

$$x^{-1} = (x^*\,x)^{-1}\,x^*. \tag{13.10.8}$$

The right side is an element of $\mathcal{B}$, because of (13.10.7), and the fact that $x^* \in \mathcal{B}$.

Let $a_1 \in \mathcal{A}$ be given, and let $\mathcal{A}_1(a_1)$ be the subalgebra of $\mathcal{A}$ generated by $e_{\mathcal{A}}$, a_1, and a_1^*. It is easy to see that $\mathcal{A}_1(a_1)$ is invariant under the involution on $\mathcal{A}$, and that the closure $\mathcal{B}(a_1)$ of $\mathcal{A}_1(a_1)$ in $\mathcal{A}$ is a subalgebra over $\mathcal{A}$ that is invariant under the involution as well. This means that

$$\sigma_{\mathcal{A}}(a_1) = \sigma_{\mathcal{B}(a_1)}(a_1), \tag{13.10.9}$$

as in (13.10.1). If a_1 is a normal element of $\mathcal{A}$, then $\mathcal{A}_1(a_1)$ and thus $\mathcal{B}(a_1)$ are commutative subalgebras of $\mathcal{A}$. In this case, the remarks in the previous section can be used for $\mathcal{B}(a_1)$, and with the spectrum of a_1 with respect to $\mathcal{A}$.

13.11 Another continuity property

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra with a nonzero multiplicative identity element $e_{\mathcal{A}}$ and involution $a \mapsto a^*$. If $a_1 \in \mathcal{A}$, then let $\mathcal{B}(a_1)$ be the closure of the subalgebra $\mathcal{A}_1(a_1)$ of $\mathcal{A}$ generated by $e_{\mathcal{A}}$, a_1, and a_1^*, as in the previous section. Thus $\mathcal{A}(a_1)$ and $\mathcal{B}(a_1)$ are invariant under the involution, as before, so that $\mathcal{B}(a_1)$ is a C^* algebra with respect to the restriction of $\|\cdot\|_{\mathcal{A}}$ to $\mathcal{B}$.

Suppose that a_1 is a normal element of $\mathcal{A}$, so that $\mathcal{A}_1(a_1)$ and $\mathcal{B}(a_1)$ are commutative subalgebras of $\mathcal{A}$, as before. If f is a continuous complex-valued function on $\sigma_{\mathcal{A}}(a_1)$, then $f(a_1) = f_{\mathcal{B}(a_1)}(a_1) \in \mathcal{B}(a_1)$ may be defined as in Section 13.9, because of (13.10.9). We may also use $f_{\mathcal{A}}(a_1)$ to denote $f(a_1)$, because $\mathcal{B}(a_1)$ is determined by a_1 and $\mathcal{A}$. Remember that $f \mapsto f_{\mathcal{A}}(a_1)$ defines an isometric algebra isomorphism from $C(\sigma_{\mathcal{A}}(a_1), \mathbf{C})$ onto $\mathcal{B}(a_1)$. Thus

$$\|f_{\mathcal{A}}(a_1)\|_{\mathcal{A}} = \|f\|_{sup,\sigma_{\mathcal{A}}(a_1)} \tag{13.11.1}$$

for every $f \in C(\sigma_{\mathcal{A}}(a_1), \mathbf{C})$, where the right side is the supremum norm of f on $\sigma_{\mathcal{A}}(a_1)$.

Now let f be any continuous complex-valued function on the complex plane. We can define $f(a_1) = f_{\mathcal{A}}(a_1) \in \mathcal{B}(a_1)$ using the restriction of f to $\sigma_{\mathcal{A}}(a_1)$. If K is a compact subset of $\mathbf{C}$ and

$$\sigma_{\mathcal{A}}(a_1) \subseteq K, \tag{13.11.2}$$

then it follows that

$$\|f_{\mathcal{A}}(a_1)\|_{\mathcal{A}} \le \|f\|_{sup,K} \tag{13.11.3}$$

where the right side is the supremum norm of f on K. Let us check that

$$a_1 \mapsto f_{\mathcal{A}}(a_1) \tag{13.11.4}$$

is a continuous mapping from

$$\{a_1 \in \mathcal{A} : a_1 \text{ is normal}\} \tag{13.11.5}$$

into $\mathcal{A}$, with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$ and its restriction to (13.11.5). This corresponds to Exercise (3) on p51 of [8].

More precisely, if r is a positive real number, then (13.11.4) is uniformly continuous on

$$\{a_1 \in \mathcal{A} : a_1 \text{ is normal, and } \|a_1\|_{\mathcal{A}} \le r\}. \tag{13.11.6}$$

To see this, one can start with the case where f is given by a polynomial in z, $\overline{z}$ with coefficients complex coefficients. In this case, one can check directly that (13.11.4) is Lipschitz on (13.11.6).

Of course, if $\|a_1\|_{\mathcal{A}} \le r$, then $\sigma_{\mathcal{A}}(a_1)$ is contained in the closed disk

$$\{\lambda \in \mathbf{C} : |\lambda| \le r\}, \tag{13.11.7}$$

as in Section 6.8. If f is any continuous complex-valued function on $\mathbf{C}$, then f can be uniformly approximated on (13.11.7) by polynomials in z and $\overline{z}$ with complex coefficients, by the Stone–Weierstrass theorem. This implies that (13.11.4) can be uniformly approximated on (13.11.6) by analogous functions associated to polynomials in z and $\overline{z}$ with complex coefficients, because of (13.11.3). It follows that (13.11.4) is uniformly continuous on (13.11.6), as desired.

13.12 Holomorphic functions in the plane

Let U be a nonempty open subset of the complex plane. It is well known that the space $\mathcal{H}(U)$ of holomorphic functions on U is a subalgebra of the space $C(U, \mathbf{C})$ of all continuous complex-valued functions on U, as an algebra over the complex numbers. Consider the space $\mathrm{Sp}(\mathcal{H}(U))$ of nonzero algebra homomorphisms from $\mathcal{H}(U)$ into $\mathbf{C}$, as in Section 12.14. If $w \in U$, then

$$h_w(f) = f(w) \tag{13.12.1}$$

defines an element of $\mathrm{Sp}(\mathcal{H}(U))$. Thus

$$w \mapsto h_w \tag{13.12.2}$$

defines a mapping from U into $\mathrm{Sp}(\mathcal{H}(U))$.

If $w \in \mathbf{C}$, then put

$$a_w(z) = z - w, \tag{13.12.3}$$

considered as an element of $\mathcal{H}(U)$, as a function of z. Note that

$$h_w(a_0) = w \tag{13.12.4}$$

for every $w \in U$, so that (13.5.3) is one-to-one. If $w \notin U$, then a_w is invertible as an element of $\mathcal{H}(U)$. This implies that

$$h(a_w) \neq 0 \tag{13.12.5}$$

for every $h \in \mathrm{Sp}(\mathcal{H}(U))$. This means that

$$h(a_0) \neq w \tag{13.12.6}$$

when $w \notin U$, so that

$$h(a_0) \in U \tag{13.12.7}$$

for every $h \in \mathcal{H}(U)$.

If $h \in \mathrm{Sp}(\mathcal{H}(U))$ satisfies

$$h(a_0) = w \tag{13.12.8}$$

for some $w \in U$, then we would like to check that

$$h = h_w. \tag{13.12.9}$$

If $f \in \mathcal{H}(U)$, then it is well known that there is a $g_w \in \mathcal{H}(U)$ such that

$$f(z) = f(w) + (z - w)\, g_w(z) \tag{13.12.10}$$

on U. Note that

$$h(a_w) = 0, \tag{13.12.11}$$

by (13.12.8). This implies that

$$h(f) = f(w) + h(a_w\, g_w) = f(w). \tag{13.12.12}$$

It follows that (13.12.2) is a one-to-one mapping from U onto $\mathcal{H}(U)$, with inverse given by

$$h \mapsto h(a_0). \tag{13.12.13}$$

Let us take $\mathrm{Sp}(\mathcal{H}(U))$ to be equipped with the weak topology associated to the functions $\widehat{f}(h) = h(f)$, $f \in \mathcal{H}(U)$, as usual. Observe that (13.12.2) is continuous as a mapping from U into $\mathrm{Sp}(\mathcal{H}(U))$ with respect to this topology, because holomorphic functions are continuous. We also have that (13.12.13) is continuous, because of the way that this topology n $\mathrm{Sp}(\mathcal{H}(U))$ is defined. This means that (13.12.2) is a homeomorphism from U onto $\mathrm{Sp}(\mathcal{H}(U))$.

13.13 Disk algebras

Let r be a positive real number, and let

$$U_r = \{z \in \mathbf{C} : |z| < r\} \tag{13.13.1}$$

be the open disk in the complex plane centered at 0 with radius r. Of course, the closure of U_r with respect to the standard Euclidean metric on $\mathbf{C}$ is the corresponding closed disk

$$\overline{U_r} = \{z \in \mathbf{C} : |z| \le r\}. \tag{13.13.2}$$

Consider the space $A(U_r)$ of continuous complex-valued functions on $\overline{U_r}$ that are holomorphic on U_r, as in Section 1.8. This is a closed linear subspace of the space $C(\overline{U_r}, \mathbf{C})$ of continuous complex-valued functions on $\overline{U_r}$, with respect to the supremum metric, as before. Note that $A(U_r)$ is a subalgebra of $C(\overline{U_r}, \mathbf{C})$, and thus a commutative Banach algebra with respect to the supremum norm.

One can use dilations on $\mathbf{C}$ to get that the $A(U_r)$'s are all isometrically isomorphic to each other, as Banach algebras. Often $A(U_1)$ is mostly considered, which is known as the *disk algebra*, as in Example 1.3.6 on p8 of [8], and Example 10.3 (c) on p230 of [162].

Suppose that $f \in A(U_r)$, and remember that f is uniformly continuous on $\overline{U_r}$, because $\overline{U_r}$ is compact with respect to the standard Euclidean metric on $\mathbf{C}$. If $0 < t < 1$, then

$$f_t(z) = f(t\,z) \tag{13.13.3}$$

defines an element of $A(U_{r/t})$, whose restriction to $\overline{U_r}$ is an element of $A(U_r)$. We also have that

$$f_t \to f \text{ uniformly on } \overline{U_r} \text{ as } t \to 1-, \tag{13.13.4}$$

because f is uniformly continuous on $\overline{U_r}$.

It is well known that any holomorphic function on U_r can be expressed as an absolutely convergent power series. The partial sums of this power series converge uniformly on any closed disk in $\mathbf{C}$ centered at 0 with radius less than r. If $0 < t < 1$, then it follows that f_t can be expressed as a power series whose partial sums converge uniformly on $\overline{U_r}$.

In particular, f_t can be approximated uniformly by polynomials in z with compelx coefficients on $\overline{U_r}$. This implies that

$$f \text{ can be approximated uniformly by polynomials on } \overline{U_r}, \tag{13.13.5}$$

because of (13.13.4).

Let us now consider the space $\mathrm{Sp}(A(\overline{U_r}))$ of nonzero algebra homomorphisms from $A(\overline{U_r})$ into $\mathbf{C}$, as in Sections 12.10 and 12.14. If $w \in \overline{U_r}$, then

$$h_w(f) = f(w) \tag{13.13.6}$$

defines an element of $\mathrm{Sp}(A(U_r))$, so that

$$w \mapsto h_w \tag{13.13.7}$$

defines a mapping from $\overline{U_r}$ into $\mathrm{Sp}(A(U_r))$. Put $a_0(z) = z$, considered as an element of $A(U_r)$, so that $h_w(a_0) = w$. Thus (13.13.7) is one-to-one on $\overline{U_r}$.

Let $h \in \mathrm{Sp}(A(U_r))$ be given, and remember that

$$|h(f)| \le \|f\|_{sup,\overline{U_r}} \tag{13.13.8}$$

for every $f \in A(U_r)$, as in Section 6.9, where the right side is the supremum norm of f on $\overline{U_r}$. In particular,

$$|h(a_0)| \le r. \tag{13.13.9}$$

If we put $w = h(a_0)$, then would like to check that

$$h(f) = h_w(f) \tag{13.13.10}$$

for every $f \in A(U_r)$. This can be verified directly when f is a polynomial. Otherwise, one can reduce to that case, using (13.13.5) and (13.13.8).

This shows that (13.13.7) maps $\overline{U_r}$ onto $\mathrm{Sp}(A(U_r))$, with inverse given by $h \mapsto h(a_0)$.

As usual, we take $\mathrm{Sp}(A(U_r))$ to be equipped with the weak topology associated to the functions $\widehat{f}(h) = h(f)$, $f \in A(U_r)$. It is easy to see that (13.13.7) is a homeomorphism from $\overline{U_r}$ onto $\mathrm{Sp}(A(U_r))$ with respect to this topology. This basically corresponds to the $n = 1$ case of Example 11.13 (c) on p271 of [162]. One can also use the maximum principle to identify $A(U_r)$ with a subalgebra of the algebra $C(\partial U_r, \mathbf{C})$ of continuous complex-valued function on the circle

$$\partial U_r = \{z \in \mathbf{C} : |z| = r\}, \tag{13.13.11}$$

which is the boundary of U_r in $\mathbf{C}$. This corresponds to Example 1.11.2 on p31 of [8].

13.14 Polydisk algebras

Let n be a positive integer. It is well known that a number of equivalent conditions may be used to characterize holomorphicity of functions of several complex variables. It will be convenient here to use the following, as in Definition 7.20 on p180 of [162]. A continuous complex-valued function f on an open set in $\mathbf{C}^n$ is *holomorphic* if it is holomorphic in each variable separately.

Let $r = (r_1, \ldots, r_n)$ be an n-tuple of positive real numbers, and let

$$U_r = \{z \in \mathbf{C}^n : |z_j| < r_j \text{ for } j = 1, \ldots, n\} \tag{13.14.1}$$

be the corresponding open polydisk in $\mathbf{C}^n$ centered at 0. The closure of U_r with respect to the standard Euclidean metric on $\mathbf{C}^n$ is the corresponding closed polydisk

$$\overline{U_r} = \{z \in \mathbf{C}^n : |z_j| \le r_j \text{ for } j = 1, \ldots, n\} \tag{13.14.2}$$

Consider the space $A(U_r)$ of continuous complex-valued functions on $\overline{U_r}$ that are holomorphic on U_r. This is a closed linear subspace of the space $C(\overline{U_r}, \mathbf{C})$ of continuous complex-valued functions on $\overline{U_r}$, with respect to the supremum metric, for essentially the same reasons as when $n = 1$. We also have that $A(\overline{U_r})$ is a subalgebra of $C(\overline{U_r}, \mathbf{C})$, and thus a commutative Banach algebra with respect to the supremum norm.

One can use linear mappings on $\mathbf{C}^n$ corresponding to dilations in each coordinate to get that the $A(U_r)$'s are all isometrically isomorphic to each other, as Banach algebras. If $r_j = 1$ for each j, then $A(U_r)$ may be called the *polydisk algebra*, as in Exercise 4 on p288 of [162], which is often mostly considered.

Let $f \in A(U_r)$ be given, and note that f is uniformly continuous on $\overline{U_r}$, because $\overline{U_r}$ is compact in $\mathbf{C}^n$. If $0 < t < 1$, then

$$f_t(z) = f(t\,z) \tag{13.14.3}$$

defines an element of $A(U_{t^{-1}\,r})$, where $t^{-1}\,r = (t^{-1}\,r_1, \ldots, t^{-1}\,r_n)$, as usual. It is easy to see that

$$f_t \to f \text{ uniformly on } \overline{U_r} \text{ as } t \to 1-, \tag{13.14.4}$$

because f is uniformly continuous on $\overline{U_r}$.

It is well known that any holomorphic function on U_r can be expressed as an absolutely convergent multiple power series. Without getting into the details too much, this can be obtained using the Cauchy integral formula in each variable separately. The main point for the moment is that f_t can be approximated uniformly by polynomials in $z_1, \ldots, z_n$ with complex coefficients on $\overline{U_r}$ when $0 < t < 1$. One can use this and (13.14.4) to get that

$$f \text{ can be approximated uniformly by polynomials on } \overline{U_r}. \tag{13.14.5}$$

This corresponds to Exercise 4 on p288 of [162].

Remember that $\mathrm{Sp}(A(U_r))$ is the space of nonzero algebra homomorphisms from $A(U_r)$ into $\mathbf{C}$, as in Sections 12.10 and 12.14. If $w \in \overline{U_r}$, then

$$h_w(f) = f(w) \tag{13.14.6}$$

defines an element of $\mathrm{Sp}(A(U_r))$, so that

$$w \mapsto h_w \tag{13.14.7}$$

defines a mapping from $\overline{U_r}$ into $\mathrm{Sp}(A(U_r))$, as before. Put

$$a_j(z) = z_j \tag{13.14.8}$$

for each $j = 1, \ldots, n$, considered as an element of $A(U_r)$. Thus

$$h_w(a_j) = w_j \tag{13.14.9}$$

for each $j = 1, \ldots, n$ and $w \in \overline{U_r}$. This implies that (13.14.7) is one-to-one on $\overline{U_r}$.

If $h \in \mathrm{Sp}(A(U_r))$, then

$$|h(f)| \le \|f\|_{sup,\overline{U_r}} \tag{13.14.10}$$

for every $f \in A(U_r)$, as in Section 6.9, where the right side is the supremum norm of f on $\overline{U_r}$. This implies that

$$|h(a_j)| \le r_j \tag{13.14.11}$$

for each $j = 1, \ldots, n$. This means that

$$w = (h(a_1), \ldots, h(a_n)) \tag{13.14.12}$$

is an element of $\overline{U_r}$.

Using this choice of w, one can check that

$$h(f) = h_w(f) \tag{13.14.13}$$

for every $f \in A(\overline{U_r})$. More precisely, if f is a polynomial in $z_1, \ldots, z_n$, then this can be obtained from the fact that h is an algebra homomorphism. Otherwise, one can approximate f by polynomials, as in (13.14.5), and use the continuity of h on $A(U_r)$, as in (13.14.10). It follows that (13.14.7) maps $\overline{U_r}$ onto $\mathrm{Sp}(A(U_r))$.

As before, we take $\mathrm{Sp}(A(U_r))$ to be equipped with the weak topology associated to the functions $\widehat{f}(h) = h(f)$, $f \in A(U_r)$. One can verify that (13.14.7) is a homeomorphism from $\overline{U_r}$ onto $\mathrm{Sp}(A(U_r))$ with respect to this topology, and the restriction of the standard Euclidean metric on $\mathbf{C}^n$ to $\overline{U_r}$. This corresponds to Example 11.13 (c) on p271 of [162].

13.15 Continuity and semisimplicity

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$, $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ be commutative Banach algebras, both real or both complex, with multiplicative identitye elements $e_{\mathcal{A}}$, $e_{\mathcal{B}}$, respectively, and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = \|e_{\mathcal{B}}\|_{\mathcal{B}} = 1$. Remember that $\mathrm{Sp}(\mathcal{A})$, $\mathrm{Sp}(\mathcal{B})$ are the spaces of nonzero algebra homomorphisms from $\mathcal{A}$, $\mathcal{B}$, respectively, into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 12.10. These homomorphisms are bounded linear funtionals on $\mathcal{A}$, $\mathcal{B}$, respectively, as in Section 6.9.

Let ϕ be an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$. If $h \in \mathrm{Sp}(\mathcal{B})$, then

$$h \circ \phi \tag{13.15.1}$$

is an algebra homomorphism from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Thus (13.15.1) is either an element of $\mathrm{Sp}(\mathcal{A})$, or it is identically equal to 0 on $\mathcal{A}$. In both cases, (13.15.1) is a bounded linear functional on $\mathcal{A}$. Of course, if $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$, then (13.15.1) is automatically nonzero.

Suppose that

$$\text{the elements of } \mathrm{Sp}(\mathcal{B}) \text{ separate points in } \mathcal{B}. \tag{13.15.2}$$

This is the same as saying that $\mathcal{B}$ is semisimple in the complex case, as in Section 6.12. Under these conditions, we have that

$$\phi \text{ is bounded as a linear mapping from } \mathcal{A} \text{ into } \mathcal{B}, \tag{13.15.3}$$

as in part (b) of Exercise (5) on p27 of [8], and Theorem 11.10 on p269 of [161]. To see this, one can use the closed graph theorem, as in [8, 162]. This corresponds to a criterion for the boundedness of a linear mapping mentioned in Section 8.7.

Suppose that

$$\text{the elements of } \mathrm{Sp}(\mathcal{A}) \text{ separate points in } \mathcal{A} \tag{13.15.4}$$

as well. If ϕ is an algebra isomorphism from $\mathcal{A}$ onto $\mathcal{B}$, then we get that

$$\phi^{-1} \text{ is bounded as a linear mapping from } \mathcal{B} \text{ into } \mathcal{A}, \tag{13.15.5}$$

as before. In particular, algebra automorphisms of $\mathcal{A}$ are bounded and have bounded inverses in this case. This corresponds to part (c) of Exercise (5) on p27 of [8], and to the corollary and remark after Theorem 11.10 on p270 of [162].

Chapter 14

Algebras, norms, and involutions

14.1 Some more remarks about involutions

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. Suppose that

$$a \mapsto a^* \tag{14.1.1}$$

is an involution on $\mathcal{A}$, which may be conjugate-linear in the complex case. If the involution is linear, then it is easy to see that

$$\sigma_{\mathcal{A}}(a^*) = \sigma_{\mathcal{A}}(a) \tag{14.1.2}$$

for every $a \in \mathcal{A}$. Similarly, if $\mathcal{A}$ is complex and the involution is conjugate-linear, then

$$\sigma_{\mathcal{A}}(a^*) = \{\overline{\lambda} : \lambda \in \sigma_{\mathcal{A}}(a)\}. \tag{14.1.3}$$

This corresponds to part (e) of Theorem 11.15 on p275 of [162], and to part (i) of 2.1 on p262 of [167].

Let $\mathrm{Sp}(\mathcal{A})$ be the set of all nonzero algebra homomorphisms from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 12.14. If $h \in \mathrm{Sp}(\mathcal{A})$ and the involution on $\mathcal{A}$ is linear, then

$$\widetilde{h}(a) = h(a^*) \tag{14.1.4}$$

defines an element of $\mathrm{Sp}(\mathcal{A})$ too. If $\mathcal{A}$ is complex and the involution is conjugate-linear, then

$$\widetilde{h}(a) = \overline{h(a^*)} \tag{14.1.5}$$

is an element of $\mathrm{Sp}(\mathcal{A})$, as mentioned in the proof of Theorem 11.16 on p276 of [162]. It is easy to see that

$$\widetilde{\widetilde{h}} = h \tag{14.1.6}$$

in both cases. In particular, this implies that

$$h \mapsto \widetilde{h} \tag{14.1.7}$$

is a one-to-one mapping from $\mathrm{Sp}(\mathcal{A})$ onto itself.

If $a \in \mathcal{A}$, then $\widehat{a}(h) = h(a)$ defines a real or complex-valued function on $\mathrm{Sp}(\mathcal{A})$, and we take $\mathrm{Sp}(\mathcal{A})$ to be equipped with the weakest topology for which these functions are continuous, as before. If $h \in \mathrm{Sp}(\mathcal{A})$, then

$$\widehat{a}(\widetilde{h}) = \widetilde{h}(a) = h(a^*) = \widehat{(a^*)}(h) \tag{14.1.8}$$

when the involution on $\mathcal{A}$ is linear, and

$$\widehat{a}(\widetilde{h}) = \widetilde{h}(a) = \overline{h(a^*)} = \overline{\widehat{(a^*)}(h)} \tag{14.1.9}$$

when $\mathcal{A}$ is complex and the involution is conjugate-linear. One can use these to get that (14.1.7) is a homeomorphism from $\mathrm{Sp}(\mathcal{A})$ onto itself in both cases.

Suppose that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra, so that every $h \in \mathrm{Sp}(\mathcal{A})$ is a bounded linear functional on $\mathcal{A}$, as in Section 6.9. This implies that

$$\widetilde{h} \text{ is a bounded linear functional on } \mathcal{A} \tag{14.1.10}$$

for every $h \in \mathrm{Sp}(\mathcal{A})$. Suppose also that $\mathcal{A}$ is commutative, and that

$$\text{the elements of } \mathrm{Sp}(\mathcal{A}) \text{ separate points in } \mathcal{A}, \tag{14.1.11}$$

which is the same as saying that $\mathcal{A}$ is semisimple in the complex case, as in Section 6.12. Under these conditions, there is a nonnegative real number C such that

$$\|a^*\|_{\mathcal{A}} \le C\, \|a\|_{\mathcal{A}} \tag{14.1.12}$$

for every $a \in \mathcal{A}$, as in Theorem 11.16 on p276 of [162].

More precisely, if the involution on $\mathcal{A}$ is linear, then it may be considered as an algebra automorphism of $\mathcal{A}$, because $\mathcal{A}$ is commutative, and (14.1.12) follows from the remarks in Section 13.15. If $\mathcal{A}$ is complex and the involution is conjugate-linear, then (14.1.12) is the same as saying that the involution is is bounded as a real-linear mapping from $\mathcal{A}$ into itself. This can be obtained from the closed graph theorem, in essentially the same way as before. In particular, one can use the criterion for the boundedness of linear mappings mentioned in Section 8.7.

14.2 Involutions and ideals

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with an involution $a \mapsto a^*$, which may be conjugate-linear in the complex case. Also let $\mathcal{I}$ be a two-sided ideal in $\mathcal{A}$, and let q be the corresponding quotient mapping from $\mathcal{A}$ onto $\mathcal{A}/\mathcal{I}$. If

$$\mathcal{I} \text{ is invariant under the involution on } \mathcal{A}, \tag{14.2.1}$$

then the involution on $\mathcal{A}$ induces a mapping from $\mathcal{A}/\mathcal{I}$ into itself, defined by

$$q(a)^* = q(a^*) \tag{14.2.2}$$

for every $a \in \mathcal{A}$. This defines an algebra involution on $\mathcal{A}/\mathcal{I}$, which is conjugate-linear in the complex case when the involution on $\mathcal{A}$ is conjugate-linear.

Suppose now that $\mathcal{A}$ is commutative and associative, and that $\mathcal{A}$ has a nonzero multiplicative identity element $e_{\mathcal{A}}$. Note that the involution on $\mathcal{A}$ takes maximal proper ideals in $\mathcal{A}$ to maximal proper ideals in $\mathcal{A}$. Remember that the Jacobson radical $\mathcal{R}(\mathcal{A})$ of $\mathcal{A}$ is the intersection of all of the maximal proper ideals in $\mathcal{A}$. It follows that

$$\mathcal{R}(\mathcal{A}) \text{ is invariant under the involution on } \mathcal{A}. \tag{14.2.3}$$

This implies that we get an induced involution on $\mathcal{A}/\mathcal{R}(\mathcal{A})$, as in (14.2.2).

Remember that $\mathrm{Sp}(\mathcal{A})$ is the set of all nonzero algebra homomorphisms from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Put

$$\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}) = \bigcap \{\ker h : h \in \mathrm{Sp}(\mathcal{A})\}, \tag{14.2.4}$$

which may be described as the *radical of $\mathcal{A}$ with respect to* $\mathrm{Sp}(\mathcal{A})$. This is interpreted as being equal to $\mathcal{A}$ when $\mathrm{Sp}(\mathcal{A}) = \emptyset$. Of course,

$$\mathcal{R}(\mathcal{A}) \subseteq \mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}), \tag{14.2.5}$$

because the kernel of every element of $\mathrm{Sp}(\mathcal{A})$ is a maximal proper ideal in $\mathcal{A}$, as in Section 6.12. If $\mathcal{A}$ is a complex Banach algebra, then

$$\mathcal{R}(\mathcal{A}) = \mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}), \tag{14.2.6}$$

as before.

Let $h \in \mathrm{Sp}(\mathcal{A})$ be given, and let $\widetilde{h} \in \mathrm{Sp}(\mathcal{A})$ be as in (14.1.4) or (14.1.5), as appropriate. In both cases, we have that

$$\ker \widetilde{h} = \{a \in \mathcal{A} : a^* \in \ker h\}. \tag{14.2.7}$$

This implies that

$$\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}) \text{ is invariant under the involution on } \mathcal{A}. \tag{14.2.8}$$

Thus we get an induced involution on $\mathcal{A}/\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A})$ too, as in (14.2.2).

If $\mathcal{I}$ is any ideal in $\mathcal{A}$, then any maximal proper ideal in $\mathcal{A}/\mathcal{I}$ corresponds to a maximal proper ideal in $\mathcal{A}$, by taking the inverse image under the quotient mapping. The maximal proper ideals in $\mathcal{A}/\mathcal{I}$ correspond exactly to the maximal proper ideals in $\mathcal{A}$ that contain $\mathcal{I}$ in this way. If

$$\mathcal{I} \subseteq \mathcal{R}(\mathcal{A}), \tag{14.2.9}$$

then every maximal proper ideal in $\mathcal{A}$ contains $\mathcal{I}$, and thus corresponds to a maximal proper ideal in $\mathcal{A}/\mathcal{I}$, so that

$$\mathcal{R}(\mathcal{A}/\mathcal{I}) = \mathcal{R}(\mathcal{A})/\mathcal{I}. \tag{14.2.10}$$

In particular,

$$\mathcal{R}(\mathcal{A}/\mathcal{R}(\mathcal{A})) = \{0\}. \tag{14.2.11}$$

Similarly, every element of $\mathrm{Sp}(\mathcal{A}/\mathcal{I})$ determines an element of $\mathrm{Sp}(\mathcal{A})$, by composition with the quotient mapping. In this case, the kernel of the corresponding element of $\mathrm{Sp}(\mathcal{A})$ contains $\mathcal{I}$, by construction. Conversely, if $h \in \mathrm{Sp}(\mathcal{A})$ and

$$\mathcal{I} \subseteq \ker h, \tag{14.2.12}$$

then h corresponds to a unique element of $\mathrm{Sp}(\mathcal{A}/\mathcal{I})$ in this way. If

$$\mathcal{I} \subseteq \mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}), \tag{14.2.13}$$

then every element of $\mathrm{Sp}(\mathcal{A})$ corresponds to a unique element of $\mathrm{Sp}(\mathcal{A}/\mathcal{I})$ in this way, and

$$\mathrm{rad}_{sp}(\mathcal{A}/\mathcal{I}) = \mathrm{rad}_{\mathrm{Sp}}(\mathcal{A})/\mathcal{I}. \tag{14.2.14}$$

Thus

$$\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}/\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A})) = \{0\}. \tag{14.2.15}$$

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra, and that $\mathrm{Sp}(\mathcal{A}) \neq \emptyset$. Remember that every $h \in \mathrm{Sp}(\mathcal{A})$ is a bounded linear functional on $\mathcal{A}$, as in Section 6.9, so that $\ker h$ is a closed set in $\mathcal{A}$, with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. This implies that $\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A})$ is a closed set in $\mathcal{A}$ too, so that $\mathcal{A}/\mathrm{rad}_{sp}(\mathcal{A})$ is a Banach algebra with respect to the corresponding quotient norm, as in Section 6.11. Note that $\mathcal{A}/\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}) \neq \{0\}$, because $\mathrm{Sp}(\mathcal{A}) \neq \emptyset$.

We also have that $\mathrm{Sp}(\mathcal{A}/\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A}))$ separates points in $\mathcal{A}/\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A})$, as in (14.2.11). It follows that the induced involution on $\mathcal{A}/\mathrm{rad}_{sp}(\mathcal{A})$ is continuous under these conditions, as in the previous section. This corresponds to an argument in the proof of Theorem 11.20 on p278 of [162].

14.3 Involutions and square roots

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. Suppose that $a \in \mathcal{A}$ satisfies $\|a\|_{\mathcal{A}} < 1$, and put

$$\phi_a(x) = (1/2)\, a - (1/2)\, x^2 \tag{14.3.1}$$

for every $x \in \mathcal{A}$, as in Section 11.15. Under these conditions, we saw before that ϕ_a defines a contraction from the closed ball $\overline{B}(0, \|a\|_{\mathcal{A}})$ in $\mathcal{A}$ into itself, so that ϕ_a has a unique fixed point x_a in $\overline{B}(0, \|a\|_{\mathcal{A}})$, by the contraction mapping theorem. We also saw that

$$\phi_a(x) = x \tag{14.3.2}$$

if and only if

$$(e_{\mathcal{A}} + x)^2 = e_{\mathcal{A}} + a. \tag{14.3.3}$$

More precisely, in the proof of the contraction theorem, we choose a point $u_1 \in \overline{B}(0, \|a\|_{\mathcal{A}})$, and we let $\{u_j\}_{j=1}^\infty$ be the sequence of elements of $\overline{B}(0, \|a\|_{\mathcal{A}})$ defined recursively by

$$u_{j+1} = \phi_a(u_j) \tag{14.3.4}$$

for each j, as in Section 7.12. This sequence converges to the unique fixed point x_a of ϕ_a in $\overline{B}(0, \|a\|_{\mathcal{A}})$, as before.

Suppose now that $x \mapsto x^*$ is an involution on $\mathcal{A}$, which may be conjugate-linear in the complex case. Suppose also that $a^* = a$, and observe that

$$\phi_a(x)^* = \phi_a(x^*) \tag{14.3.5}$$

for every $x \in \mathcal{A}$. We can take $u_1 = 0$ in the preceding paragraph, to get that

$$u_j^* = u_j \tag{14.3.6}$$

for each j. If the involution is continuous on $\mathcal{A}$, then it follows that

$$x_a^* = x_a. \tag{14.3.7}$$

Self-adjoint square roots for a larger class of self-adjoint elements are considered in the complex case in Theorems 11.20 and 11.26 on p278, 281 of [162], respectively, without asking that the involution be continuous. The argument for commutative Banach algebras will be discussed in the next section. An extension for Banach algebras that may not be commutative will be discussed in Section 14.9.

Remember that $e_{\mathcal{A}}^* = e_{\mathcal{A}}$, as in Section 6.4, so that

$$(e_{\mathcal{A}} + a)^* = e_{\mathcal{A}} + a. \tag{14.3.8}$$

We also have that

$$e_{\mathcal{A}} + a \text{ is invertible in } \mathcal{A}, \tag{14.3.9}$$

because $\|a\|_{\mathcal{A}} < 1$, as in Section 6.5. These conditions are used in the statement dicsussed in the next section.

14.4 Self-adjoint square roots

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a complex commutative Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and let $v \mapsto v^*$ be an involution on $\mathcal{A}$, which may be conjugate-linear. Also let $\{y_j\}_{j=1}^\infty$ be a sequence of elements of $\mathcal{A}$ that converges to $y \in \mathcal{A}$, and suppose that

$$y_j^* = y_j \tag{14.4.1}$$

for each j. If the involution is continuous on $\mathcal{A}$, then we get that

$$y^* = y, \tag{14.4.2}$$

and thus

$$(y^2)^* = y^2. \tag{14.4.3}$$

We would like to show that if

$$y^2 \text{ is invertible in } \mathcal{A}, \tag{14.4.4}$$

and if (14.4.3) holds then (14.4.2) holds, without asking that the involution be continuous on $\mathcal{A}$. This corresponds to part of the proof of Theorem 11.20 on p278 of [162].

Let $\mathrm{rad}(\mathcal{A})$ be the radical of $\mathcal{A}$, as in Section 6.12, which is the same as the Jacobson radical $\mathcal{R}(\mathcal{A})$ of $\mathcal{A}$, as well as the radical $\mathrm{rad}_{\mathrm{Sp}}(\mathcal{A})$ of $\mathcal{A}$ with respect to $\mathrm{Sp}(\mathcal{A})$ in this case, as before. Remember that $\mathrm{Sp}(\mathcal{A}) \neq \emptyset$ under these conditions, and that $\mathrm{rad}(\mathcal{A})$ is a proper closed ideal in $\mathcal{A}$. Thus $\mathcal{A}/\mathrm{rad}(\mathcal{A})$ is a Banach algebra with respect to the corresponding quotient norm, as in Section 6.11. Let q be the natural quotient mapping from $\mathcal{A}$ onto $\mathcal{A}/\mathrm{rad}(\mathcal{A})$, so that $q(e_{\mathcal{A}})$ is the multiplicative identity element in $\mathcal{A}/\mathrm{rad}(\mathcal{A})$, as before. We also have that the quotient norm of $q(e_{\mathcal{A}})$ is equal to 1, as before.

The involution on $\mathcal{A}$ induces an involution on $\mathcal{A}/\mathrm{rad}(\mathcal{A})$, with

$$q(u)^* = q(u^*) \tag{14.4.5}$$

for every $u \in \mathcal{A}$, as in Section 14.2. Remember that the induced involution on $\mathcal{A}/\mathrm{rad}(\mathcal{A})$ is continuous, as before. Of course, $\{q(y_j)\}_{j=1}^\infty$ converges to $q(y)$ in $\mathcal{A}/\mathrm{rad}(\mathcal{A})$, because $\{y_j\}_{j=1}^\infty$ converges to y in $\mathcal{A}$, by hypothesis. This implies that

$$q(y)^* = \lim_{j\to\infty} q(y_j)^* = \lim_{j\to\infty} q(y_j^*) = \lim_{j\to\infty} q(y_j) = q(y), \tag{14.4.6}$$

using (14.4.1) in the third step. It follows that

$$y^* - y \in \mathrm{rad}(\mathcal{A}), \tag{14.4.7}$$

because of (14.4.5).

Put $w = (1/2)\,(y + y^*)$ and $z = (1/2)\,(y - y^*)$, so that

$$y = w + z,\ w^* = w,\ z^* = -z, \tag{14.4.8}$$

as in Section 7.5. Note that

$$z \in \mathrm{rad}(\mathcal{A}), \tag{14.4.9}$$

by (14.4.7). Of course,

$$y^2 = w^2 + z^2 + 2\,w\,z. \tag{14.4.10}$$

We also have that

$$(w^2)^* = w^2,\ (z^2)^* = z^2,\ (w\,z)^* = -w\,z. \tag{14.4.11}$$

Using our hypothesis (14.4.3), we get that

$$w\,z = 0. \tag{14.4.12}$$

We would like to show that $z = 0$, to get (14.4.2). To do this, it suffices to show that w is invertible in $\mathcal{A}$, because of (14.4.12). Suppose for the sake of a contradiction that w is not invertible in $\mathcal{A}$, so that there is an algebra homomorphism h from $\mathcal{A}$ into $\mathbf{C}$ such that

$$h(w) = 0 \tag{14.4.13}$$

and $h(e_{\mathcal{A}}) = 1$, as in Section 6.12. Note that

$$h(y^2) \neq 0, \tag{14.4.14}$$

by (14.4.4), and that

$$h(z) = 0, \tag{14.4.15}$$

by (14.4.9). However, (14.4.13) nad (14.4.15) imply that

$$h(y^2) = 0, \tag{14.4.16}$$

because of (14.4.10).

14.5 Centralizers in associative algebras

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers. If E is a nonempty subset of $\mathcal{A}$, then the *centralizer* of E in $\mathcal{A}$ is defined by

$$C(E) = C_{\mathcal{A}}(E) = \{a \in \mathcal{A} : a\,x = x\,a \text{ for every } x \in E\}. \tag{14.5.1}$$

It is easy to see that

$$C_{\mathcal{A}}(E) \text{ is a subalgebra of } \mathcal{A}. \tag{14.5.2}$$

If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then

$$e_{\mathcal{A}} \in C_{\mathcal{A}}(E). \tag{14.5.3}$$

If $\|\cdot\|_{\mathcal{A}}$ is a submultiplicatve norm on $\mathcal{A}$, then

$$C_{\mathcal{A}}(E) \text{ is a closed set in } \mathcal{A}, \tag{14.5.4}$$

with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. We also have that

$$E \subseteq C_{\mathcal{A}}(C_{\mathcal{A}}(E)). \tag{14.5.5}$$

This corresponds to (a) and (b) in Section 11.21 on p280 of [162].

Suppose that $\mathcal{A}$ has a nonzero multiplicative identity element $e_{\mathcal{A}}$, and that $a \in C_{\mathcal{A}}(E)$ is invertible in $\mathcal{A}$. Observe that

$$a^{-1} \in C_{\mathcal{A}}(E). \tag{14.5.6}$$

This implies that

$$\sigma_{\mathcal{A}}(y) = \sigma_{C_{\mathcal{A}}(E)}(y) \tag{14.5.7}$$

for every $y \in C_{\mathcal{A}}(E)$. This corresponds to part of Theorem 11.22 on p280 of [162], and its proof.

If E_1, E_2 are nonempty subsets of $\mathcal{A}$ with

$$E_1 \subseteq E_2, \tag{14.5.8}$$

then

$$C_{\mathcal{A}}(E_2) \subseteq C_{\mathcal{A}}(E_1). \tag{14.5.9}$$

Note that the elements of E commute with each other if and only if

$$E \subseteq C_{\mathcal{A}}(E). \tag{14.5.10}$$

This implies that

$$C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \subseteq C_{\mathcal{A}}(E), \tag{14.5.11}$$

as in (14.5.9).

It is easy to see that

$$C_{\mathcal{A}}(E_1) \text{ is a commutative subalgebra of } \mathcal{A} \tag{14.5.12}$$

when

$$C_{\mathcal{A}}(E_1) \subseteq E_1. \tag{14.5.13}$$

This means that

$$C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \text{ is a commutative subalgebra of } \mathcal{A} \tag{14.5.14}$$

when the elements of E commute with each other, because of (14.5.11). This corresponds to (c) in Section 11.21 on p280 of [162].

If E is any nonempty subset of $\mathcal{A}$, then

$$\mathcal{B} = C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \tag{14.5.15}$$

is a subalgebra of $\mathcal{A}$ that contains E, as before. If $\mathcal{A}$ has a nonzero multiplicative identity element $e_{\mathcal{A}}$, then $e_{\mathcal{A}} \in \mathcal{B}$, and

$$\sigma_{\mathcal{A}}(b) = \sigma_{\mathcal{B}}(b) \tag{14.5.16}$$

for every $b \in \mathcal{B}$, as in (14.5.7). In particular, this holds when $b \in E$. If the elements of E commute with each other, and E is maximal with respect to inclusion among subsets of $\mathcal{A}$ with this property, then

$$\mathcal{B} = E, \tag{14.5.17}$$

by (14.5.5) and (14.5.14).

14.6 Spectra of sums and products

If E_1, E_2 are nonempty subsets of $\mathbf{R}$ or $\mathbf{C}$, then put

$$E_1 + E_2 = \{t_1 + t_2 : t_1 \in E_1,\, t_2 \in E_2\}, \tag{14.6.1}$$

$$E_1 - E_2 = \{t_1 - t_2 : t_1 \in E_1,\, t_2 \in E_2\}, \tag{14.6.2}$$

and

$$E_1\, E_2 = \{t_1\, t_2 : t_1 \in E_1,\, t_2 \in E_2\}, \tag{14.6.3}$$

as usual.

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$, and let $\mathcal{B}$ be a commutative subalgebra of $\mathcal{A}$ that contains $e_{\mathcal{A}}$. Also let h be a nonzero algebra homomorphism from $\mathcal{B}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. If $x, y \in \mathcal{B}$, then

$$h(x + y) = h(x) + h(y) \in \sigma_{\mathcal{B}}(x) + \sigma_{\mathcal{B}}(y) \tag{14.6.4}$$

and

$$h(x\, y) = h(x)\, h(y) \in \sigma_{\mathcal{B}}(x)\, \sigma_{\mathcal{B}}(y). \tag{14.6.5}$$

More precisely, the second step in each of (14.6.4) and (14.6.5) is as in Section 6.9.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a complex Banach algebra with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and that $\mathcal{B}$ is a closed subalgebra of $\mathcal{A}$, with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. Thus $\mathcal{B}$ is a commutative complex Banach algebra with respect to the restriction of $\|\cdot\|_{\mathcal{A}}$ to $\mathcal{B}$. If $b \in \mathcal{B}$, then $\sigma(b)$ consists of all complex numbers of the form $h(b)$, where h is a nonzero algebra homomorphism from $\mathcal{B}$ into $\mathbf{C}$, as in Section 6.12. It follows that

$$\sigma_{\mathcal{B}}(x + y) \subseteq \sigma_{\mathcal{B}}(x) + \sigma_{\mathcal{B}}(y) \tag{14.6.6}$$

and

$$\sigma_{\mathcal{B}}(x\, y) \subseteq \sigma_{\mathcal{B}}(x)\, \sigma_{\mathcal{B}}(y) \tag{14.6.7}$$

for every $x, y \in \mathcal{B}$, by (14.6.4) and (14.6.5). Of course, if (14.5.16) holds for every $b \in \mathcal{B}$, then we get that

$$\sigma_{\mathcal{A}}(x + y) \subseteq \sigma_{\mathcal{A}}(x) + \sigma_{\mathcal{A}}(y) \tag{14.6.8}$$

and

$$\sigma_{\mathcal{A}}(x\, y) \subseteq \sigma_{\mathcal{A}}(x)\, \sigma_{\mathcal{A}}(y) \tag{14.6.9}$$

for every $x, y \in \mathcal{B}$.

Let x, y be any two commuting elements of $\mathcal{A}$, and put $E = \{x, y\}$. If $\mathcal{B}$ is as in (14.5.15), then $\mathcal{B}$ is a closed commutative subalgebra of $\mathcal{A}$ that contains $e_{\mathcal{A}}$, x, and y, and (14.5.16) holds, as before. Thus (14.6.8) and (14.6.9) hold under these conditions, as in Theorem 11.23 on p280 of [162].

14.7 Invertibility of L_a, R_a

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$. If $a \in \mathcal{A}$, then let L_a and R_a be the corresponding left and right multiplication operators on $\mathcal{A}$, as in Sections 6.3 and 6.4. If a is invertible in $\mathcal{A}$, then L_a and R_a are invertible as linear mappings from $\mathcal{A}$ into itself, with

$$L_a^{-1} = L_{a^{-1}}, \ R_a^{-1} = R_{a^{-1}}. \tag{14.7.1}$$

If $b \in \mathcal{A}$, then it follows that the spectra of L_b and R_b in the algebra $\mathcal{L}(\mathcal{A})$ of linear mappings from $\mathcal{A}$ into itself satisfy

$$\sigma_{\mathcal{L}(\mathcal{A})}(L_b), \sigma_{\mathcal{L}(\mathcal{A})}(R_b) \subseteq \sigma_{\mathcal{A}}(b). \tag{14.7.2}$$

Similarly, if $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$, then the spectra of L_b, R_b in the algebra $\mathcal{BL}(\mathcal{A})$ of bounded linear mappings from $\mathcal{A}$ into itself satisfy

$$\sigma_{\mathcal{BL}(\mathcal{A})}(L_b), \sigma_{\mathcal{BL}(\mathcal{A})}(R_b) \subseteq \sigma_{\mathcal{A}}(b). \tag{14.7.3}$$

In fact,

$$\sigma_{\mathcal{BL}(\mathcal{A})}(L_b), \sigma_{\mathcal{BL}(\mathcal{A})}(R_b) = \sigma_{\mathcal{A}}(b), \tag{14.7.4}$$

as indicated in the proof of the corollary to Theorem 11.23 on p281 of [162]. Similarly,

$$\sigma_{\mathcal{L}(\mathcal{A})}(L_b), \sigma_{\mathcal{L}(\mathcal{A})}(R_b) = \sigma_{\mathcal{A}}(b), \tag{14.7.5}$$

without using a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$ on $\mathcal{A}$.

To see this, it suffices to show that if L_a or R_a is invertible in $\mathcal{L}(\mathcal{A})$ for some $a \in \mathcal{A}$, then a is invertible in $\mathcal{A}$. In particular, if $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$, and L_a or R_a is invertible in $\mathcal{BL}(\mathcal{A})$, then a is invertible in $\mathcal{A}$.

Put

$$L_{\mathcal{A}} = \{L_a : a \in \mathcal{A}\}, \ R_{\mathcal{A}} = \{R_a : a \in \mathcal{A}\}, \tag{14.7.6}$$

which are subalgebras of $\mathcal{L}(\mathcal{A})$. Note that

$$L_a \circ R_b = R_b \circ L_a \tag{14.7.7}$$

on $\mathcal{A}$ for every $a, b \in \mathcal{A}$, because $\mathcal{A}$ is associative, which was mentioned in Section 10.6 when $a = b$. In fact,

$$C_{\mathcal{L}(\mathcal{A})}(L_{\mathcal{A}}) = R_{\mathcal{A}}, \ C_{\mathcal{L}(\mathcal{A})}(R_{\mathcal{A}}) = L_{\mathcal{A}}, \tag{14.7.8}$$

using the notation for centralizers in Section 14.5. Indeed, if T is a linear mapping from $\mathcal{A}$ into itself that commutes with all left multiplication operators, then one can check that T is the same as right multiplication by $T(e_{\mathcal{A}})$. Similarly, if T commutes with all right multilpication oprators, then T is the same as left moltiplication by $T(e_{\mathcal{A}})$.

If $\mathcal{A}$ is equipped with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$, then $L_{\mathcal{A}}$, $R_{\mathcal{A}}$ may be considered as subalgebras of $\mathcal{BL}(\mathcal{A})$. In this case, we have that

$$C_{\mathcal{BL}(\mathcal{A})}(L_{\mathcal{A}}) = R_{\mathcal{A}},\ C_{\mathcal{BL}(\mathcal{A})}(R_{\mathcal{A}}) = L_{\mathcal{A}}, \tag{14.7.9}$$

as in (14.7.8).

If L_a is invertible in $\mathcal{L}(\mathcal{A})$ for some $a \in \mathcal{A}$, then L_a^{-1} commutes with all right multiplication operators on $\mathcal{A}$. This implies that $L_a^{-1} = L_b$ for some $b \in \mathcal{A}$, and one can use this to get that b is the inverse of a in $\mathcal{A}$. Similarly, if R_a is invertible in $\mathcal{L}(\mathcal{A})$ for some $a \in \mathcal{A}$, then $R_a^{-1} = R_b$ for some $b \in \mathcal{A}$, which implies that b is the inverse of a in $\mathcal{A}$.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a complex Banach algebra, so that $\mathcal{BL}(\mathcal{A})$ is a complex Banach algebra with respect to the corresponding operator norm. Let $a \in \mathcal{A}$ be given, and put $\delta_a = L_a - R_a$, as in Section 10.6. Observe that

$$\sigma_{\mathcal{BL}(\mathcal{A})}(\delta_a) \subseteq \sigma_{\mathcal{BL}(\mathcal{A})}(L_a) + \sigma_{\mathcal{BL}(\mathcal{A})}(-R_a), \tag{14.7.10}$$

as in (14.6.8), because L_a and R_a commute with each other. This implies that

$$\sigma_{\mathcal{BL}(\mathcal{A})}(\delta_a) \subseteq \sigma_{\mathcal{BL}(\mathcal{A})}(L_a) - \sigma_{\mathcal{BL}(\mathcal{A})}(R_a). \tag{14.7.11}$$

It follows that

$$\sigma_{\mathcal{BL}(\mathcal{A})}(\delta_a) \subseteq \sigma_{\mathcal{A}}(a) - \sigma_{\mathcal{A}}(a), \tag{14.7.12}$$

because of (14.7.3), as in the corollary to Theorem 11.23 on p281 of [162].

14.8 Involutions and centralizers

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with an involution $a \mapsto a^*$, which may be conjugate-linear in the complex case. Also let E be a nonempty subset of $\mathcal{A}$, so that the centralizer $C_{\mathcal{A}}(E)$ of E in $\mathcal{A}$ may be defined as in Section 14.5. If

$$E \text{ is invariant under the involution on } \mathcal{A}, \tag{14.8.1}$$

then

$$C_{\mathcal{A}}(E) \text{ is invariant under the involution on } \mathcal{A}. \tag{14.8.2}$$

Indeed, if $a \in C_{\mathcal{A}}(E)$ and (14.8.1) holds, then

$$a\,x^* = x^*\,a \tag{14.8.3}$$

for every $x \in E$. This implies that

$$a^*\,x = x\,a^* \tag{14.8.4}$$

for every $x \in E$, so that $a^* \in C_{\mathcal{A}}(E)$.

It follows that

$$C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \text{ is invariant under the involution on } \mathcal{A} \tag{14.8.5}$$

under these conditions. Remember that

$$\mathcal{B} = C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \tag{14.8.6}$$

is a subalgebra of $\mathcal{A}$ that contains E, as in Section 14.5. Of course, if $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then $e_{\mathcal{A}}$ is contained in the centralizer of any nonempty subset of $\mathcal{A}$, and in $\mathcal{B}$ in particular. If the elements of E commute with each other, then $\mathcal{B}$ is a commutative subalgebra of $\mathcal{A}$, as before. If E is a subset of $\mathcal{A}$ whose elements commute with each other, and if E is invariant under the involution on $\mathcal{A}$ and maximal with respect to inclusion among subsets of $\mathcal{A}$ with these properties, then it follows that

$$\mathcal{B} = E. \tag{14.8.7}$$

Theorem 11.25 on p281 of [162] discusses some properties of maximal subsets E of Banach algebras with involution such that the elements of E commute with each other and E is invariant under the involution. These properties can be obtained from (14.8.7) and the remarks in Section 14.5, which use similar arguments. In particular, E is a closed set under these conditions, although the involution is not asked to be continuous, as in [162].

14.9 Another criterion for self-adjointness

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a complex Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and an involution $a \mapsto a^*$ that may be conjugate-linear. Also let $\{y_j\}_{j=1}^{\infty}$ be a sequence of self-adjoint elements of $\mathcal{A}$ that converges to an element y of $\mathcal{A}$, and suppose that y^2 is self-adjoint and invertible in $\mathcal{A}$. Let us ask as well that

$$y_j\, y_l = y_l\, y_j \tag{14.9.1}$$

for all j, l, which implies that

$$y_j\, y = y\, y_j \tag{14.9.2}$$

for each j. We would like to show that y is self-adjoint under these conditions. Of course, if the involution on $\mathcal{A}$ is continuous, then the self-adjointness of y follows from the self-adjointness of the y_j's.

Remember that this was already shown in Section 14.4 when $\mathcal{A}$ is commutative. We would like to reduce to the previous result, as in Theorem 11.26 on p281 in [162], and its proof. To do this, we take

$$E = \{y_j : j \geq 1\}, \tag{14.9.3}$$

and $\mathcal{B}$ to be as in (14.8.6). The elements of E commute with each other and are self-adjoint, by hypothesis, so that E is invariant under the involution on $\mathcal{A}$. This implies that $\mathcal{B}$ is a commutative subalgebra of $\mathcal{A}$ that is invariant under the involution on $\mathcal{A}$, as before.

We also have that $\mathcal{B}$ contains $e_{\mathcal{A}}$ and is a closed set with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, as in Section 14.5. Note that $y \in \mathcal{B}$, so that $y^2 \in \mathcal{B}$ as well. This implies that

$$(y^2)^{-1} \in \mathcal{B}, \tag{14.9.4}$$

because y^2 is invertible in $\mathcal{A}$, and $\mathcal{B}$ is the centralizer of a subset of $\mathcal{A}$, as in Section 14.5. Thus we can use the argument in Section 14.4, with $\mathcal{B}$ in place of $\mathcal{A}$, to get that y is self-adjoint.

In practive, we may start with an invertible self-adjoint element b of $\mathcal{A}$ with some additional properties, and find a square root of b as a limit of a sequence $\{y_j\}_{j=1}^{\infty}$ of self-adjoint elements of $\mathcal{A}$. The y_j's may be elements of the subalgebra of $\mathcal{A}$ generated by $e_{\mathcal{A}}$ and b, which implies that they commute with each other. In this case, we can take

$$E = \{b\} \tag{14.9.5}$$

in the previous argument. This is a bit closed to the formulations in [162].

This includes the analogous question for complex Banach algebras in Section 14.3 too. In the previous notation, $b = e_{\mathcal{A}} + a$, and $y_j = e_{\mathcal{A}} + u_j$ for each j. Note that the inverse of b is automatically contained in any closed subalgebra of $\mathcal{A}$ that contains $e_{\mathcal{A}}$ and a or equivalently b, because $\|a\|_{\mathcal{A}} < 1$, by hypothesis.

14.10 A nonnegativity condition

Let $\mathcal{A}$ be an associative algebra over the complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$ and a conjugate-linear involution $a \mapsto a^*$. Let us say that $a \in \mathcal{A}$ is *nonnegative* if $a^* = a$ and

$$\sigma_{\mathcal{A}}(a) \subseteq \{\lambda \in \mathbf{R} : \lambda \geq 0\}. \tag{14.10.1}$$

This may be expressed by

$$a \geq 0, \tag{14.10.2}$$

as in Definition 11.27 on p282 of [162].

Let X be a nonempty metric or topological space, and consider the algebra $C(X, \mathbf{C})$ of all continuous complex-valued functions on X. The complex-conjugate of such a function is clearly continuous on X too, which defines a conjugate-linear involution on $C(X, \mathbf{C})$. An element of $C(X, \mathbf{C})$ is self-adjoint with respect to this involution if and only if it is real-valued on X. The spectrum of an element of $C(X, \mathbf{C})$ is its image in the complex plane, as mentioned in Section 8.15. The nonnegative elements of $C(X, \mathbf{C})$ in the sense described in the preceding paragraph are the same as the elements of $C(X, \mathbf{C})$ that are real-valued and nonnegative on X.

Similarly, complex-conjugation defines a conjugate-linear involution on the algebra $C_b(X, \mathbf{C})$ of all bounded continuous complex-valued functions on X. This is a subalgebra of $C(X, \mathbf{C})$ that is invariant under complex-conjugation. In particular, an element of $C_b(X, \mathbf{C})$ is self-adjoint with respect to this involution if and only if it is real-valued on X, as before. The spectrum of an element of

$C_b(X, \mathbf{C})$ is the closure of its image in the complex plane, as in Section 12.12. The nonnegative elements of $C_b(X, \mathbf{C})$ in the sense described earlier are the same as the elements of $C_b(X, \mathbf{C})$ that are real-valued and nonnegative on X.

If $a \in \mathcal{A}$ is self-adjoint, then a satisfies

$$a \geq 0 \text{ and } -a \geq 0 \tag{14.10.3}$$

if and only if

$$\sigma_{\mathcal{A}}(a) \subseteq \{0\}. \tag{14.10.4}$$

This implies that $a = 0$ when $\mathcal{A}$ is a semisimple commutative Banach algebra, as in Section 6.12. This also implies that $a = 0$ when $\mathcal{A}$ is a C^* algebra and a is normal, as in (14.12.1).

It is perhaps worth emphasizing that (14.10.1) includes the condition that

$$\sigma_{\mathcal{A}}(a) \subseteq \mathbf{R}. \tag{14.10.5}$$

If $\mathcal{A}$ is a C^* algebra, then this holds when a is self-adjoint, as in Section 12.8.

If a is any element of $\mathcal{A}$, then

$$\sigma_{\mathcal{A}}(a^2) = \{\lambda^2 : \lambda \in \sigma_{\mathcal{A}}(a)\}, \tag{14.10.6}$$

as in Section 8.13. If (12.8.1) holds, then it follows that

$$\sigma_{\mathcal{A}}(a^2) \subseteq \{\mu \in \mathbf{R} : \mu \geq 0\}. \tag{14.10.7}$$

If a is also self-adjoint, then a^2 is self-adjoint as well, and we get that

$$a^2 \geq 0. \tag{14.10.8}$$

Some additional properties of nonnegative elements of Banach algebras will be discussed in the next section. Afterwards, we shall consider nonnegativity in C^* algebras.

14.11 More on nonnegativity

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a complex Banach algebra, with a multiplicative identity element $e_{\mathcal{A}}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$ and a conjugate-linear involution, as before. Suppose for the moment that $\mathcal{A}$ is commutative, and let $\mathrm{Sp}(\mathcal{A})$ be the set of nonzero algebra homomorphisms from $\mathcal{A}$ into $\mathbf{C}$, as usual. Remember that the Gelfand transform of a is the complex-valued function defined on $\mathrm{Sp}(\mathcal{A})$ by $\widehat{a}(h) = h(a)$, as in Section 12.10. We have seen that the spectrum of $a \in \mathcal{A}$ is the same as the image of $\widehat{a}$ on $\mathrm{Sp}(\mathcal{A})$ in $\mathbf{C}$, as in Sections 6.9 and 6.12. Thus (14.10.5) is the same as saying that $\widehat{a}$ is real-valued on $\mathrm{Sp}(\mathcal{A})$, and (14.10.1) is the same as saying that $\widehat{a}$ is nonnegative on $\mathrm{Sp}(\mathcal{A})$ too.

Suppose for the moment again that a, b are commuting elements of $\mathcal{A}$. If a satisfies (14.10.5) and

$$\sigma_{\mathcal{A}}(b) \subseteq \mathbf{R}, \tag{14.11.1}$$

then

$$\sigma_{\mathcal{A}}(a+b) \subseteq \mathbf{R} \tag{14.11.2}$$

and

$$\sigma_{\mathcal{A}}(a\,b) \subseteq \mathbf{R}, \tag{14.11.3}$$

as in Section 14.6. Similarly, if a satisfies (14.10.1) and

$$\sigma_{\mathcal{A}}(b) \subseteq \{\lambda \in \mathbf{R} : \lambda \geq 0\}, \tag{14.11.4}$$

then

$$\sigma_{\mathcal{A}}(a+b) \subseteq \{\lambda \in \mathbf{R} : \lambda \geq 0\} \tag{14.11.5}$$

and

$$\sigma_{\mathcal{A}}(a\,b) \subseteq \{\lambda \in \mathbf{R} : \lambda \geq 0\}. \tag{14.11.6}$$

If a and b are self-adjoint too, then $a+b$ and $a\,b$ are self-adjoint, so that

$$a+b \geq 0 \tag{14.11.7}$$

and

$$a\,b \geq 0. \tag{14.11.8}$$

Suppose for the moment that $a \in \mathcal{A}$ is normal, so that a commutes with a^*. We can express a as

$$a = a_1 + i\,a_2, \tag{14.11.9}$$

where $a_1, a_2 \in \mathcal{A}$ are self-adjoint, because the involution on $\mathcal{A}$ is conjugate-linear, as in Section 7.5. Note that a_1 and a_2 commute, because a is normal. This implies that

$$a^*\,a = a_1^2 + a_2^2. \tag{14.11.10}$$

Of course, $a^*\,a$ is self-adjoint, and in fact

$$a^*\,a \geq 0, \tag{14.11.11}$$

because of (14.10.8) and (14.11.7).

If a is any element of $\mathcal{A}$ and $t \in \mathbf{C}$, then it is easy to see that

$$\sigma_{\mathcal{A}}(t\,e_{\mathcal{A}} - a) = t - \sigma_{\mathcal{A}}(a) = \{t - \lambda : \lambda \in \sigma_{\mathcal{A}}(a)\}. \tag{14.11.12}$$

Suppose that (14.10.5) holds, so that

$$\sigma_{\mathcal{A}}(a) \subseteq [-\|a\|_{\mathcal{A}}, \|a\|_{\mathcal{A}}], \tag{14.11.13}$$

as in Section 6.8. If $t \in \mathbf{R}$, then it follows that

$$\sigma_{\mathcal{A}}(t\,e_{\mathcal{A}} - a) \subseteq [t - \|a\|_{\mathcal{A}}, t + \|a\|_{\mathcal{A}}]. \tag{14.11.14}$$

In particular, this means that

$$\sigma_{\mathcal{A}}(t\,e_{\mathcal{A}} - a) \subseteq \{\lambda \in \mathbf{R} : \lambda \geq 0\} \tag{14.11.15}$$

when $t \geq \|a\|_{\mathcal{A}}$.

Suppose that $b \in \mathcal{A}$, $t \in \mathbf{R}$, and

$$\|t\,e_{\mathcal{A}} - b\|_{\mathcal{A}} \leq t. \tag{14.11.16}$$

If (14.11.1) holds, then (14.11.4) holds as well. This is the same as (14.11.15), with $a = t\,e_{\mathcal{A}} - b$.

14.12 Sums in C^* algebras

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra, with nonzero multiplicative identity element $e_{\mathcal{A}}$. If a is a normal element of $\mathcal{A}$, then

$$\|a\|_{\mathcal{A}} = r_{\mathcal{A}}(a) = \max\{|\lambda| : \lambda \in \sigma_{\mathcal{A}}(a)\}, \tag{14.12.1}$$

where the first step is as in Section 7.7, and the second step is as in Section 6.14. In particular, this holds when a is self-adjoint.

Suppose that u, v are nonnegative elements of $\mathcal{A}$, in the sense defined in Section 14.10. We would like to show that

$$u + v \ge 0 \tag{14.12.2}$$

too. This corresponds to Lemma 4.8.1 on p126 of [8], part (d) of Theorem 11.28 on p282 of [162], and part of part (i) of Theorem 3.1 on p267 of [167]. Of course, $u + v$ is self-adjoint, because u, v are self-adjoint, by hpothesis, and so we want to show that

$$\sigma_{\mathcal{A}}(u + v) \subseteq \{\lambda \in \mathbf{R} : \lambda \ge 0\}. \tag{14.12.3}$$

This was discussed in the previous section when u and v commute.

Note that

$$\sigma_{\mathcal{A}}(u) \subseteq [0, \|u\|_{\mathcal{A}}], \tag{14.12.4}$$

because of (14.11.13) and the hypothesis that $u \ge 0$. This implies that

$$\sigma_{\mathcal{A}}(\|u\|_{\mathcal{A}}\, e_{\mathcal{A}} - u) \subseteq [0.\|u\|_{\mathcal{A}}], \tag{14.12.5}$$

as in (14.11.12). It follows that

$$\big\|\|u\|_{\mathcal{A}}\, e_{\mathcal{A}} - u\big\|_{\mathcal{A}} \le \|u\|_{\mathcal{A}}, \tag{14.12.6}$$

by (14.12.1). Similarly,

$$\big\|\|v\|_{\mathcal{A}}\, e_{\mathcal{A}} - v\big\|_{\mathcal{A}} \le \|v\|_{\mathcal{A}}. \tag{14.12.7}$$

Using (14.12.6) and (14.12.7), we get that

$$\begin{aligned} \|(\|u\|_{\mathcal{A}} + \|v\|_{\mathcal{A}})\, e_{\mathcal{A}} - (u + v)\|_{\mathcal{A}} &\le \big\|\|u\|_{\mathcal{A}}\, e_{\mathcal{A}} - u\big\|_{\mathcal{A}} + \big\|\|v\|_{\mathcal{A}}\, e_{\mathcal{A}} - v\big\|_{\mathcal{A}} \\ &\le \|u\|_{\mathcal{A}} + \|v\|_{\mathcal{A}}. \end{aligned} \tag{14.12.8}$$

Put $t = \|u\|_{\mathcal{A}} + \|v\|_{\mathcal{A}}$, so that (14.12.8) is the same as saying that

$$\|t\, e_{\mathcal{A}} - (u + v)\|_{\mathcal{A}} \le t. \tag{14.12.9}$$

This implies (14.12.3), as mentioned at the end of the previous section, with $b = u + v$.

14.13 Nonnegativity of $a^* a$

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra again, with a nonzero multiplicative identity element $e_{\mathcal{A}}$. If $a \in \mathcal{A}$, then $a^* a$ is self-adjoint, and it is well known that

$$a^* a \geq 0, \tag{14.13.1}$$

in the sense of Section 14.10. This is Theorem 4.8.3 on p127 of [8], part (e) of Theorem 11.28 on p282 of [162], and part of part (iii) of Theorem 3.1 on p267 of [167]. A couple of proofs of this will be discussed in the next two sections. Remember that this was discussed in Section 14.11 when a is normal.

Suppose that

$$\sigma_{\mathcal{A}}(a^* a) \subseteq \{\lambda \in \mathbf{R} : \lambda \leq 0\}, \tag{14.13.2}$$

which means that

$$-a^* a \geq 0. \tag{14.13.3}$$

If (14.13.1) holds, then

$$\sigma_{\mathcal{A}}(a^* a) = \{0\}, \tag{14.13.4}$$

as in Section 14.10. This implies that

$$a^* a = 0, \tag{14.13.5}$$

as in (14.12.1), because $a^* a$ is self-adjoint. This means that

$$a = 0, \tag{14.13.6}$$

because $\mathcal{A}$ is a C^* algebra.

Lemma 4.8.2 on p127 of [8] states that (14.13.2) implies (14.13.6), without using (14.13.1). In fact, this is used to get (14.13.1) afterwards, as in the next section.

To see this, remember that the nonzero elements of the spectra of $a^* a$ and $a\, a^*$ are the same, as in Section 11.5. Thus (14.13.2) implies that

$$\sigma_{\mathcal{A}}(a\, a^*) \subseteq \{\lambda \in \mathbf{R} : \lambda \leq 0\}. \tag{14.13.7}$$

It follows that

$$\sigma_{\mathcal{A}}(a^* a + a\, a^*) \subseteq \{\lambda \in \mathbf{R} : \lambda \leq 0\}, \tag{14.13.8}$$

as in the previous section.

Let us express a as $a_1 + i\, a_2$, where $a_1, a_2 \in \mathcal{A}$ are self-adjoint, as in Section 7.5. Observe that

$$a^* a + a\, a^* = 2\, a_1^2 + 2\, a_2^2. \tag{14.13.9}$$

This means that

$$-2\, a_1^2 - 2\, a_2^2 \geq 0, \tag{14.13.10}$$

by (14.13.8).

Note that $a_1^2, a_2^2 \geq 0$, because a_1, a_2 are self-adjoint, as in Section 14.10. This implies that $2\,a_1^2, 2\,a_2^2 \geq 0$ as well. It follows that

$$-2\,a_1^2 = (-2\,a_1^2 - 2\,a_2^2) + 2\,a_2^2 \geq 0, \tag{14.13.11}$$

as in the previous section. This means that $2\,a_1^2 = 0$, as before. Thus $a_1 = 0$, because a_1 is self-adjoint, and $\mathcal{A}$ is a C^* algebra.

Similarly, $a_2 = 0$. This shows that (14.13.6) holds, because $a = a_1 + i\,a_2$.

Observe that (14.13.9) implies that

$$a^*\,a + a\,a^* \geq 0, \tag{14.13.12}$$

using the result in the previous section. This could also be obtained from (14.13.1) and the analogous property of

$$a\,a^* = (a^*)^*\,a^*. \tag{14.13.13}$$

14.14 A proof of $a^*\,a \geq 0$

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra with a nonzero multiplicative identity element $e_{\mathcal{A}}$, and let $a \in \mathcal{A}$ be given. We would like to show that $a^*\,a \geq 0$, as mentioned in the previous section, following the proof of Theorem 4.8.3 on p127 of [8]. Let

$$\mathcal{A}_1(a^*\,a) \tag{14.14.1}$$

be the subalgebra of $\mathcal{A}$ generated by $e_{\mathcal{A}}$ and $a^*\,a$, and let

$$\mathcal{B}(a^*\,a) \tag{14.14.2}$$

be the closure of $\mathcal{A}_1(a^*\,a)$ in $\mathcal{A}$ with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. Note that $\mathcal{A}_1(a^*\,a)$ is a commutative subalgebra of $\mathcal{A}$ that is invariant under the involution, because $a^*\,a$ is self-adjoint. This implies that $\mathcal{B}(a^*\,a)$ has the same properties, so that it is a commutative C^* algebra with respect to the restriction of $\|\cdot\|_{\mathcal{A}}$ to $\mathcal{B}(a^*\,a)$.

Remember that the spectrum of $a^*\,a$ as an element of $\mathcal{B}(a^*\,a)$ is the same as $\sigma_{\mathcal{A}}(a^*\,a)$, as in Section 13.10. If f is a continuous complex-valued function on $\sigma_{\mathcal{A}}(a^*\,a)$, then $f(a^*\,a) \in \mathcal{B}(a^*\,a)$ may be defined as in Section 13.9. The mapping

$$f \mapsto f(a^*\,a) \tag{14.14.3}$$

defines an isometric algebra isomorphism from $C(\sigma_{\mathcal{A}}(a^*\,a), \mathbf{C})$ onto $\mathcal{B}(a^*\,a)$, as before. We also have that

$$(\overline{f})(a^*\,a) = f(a^*\,a)^* \tag{14.14.4}$$

for every $f \in C(\sigma_{\mathcal{A}}(a^*\,a), \mathbf{C})$. This means that

$$f(a^*\,a)^* = f(a^*\,a) \tag{14.14.5}$$

when f is real-valued on $\sigma_{\mathcal{A}}(a)$.

Note that

$$\sigma_{\mathcal{A}}(a^* a) \subseteq \mathbf{R}, \tag{14.14.6}$$

because $a^* a$ is self-adjoint. Consider the continuous real-valued functions f, g defined on $\sigma_{\mathcal{A}}(a^* a)$ by

$$\begin{aligned} f(t) &= \sqrt{t} \quad \text{when } t \geq 0 \\ &= 0 \quad \text{when } t < 0 \end{aligned} \tag{14.14.7}$$

and

$$\begin{aligned} g(t) &= 0 \quad \text{when } t > 0 \\ &= \sqrt{-t} \quad \text{when } t \leq 0. \end{aligned} \tag{14.14.8}$$

Thus

$$f(t)^2 - g(t)^2 = t \tag{14.14.9}$$

and

$$f(t)\, g(t) = 0 \tag{14.14.10}$$

for every $t \in \sigma_{\mathcal{A}}(a^* a)$. Put

$$x = f(a^* a),\ y = g(a^* a), \tag{14.14.11}$$

which are self-adjoint elements of $\mathcal{B}(a^* a)$, as in the previous paragraph. These elements satisfy

$$x^2 - y^2 = a^* a \tag{14.14.12}$$

and

$$x\, y = y\, x = 0, \tag{14.14.13}$$

because of (14.14.9) and (14.14.10).

Consider

$$(a\, y)^* (a\, y) = y\, a^* a\, y = y\,(x^2 - y^2)\, y = -y^4, \tag{14.14.14}$$

as on p128 of [8]. Observe that

$$y^4 = (y^2)^2 \geq 0, \tag{14.14.15}$$

as in Section 14.10, so that

$$\sigma_{\mathcal{A}}((a\, y)^* (a\, y)) \subseteq \{\lambda \in \mathbf{R} : \lambda \leq 0\}. \tag{14.14.16}$$

This implies that

$$a\, y = 0, \tag{14.14.17}$$

as in the previous section. This means that

$$y^4 = 0, \tag{14.14.18}$$

because of (14.14.14). It follows that $y^2 = 0$, because $\mathcal{A}$ is a C^* algebra, and y is self-adjoint. Similarly,

$$y = 0. \tag{14.14.19}$$

Thus

$$a^* a = x^2 \geq 0, \tag{14.14.20}$$

as in Section 14.10.

14.15 Some remarks about related arguments

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra with a nonzero multiplicative identity element $e_{\mathcal{A}}$ again, and let $a \in \mathcal{A}$ be given. The fact that $a^*\, a \geq 0$ corresponds to part (e) of Theorem 11.28 on p282 of [162], as mentioned in Section 14.13. The proof is basically similar to the one discussed in the previous two sections, aside from one main difference near the beginning.

One starts by choosing a commutative subalgebra $\mathcal{B}$ of $\mathcal{A}$ that contains $e_{\mathcal{A}}$ and $a^*\, a$, is invariant under the involution, is closed with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, and for which

$$\sigma_{\mathcal{A}}(b) = \sigma_{\mathcal{B}}(b) \tag{14.15.1}$$

for every $b \in \mathcal{B}$. In fact, this last property holds automatically under these conditions, as in Section 13.10. This was discussed afterwards in Theorem 11.29 on p283 of [162], and $\mathcal{B}$ was taken instead to be a maximal subset of $\mathcal{A}$ whose elements commute with each other, and which is invariant under the involution. This is related to the remarks in Section 14.8. Using the result in Section 13.10, one can take $\mathcal{B} = \mathcal{B}(a^*\, a)$ as in the previous section.

Remember that $\mathcal{B}$ is isometrically isomorphic to the algebra of continuous complex-valued functions on the compact Hausdorff space $\mathrm{Sp}(\mathcal{B})$, as in Section 12.11. One can use this to deal with functions of $a^*\, a$ as elements of $\mathcal{B}$, instead of the version in Section 13.9, as in the previous section.

Note that nonnegativity of a self-adjoint element b of $\mathcal{B}$ is the same with respect to $\mathcal{A}$ and $\mathcal{B}$, because of (14.15.1). If b is any element of $\mathcal{B}$, then

$$\sigma_{\mathcal{B}}(b) = \widehat{b}(\mathrm{Sp}(\mathcal{B})), \tag{14.15.2}$$

where $\widehat{b}$ is the Gelfand transform of b. This is the same as the spectrum of $\widehat{b}$ in $C(\mathrm{Sp}(\mathcal{B}), \mathbf{C})$. The nonnegativity of b in $\mathcal{B}$ is equivalent to the nonnegativity of $\widehat{b}$ in $C(\mathrm{Sp}(\mathcal{B}), \mathbf{C})$, which means that $\widehat{b}$ is real-valued and nonnegative on $\mathrm{Sp}(\mathcal{B})$.

The nonnegativity of $a^*\, a$ is also part of part (iii) of Theorem 3.1 on p267 of [167], as mentioned in Section 14.13. The proof is basically similar to the one in the previous two sections as well, and some additional related facts are discussed too.

Chapter 15

Norms, weights, and power series

15.1 Norms, isometries, and $G(\mathcal{A})$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$, and remember that $G(\mathcal{A})$ is the group of invertible elements of $\mathcal{A}$. Of course, if $x \in G(\mathcal{A})$, then 0 is not an element of either $\sigma_{\mathcal{A}}(x)$ or $\sigma_{\mathcal{A}}(x^{-1})$. One can check that

$$\sigma_{\mathcal{A}}(x^{-1}) = \{1/\lambda : \lambda \in \sigma_{\mathcal{A}}(x)\} \tag{15.1.1}$$

in this case. Suppose that $\|\cdot\|_{\mathcal{A}}$ is a submultiplicative norm on $\mathcal{A}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Observe that

$$1 = \|e_{\mathcal{A}}\|_{\mathcal{A}} \le \|x\|_{\mathcal{A}}\,\|x^{-1}\|_{\mathcal{A}} \tag{15.1.2}$$

for every $x \in G(\mathcal{A})$.

One can check that

$$\{u \in G(\mathcal{A}) : \|u\|_{\mathcal{A}}, \|u^{-1}\|_{\mathcal{A}} \le 1\} \tag{15.1.3}$$

is a subgroup of $G(\mathcal{A})$. This is the same as

$$\{u \in G(\mathcal{A}) : \|u\|_{\mathcal{A}} = \|u^{-1}\|_{\mathcal{A}} = 1\}, \tag{15.1.4}$$

because of (15.1.2). It is easy to see that this is a relatively closed set in $G(\mathcal{A})$, because $x \mapsto x^{-1}$ is continuous on $G(\mathcal{A})$. If $\mathcal{A}$ is a Banach algebra, then one can verify that

$$\text{(15.1.3) is a closed set in } \mathcal{A}, \tag{15.1.5}$$

using a remark in Section 6.7.

If u is an element of (15.1.3) and $\lambda \in \sigma_{\mathcal{A}}(u)$, then $\lambda \ne 0$, as before, and $|\lambda| \le 1$, as in Section 6.8. Similarly,

$$|1/\lambda| \le 1, \tag{15.1.6}$$

because $1/\lambda \in \sigma_{\mathcal{A}}(u^{-1})$, as in (15.1.1). This means that

$$|\lambda| = 1. \tag{15.1.7}$$

If u is an element of (15.1.3) and $a \in \mathcal{A}$, then

$$\|u\,a\|_{\mathcal{A}} \le \|u\|_{\mathcal{A}}\,\|a\|_{\mathcal{A}} = \|a\|_{\mathcal{A}} \tag{15.1.8}$$

and

$$\|a\|_{\mathcal{A}} = \|u^{-1}\,u\,a\|_{\mathcal{A}} \le \|u^{-1}\|_{\mathcal{A}}\,\|u\,a\|_{\mathcal{A}} = \|u\,a\|_{\mathcal{A}}. \tag{15.1.9}$$

This implies that

$$\|u\,a\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}. \tag{15.1.10}$$

Similarly,

$$\|a\,u\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}. \tag{15.1.11}$$

Conversely, if $u \in \mathcal{A}$ and (15.1.10) or (15.1.11) holds with $a = e_{\mathcal{A}}$, then $\|u\|_{\mathcal{A}} = 1$. If $u \in G(\mathcal{A})$ and (15.1.10) or (15.1.11) holds with $a = u^{-1}$, then $\|u^{-1}\|_{\mathcal{A}} = 1$.

Let $V, W \ne \{0\}$ be vector spaces, both real or both complex, with norms $\|\cdot\|_V$, $\|\cdot\|_W$, respectively, and let T be a one-tone bounded linear mapping from V onto W, with bounded inverse. Remember that $\|T^{-1}\|_{op,WV}$ is the smallest nonnegative real number such that

$$\|T^{-1}(w)\|_V \le \|T^{-1}\|_{op,WV}\,\|w\|_W \tag{15.1.12}$$

for every $w \in W$. Equivalently, $\|T^{-1}\|_{op,WV}$ is the smallest nonnegative real number such that

$$\|v\|_V \le \|T^{-1}\|_{op,WV}\,\|T(v)\|_W \tag{15.1.13}$$

for every $v \in V$. It follows that T is an isometric linear mapping from V onto W if and only if

$$\|T\|_{op,VW},\ \|T^{-1}\|_{op,WV} \le 1, \tag{15.1.14}$$

in which case $\|T\|_{op,VW} = \|T^{-1}\|_{op,WV} = 1$. In particular, if $\mathcal{A}$ is the algebra $\mathcal{BL}(V)$ of bounded linear mappings from V into itself, then (15.1.3) is the group of isometric linear mappings from V onto itself.

15.2 Some remarks about unitary elements

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a C^* algebra, with nonzero multiplicative identity element $e_{\mathcal{A}}$. An element u of $\mathcal{A}$ is said to be *unitary* if

$$u^*\,u = u\,u^* = e_{\mathcal{A}}, \tag{15.2.1}$$

as on p261 of [167]. Of course, this is the same as saying that u is invertible in $\mathcal{A}$, with

$$u^{-1} = u^*. \tag{15.2.2}$$

The collection $U(\mathcal{A})$ of unitary elements of $\mathcal{A}$ is a subgroup of the group $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$. This was mentioned in Section 7.5 for arbitrary associative algebras with involutions.

Using (15.2.1), we get that

$$\|u\|_{\mathcal{A}} = 1. \tag{15.2.3}$$

Similarly,

$$\|u^{-1}\|_{\mathcal{A}} = 1, \tag{15.2.4}$$

because $(u^{-1})^* \, u^{-1} = u \, u^{-1} = e_{\mathcal{A}}$. This shows that $U(\mathcal{A})$ is contained in (15.1.3).

Suppose now that $u \in G(\mathcal{A})$ satisfies $\|u\|_{\mathcal{A}}, \|u^{-1}\|_{\mathcal{A}} \leq 1$, so that

$$\|u^* \, u\|_{\mathcal{A}} = \|u\|_{\mathcal{A}}^2 \leq 1 \tag{15.2.5}$$

and

$$\|(u^{-1})^*\|_{\mathcal{A}} = \|u^{-1}\|_{\mathcal{A}} \leq 1. \tag{15.2.6}$$

This uses the fact that the involution on a C^* algebra preserves the norm, as in Section 7.7. It follows that

$$\|u^{-1} \, (u^{-1})^*\|_{\mathcal{A}} = \|((u^{-1})^*)^* \, (u^{-1})^*\|_{\mathcal{A}} = \|(u^{-1})^*\|_{\mathcal{A}}^2 \leq 1. \tag{15.2.7}$$

Observe that

$$(u^* \, u)^{-1} = u^{-1} \, (u^*)^{-1} = u^{-1} \, (u^{-1})^*, \tag{15.2.8}$$

so that

$$\|(u^* \, u)^{-1}\|_{\mathcal{A}} \leq 1. \tag{15.2.9}$$

This shows that $u^* \, u$ is an element of (15.1.3), so that

$$\sigma_{\mathcal{A}}(u^* \, u) \subseteq \{\lambda \in \mathbf{C} : |\lambda| = 1\}, \tag{15.2.10}$$

as in (15.1.7). Of course, $u^* \, u$ is self-adjoint, so that

$$\sigma_{\mathcal{A}}(u^* \, u) \subseteq \mathbf{R}, \tag{15.2.11}$$

as in Section 12.8. Combining this with (15.2.10), we get that

$$\sigma_{\mathcal{A}}(u^* \, u) \subseteq \{1, -1\}. \tag{15.2.12}$$

This implies that

$$\sigma_{\mathcal{A}}(u^* \, u) = \{1\}, \tag{15.2.13}$$

because $u^* \, u \geq 0$, as in Section 14.13.

This means that

$$\sigma_{\mathcal{A}}(u^* \, u - e_{\mathcal{A}}) = \{0\}, \tag{15.2.14}$$

as in Section 14.11. It follows that $u^* \, u = e_{\mathcal{A}}$, because $u^* \, u - e_{\mathcal{A}}$ is self-adjoint, as in Section 14.12. Thus $u \in U(\mathcal{A})$, because $u \in G(\mathcal{A})$, by hypothesis. This shows that $U(\mathcal{A})$ is equal to (15.1.3) in this case.

15.3 Weighted ℓ^p spaces

Let X be a nonempty set, let p be a positive real number, and let Z be a vector space over the real or complex numbers with a norm $\|\cdot\|_Z$. Also let w be a positive real-valued function on X, and consider the space

$$\ell^p_w(X, Z) \tag{15.3.1}$$

of Z-valued functions f on X such that

$$\|f(x)\|_Z\, w(x) \tag{15.3.2}$$

is p-summable as a nonnegative real-valued function on X. Equivalently, this means that

$$\|f(x)\|_Z^p\, w(x)^p \tag{15.3.3}$$

is summable on X. Sometimes one may consider w^p as the relevant weight here, and use somewhat different notation to reflect that. This version can be more convenient in some ways, in connection with the depedence on p.

Similarly, let $\ell^\infty_w(X, Z)$ be the space of Z-valued functions f on X such that (15.3.2) is bounded on X. If $f \in \ell^p_w(X, Z)$, $0 < p \le \infty$, then put

$$\|f\|_{p,w} = \|f\|_{\ell^p_w(X,Z)} = \Big(\sum_{x \in X} \|f(x)\|_Z^p\, w(x)^p\Big)^{1/p} \tag{15.3.4}$$

when $p < \infty$, and

$$\|f\|_{\infty,w} = \|f\|_{\ell^\infty_w(X,Z)} = \sup_{x \in X} (\|f(x)\|_Z\, w(x)) \tag{15.3.5}$$

when $p = \infty$. Of course, $\ell^p_w(X, Z)$ is the same as $\ell^p(X, Z)$, as in Section 2.6, for every $p > 0$ when $w \equiv 1$ on X, in which case $\|f\|_{p,w}$ is the same as $\|f\|_p = \|f\|_{\ell^p(X,Z)}$. Otherwise,

$$f \in \ell^p_w(X, Z) \text{ if and only if } w\, f \in \ell^p(X, Z), \tag{15.3.6}$$

so that

$$f \mapsto w\, f \tag{15.3.7}$$

defines a one-to-one mapping from $\ell^p_w(X, Z)$ onto $\ell^p(X, Z)$. In particular, $\ell^p_w(X, Z)$ is a linear subspace of the space of all Z-valued functions on X, because of the analogous statement for $\ell^p(X, Z)$.

Equivalently,

$$\|f\|_{p,w} = \|w\, f\|_p \tag{15.3.8}$$

for every $f \in \ell^p_w(X, Z)$ and $p > 0$. If $p \ge 1$, then $\|\cdot\|_{p,w}$ defines a norm on $\ell^p_w(X, Z)$, because of the analogous statement for $\ell^p(X, Z)$. If $0 < p \le 1$, then (15.3.4) satisfies the usual homogeneity property of a norm, and

$$\|f + g\|_{p,w}^p \le \|f\|_{p,w}^p + \|g\|_{p,w}^p \tag{15.3.9}$$

for all $f, g \in \ell^p_w(X, Z)$, because of the analogous statement for $\ell^p(X, Z)$. It follows that

$$\|f - g\|^p_{p,w} \tag{15.3.10}$$

defines a metric on $\ell^p_w(X, Z)$ when $p \leq 1$, as before. If Z is complete with respect to the metric associated to $\| \cdot \|_Z$, then

$$\ell^p_w(X, Z) \text{ is complete} \tag{15.3.11}$$

with respect to the appropriate metric for each $p > 0$, because of the analogous statement for $\ell^p(X, Z)$.

Remember that $c_{00}(X, Z)$ is the space of Z-valued functions on X with finite support, as in Section 2.3. Let

$$c_{0,w}(X, Z) \tag{15.3.12}$$

be the space of all Z-valued functions on X such that (15.3.2) vanishes at infinity on X. This is the same as $c_0(X, Z)$ when $w \equiv 1$ on X, and otherwise

$$f \in c_{0,w}(X, Z) \text{ if and only if } w\, f \in c_0(X, Z). \tag{15.3.13}$$

Thus (15.3.7) is a one-to-one mapping from $c_{0,w}(X, Z)$ onto $c_0(X, Z)$, and

$$c_{0,w}(X, Z) \text{ is a linear subspace of } \ell^\infty_w(X, Z), \tag{15.3.14}$$

because of the analogous statement for $c_0(X, Z)$. In fact,

$$c_{0,w}(X, Z) \text{ is the same as the closure of } c_{00}(X, Z) \text{ in } \ell^\infty_w(X, Z) \tag{15.3.15}$$

with respect to the metric associated to (15.3.5), because of the analogous statement for $c_0(X, Z)$.

If $0 < p_1 \leq p_2 \leq +\infty$, then

$$\ell^{p_1}_w(X, Z) \subseteq \ell^{p_2}_w(X, Z), \tag{15.3.16}$$

and

$$\|f\|_{p_2,w} \leq \|f\|_{p_1,w} \tag{15.3.17}$$

for all $f \in \ell^{p_1}_w(X, Z)$, because of the analogous staements for $\ell^p(X, Z)$. If $0 < p < \infty$, then

$$c_{00}(X, Z) \subseteq \ell^p_w(X, Z) \subseteq c_{0,w}(X, Z), \tag{15.3.18}$$

where the second inclusion follows from the analogous statement for $\ell^p(X, Z)$, as usual. We also have that

$$c_{00}(X, Z) \text{ is dense in } \ell^p_w(X, Z) \tag{15.3.19}$$

with respect to the appropriate metric when $p < \infty$, because of the analogous statement for $\ell^p(X, Z)$.

15.4 Polynomials with vector coefficients

Let V be a vector space over the real or complex numbers, and let T be an indeterminate. A *formal polynomial in T with coefficients in V* may be expressed as

$$f(T) = \sum_{j=0}^{n} f_j\, T^j, \tag{15.4.1}$$

with $f_j \in V$ for each j. The coefficients f_j should be considered as being defined for all $j \geq 0$, with $f_j = 0$ when $j > n$. The space of all of these formal polynomials may be denoted $V[T]$. This may be defined more precisely as the space

$$c_{00}(\mathbf{Z}_+ \cup \{0\}, V) \tag{15.4.2}$$

of all V-valued functions on the set $\mathbf{Z}_+ \cup \{0\}$ of nonnegative integers with finite support.

Note that $V[T]$ is a vector space over the real or complex numbers, as appropriate, with respect to termwise addition and scalar multiplication of formal polynomials. Of course, this corresponds to pointwise addition and scalar multiplication on (15.4.2). If $f(T) \in V[T]$ is as in (15.4.1) and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then

$$f(t) = \sum_{j=0}^{n} f_j\, t^j \tag{15.4.3}$$

defines an element of V. Clearly

$$f(T) \mapsto f(t) \tag{15.4.4}$$

is a linear mapping from $V[T]$ into V.

It is sometimes convenient to identify $v \in V$ with the formal polynomial in T for which the coefficient of T^j is equal to v when $j = 0$, and to 0 otherwise. Thus V corresponds to a linear subspace of $V[T]$ in this way.

If $\mathcal{A}$ is an algebra in the strict sense over the real or complex numbers, then $\mathcal{A}[T]$ may be defined initially as a vector space over the real or complex numbers, as appropriate, as before. Multiplication on $\mathcal{A}$ can be used to define a bilinear operation of multiplication on $\mathcal{A}[T]$, with

$$(a\, T^j)\,(b\, T^l) = (a\, b)\, T^{j+l} \tag{15.4.5}$$

for all $a, b \in \mathcal{A}$ and $j, l \geq 0$. This makes $\mathcal{A}[T]$ an algebra in the strict sense as well. If $\mathcal{A}$ is commutative or associative, then one can check that $\mathcal{A}[T]$ has the same property. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then (15.4.4) defines an algebra homomorphism from $\mathcal{A}[T]$ into $\mathcal{A}$.

Let us identify $a \in \mathcal{A}$ with the formal polynomial in T whose constant term is equal to a, and for which the coefficient of T^j is 0 when $j \geq 1$, as before. It is easy to see that $\mathcal{A}$ corresponds to a subalgebra of $\mathcal{A}[T]$ in this way. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then the corresponding formal polynomial is the multiplicative identity element in $\mathcal{A}[T]$.

Suppose for the moment that $\mathcal{A} = \mathbf{R}$ or $\mathbf{C}$, considered as an algebra over itself. In this case, $\mathcal{A}[T]$ is the same as $\mathbf{R}[T]$ or $\mathbf{C}[T]$, as in Section 13.3, with $n = 1$. We may identify $\mathbf{R}$ and $\mathbf{C}$ with subalgebras of $\mathbf{R}[T]$ and $\mathbf{C}[T]$, respectively, as in the preceding paragraph.

One can also consider formal polynomials in n commuting indeterminates $T_1, \dots, T_n$ with coefficients in a vector space V or an algebra $\mathcal{A}$ for any positive integer n, as in Section 13.3 for polynomials with real or complex coefficients. If $n \geq 2$, then a formal polynomial in $T_1, \dots, T_n$ corresponds to a formal polynomial in T_n whose coefficients are formal polynomials in $T_1, \dots, T_{n-1}$.

Let V be a vector space over the real or complex numbers again. If $a(T)$ is an element of $\mathbf{R}[T]$ or $\mathbf{C}[T]$, as appropriate, and $f(T) \in V[T]$, then one can define

$$a(T)\, f(T) \in V[T] \tag{15.4.6}$$

in an obvious way.

Let $f(T) \in V[T]$ be as in (15.4.1). We may say that $f(T)$ has *degree* less than or equal to n in this case, or equal to n if $f_n \neq 0$. If $n \geq 1$ and $t_0 \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then

$$f(T) = f(t_0) + (T - t_0)\, g(T) \tag{15.4.7}$$

for some $g(T) \in V[T]$ of degree less than or equal to $n - 1$. In particular, if $f(t_0) = 0$, then

$$f(T) = (T - t_0)\, g(T). \tag{15.4.8}$$

It follows that $f(t)$ has at most n zeros in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, unless $f(T) = 0$.

15.5 Formal power series

Let V be a vector space over the real or complex numbers, and let T be an indeterminate again. A *formal power series in T with coefficients in V* may be expressed as

$$f(T) = \sum_{j=0}^{\infty} f_j\, T^j, \tag{15.5.1}$$

with $f_j \in V$ for each j. The space of all of these formal power series may be denoted $V[[T]]$, which may be defined more precisely as the space of all V-valued functions on $\mathbf{Z}_+ \cup \{0\}$. This is a vector space over the real or complex numbers, as appropriate, with respect to termwise addition and scalar multiplication, which corresponds to pointwise addition and scalar multiplication of V-valued functions on $\mathbf{Z}_+ \cup \{0\}$. The space $V[T]$ of formal polynomials in T with coefficients in V may be considered as a linear subspace of $V[[T]]$.

If $\mathcal{A}$ is an algebra in the strict sense over the real or complex numbers, then $\mathcal{A}[[T]]$ may be defined initially as a vector space over the real or complex numbers, as appropriate, as in the preceding paragraph. Multiplication in $\mathcal{A}[[T]]$

may be defined using Cauchy products, as in Section 9.10, as follows. Let $f(T), g(T) \in \mathcal{A}[[T]]$ be given, with $f(T)$ as in (15.5.1), and

$$g(T) = \sum_{l=1}^{\infty} g_l \, T^l. \tag{15.5.2}$$

Put

$$h_n = \sum_{j=0}^{n} f_j \, g_{n-j} \tag{15.5.3}$$

for each $n \geq 0$, and

$$h(T) = \sum_{n=0}^{\infty} h_n \, T^n. \tag{15.5.4}$$

The product of $f(T)$ and $g(T)$ in $\mathcal{A}[[T]]$ is defined by

$$f(T) \, g(T) = h(T). \tag{15.5.5}$$

It is easy to see that this definition of multiplication is bilinear, so that $\mathcal{A}[[T]]$ is an algebra in the strict sense. One can check that this definition of multiplication agrees with multiplication on $\mathcal{A}[T]$, so that $\mathcal{A}[T]$ is a subalgebra of $\mathcal{A}[[T]]$. In particular, we may consider $\mathcal{A}$ as a subalgebra of $\mathcal{A}[[T]]$. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then the corresponding element of $\mathcal{A}[T]$ is the multiplicative identity element in $\mathcal{A}[[T]]$. Note that

$$f(T) \mapsto f_0 \tag{15.5.6}$$

defines an algebra homomorphism from $\mathcal{A}[[T]]$ onto $\mathcal{A}$.

Suppose that $f(T), g(T) \in \mathcal{A}[[T]]$ are as in (15.5.1) and (15.5.2) again, with

$$f_j = 0 \text{ when } j \leq j_0 \tag{15.5.7}$$

and

$$g_l = 0 \text{ when } l \leq l_0 \tag{15.5.8}$$

for some $j_0, l_0 \geq 0$. If h_n is as in (15.5.3), then

$$h_n = 0 \text{ when } n \leq j_0 + l_0. \tag{15.5.9}$$

If $\mathcal{A}$ is commutative or associative, then one can check that $\mathcal{A}[[T]]$ has the same property. One way to do this is to reduce to the case of formal polynomials using the previous remark. In particular, $\mathbf{R}[[T]]$ and $\mathbf{C}[[T]]$ are commutative associative algebras over the real and complex numbers, respectively.

Let V be a vector space over the real or complex numbers again, and let $f(T) \in \mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate, be given, as in (15.5.1). If $g(T) \in V[[T]]$ is as in (15.5.2), then (15.5.3) defines an element of V for each $n \geq 0$, so that (15.5.4) defines an element of $V[[T]]$. Thus one can define $f(T) \, g(T) \in V[[T]]$ as in (15.5.5). If $f(T) \in \mathbf{R}[T]$ or $\mathbf{C}[T]$, as appropriate, and $g(T) \in V[T]$, then $f(T) \, g(T) \in V[T]$, as in the previous section. We also have that (15.5.7) and (15.5.8) imply (15.5.9), as before.

15.6 Invertibility in $\mathcal{A}[[T]]$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a nonzero multiplicative identity element $e_{\mathcal{A}}$, and let T be an indeterminate. Also let $a(T) \in \mathcal{A}[[T]]$ be given, so that $a(T)\,T$ is a formal power series in T with constant term equal to 0. If n is a nonnegative integer, then consider

$$\sum_{j=0}^{n} (a(T)\,T)^j = \sum_{j=0}^{n} a(T)^j\,T^j \tag{15.6.1}$$

where

$$(a(T)\,T)^j = a(T)^j\,T^j \tag{15.6.2}$$

is interpreted as being $e_{\mathcal{A}}$ when $j = 0$. If l is a nonnegative integer, then the coefficient of T^l in (15.6.2) is equal to 0 when $j > l$. It follows that the coefficient of T^l in (15.6.1) does not depend on n when $n \geq l$.

We would like to define

$$\sum_{j=0}^{\infty} (a(T)\,T)^j = \sum_{j=0}^{\infty} a(T)^j\,T^j \tag{15.6.3}$$

as an element of $\mathcal{A}[[T]]$. Namely, for each $l \geq 0$, the coefficient of T^l in (15.6.3) is defined to be the same as in (15.6.1) when $n \geq l$.

Note that

$$\begin{aligned} (e_{\mathcal{A}} - a(T)\,T) \sum_{j=0}^{n} (a(T)\,T)^j &= \Big(\sum_{j=0}^{n} (a(T)\,T)^j\Big)\,(e_{\mathcal{A}} - a(T)\,T) \\ &= e_{\mathcal{A}} - (a(T)\,T)^{n+1} \end{aligned} \tag{15.6.4}$$

for every $n \geq 0$, as in Section 6.5. This implies that

$$(e_{\mathcal{A}} - a(T)\,T) \sum_{j=0}^{\infty} (a(T)\,T)^j = \Big(\sum_{j=0}^{\infty} (a(T)\,T)^j\Big)\,(e_{\mathcal{A}} - a(T)\,T) = e_{\mathcal{A}}. \tag{15.6.5}$$

More precisely, if $n \geq r \geq 0$, then the coefficient of T^r in each of these three expressions is the same as for the corresponding expression in (15.6.4). In fact, the coefficient of T^r in the first two expressions in (15.6.5) only involves the coefficients of T^l in (15.6.3) for $0 \leq l \leq r$, which are the same as for (15.6.1), because $r \leq n$. It follows that

$$e_{\mathcal{A}} - a(T)\,T \text{ is invertible in } \mathcal{A}[[T]], \tag{15.6.6}$$

with inverse equal to (15.6.3).

Suppose that $f(T) \in \mathcal{A}[[T]]$ is as in (15.5.1), with f_0 an invertible element of $\mathcal{A}$. In this case, $f(T)$ can be expressed as

$$f(T) = f_0\,(e_{\mathcal{A}} - a(T)\,T) \tag{15.6.7}$$

for some $a(T) \in \mathcal{A}[[T]]$. This implies that $f(T)$ is invertible in $\mathcal{A}[[T]]$, with

$$f(T)^{-1} = (e_{\mathcal{A}} - a(T)\,T)^{-1}\, f_0^{-1}. \tag{15.6.8}$$

Conversely, if $f(T)$ is invertible in $\mathcal{A}[[T]]$, then one can check that f_0 is invertible in $\mathcal{A}$, using the homomorphism (15.5.6).

15.7 Some spaces of power series

Let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$, and let r be a positive real number. Put

$$w_r(j) = r^j \tag{15.7.1}$$

for each nonnegative integer j, which is interpreted as being equal to 1 when $j = 0$, as usual. This is a positive real-valued function on the set $\mathbf{Z}_+ \cup \{0\}$ of nonnegative integers, which can be used to define the spaces $\ell^p_{w_r}(\mathbf{Z}_+ \cup \{0\}, V)$ for $0 < p \le \infty$ as in Section 15.3, as well as the spaces $c_{0,w_r}(\mathbf{Z}_+ \cup \{0\}, V)$.

Suppose that

$$0 < r < t < +\infty, \tag{15.7.2}$$

so that

$$w_r \le w_t \tag{15.7.3}$$

on $\mathbf{Z}_+ \cup \{0\}$. This implies that

$$\ell^p_{w_t}(\mathbf{Z}_+ \cup \{0\}, V) \subseteq \ell^p_{w_r}(\mathbf{Z}_+ \cup \{0\}, V), \tag{15.7.4}$$

with

$$\|a\|_{p,w_r} \le \|a\|_{p,w_t} \tag{15.7.5}$$

for every $a \in \ell^p_{w_t}(\mathbf{Z}_+ \cup \{0\}, V)$. Similarly,

$$c_{0,w_t}(\mathbf{Z}_+ \cup \{0\}, V) \subseteq c_{0,w_r}(\mathbf{Z}_+ \cup \{0\}, V). \tag{15.7.6}$$

If $p < \infty$, then

$$\ell^\infty_{w_t}(\mathbf{Z}_+ \cup \{0\}, V) \subseteq \ell^p_{w_r}(\mathbf{Z}_+ \cup \{0\}, V). \tag{15.7.7}$$

More precisely, if $a \in \ell^\infty_{w_t}(\mathbf{Z}_+ \cup \{0\}, V)$, then

$$\begin{aligned} \|a\|_{p,w_r} &= \Big(\sum_{j=0}^\infty \|a(j)\|_V^p\, w_r(j)^p\Big)^{1/p} \le \|a\|_{\infty,w_t} \Big(\sum_{j=0}^\infty w_r(j)^p\, w_t(j)^{-p}\Big)^{1/p} \\ &= \|a\|_{\infty,w_t} \Big(\sum_{j=0}^\infty (r/t)^{j\,p}\Big)^{1/p} = (1 - (r/t)^p)^{-1/p}\, \|a\|_{\infty,w_t}. \end{aligned} \tag{15.7.8}$$

Let T be an indeterminate, so that $V[[T]]$ is the space of formal power series $f(T) = \sum_{j=0}^\infty f_j\, T^j$ in T with coefficients in V, as in Section 15.5. In this case, we may also use f to refer to the V-valued function on $\mathbf{Z}_+ \cup \{0\}$ defined by the coefficients f_j of $f(T)$. If $0 < p \le \infty$, then let $V_r^p[[T]]$ be the space of $f(T) \in V[[T]]$ such that

$$f \in \ell^p_{w_r}(\mathbf{Z}_+ \cup \{0\}, V), \tag{15.7.9}$$

and put

$$\|f(T)\|_{p,r} = \|f(T)\|_{V_r^p[[T]]} = \|f\|_{p,w_r}. \tag{15.7.10}$$

Similarly, let $V_{0,r}[[T]]$ be the space of $f(T) \in V[[T]]$ such that

$$f \in c_{0,w_r}(\mathbf{Z}_+ \cup \{0\}, V). \tag{15.7.11}$$

If $p < \infty$, then

$$V[T] \subseteq V_r^p[[T]] \subseteq V_{0,r}[[T]], \tag{15.7.12}$$

as in Section 15.3. In fact, $V[T]$ is dense in $V_r^p[[T]]$ when $p < \infty$, and $V[T]$ is dense in $V_{0,r}[[T]]$, with respect to the appropriate metrics, as before. If $0 < p_1 \le p_2 \le \infty$, then

$$V_r^{p_1}[[T]] \subseteq V_r^{p_2}[[T]], \tag{15.7.13}$$

as before. If $0 < r_1 < r_2 < \infty$, then

$$V_{r_2}^p[[T]] \subseteq V_{r_1}^p[[T]] \tag{15.7.14}$$

and

$$V_{0,r_2}[[T]] \subseteq V_{0,r_1}[[T]], \tag{15.7.15}$$

as in (15.7.4) and (15.7.6), respectively. In this case, if $p < \infty$, then

$$V_{r_2}^\infty[[T]] \subseteq V_{r_1}^p[[T]], \tag{15.7.16}$$

as in (15.7.7).

If $0 < \rho \le \infty$, then put

$$V_\rho[[T]] = \bigcap_{0<r<\rho} V_r^p[[T]]. \tag{15.7.17}$$

It is easy to see that the right side does not depend on $p > 0$, because of (15.7.13) and (15.7.16).

Of course, $\mathbf{R}$ and $\mathbf{C}$ may be considered as one-dimensional vector spaces over themselves, with their standard absolute value functions as norms. If $V = \mathbf{R}$ or $\mathbf{C}$, then we may use $(\mathbf{R})_r^p[[T]]$ or $(\mathbf{C})_r^p[[T]]$, as appropriate, for $V_r^1[[T]]$. Similarly, we may use $(\mathbf{R})_{0,r}[[T]]$ or $(\mathbf{C})_{0,r}[[T]]$, as appropriate, for $V_{0,r}[[T]]$. We may also use $(\mathbf{R})_\rho[[T]]$ or $(\mathbf{C})_\rho[[T]]$, as appropriate, for $V_\rho[[T]]$ in this case.

15.8 Absolute convergence of power series

Let $(V, \|\cdot\|_V)$ be a Banach space over the real or complex numbers, let T be an indeterminate, and let r be a positive real number. Also let $f(T) = \sum_{j=0}^\infty f_j\, T^j \in V_r^1[[T]]$ be given, and let t be a real or complex number, as appropriate, with $|t| \le r$. Under these conditions,

$$f(t) = \sum_{j=0}^\infty f_j\, t^j \tag{15.8.1}$$

defines an element of V, because the sum on the right converges absolutely, as in Section 1.7. More precisely,

$$\|f(t)\|_V \le \sum_{j=0}^\infty \|f_j\|_V\, |t|^j \le \sum_{j=0}^\infty \|f_j\|_V\, r^j = \|f(T)\|_{1,r}. \tag{15.8.2}$$

Thus (15.8.1) defines a V-valued function on

$$\{t \in \mathbf{R} : |t| \le r\} = [-r, r] \tag{15.8.3}$$

or

$$\{t \in \mathbf{C} : |t| \le r\}, \tag{15.8.4}$$

as appropriate.

The partial sums

$$\sum_{j=0}^n f_j\, t^j \tag{15.8.5}$$

converge to (15.8.1) uniformly on (15.8.3) or (15.8.4), as appropriate, with respect to the metric on V associated to $\|\cdot\|_V$, as in Weierstrass' well-known criterion for uniform convergece. One can check that the partial sums (15.8.5) are continuous as V-valued functions on (15.8.3) or (15.8.4), as appropriate. It follows that (15.8.1) is continuous as a V-valued function on (15.8.3) or (15.8.4), as appropriate. This defines a bounded linear mapping from $V_r^1[[T]]$ into the space of continuous V-valued functions on (15.8.3) or (15.8.4), as appropriate, with respect to the corresponding supremum norm, because of (15.8.2). One may also consider the right side of (15.8.1) as an absolutely convergent sum in this space of continuous V-valued functions with respect to the supremum norm.

Suppose now that $0 < \rho \le \infty$, and that $f(T) \in V_\rho[[T]]$. If $0 < r < \rho$, then $f(T) \in V_r^1[[T]]$, as in (15.7.17). If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| < \rho$, then (15.8.1) defines an element of V, as before. This defines a V-valued function on

$$\{t \in \mathbf{R} : |t| < \rho\} = (-\rho, \rho) \tag{15.8.6}$$

or

$$\{t \in \mathbf{C} : |t| < \rho\}, \tag{15.8.7}$$

as appropriate. This function is continuous with respect to the metric on V associated to $\|\cdot\|_V$. This can be obtained from the continuity of the restriction of the function to (15.8.3) or (15.8.4), as appropriate, when $0 < r < \rho$.

15.9 Products and absolute convergence

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$, let T be an indeterminate, and let r be a positive real number. Suppose that $f(T) = \sum_{j=0}^{\infty} f_j\, T^j$ and $g(T) = \sum_{l=0}^{\infty} g_l\, T^l$ are elements of $\mathcal{A}^1_r[[T]]$. Let $h(T) = f(T)\, g(T)$ be their product in $\mathcal{A}[[T]]$, as in Section 15.5, with $h(T) = \sum_{n=0}^{\infty} h_n\, T^n$. Observe that

$$\|h_n\|_{\mathcal{A}} \le \sum_{j=0}^{n} \|f_j\|_{\mathcal{A}}\, \|g_{n-j}\|_{\mathcal{A}} \tag{15.9.1}$$

for each $n \ge 0$, so that

$$\|h_n\|_{\mathcal{A}}\, r^n \le \sum_{j=0}^{n} (\|f_j\|_{\mathcal{A}}\, r^j)\, (\|g_{n-j}\|_{\mathcal{A}}\, r^{n-j}). \tag{15.9.2}$$

The right side of (15.9.2) is the same as the nth term of the Cauchy product of the sums corresponding to $\|f(T)\|_{1,r}$ and $\|g(T)\|_{1,r}$, as in Section 9.10.

It follows that

$$h(T) \in \mathcal{A}^1_r[[T]] \tag{15.9.3}$$

too, with

$$\|h(T)\|_{1,r} \le \|f(T)\|_{1,r}\, \|g(T)\|_{1,r}. \tag{15.9.4}$$

Suppose that $\mathcal{A}$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, and let t be a real or complex number, as appropriate, with $|t| \le r$. The series corresponding to $h(t)$ as in the previous section is the same as the Cauchy product of the series corresponding to $f(t)$ and $g(t)$, as in Section 9.11. This implies that

$$h(t) = f(t)\, g(t), \tag{15.9.5}$$

as before. If $0 < \rho \le \infty$ and $f(T), g(T) \in \mathcal{A}_\rho[[T]]$, then we get that

$$h(T) \in \mathcal{A}_\rho[[T]], \tag{15.9.6}$$

and that (15.9.5) holds for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|t| < \rho$.

In particular,

$$\mathcal{A}^1_r[[T]] \text{ and } \mathcal{A}_\rho[[T]] \text{ are subalgebras of } \mathcal{A}[[T]]. \tag{15.9.7}$$

If $\mathcal{A}$ is a Banach algebra, then $\mathcal{A}^1_r[[T]]$ is a Banach algebra with respect to $\|\cdot\|_{1,r}$.

Now let V be a vector space over the real or complex numbers with a norm $\|\cdot\|_V$, and let $g(T)$ be an element of $V^1_r[[T]]$. Also let $f(T)$ be an element of $(\mathbf{R})^1_r[[T]]$ or $(\mathbf{C})^1_r[[T]]$, as appropriate. Thus $h(T) = f(T)\, g(T)$ may be defined as an element of $V[[T]]$, as in Section 15.5. As before, we have that

$$\|h_n\|_V \le \sum_{j=0}^{n} |f_j|\, \|g_{n-j}\|_V \tag{15.9.8}$$

for each $n \geq 0$, so that

$$\|h_n\|_V \, r^n \leq \sum_{j=0}^{n} (|f_j| \, r^j) \, (\|g_{n-j}\|_V \, r^{n-j}). \tag{15.9.9}$$

The right side of (15.9.9) is the same as the nth term of the Cauchy product of the sums corresponding to $\|f(T)\|_{1,r}$ and $\|g(T)\|_{1,r}$.

This implies that

$$h(T) \in V_r^1[[T]], \tag{15.9.10}$$

and that (15.9.4) holds, as before. Suppose that V is a Banach space, and that $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, satisfies $|t| \leq r$. The series corresponding to $h(t)$ as in the previous section may be considered as the Cauchy product of the series corresponding to $f(t)$ and $g(t)$, as before. It follows that (15.9.5) holds in this case as well. Similarly, if $0 < \rho \leq \infty$, $f(T) \in (\mathbf{R})_\rho[[T]]$ or $(\mathbf{C})_\rho[[T]]$, as appropriate, and $g(T) \in V_\rho[[T]]$, then

$$h(T) \in V_\rho[[T]], \tag{15.9.11}$$

and (15.9.5) holds for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|t| < \rho$.

15.10 Differentiating power series

Let V be a vector space over the real or complex numbers, and let T be an indeterminate. If $f(T) = \sum_{j=0}^{\infty} f_j \, T^j \in V[[T]]$, then the *derivative* $f'(T)$ of $f(T)$ in $V[[T]]$ is defined by

$$f'(T) = \sum_{j=1}^{\infty} j \, f_j \, T^{j-1} = \sum_{j=0}^{\infty} (j+1) \, f_{j+1} \, T^j. \tag{15.10.1}$$

Note that

$$f(T) \mapsto f'(T) \tag{15.10.2}$$

is a linear mapping from $V[[T]]$ into itself that sends $V[T]$ into itself.

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers. One can check that

$$(f(T) \, g(T))' = f'(T) \, g(T) + f(T) \, g'(T) \tag{15.10.3}$$

for all $f(T), g(T) \in \mathcal{A}[[T]]$. Similarly, this holds when $g(T) \in V[[T]]$ and $f(T)$ is an element of $\mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate.

Suppose now that $(V, \| \cdot \|_V)$ is a Banach space, and let r be a positive real number. Suppose also that $f(T) \in V[[T]]$ satisfies

$$f'(T) \in V_r^1[[T]], \tag{15.10.4}$$

and note that

$$\|f'(T)\|_{1,r} = \sum_{j=1}^{\infty} j \, \|f_j\|_V \, r^{j-1}. \tag{15.10.5}$$

It is easy to see that

$$f(T) \in V_r^1[[T]] \tag{15.10.6}$$

in this case. Let t_1, t_2 be real or complex numbers, as appropriate, such that

$$|t_1|,\ |t_2| \le r. \tag{15.10.7}$$

Thus $f(t_1), f(t_2) \in V$ may be defined as in Section 15.8, and we have that

$$\begin{aligned} \|f(t_1) - f(t_2)\|_V &= \Big\|\sum_{j=0}^\infty f_j\, t_1^j - \sum_{j=0}^\infty f_j\, t_2^j\Big\|_V = \Big\|\sum_{j=1}^\infty f_j\, (t_1^j - t_2^j)\Big\|_V \\ &\le \sum_{j=1}^\infty \|f_j\|_V\, |t_1^j - t_2^j| \le \sum_{j=1}^\infty j\, \|f_j\|_V\, r^{j-1}\, |t_1 - t_2|, \end{aligned} \tag{15.10.8}$$

where the last step is as in Section 10.1.

This implies that

$$\|f(t_1) - f(t_2)\|_V \le \|f'(T)\|_{1,r}\, |t_1 - t_2| \tag{15.10.9}$$

when t_1, t_2 satisfy (15.10.7). This is a Lipschitz condition for $f(t)$ on the set of $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|t| \le r$. If $f(T) \in V[T]$, then completeness of V is not needed here.

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le r$, then

$$f'(t) = \sum_{j=1}^\infty j\, f_j\, t^{j-1} = \sum_{j=0}^\infty (j+1)\, f_{j+1}\, t^j \tag{15.10.10}$$

defines an element of V, as in Sections 1.7 and 15.8. Of course, this corresponds to differentiating the power series for $f(t)$ termwise. One can use the same type of Lipschitz conditions as before to show that the derivative of the sum is equal to the sum of the derivatives, as mentioned in Section 10.1 for power series with real or complex coefficients.

In the complex case, (15.10.10) should be interpreted as a suitable complex derivative, as before. In particular, $f(t)$ is holomorphic as a V-valued function in a suitable sense.

If $0 < \rho \le \infty$ and $f(T) \in V_\rho[[T]]$, then

$$f'(T) \in V_\rho[[T]] \tag{15.10.11}$$

too, by standard arguments. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| < \rho$, then $f(t)$ and $f'(t)$ are defined as elements of V, and $f'(t)$ corresponds to the derivative of $f(t)$, as before. In the complex case, $f(t)$ is holomorphic as a V-valued function on

$$\{t \in \mathbf{C} : |t| < \rho\}. \tag{15.10.12}$$

15.11 Coefficients in R, C

Let T be an indeterminate, and let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. If $f(T)$ is a formal polynomial in T with real or complex coefficients, as appropriate, and $x \in \mathcal{A}$, then we may define $f(x)$ as an element of $\mathcal{A}$, as in Section 8.13. We also have that

$$f(T) \mapsto f(x) \tag{15.11.1}$$

defines an algebra homomorphism from $\mathbf{R}[T]$ or $\mathbf{C}[T]$, as appropriate, into $\mathcal{A}$, as before. This corresponds to a remark in Section 13.4 with $n = 1$ as well.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra, $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and that $f(T)$ is a formal power series in T with real or complex coefficients, as appropriate. Let r be a positive real number, and suppose that

$$f(T) \in (\mathbf{R})^1_r[[T]] \text{ or } (\mathbf{C})^1_r[[T]], \tag{15.11.2}$$

as appropriate. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \le r$, then $f(x)$ may be defined as an element of $\mathcal{A}$ as in Section 9.14, and we have that

$$\|f(x)\|_{\mathcal{A}} \le \|f(T)\|_{1,r}, \tag{15.11.3}$$

as before. We also have that (15.11.1) defines an algebra homomorphism from $(\mathbf{R})^1_r[[T]]$ or $(\mathbf{C})^1_r[[T]]$, as appropriate, into $\mathcal{A}$, as in Section 9.13.

If

$$f'(T) \in (\mathbf{R})^1_r[[T]] \text{ or } (\mathbf{C})^1_r[[T]], \tag{15.11.4}$$

as appropriate, then (15.11.2) holds, as in the previous section. If $x, y \in \mathcal{A}$ and $\|x\|_{\mathcal{A}}, \|y\|_{\mathcal{A}} \le r$, then we get that

$$\|f(x) - f(y)\|_{\mathcal{A}} \le \|f'(T)\|_{1,r}\, \|x - y\|_{\mathcal{A}}, \tag{15.11.5}$$

as in Section 10.1.

Suppose that $0 < \rho \le \infty$, and that

$$f(T) \in (\mathbf{R})_\rho[[T]] \text{ or } (\mathbf{C})_\rho[[T]], \tag{15.11.6}$$

as appropriate. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} < \rho$, then $f(x)$ may be defined as an element of $\mathcal{A}$ as in Section 9.14 again. Of course, we can take $r > 0$ such that

$$\|x\|_{\mathcal{A}} \le r < \rho, \tag{15.11.7}$$

so that (15.11.2) holds. Under these conditions, (15.11.1) defines an algebra homomorphism from $(\mathbf{R})_\rho[[T]]$ or $(\mathbf{C})_\rho[[T]]$, as appropriate, into $\mathcal{A}$, as before.

Let $\mathcal{A}$ be any associative algebra over $\mathbf{R}$ or $\mathbf{C}$ with a multiplicative identity element $e_{\mathcal{A}}$ again, and suppose that $f(T) = \sum_{j=0}^\infty f_j\, T^j \in \mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate. If $x \in \mathcal{A}$ is nilpotent, then

$$f(x) = \sum_{j=0}^\infty f_j\, x^j \tag{15.11.8}$$

reduces to a finite sum in $\mathcal{A}$, and thus defines an element of $\mathcal{A}$. One can check that (15.11.1) defines an algebra homomorphism from $\mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate, into $\mathcal{A}$, as usual.

Remember that $\mathcal{A}[[T]]$ is an associative algebra over the real or complex numbers, as appropriate, with a multiplicative identity element, as in Section 15.5. If $f(T) \in \mathbf{R}[T]$ or $\mathbf{C}[T]$, as appropriate, and $a(T) \in \mathcal{A}[[T]]$, then

$$f(a(T)) \tag{15.11.9}$$

may be defined as an element of $\mathcal{A}[[T]]$, as in Section 8.13. The mapping

$$f(T) \mapsto f(a(T)) \tag{15.11.10}$$

is an algebra homomorphism from $\mathbf{R}[T]$ or $\mathbf{C}[T]$, as appropriate, into $\mathcal{A}[[T]]$, as before. If $a(T) \in \mathcal{A}[T]$, then (15.11.9) is an element of $\mathcal{A}[T]$ too.

Suppose that the constant term in $a(T) \in \mathcal{A}[[T]]$ is equal to 0, and that $f(T) \in \mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate. Note that the coefficient of T^l in $a(T)^j$ is equal to 0 when $j > l$. This implies that the coefficient of T^l in

$$\sum_{j=0}^{n} f_j \, a(T)^j \tag{15.11.11}$$

does not depend on n when $n \geq l$. Let us define

$$f(a(T)) = \sum_{j=0}^{\infty} f_j \, a(T)^j \tag{15.11.12}$$

to be the formal power series in T with coefficients in $\mathcal{A}$ such that the coefficient of T^l is equal to the coefficient of T^l in (15.11.11) when $n \geq l$. One can check that (15.11.10) is an algebra homomorphism from $\mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate, into $\mathcal{A}[[T]]$ under these conditions.

15.12 Some formal compositions

Let T be an indeterminate, and let $f(T)$ and $g(T)$ be formal polynomials in T with real or complex coefficients. Thus $f(g(T))$ may be defined as a formal polynomial with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in the previous section. Let us put

$$(f \circ g)(T) = f(g(T)), \tag{15.12.1}$$

which corresponds formally to composing $g(T)$ with $f(T)$.

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{A}}$. If $x \in \mathcal{A}$, then $g(x)$ may be defined as an element of $\mathcal{A}$ as in Section 8.13. Similarly, $f(g(x))$ and $(f \circ g)(x)$ may be defined as elements of $\mathcal{A}$. One can check that

$$(f \circ g)(x) = f(g(x)) \tag{15.12.2}$$

under these conditions. In particular, one can take $\mathcal{A} = \mathbf{R}[T]$ or $\mathbf{C}[T]$, as appropriate, to get that formal compositions of formal polynomials in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$ is associative.

Let $a(T)$ be a formal power series in T with coefficients in $\mathcal{A}$, so that $g(a(T))$ may be defined as an element of $\mathcal{A}[[T]]$, as in the previous section. As before, we may put

$$(g \circ a)(T) = g(a(T)), \tag{15.12.3}$$

because it also corresponds to formally composing $a(T)$ with $g(T)$. Similarly,

$$((f \circ g) \circ a)(T) = (f \circ g)(a(T)) \tag{15.12.4}$$

and

$$(f \circ (g \circ a))(T) = f((g \circ a)(T)) \tag{15.12.5}$$

are defined as elements of $\mathcal{A}[[T]]$. Note that

$$(f \circ g)(a(T)) = f(g(a(T))), \tag{15.12.6}$$

as in (15.12.2). This implies that

$$((f \circ g) \circ a)(T) = (f \circ (g \circ a))(T), \tag{15.12.7}$$

because of (15.12.3).

Suppose for the moment that $a(T) \in \mathcal{A}[T]$, so that (15.12.3) is an element of $\mathcal{A}[T]$ too, as before. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $a(t)$ and $(g \circ a)(t)$ may be defined as elements of $\mathcal{A}$ as in Section 15.4. Similarly, $g(a(t))$ is defined as an element of $\mathcal{A}$, because $a(t) \in \mathcal{A}$, as in Section 8.13. One can verify that

$$(g \circ a)(t) = g(a(t)). \tag{15.12.8}$$

If $a(T) \in \mathcal{A}[[T]]$, then $a(0)$ and $(g \circ a)(0)$ may be interpreted as the constant terms of these power series, and one can check that (15.12.8) holds, with $t = 0$.

Suppose now that $g(T)$ is a formal power series in T with real or complex coefficients, so that (15.12.1) is a formal power series in T with real or complex coefficients too, as appropriate. If $x \in \mathcal{A}$ is nilpotent, then $g(x)$ and $(f \circ g)(x)$ may be defined as elements of $\mathcal{A}$ as in the previous section, and $f(g(x))$ may be defined as an element of $\mathcal{A}$ as well. One can check that (15.12.2) also holds in this case.

15.13 Some more formal compositions

Let T be an indeterminate, and let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{E}}$. If g is a formal power series in T with real or complex coefficients, and $x \in \mathcal{A}$ is nilpotent, then $g(x)$ may be defined as an element of $\mathcal{A}$, as in Section 15.11. Suppose that

$$\text{the constant term in } g(T) \text{ is equal to } 0. \tag{15.13.1}$$

In this case, one can check that

(15.13.2) $$g(x) \text{ is nilpotent.}$$

Let $f(T)$ be another formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, and remember that $f(g(T))$ may be defined as an element of $\mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$ under these conditions, as appropriate, as in Section 15.11. This corresponds to formally composing $g(T)$ with $f(T)$, as before, and thus may be expressed as in (15.12.1) again. Note that $(f \circ g)(x)$ and $f(g(x))$ are defined as elements of $\mathcal{A}$, because x and $g(x)$ are nilpotent. One can verify that (15.12.2) holds in this case too.

Let $a(T)$ be a formal power series in T with coefficients in $\mathcal{A}$, and suppose that

(15.13.3) $$\text{the constant term in } a(T) \text{ is equal to } 0.$$

If $g(T)$ is a formal power series in $\mathbf{R}$ or $\mathbf{R}$, as appropriate, then $g(a(T))$ may be defined as a formal power series in T with coefficients in $\mathcal{A}$ as in Section 15.11. This corresponds to formally composing $a(T)$ and $g(T)$, as usual, and may be expressed as in (15.12.3).

If $f(T)$ is a formal polynomial in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, then $f(g(T))$ is a formal power series with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, which may be expressed as in (15.12.1), as before. Similarly, (15.12.4) and (15.12.5) may be defined as formal power series in T with coefficients in $\mathcal{A}$ under these conditions. One can check that (15.12.6) holds in this case too, using some of the remarks in Section 15.11. It follows that (15.12.7) holds, as before.

Suppose now that the constant term in $g(T)$ is equal to 0 too. Using this, it is easy to see that

(15.13.4) $$\text{the constant term in } g(a(T)) \text{ is equal to } 0,$$

because of (15.13.3). If $f(T)$ is a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, then $f(g(T))$ may be defined as a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 15.11, and which may be expressed as in (15.12.1). Similarly, (15.12.4) and (15.12.5) may be defined as formal power series in T with coefficients in $\mathcal{A}$, as in Section 15.11. In order to verify (15.12.6), one can approximate $f(T)$ by formal polynomials in T, for which the analogous statement was mentioned in the preceding paragraph. This implies that (15.12.7) holds under these conditions as well.

15.14 Polynomials of power series

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let T be an indeterminate. If $f(T) \in \mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate, and $a(T) \in \mathcal{A}[[T]]$, then we may be able to define

$f(a(T))$ as an element of $\mathcal{A}[[T]]$ in a reasonable way under suitable conditions. In this case, we may put

$$(f \circ a)(T) = f(a(T)), \tag{15.14.1}$$

because it corresponds formally to composing $a(T)$ and $f(T)$, as before. In particular, we would like to have

$$(f \circ a)(t) = f(a(t)) \tag{15.14.2}$$

for suitable $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. Similarly, if the coefficients of $a(T)$ are in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, then we would like to have

$$(f \circ a)(x) = f(a(x)) \tag{15.14.3}$$

for suitable $x \in \mathcal{A}$.

Of course, we have seen some versions of this in the previous two sections. In this section and the next one, we would like to consider some more versions of this, involving convergence of power series.

Suppose for the rest of this section that $f(T)$ is a formal polynomial in T with real or complex coefficients, as appropriate. Thus $f(a(T))$ is defined as an element of $\mathcal{A}[[T]]$ for each $a(T) \in \mathcal{A}[[T]]$, as in Section 15.11. Let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$, let r_a be a positive real number, and suppose that

$$a(T) \in \mathcal{A}^1_{r_a}[[T]]. \tag{15.14.4}$$

This implies that

$$f(a(T)) \in \mathcal{A}^1_{r_a}[[T]], \tag{15.14.5}$$

because $\mathcal{A}^1_{r_a}[[T]]$ is a subalgebra of $\mathcal{A}[[T]]$, as in Section 15.9.

Suppose that $\mathcal{A}$ is a Banach algebra with respect to $\|\cdot\|_{\mathcal{A}}$, and that t is a real or complex number, as appropriate, that satisfies

$$|t| \le r_a. \tag{15.14.6}$$

This means that $a(t)$ and $(f \circ a)(t)$ may be defined as elements of $\mathcal{A}$, as in Section 15.8, so that $f(a(t))$ may be defined as an element of $\mathcal{A}$ as well. One can check that (15.14.2) holds, by reducing to the case where $f(T) = T^j$ for some j, as usual.

Similarly, let $0 < \rho_a \le \infty$ be given, and suppose that

$$a(T) \in \mathcal{A}_{\rho_a}[[T]]. \tag{15.14.7}$$

This implies that

$$f(a(T)) \in \mathcal{A}_{\rho_a}[[T]], \tag{15.14.8}$$

because $\mathcal{A}_{\rho_a}[[T]]$ is a subalgebra of $\mathcal{A}[[T]]$, as before. If $\mathcal{A}$ is a Banach algebra and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, satisfies

$$|t| < \rho_a, \tag{15.14.9}$$

then $a(t)$, $(f \circ a)(t)$, and $f(a(t))$ may be defined as elements of $\mathcal{A}$. We also have that (15.14.2) holds in this case, by the same type of argument as before, or by reducing to the previous case.

Suppose now that $a(T)$ is a formal power series in T with real or complex coefficients, as appropriate, so that $f(a(T))$ is a formal power series in T with real or complex coefficients, as appropriate, too. Let r_a be a positive real number again, and suppose for the moment that

$$a(T) \in (\mathbf{R})^1_{r_a}[[T]] \text{ or } (\mathbf{C})^1_{r_a}[[T]], \tag{15.14.10}$$

as appropriate, so that

$$(f \circ a)(T) = f(a(T)) \in (\mathbf{R})^1_{r_a}[[T]] \text{ or } (\mathbf{C})^1_{r_a}[[T]], \tag{15.14.11}$$

as appropriate. If $\mathcal{A}$ is a Banach algebra, and $x \in \mathcal{A}$ satisfies

$$\|x\|_{\mathcal{A}} \le r_a, \tag{15.14.12}$$

then $a(x)$ and $(f \circ a)(x)$ may be defined as elements of $\mathcal{A}$ as in Section 9.14, so that $f(a(x))$ is defined as an element of $\mathcal{A}$ too. One can verify that (15.14.3) holds under these conditions, by reducing to the case where $f(T) = T^j$ for some j, as before.

Similarly, let $0 < \rho_a \le \infty$ be given again, and suppose that

$$a(T) \in (\mathbf{R})_\rho[[T]] \text{ or } (\mathbf{C})_\rho[[T]], \tag{15.14.13}$$

as appropriate, so that

$$(f \circ a)(T) = f(a(T)) \in (\mathbf{R})_\rho[[T]] \text{ or } (\mathbf{C})_\rho[[T]], \tag{15.14.14}$$

as appropriate. If $\mathcal{A}$ is a Banach algebra and $x \in \mathcal{A}$ satisfies

$$\|x\|_{\mathcal{A}} < \rho_a, \tag{15.14.15}$$

then $a(x)$, $(f \circ a)(x)$, and $f(a(x))$ may be defined as elements of $\mathcal{A}$. We also get that (15.14.3) holds, by reducing to the previous case.

15.15 Some compositions of power series

Let $(\mathcal{A}, \| \cdot \|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and let r_a, r_f be positive real numbers. Also let

$$a(T) \in \mathcal{A}^1_{r_a}[[T]] \tag{15.15.1}$$

and

$$f(T) \in (\mathbf{R})^1_{r_f}[[T]] \text{ or } (\mathbf{C})^1_{r_f}[[T]], \tag{15.15.2}$$

as appropriate, be given. Remember that $\mathcal{A}^1_{r_a}[[T]]$ is a Banach algebra with respect to the appropriate norm, as in Section 15.9.

If

$$\|a(T)\|_{1,r_a} \le r_f, \tag{15.15.3}$$

then we may define

$$f(a(T)) \in \mathcal{A}^1_{r_a}[[T]], \tag{15.15.4}$$

as in Section 9.14. The mapping

$$f(T) \mapsto f(a(T)) \tag{15.15.5}$$

is an algebra homomorphism from $(\mathbf{R})^1_{r_f}[[T]]$ or $(\mathbf{C})^1_{r_f}[[T]]$, as appropriate, into $\mathcal{A}^1_{r_a}[[T]]$, as in Sections 9.13 and 15.11. Of course, this definition of $f(a(T))$ is the same as the one in the previous section when $f(T)$ is a formal polynomial in T. We may express $f(a(T))$ as $(f \circ a)(T)$, as in (15.14.1).

We also have that

$$\|f(a(T))\|_{1,r_a} \le \|f(T)\|_{1,r_f}, \tag{15.15.6}$$

as in Sections 9.14 and 15.11. More precisely, the left side is the norm of $f(a(T))$ in $\mathcal{A}^1_{r_a}[[T]]$, and the right side refers to the norm of $f(T)$ in $(\mathbf{R})^1_{r_f}[[T]]$ or $(\mathbf{C})^1_{r_f}[[T]]$, as appropriate. If $f(T) = \sum_{j=0}^\infty f_j\, T^j$, as usual, then

$$\lim_{n\to\infty} \sum_{j=0}^n f_j\, (a(T))^j = f(a(T)), \tag{15.15.7}$$

with respect to the metric on $\mathcal{A}^1_{r_a}[[T]]$ associated to its norm, by construction.

If t is a real or complex number, as appropriate, with $|t| \le r_a$, then $a(t)$ is defined as an element of $\mathcal{A}$ as in Section 15.8, with

$$\|a(t)\|_{\mathcal{A}} \le \|a(T)\|_{1,r_a}. \tag{15.15.8}$$

This implies that

$$\|a(t)\|_{\mathcal{A}} \le r_f, \tag{15.15.9}$$

because of (15.15.3), so that $f(a(t))$ may be defined as an element of $\mathcal{A}$ as in Section 9.14. Similarly, $(f \circ a)(t)$ is defined as an element of $\mathcal{A}$ as in Section 15.8. One can check that (15.14.2) holds, by approximating $f(T)$ by formal polynomials, for which the analogous statement was mentioned in the previous section.

Suppose for the moment that constant term in $a(T)$ is equal to 0, so that $f(a(T))$ may be defined as a formal power series in T with coefficients in $\mathcal{A}$ as in Section 15.11. One can check that this is equivalent to the definition in (15.15.4), as a formal power series in T, using (15.15.7).

Let r_a be a positive real number again, and suppose now that

$$a(T) \in (\mathbf{R})^1_{r_a}[[T]] \text{ or } (\mathbf{C})^1_{r_a}[[T]], \tag{15.15.10}$$

as appropriate. If $a(T)$ satisfies (15.15.3), then we may define

$$f(a(T)) \in (\mathbf{R})^1_{r_a}[[T]] \text{ or } (\mathbf{C})^1_{r_a}[[T]], \tag{15.15.11}$$

as appropriate, as in Section 9.14 again. As before, (15.15.5) is an algebra homomorphism from $(\mathbf{R})^1_{r_f}[[T]]$ or $(\mathbf{C})^1_{r_f}[[T]]$ into $(\mathbf{R})^1_{r_a}[[T]]$ or $(\mathbf{C})^1_{r_a}[[T]]$, as appropriate. If $f(T)$ is a formal polynomial in T, then this definition of $f(a(T))$ is the same as the one in the previous section, as before.

We also get that (15.15.6) holds, where now the left side is the norm of $f(a(T))$ in $(\mathbf{R})^1_{r_a}[[T]]$ or $(\mathbf{C})^1_{r_a}[[T]]$, as appropriate. This implies $f(a(T))$ may be given as in (15.15.7), as before.

If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \le r_a$, then $a(x)$ may be defined as an element of $\mathcal{A}$ as in Section 9.14, with

$$\|a(x)\|_{\mathcal{A}} \le \|a(T)\|_{1,r_a}. \tag{15.15.12}$$

This means that

$$\|a(x)\|_{\mathcal{A}} \le r_f, \tag{15.15.13}$$

by (15.15.3), so that $f(a(x))$ may be defined as an element of $\mathcal{A}$ in the same way. We may define $(f \circ a)(x)$ as an element of $\mathcal{A}$ too, where $(f \circ a)(T) = f(a(T))$, as before. One can use (15.15.7) to get that (15.14.3) holds, because of the analogous statement when $f(T)$ is a formal polynomial, as in the previous section. In particular, if $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le r_a$, then (15.14.2) holds, as before.

Part IV

Algebras, norms, and operators, 3

Chapter 16

Algebras, norms, and power series

16.1 More on differentiation

Let $\mathcal{A}$ be a commutative associative algebra over the real or complex numbers, and let T be an indeterminate. Remember that the space $\mathcal{A}[[T]]$ of formal power series in T with coefficients in $\mathcal{A}$ is a commutative associative algebra over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as well, as in Section 15.5. If $a(T) \in \mathcal{A}[[T]]$ and j is a positive integer, then we can define $a(T)^j$ as formal power series in T with coefficients in $\mathcal{A}$ in the usual way. The derivative of this power series is given by

$$(a(T)^j)' = j\, a(T)^{j-1}\, a'(T), \tag{16.1.1}$$

because of the product rule, as in Section 15.10. Of course, the factor of $a(T)^{j-1}$ on the right is not needed when $j = 1$.

Suppose now that $\mathcal{A}$ also has a multiplicative identity element $e_{\mathcal{A}}$, which may be considered as the multiplicative identity element in $\mathcal{A}[[T]]$ too, as in Section 15.5. If f is a formal polynomial in T with real or complex coefficients, as appropriate, then $f(a(T))$ may be defined as an element of $\mathcal{A}[[T]]$, as in Sections 8.13 and 15.11. Note that the derivative $f'(T)$ of $f(T)$ is a formal polynomial in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, so that

$$f'(a(T)) \tag{16.1.2}$$

may be defined as an element of $\mathcal{A}[[T]]$, as before. It is easy to see that

$$f(a(T))' = f'(a(T))\, a'(T), \tag{16.1.3}$$

as in the *chain rule*, using (16.1.1).

Suppose that the constant term in $a(T)$ is equal to 0, and that $f(T)$ is a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Remember that $f(a(T))$ may be defined as a formal power series in T with coefficients in $\mathcal{A}$, as in Section 15.11. Similarly, (16.1.2) may be defined as an element of $\mathcal{A}[[T]]$. One can check that (16.1.3) also holds in this case, by approximating $f(T)$ by formal polynomials in T.

16.2 Some norms of derivatives

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$, let T be an indeterminate, and let r be a positive real number. Suppose that $a(T)$, $b(T)$ are elements of the space $\mathcal{A}_1^r[[T]]$ defined in Section 15.7. This implies that

$$a(T)\,b(T) \in \mathcal{A}_r^1[[T]], \tag{16.2.1}$$

with

$$\|a(T)\,b(T)\|_{1,r} \le \|a(T)\|_{1,r}\,\|b(T)\|_{1,r}, \tag{16.2.2}$$

as in Section 15.9. Remember that $\|\cdot\|_{1,r} = \|\cdot\|_{\mathcal{A}_r^1[[T]]}$ is as in Section 15.7.

We also have that

$$(a(T)\,b(T))' = a'(T)\,b(T) + a(T)\,b'(T), \tag{16.2.3}$$

as in Section 15.10. If

$$a'(T), b'(T) \in \mathcal{A}_r^1[[T]], \tag{16.2.4}$$

then we get that

$$(a(T)\,b(T))' \in \mathcal{A}_r^1[[T]], \tag{16.2.5}$$

with

$$\|(a(T)\,b(T))'\|_{1,r} \le \|a'(T)\|_{1,r}\,\|b(T)\|_{1,r} + \|a(T)\|_{1,r}\,\|b'(T)\|_{1,r}. \tag{16.2.6}$$

Remember that $a(T) \in \mathcal{A}_r^1[[T]]$ when $a'(T) \in \mathcal{A}_r^1[[T]]$, as mentioned in Section 15.10. Thus

$$\{a(T) \in \mathcal{A}[[T]] : a'(T) \in \mathcal{A}_r^1[[T]]\} \tag{16.2.7}$$

is a subalgebra of $\mathcal{A}_r^1[[T]]$.

Suppose now that $\mathcal{A}$ is associative too, and let j be a positive integer. If $a'(T) \in \mathcal{A}_r^1[[T]]$, then

$$(a(T)^j)' \in \mathcal{A}_r^1[[T]], \tag{16.2.8}$$

as in the preceding paragraph. More precisely, one can check that

$$\|(a(T)^j)'\|_{1,r} \le j\,\|a(T)\|_{1,r}^{j-1}\,\|a'(T)\|_{1,r}, \tag{16.2.9}$$

using induction. This can be obtained from (16.1.1) when $\mathcal{A}$ is commutative as well.

Suppose that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, and let $f(T)$ be a formal polynomial with real or complex coefficients. Observe that

$$f(a(T))' \in \mathcal{A}_r^1[[T]]. \tag{16.2.10}$$

If

$$\|a(T)\|_{1.r} \le r_0 \tag{16.2.11}$$

for some positive real number r_0, then one can check that

$$\|f(a(T))'\|_{1,r} \le \|f'(T)\|_{1,r_0} \, \|a'(T)\|_{1,r}, \tag{16.2.12}$$

where $\|f'(T)\|_{1,r_0}$ is as in Section 15.7. If $\mathcal{A}$ is commutative, then this can be obtained from (16.1.3) and the fact that

$$\|f'(a(T))\|_{1,r} \le \|f'(T)\|_{1,r_0}. \tag{16.2.13}$$

16.3 More on compositions

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, let T be an indeterminate, and let r_a, r_f be positive real numbers. Also let $a(T)$ and $f(T)$ be formal power series in T with coefficients in $\mathcal{A}$ and $\mathbf{R}$ or $\mathbf{C}$, as appropriate, respectively. Suppose that

$$a'(T) \in \mathcal{A}^1_{r_a}[[T]], \tag{16.3.1}$$

where $a'(T)$ is as in Section 15.10 and $\mathcal{A}^1_{r_a}[[T]]$ is as in Section 15.7. This implies that $a(T) \in \mathcal{A}^1_{r_a}[[T]]$, as mentioned in Section 15.10, and we ask that

$$\|a(T)\|_{1,r_a} \le r_f, \tag{16.3.2}$$

where $\|a(T)\|_{1,r_a} = \|a(T)\|_{\mathcal{A}^1_{r_a}[[T]]}$ is as in Section 15.7.

Similarly, suppose that

$$f'(T) \in (\mathbf{R})^1_{r_f}[[T]] \text{ or } (\mathbf{C})^1_{r_f}[[T]], \tag{16.3.3}$$

as appropriate, which implies that $f(T)$ is an element of $(\mathbf{R})^1_{r_f}[[T]]$ or $(\mathbf{C})^1_{r_f}[[T]]$, as appropriate, as before. This permits us to define $f(a(T))$ as an element of $\mathcal{A}^1_{r_a}[[T]]$, as in Section 15.15. We can define

$$f'(a(T)) \in \mathcal{A}^1_{r_a}[[T]] \tag{16.3.4}$$

in the same way. Note that

$$\|f'(a(T))\|_{1,r_a} \le \|f'(T)\|_{1,r_f}, \tag{16.3.5}$$

as before.

Under these conditions, we would like to check that

$$f(a(T))' \in \mathcal{A}^1_{r_a}[[T]], \tag{16.3.6}$$

with

$$\|f(a(T))'\|_{1,r_a} \le \|f'(T)\|_{1,r_f} \, \|a'(T)\|_{1,r_a}. \tag{16.3.7}$$

This follows from the remarks in the previous section when $f(T)$ is a formal polynomial in T. Otherwise, we can approximate $f(T)$ by its partial sums, which are formal polynomials in T, as follows.

Remember that if $f(T) = \sum_{j=0}^{\infty} f_j\, T^j$, as usual, then

$$\sum_{j=0}^{n} f_j\, a(T)^j \tag{16.3.8}$$

converges to $f(a(T))$ as $n \to \infty$, with respect to the metric on $\mathcal{A}^1_{r_a}[[T]]$ associated to $\| \cdot \|_{1,r_a}$, as in Section 15.15. In particular, this implies that for each nonnegative integer l, the coefficient of T^l in (16.3.8) converges to the coefficient of T^l in $f(a(T))$ as $n \to \infty$, as a sequence of real or complex numbers, as appropriate. This means that the coefficients of $f(a(T))'$ are the same as the limits of the corresponding coefficients of the formal derivative of (16.3.8) as $n \to \infty$.

The analogue of (16.3.7) for formal polynomials implies that

$$\Big\| \Big(\sum_{j=0}^{n} f_j\, a(T)^j \Big)' \Big\|_{1,r_a} \le \Big\| \Big(\sum_{j=0}^{n} f_j\, T^j \Big)' \Big\|_{1,r_f} \|a'(T)\|_{1,r_a} \tag{16.3.9}$$

for each $n \ge 0$. It is easy to see that

$$\Big\| \Big(\sum_{j=0}^{n} f_j\, T^j \Big)' \Big\|_{1,r_f} \le \|f'(T)\|_{1,r_j} \tag{16.3.10}$$

for each n, because of the way that the norm is defined. It follows that

$$\Big\| \Big(\sum_{j=0}^{n} f_j\, a(T)^j \Big)' \Big\|_{1,r_a} \le \|f'(T)\|_{1,r_j}\, \|a'(T)\|_{1,r_a} \tag{16.3.11}$$

for each n. One can use this to get the desired properties of $f(a(T))'$, because of the remarks in the preceding paragraph.

In fact, the formal derivative of (16.3.8) converges to $f(a(T))'$ with respect to the metric on $\mathcal{A}^1_{r_a}[[T]]$ associated to $\| \cdot \|_{1,r_a}$. This can be obtained from (16.3.7), with $f(T)$ replaced with

$$f(T) - \sum_{j=0}^{n} f_j\, T^j = \sum_{j=n+1}^{\infty} f_j\, T^j. \tag{16.3.12}$$

If $\mathcal{A}$ is commutative, then

$$f(a(T))' = f'(a(T))\, a'(T). \tag{16.3.13}$$

This was mentioned in Section 16.1 when $f(T)$ is a formal polynomial in T, and when the constant term in $a(T)$ is equal to 0. Here one can reduce to the case of formal polynomials by approximating $f(T)$ by its partial sums again.

16.4 Some differential equations

Let T be an indeterminate, and let $a(T)$, $b(T)$ be formal power series in T with coefficients in the real or complex numbers. Consider the ordinary differential equation

$$f'(T) = a(T)\, f(T) + b(T), \tag{16.4.1}$$

where $f(T)$ is another formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$. It is easy to see that there is a unique solution to this equation with a prescribed constant term. Indeed, (16.4.1) determines the coefficients of $f(T)$ after the constant term in terms of the previous coefficients of $f(T)$ and the coefficients of $a(T)$ and $b(T)$.

The *exponential function* can be defined as a formal power series in T by

$$\exp(T) = \sum_{j=0}^{\infty} (1/j!)\, T^j, \tag{16.4.2}$$

as in Section 10.4. Of course,

$$\exp'(T) = \exp(T), \tag{16.4.3}$$

as usual. This and the condition that the constant term in (16.4.2) be equal to 1 determines $\exp(T)$ uniquely, as in the preceding paragraph.

Similarly, consider

$$\log(1+T) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}\, T^j, \tag{16.4.4}$$

which is the formal power series corresponding to the *natural logarithm* of $1+T$. Observe that

$$\log'(1+T) = \sum_{j=1}^{\infty} (-1)^{j+1}\, T^{j-1} = \sum_{j=0}^{\infty} (-1)^j\, T^j = (1+T)^{-1}, \tag{16.4.5}$$

where the third step is as in Section 15.6. As before, (16.4.4) is uniquely determined by this and the condition that the constant term be equal to 0.

Because the constant term in (16.4.4) is equal to 0,

$$f(T) = \exp(\log(1+T)) \tag{16.4.6}$$

may be defined as a formal power series in T as in Section 15.11. Of course, we should have that

$$f(T) = 1 + T. \tag{16.4.7}$$

This could be obtained from the usual properties of the exponential and logarithm functions on the real line and the set $\mathbf{R}_+$ of positive real numbers, respectively.

Alternatively,

$$\begin{aligned} f'(T) &= \exp'(\log(1+T))\,\log'(1+T) \\ &= \exp(\log(1+T))\,(1+T)^{-1} = f(T)\,(1+T)^{-1}, \end{aligned} \tag{16.4.8}$$

where the first step is as in Section 16.1. The right side of (16.4.7) satisfies the same differential equation. This implies (16.4.7), because both sides have the same constant term.

16.5 More on $\log(1+T)$

Let T be an indeterminate, and let r be a positive real number strictly less than 1. It is easy to see that

$$\log(1+T) \in (\mathbf{R})^1_r[[T]], \tag{16.5.1}$$

in the notation of Section 15.7. More precisely,

$$\|\log(1+T)\|_{1,r} = \|\log(1+T)\|_{(\mathbf{R})^1_r[[T]]} = \sum_{j=1}^{\infty} \frac{r^j}{j} = -\log(1-r), \tag{16.5.2}$$

using the standard power series representation for $\log(1-t)$ when $t \in \mathbf{R}$ satisfies $|t| < 1$ in the last step.

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \le r$, then we put

$$\log(e_{\mathcal{A}} + x) = \log_{\mathcal{A}}(e_{\mathcal{A}} + x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}\, x^j, \tag{16.5.3}$$

as in Sections 9.14 and 15.11. This is an element of $\mathcal{A}$ with

$$\|\log(e_{\mathcal{A}} + x)\|_{\mathcal{A}} \le \|\log(1+T)\|_{1,r} = -\log(1-r), \tag{16.5.4}$$

as before.

Note that

$$\exp(T) \in (\mathbf{R})^1_{r_1}[[T]] \tag{16.5.5}$$

for every positive real number r_1, with

$$\|\exp(T)\|_{1,r_1} = \sum_{j=0}^{\infty} \frac{(r_1)^j}{j!} = \exp(r_1). \tag{16.5.6}$$

Remember that the exponential function on $\mathcal{A}$ was discussed in Section 10.4.

In fact, we have that

$$\exp(\log(e_{\mathcal{A}} + x)) = e_{\mathcal{A}} + x \tag{16.5.7}$$

under these conditions. This uses (16.4.7) and the remarks in Section 15.15. This is related to part (b) of Theorem 10.30 on p246 of [162].

Similarly, let $\mathcal{A}$ be any associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and suppose that $x \in \mathcal{A}$ is nilpotent. This implies that $\log(e_{\mathcal{A}} + x)$ may be defined as an element of $\mathcal{A}$ as in (16.5.3), as in Section 15.11. In fact,

$$\log(e_{\mathcal{A}} + x) \text{ is nilpotent in } \mathcal{A}, \tag{16.5.8}$$

as in Section 15.13. It follows that

$$\exp(\log(e_{\mathcal{A}} + x)) \tag{16.5.9}$$

may be defined as an element of $\mathcal{A}$ too, as in Sections 11.1 and 15.11. We also have that (16.5.7) holds, as in Section 15.13 again.

16.6 Some more logarithms

Let T be an indeterminate, and consider the formal power series

$$\exp(-T) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} T^j. \tag{16.6.1}$$

This is the same as formally composing $-T$ with $\exp(T)$. One can check that

$$\exp(-T)\,\exp(T) = \exp(T)\,\exp(-T) = 1, \tag{16.6.2}$$

using the binomial theorem, as in Section 10.4. This means that $\exp(-T)$ is the multiplicative inverse of $\exp(T)$ in the algebra of formal power series in T with coefficients in $\mathbf{R}$, as in Section 15.6.

Let $a(T)$ be a formal power series in T with real or complex coefficients, and suppose that the constant term in $a(T)$ is equal to 0. This implies that

$$f(T) = \log(1 + a(T)) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} a(T)^j \tag{16.6.3}$$

may be defined as a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 15.11. Note that the constant term in (16.6.3) is equal to 0 too. We also have that

$$f'(T) = a'(T)\,(1 + a(T))^{-1}, \tag{16.6.4}$$

as in Section 16.1. In fact, (16.6.3) is uniquely determined by these two properties.

In particular, we can take

$$a(T) = \exp(T) - 1 = \sum_{l=1}^{\infty} (1/l!)\,T^l. \tag{16.6.5}$$

Note that

$$a'(T) = \exp'(T) = \exp(T). \tag{16.6.6}$$

Thus

$$f'(T) = \exp(T)\,\exp(T)^{-1} = 1, \tag{16.6.7}$$

as in (16.6.4). It follows that

$$f(T) = T, \tag{16.6.8}$$

because the constant term in $f(T)$ is equal to 0 as before. This means that

$$\log(\exp(T)) = \log(1 + (\exp(T) - 1)) = T, \tag{16.6.9}$$

where the first step may be considered as the definition of the left side.

If r_1 is a positive real number, then $\exp(T) - 1 \in (\mathbf{R})^1_{r_1}[[T]]$, as in (16.5.5). More precisely,

$$\|\exp(T) - 1\|_{1,r_1} = \sum_{l=1}^{\infty} (1/l!)\, r_1^l = \exp(r_1) - 1. \tag{16.6.10}$$

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $x \in \mathcal{A}$, then

$$\|exp(x) - e_{\mathcal{A}}\|_{\mathcal{A}} \le \exp(\|x\|_{\mathcal{A}}) - 1, \tag{16.6.11}$$

as in Sections 9.14 and 15.11. If

$$\exp(\|x\|_{\mathcal{A}}) - 1 < 1, \tag{16.6.12}$$

then we would like to check that

$$\log(\exp(x)) = x. \tag{16.6.13}$$

More precisely, the left side is interpreted as being

$$\log(e_{\mathcal{A}} + (\exp(x) - e_{\mathcal{A}})), \tag{16.6.14}$$

as defined as in the previous section.

Put

$$r = \exp(\|x\|_{\mathcal{A}}) - 1, \tag{16.6.15}$$

so that $r < 1$, by hypothesis. Remember that $\log(1 + T) \in \mathbf{R})^1_r[[T]]$, as in (16.5.1). One can get (16.6.14) using (16.6.9) and the remarks in Section 15.15. This also uses (16.6.10), with $r_1 = \|x\|_{\mathcal{A}}$.

Now let $\mathcal{A}$ be any associative algebra over the real or complex numbers witha multiplicative identity element $e_{\mathcal{A}}$, and suppose that $x \in \mathcal{A}$ is nilpotent. This means that $\exp(x)$ can be defined as an element of $\mathcal{A}$ as in Sections 11.1 and 15.11, and that

$$\exp(x) - e_{\mathcal{A}} \text{ is nilpotent in } \mathcal{A}, \tag{16.6.16}$$

as in Section 15.13. If we interpret $\log(\exp(x))$ as in (16.6.14), then this can be defined as an element of $\mathcal{A}$ as in Section 15.11. Under these conditions, (16.6.13) follows from (16.6.9), as in Section 15.13.

16.7 A basic type of composition

Let T be an indeterminate, and let $a(T) = \sum_{l=0}^{\infty} a_l \, T^l$ be a formal power series in T with coefficients in the real or complex numbers. Also let $\mathcal{A}$ be an associative algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{A}}$. If $x \in \mathcal{A}$, then

$$a_x(T) = \sum_{l=0}^{\infty} a_l \, x^l \, T^l \tag{16.7.1}$$

is a formal power series in T with coefficients in $\mathcal{A}$. This may also be expressed as $a(x\,T)$, which may be considered as the formal power series in T with coefficients in $\mathcal{A}$ obtained by composing $x\,T$ with $a(T)$ as in Section 15.11. Note that

$$a(T) \mapsto a_x(T) = a(x\,T) \tag{16.7.2}$$

is an algebra homomorphism from $\mathbf{R}[[T]]$ or $\mathbf{C}[[T]]$, as appropriate, into $\mathcal{A}[[T]]$, as before.

Suppose for the moment that $a(T)$ is a formal polynomial in T, so that $a_x(T)$ is a formal polynomial in T as well. In this case, $a_x(t)$ and $a(t\,x)$ are defined as elements of $\mathcal{A}$ for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and it is easy to see that

$$a_x(t) = a(t\,x). \tag{16.7.3}$$

Similarly, if $a(T)$ is a formal power series in T, and $x \in \mathcal{A}$ is nilpotent, then

$$a_x(T) = a(x\,T) \in \mathbf{R}[T] \text{ or } \mathbf{C}[T], \tag{16.7.4}$$

as appropriate. If $t \in \mathbf{R}$ or $\mathbf{C}$, then $t\,x$ is nilpotent in $\mathcal{A}$ too, so that $a(t\,x)$ may be defined as an element of $\mathcal{A}$, as in Section 15.11. Under these conditions, $a_x(t)$ defines an element of $\mathcal{A}$ as well, and (16.7.3) holds.

Now let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{A}}$ with $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Suppose that $x \in \mathcal{A}$ has the property that

$$\sum_{l=0}^{\infty} |a_l| \, \|x^l\|_{\mathcal{A}} \tag{16.7.5}$$

converges as an infinite series of nonnegative real numbers. In this case, we can put

$$a(x) = \sum_{l=0}^{\infty} a_l \, x^l, \tag{16.7.6}$$

where the series on the right converges in $\mathcal{A}$, because it converges absolutely.

Observe that

$$a_x(T) \in \mathcal{A}_1^1[[T]], \tag{16.7.7}$$

in the notation of Section 15.7, with

$$\|a_x(T)\|_{1,1} = \|a_x(T)\|_{\mathcal{A}^1_1[[T]]} = \sum_{l=0}^{\infty} |a_l| \, \|x^l\|_{\mathcal{A}}. \tag{16.7.8}$$

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|t| \leq 1$, then

$$a_x(t) = \sum_{l=0}^{\infty} a_l \, t^l \, x^l \tag{16.7.9}$$

defines an element of $\mathcal{A}$, as in Section 15.8. Of course,

$$a_x(1) = a(x), \tag{16.7.10}$$

by construction.

16.8 Some more compositions

Let us return to the same notation and hypotheses as at the beginning of the previous section. Suppose for the moment that $f(T)$ is a formal polynomial in T with coefficients in the real or complex numbers, as appropriate. Remember that $(f \circ a)(T) = f(a(T))$ may be defined as a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Sections 15.11 and 15.12. Thus

$$(f \circ a)_x(T) = (f \circ a)(x \, T) \tag{16.8.1}$$

may be defined as a formal power series in T with coefficients in $\mathcal{A}$, as in the previous section. We can also define

$$(f \circ a_x)(T) = f(a_x(T)) \tag{16.8.2}$$

as a formal power series in T with coefficients in $\mathcal{A}$, as in Sections 15.11 and 15.12. It is easy to see that (16.8.1) and (16.8.2) are the same. We may express this formal power series as

$$f(a(x \, T)). \tag{16.8.3}$$

The type of formal composition mentioned in the previous section may be considered as a version of one discussed in Section 15.13. This means that the equality of (16.8.1) and (16.8.2) is basically an associativity property of formal compositions like these, which were discussed previously.

Suppose for the moment again that the constant term in $a(T)$ is equal to 0, and let $f(T)$ be a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. In this case, $(f \circ a)(T) = f(a(T))$ may be defined as a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Sections 15.11 and 15.13. This means that (16.8.1) may be defined as a formal power series in T with coefficients in $\mathcal{A}$, as in the previous section again. Note that

$$\text{the constant term in } a_x(T) = a(x \, T) \text{ is equal to } 0, \tag{16.8.4}$$

so that (16.8.2) may be defined as a formal power series in T with coefficients in $\mathcal{A}$ as in Sections 15.11 and 15.13. One can check that (16.8.1) and (16.8.2) are the same under these conditions too. This can be seen by approximating $f(T)$ by formal polynomials in T, as in Section 15.11. This may also be considered as an instance of an associativity property of formal compositions like these, as in Section 15.13. The resulting formal power series may be expressed as in (16.8.3), as before.

Suppose for the moment that $a(T)$ and $f(T)$ are formal polynomials in T, so that

$$(f \circ a)(T),\ a_x(T),\ \text{and } (f \circ a_x)(T) \tag{16.8.5}$$

are formal polynomials in T too. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then

$$(f \circ a_x)(t) = (f \circ a)_x(t) = (f \circ a)(t\,x) = f(a(t\,x)), \tag{16.8.6}$$

where the second step is as in (16.7.3), and the third step is as in Section 15.12. Alternatively,

$$(f \circ a_x)(t) = f(a_x(t)) = f(a(t\,x)), \tag{16.8.7}$$

where the first step is as in Section 15.12, and the second step is as in (16.7.3).

Let $a(T)$ be a formal power series in T again, and suppose that $x \in \mathcal{A}$ is nilpotent, so that $a_x(T)$ is a formal polynomial in T, as in (16.7.4). If $f(T)$ is a formal polynomial in T, then $(f \circ a)(T)$ is a formal power series in T, and

$$(f \circ a_x)(T),\ (f \circ a)_x(T) \text{ are formal polynomials in } T. \tag{16.8.8}$$

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $t\,x$ is nilpotent in $\mathcal{A}$, $a(t\,x)$ is defined as an element of $\mathcal{A}$, and (16.8.6) holds, where the second step is as in the previous section, and the third step is as in Section 15.12. Alternatively, (16.8.7) holds, where the first step is as in Section 15.12, and the second step is as in the previous section.

Suppose that the constant term in $a(T)$ is equal to 0. If $x \in \mathcal{A}$ is nilpotent, then it is easy to see that

$$a_x(T) = a(x\,T) \text{ is nilpotent in } \mathcal{A}[T]. \tag{16.8.9}$$

If $f(T)$ is a formal power series in T again, then it follows that

$$(f \circ a_x)(T) = f(a_x(T)) \in \mathcal{A}[T]. \tag{16.8.10}$$

More precisely, $f(a_x(T))$ may be defined initially as a formal power series in T, because of (16.8.4), as in Section 15.11. It is easy to see that (16.8.10) holds in this case, because of (16.8.9). However, one can also define $f(a_x(T))$ as an element of $\mathcal{A}[T]$ more directly, because of (16.8.9), as in Section 15.11. One can check that these two ways of defining $f(a(T))$ are equivalent.

Note that $(f \circ a)(T)$ is defined as a formal power series in T, and that $(f \circ a)_x(T)$ is a formal polynomial in T, as before. If $t \in \mathbf{R}$ or $\mathbf{C}$, then $t\,x$ is nilpotent in $\mathcal{A}$, and $a(t\,x)$ is nilpotent as well, because the constant term in $a(T)$

is equal to 0, as in Section 15.13. Thus $f(a(t\,x))$ is defined as an element of $\mathcal{A}$, as in Section 15.11. Under these conditions, (16.8.6) holds, where the second step is as in the previous section again, and the third step is as in Section 15.13.

To get the first step in (16.8.7), one can approximate $f(T)$ by a formal polynomial, because of (16.8.9). The analogous statement for formal polynomials was mentioned in Section 15.12, as before. The second step in (16.8.7) is as in the previous section, as before.

16.9 Some compositions with convergence

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and let T be an indeterminate. Also let $a(T) = \sum_{l=0}^{\infty} a_l\, T^l$ be a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, let x be an element of $\mathcal{A}$, and suppose that $\sum_{l=0}^{\infty} |a_l|\, \|x^l\|_{\mathcal{A}}$ converges, as an infinite series of nonnegative real numbers. Thus

$$a_x(T) = \sum_{l=0}^{\infty} a_l\, x^l\, T^l \in \mathcal{A}_1^1[[T]], \tag{16.9.1}$$

in the notation of Section 15.7, as in Section 16.7.

Let $f(T)$ be a formal polynomial in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Observe that

$$f(a_x(T)) \in \mathcal{A}_1^1[[T]], \tag{16.9.2}$$

as in Section 15.14. This may be expressed as $(f \circ a_x)(T)$, as in (16.8.2). If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and

$$|t| \leq 1, \tag{16.9.3}$$

then $a_x(t)$ and $(f \circ a_x)(t)$ may be defined as elements of $\mathcal{A}$, as in Section 15.8. This implies that $f(a_x(t))$ may be defined as an element of $\mathcal{A}$, and

$$(f \circ a_x)(t) = f(a_x(t)), \tag{16.9.4}$$

as in Section 15.14.

Remember that $(f \circ a)(T) = f(a(T))$ is a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, which may be expressed as

$$(f \circ a)(T) = \sum_{m=0}^{\infty} (f \circ a)_m\, T^m. \tag{16.9.5}$$

We also have that $(f \circ a_x)(T) = f(a_x(T))$ is a formal power series in T with coefficients in $\mathcal{A}$, which may be expressed as

$$(f \circ a_x)(T) = (f \circ a)_x(T) = \sum_{m=0}^{\infty} (f \circ a)_m\, x^m\, T^m, \tag{16.9.6}$$

where the first step is as in the previous section. Note that

$$\sum_{m=0}^{\infty} |(f \circ a)_m| \, \|x^m\|_{\mathcal{A}} \tag{16.9.7}$$

converges as an infinite series of nonnegative real numbers, as in (16.9.2). If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, satisfies (16.9.3), then we get that

$$f(a_x(t)) = (f \circ a_x)(t) = \sum_{m=0}^{\infty} (f \circ a)_m \, t^m \, x^m, \tag{16.9.8}$$

using (16.9.4) in the first step. Of course, the right side converges absolutely in this case, because (16.9.7) converges.

16.10 Some more convergence conditions

Let us return to the same notation and hypotheses as at the beginning of the previous section. Now let $f(T) = \sum_{j=0}^{\infty} f_j \, T^j$ be a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with

$$f(T) \in (\mathbf{R})^1_{r_f}[[T]] \text{ or } (\mathbf{C})^1_{r_f}[[T]], \tag{16.10.1}$$

as appropriate, for some positive real number r_f, in the notation of Section 15.7. Suppose that

$$\|a_x(T)\|_{1,1} = \|a_x(T)\|_{\mathcal{A}^1_1[[T]]} = \sum_{l=0}^{\infty} |a_l| \, \|x^l\|_{\mathcal{A}} \le r_f. \tag{16.10.2}$$

Under these conditions, we may define

$$f(a_x(T)) \in \mathcal{A}^1_1[[T]], \tag{16.10.3}$$

as in Section 15.15. We also have that

$$\|f(a_x(T))\|_{1,1} \le \|f(T)\|_{1,r_f}, \tag{16.10.4}$$

as before. Remember that the left side is the norm of $f(a_x(T))$ in $\mathcal{A}^1_1[[T]]$, and that the right side is the norm of $f(T)$ in $(\mathbf{R})^1_{r_f}[[T]]$ or $(\mathbf{C})^1_{r_f}[[T]]$, as appropriate. We may express $f(a_x(T))$ as $(f \circ a_x)(T)$, as before. If $f(T)$ is a formal polynomial in T, then this is equivalent to the definition of $f(a(T))$ as in the previous section, as mentioned in Section 15.15.

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then $a_x(t)$ may be defined as an element of $\mathcal{A}$ as in Section 15.8, with

$$\|a_x(t)\|_{\mathcal{A}} \le \|a_x(T)\|_{1,1}. \tag{16.10.5}$$

It follows that

$$\|a_x(t)\|_{\mathcal{A}} \le r_f, \tag{16.10.6}$$

because of (16.10.2), so that $f(a_x(t))$ may be defined as an element of $\mathcal{A}$ as in Section 9.14. Similarly, $(f \circ a_x)(t)$ may be defined as an element of $\mathcal{A}$ as in Section 15.8. In fact, we have that

$$(f \circ a_x)(t) = f(a_x(t)), \tag{16.10.7}$$

as in Section 15.15.

Suppose that the constant term in $a(T)$ is equal to 0, so that $(f \circ a)(T) = f(a(T))$ may be defined as a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Sections 15.11 and 15.13. This implies that the constant term in $a_x(T)$ is equal to 0, so that $(f \circ a_x)(T) = f(a_x(T))$ may be defined as a formal power series in T with coefficients in $\mathcal{A}$, as in Sections 15.11 and 15.13 again. This is the same as in (16.10.3), as a formal power series in T, as mentioned in Section 15.15. Remember that $(f \circ a)_x(T)$ may be defined as a formal power series in T with coefficients in $\mathcal{A}$ as in Section 16.7. We also have that

$$(f \circ a)_x(T) = (f \circ a_x)(T), \tag{16.10.8}$$

as in Section 16.8.

If $(f \circ a)(T)$ is as in (16.9.5), then $(f \circ a_x)(T)$ is as in (16.9.6), as before. This implies that (16.9.7) converges as an infinite series of nonnegative real numbers, because of (16.10.3). In fact,

$$\sum_{m=0}^{\infty} |(f \circ a)_m| \, \|x^m\|_{\mathcal{A}} = \|(f \circ a_x)(T)\|_{1,1} \le \|f(T)\|_{1,r_j}, \tag{16.10.9}$$

as in (16.10.4). If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then (16.9.8) holds, where the first step is as in (16.10.7).

16.11 Some more exponentials of logarithms

Let T be an indeterminate, and let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$. Put

$$a(T) = \log(1+T) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} T^l, \tag{16.11.1}$$

as in Section 16.4. If $x \in \mathcal{A}$, then

$$a_x(T) = a(x\,T) = \log(1 + x\,T) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} x^l \, T^l \tag{16.11.2}$$

is a formal power series in T with coefficients in $\mathcal{A}$, as in Section 16.7.

Note that the constant term in (16.11.2) is equal to 0, so that

$$\exp(a_x(T)) \tag{16.11.3}$$

may be defined as a formal power series in T with coefficients in $\mathcal{A}$ as in Section 15.11. Remember that

$$\exp(\log(1+T)) = 1 + T, \tag{16.11.4}$$

as in Section 16.4. Using this, we get that

$$\exp(a_x(T)) = 1 + x\,T. \tag{16.11.5}$$

as in Section 16.8.

Suppose now that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and that $x \in \mathcal{A}$ has the property that

$$\sum_{l=1}^{\infty} (1/l)\,\|x^l\|_{\mathcal{A}} \tag{16.11.6}$$

converges as an infinite series of nonnegative real numbers. This means that

$$a_x(T) \in \mathcal{A}_1^1[[T]], \tag{16.11.7}$$

in the notation of Section 15.7, with

$$\|a_x(T)\|_{1,1} = \|a_x(T)\|_{\mathcal{A}_1^1[[T]]} = \sum_{l=1}^{\infty} (1/l)\,\|x^l\|_{\mathcal{A}}. \tag{16.11.8}$$

Remember that

$$f(T) = \exp(T) \in (\mathbf{R})_{r_1}^1[[T]] \tag{16.11.9}$$

for every $r_1 > 0$, as in Section 16.5. Thus

$$(f \circ a_x)(T) = f(a_x(T)) = \exp(a_x(T)) \in \mathcal{A}_1^1[[T]] \tag{16.11.10}$$

may be defined as in Section 15.15. This is the same as (16.11.5), as a formal power series in T with coefficients in $\mathcal{A}$, as before.

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then $a_x(t)$ may be defined as an element of $\mathcal{A}$ as in Section 15.8. This means that

$$f(a_x(t)) = \exp(a_x(t)) \tag{16.11.11}$$

may be defined as an element of $\mathcal{A}$ too, as in Sections 9.14 and 10.4. Under these conditions, we have that

$$\exp(a_x(t)) = f(a_x(t)) = (f \circ a_x)(t) = e_{\mathcal{A}} + t\,x, \tag{16.11.12}$$

as in the previous section. We may consider $a_x(t)$ as a definition of the *logarithm* of $e_{\mathcal{A}} + t\,x$ in $\mathcal{A}$ in this case, as in Section 16.5. This is related to part (b) of Theorem 10.30 on p246 of [162].

16.12 The binomial theorem

Let n be a positive integer, and let

$$\binom{n}{j} = \frac{n!}{j!\,(n-j)!} \tag{16.12.1}$$

be the usual *binomial coefficient* for $j = 0, 1, \ldots, n$. Here "l factorial" $l!$ is defined in the usual way for each positive integer l, and is interpreted as being equal to 1 when $l = 0$, so that $\binom{0}{0}$ is interpreted as being equal to 1 as well. It is well known that

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j\, y^{n-j} \tag{16.12.2}$$

when x and y are commuting elements of an associative algebra $\mathcal{A}$ over the real or complex numbers, or a ring, for that matter, by the *binomial theorem.* If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then both sides of the equation may be interpreted as being equal to $e_{\mathcal{A}}$ when $n = 0$.

It is easy to see that $(x+y)^n$ can be expressed as a sum of 2^n terms, each of which is of the form

$$x^j\, y^{n-j} \tag{16.12.3}$$

for some j. Thus $(x+y)^n$ may be expressed as a sum of terms of this form with positive integer coefficients. The coefficients do not depend on x, y, or $\mathcal{A}$, and are the same as the number of subsets of $\{1, \ldots, n\}$ with j elements for each j.

The expression (16.12.1) for these coefficients can be verified by induction. Alternatively, one can take $x, y \in \mathbf{R}$, and consider the derivatives of $(x+y)^n$ in x or y. More precisely, one can take $y = 1$, and consider the derivatives of $(x+1)^n$ at $x = 0$.

In particular, the binomial coefficients are uniquely determined by the condition that

$$(x+1)^n = \sum_{j=0}^{n} \binom{n}{j} x^j \tag{16.12.4}$$

on $\mathbf{R}$. In fact, $\binom{n}{j}$ is uniquely determined for $j = 0, 1, \ldots, n$ by the condition that (16.12.4) holds for $n+1$ elements x of $\mathbf{R}$, because $(x+1)^n$ is a polynomial of degree n.

Let T be an indeterminate, and for each positive integer l, consider the formal polynomial

$$\binom{T}{l} = \frac{T\,(T-1)\cdots(T-l+1)}{l!} \tag{16.12.5}$$

of degree l in T with rational coefficients, as in Exercise 8 on p 74 of [44]. If $l = 0$, then this is interpreted as the constant polynomial equal to 1. This corresponds to the polynomial function

$$\binom{t}{l} = \frac{t\,(t-1)\cdots(t-l+1)}{l!} \tag{16.12.6}$$

of t in $\mathbf{R}$, $\mathbf{C}$, or any commutative associative algebra over the real or rational numbers with a multiplicative identity element. If $t \geq l$ is an integer, then this is the same as the usual binomial coefficient, as before. If $t < l$ is a nonnegative integer, then this is equal to 0.

If t_1 and t_2 are nonnegative integers, then

$$(x+1)^{t_1+t_2} = (x+1)^{t_1}\,(1+x)^{t_2} \tag{16.12.7}$$

for every $x \in \mathbf{R}$. One can check that

$$\binom{t_1+t_2}{l} = \sum_{j=0}^{l} \binom{t_1}{j} \binom{t_2}{l-j} \tag{16.12.8}$$

for every nonnegative integer l, using the appropriate binomial expansions for both sides of (16.12.7).

If t_2 is fixed, then both sides of (16.12.8) are polynomials in t_1. It follows that (16.12.8) holds for all $t_1 \in \mathbf{R}$, because it holds for all nonnegative integers t_1, as before.

If $t_1 \in \mathbf{R}$ is fixed, then both sides of (16.12.8) are polynomials in t_2 with real coefficients. This implies that (16.12.8) holds for all $t_2 \in \mathbf{R}$, because it holds for all nonnegative integers t_2, as in the preceding paragraph.

Thus (16.12.8) holds for all $t_1, t_2 \in \mathbf{R}$. The same argument could be used to get that (16.12.8) holds for all $t_1, t_2 \in \mathbf{C}$.

Let T_1, T_2 be commuting indeterminates, and observe that

$$\binom{T_1+T_2}{l} \tag{16.12.9}$$

defines a formal polynomial in T_1, T_2 with rational coefficients, as in Section 13.3. In fact,

$$\binom{T_1+T_2}{l} = \sum_{j=0}^{l} \binom{T_1}{j} \binom{T_2}{l-j}, \tag{16.12.10}$$

as formal polynomials in T_1, T_2 for every nonnegative integer l. This means that the coefficients of the monomials in T_1, T_2 on both sides of the equation are the same.

Indeed, the corresponding polynomial functions on $\mathbf{R}^2$ are the same, as in (16.12.8). To get that the coefficients are the same, one can consider the derivatives of these polynomials at $(0,0)$. This corresponds to part (i) of Exercise 8 on p74 of [44]. It follows that (16.12.8) holds for all t_1, t_2 in any commutative associative algebra over the real or rational numbers with a multiplicative identity element.

16.13 Binomial series

Let X be an indeterminate. If α is a real or complex number, then

$$b_\alpha(X) = \sum_{j=0}^{\infty} \binom{\alpha}{j} X^j \tag{16.13.1}$$

is a formal power series in X with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. This is the *binomial series* in X associated to α. If α is a nonnegative integer, then

$$b_\alpha(X) = (1 + X)^\alpha, \tag{16.13.2}$$

as before. Otherwise, this may be considered as a definition of the right side.

If β is another real or complex number, then

$$b_{\alpha+\beta}(X) = b_\alpha(X)\, b_\beta(X). \tag{16.13.3}$$

This follows from (16.12.8), with $t_1 = \alpha$ and $t_2 = \beta$. This means that

$$(1 + X)^{\alpha+\beta} = (1 + X)^\alpha\, (1 + X)^\beta, \tag{16.13.4}$$

using the notation in (16.13.2).

Observe that

$$\binom{-1}{l} = (-1)^l \tag{16.13.5}$$

for each nonnegative integer l. This implies that (16.13.2) is the same as the usual definition of $(1 + X)^{-1}$, as in Section 15.6, when $\alpha = -1$. Similarly, if α is a negative integer, then $(1+X)^{-\alpha}$ is a formal polynomial in X, and $(1+X)^\alpha$ may be defined as its multiplicative inverse in $\mathbf{R}[[X]]$. One can check that this is the same as (16.13.2), using (16.13.4).

Of course,

$$b'_\alpha(X) = \sum_{j=1}^{\infty} j \binom{\alpha}{j} X^{j-1} = \sum_{j=0}^{\infty} (j + 1) \binom{\alpha}{j+1} X^j, \tag{16.13.6}$$

as in Section 15.10. If l is a positive integer, then it is easy to see that

$$l \binom{\alpha}{l} = \alpha \binom{\alpha - 1}{l - 1}. \tag{16.13.7}$$

This implies that

$$b'_\alpha(X) = \sum_{j=0}^{\infty} \alpha \binom{\alpha - 1}{j} X^j = \alpha\, b_{\alpha-1}(X). \tag{16.13.8}$$

If l is a nonnegative integer, then

$$\binom{\alpha}{l} = \sum_{j=0}^{l} \binom{1}{j} \binom{\alpha - 1}{l - j}, \tag{16.13.9}$$

by (16.12.8), with $t_1 = 1$, and $t_2 = \alpha - 1$. This means that

$$\binom{\alpha}{l} = \binom{\alpha - 1}{l} + \binom{\alpha - 1}{l - 1} \tag{16.13.10}$$

when $l \geq 1$.

It is easy to see that

$$(1 + X)\, b'_\alpha(X) = \sum_{j=0}^{\infty} (j+1) \binom{\alpha}{j+1} X^j + \sum_{j=1}^{\infty} j \binom{\alpha}{j} X^j. \tag{16.13.11}$$

It follows that

$$(1 + X)\, b'_\alpha(X) = \sum_{j=0}^{\infty} \alpha \binom{\alpha - 1}{j} X^j + \sum_{j=1}^{\infty} \alpha \binom{\alpha - 1}{j - 1} X^j, \tag{16.13.12}$$

because of (16.13.7). One can use this and (16.13.10) to obtain that

$$(1 + X)\, b'_\alpha(X) = \alpha\, b_\alpha(X), \tag{16.13.13}$$

which is another way to look at (16.13.8). In fact, $b_\alpha(X)$ is uniquely determined by (16.13.13) and the condition that its constant term be equal to 1, as in Section 16.4.

Observe that

$$f_\alpha(X) = \exp(\alpha \log(1 + X)) \tag{16.13.14}$$

defines a formal power series in X with real or complex coefficients, as appropriate, where $\log(1+X)$ is as in Section 16.4. More precisely, the constant term in $\log(1 + X)$ is equal to 0, so that the formal composition may be defined as in Section 15.11. We also have that

$$\begin{aligned} f'_\alpha(X) &= \exp'(\alpha \log(1 + X))\, \alpha\, \log'(1 + X) \\ &= \exp(\alpha \log(1 + X))\, \alpha\, (1 + X)^{-1} = \alpha\, (1 + X)^{-1} f_\alpha(X), \end{aligned} \tag{16.13.15}$$

where the first step is as in Section 16.1, and the second step is as in Section 16.4. It is easy to see that the constant term in $f_\alpha(X)$ is equal to 1, because the constant term in $\log(1 + X)$ is equal to 0. It follows that

$$b_\alpha(X) = f_\alpha(X), \tag{16.13.16}$$

as in the preceding paragraph.

16.14 Using the ratio test

If α is a real or complex number that it not a nonnegative integer, then it is easy to see that

$$\binom{\alpha}{l} \neq 0 \tag{16.14.1}$$

for each nonnegative integer l. In this case,

$$\binom{\alpha}{l}\binom{\alpha}{l+1}^{-1} = \frac{\alpha - l}{l+1} \tag{16.14.2}$$

for every $l \geq 0$. This implies that

$$\lim_{l \to \infty}\left(\binom{\alpha}{l}\binom{\alpha}{l+1}^{-1}\right) = 1, \tag{16.14.3}$$

by a standard argument. If r is a nonnegative real number with $r < 1$, then it follows that

$$\sum_{j=0}^{\infty}\left|\binom{\alpha}{j}\right| r^j \tag{16.14.4}$$

converges as an infinite series of nonnegative real numbers, by the ratio test. If X is an indeterminate, then we get that

$$b_\alpha(X) \in (\mathbf{R})_1[[X]] \text{ or } (\mathbf{C})_1[[X]] \tag{16.14.5}$$

when $\alpha \in \mathbf{R}$ or $\mathbf{C}$, respectively, in the notation of Section 15.7.

If x is a real or complex number with $|x| < 1$, then put

$$b_\alpha(x) = \sum_{j=0}^{\infty}\binom{\alpha}{j} x^j, \tag{16.14.6}$$

where the series on the right converges absolutely, by the comparison test. Of course, if α is a nonnegative integer, then the right side reduces to a finite sum. If β is another real or complex number, then

$$b_{\alpha+\beta}(x) = b_\alpha(x)\, b_\beta(x), \tag{16.14.7}$$

because of (16.13.3), as in Section 9.10.

It is well known that

$$b_\alpha(x) = (1+x)^\alpha \tag{16.14.8}$$

under these conditions, where the right side can be defined in other ways. One can use the principal branch of the complex logarithm to define the right side by

$$\exp(\alpha \log(1+x)), \tag{16.14.9}$$

which is a holomorphic function on a larger open set in the complex plane. The power series expansion for $b_\alpha(x)$ corresponds exactly to the Taylor series for this function at $x = 0$.

Of course, $b_\alpha(x)$ is a holomorphic function on the open unit disk, because it is defined by an absolutely convergent power series there. We have that

$$b_\alpha'(x) = \alpha\, b_{\alpha-1}(x) \tag{16.14.10}$$

on the open unit disk, because of (16.13.8). Alternatively, we have that

$$(1 + x)\, b'_\alpha(x) = \alpha\, b_\alpha(x) \tag{16.14.11}$$

on the open unit disk, as in (16.13.13).

Similarly, we may consider $b_\alpha(x)$ as a smooth function on the open interval $(-1, 1)$ in the real line. One can obtain (16.14.8) from (16.14.11), as in Exercise 22 on p201 of [159].

16.15 Some related estimates

If α is a real number and $\alpha \le 0$, then it is easy to see that

$$(-1)^l \binom{\alpha}{l} \ge 0 \tag{16.15.1}$$

for every nonnegative integer l. If α is any real or complex number, then

$$|\alpha - l| \le |\alpha| + l = \big|-|\alpha| - l\big| \tag{16.15.2}$$

for every nonnegative integer l. This implies that

$$\left|\binom{\alpha}{l}\right| \le \left|\binom{-|\alpha|}{l}\right| = (-1)^l \binom{-|\alpha|}{l} \tag{16.15.3}$$

for every nonnegative integer l.

If r is a nonnegative real number with $r < 1$, then we get that

$$\sum_{l=0}^{\infty} \left|\binom{\alpha}{l}\right| r^l \le \sum_{l=0}^{\infty} \left|\binom{-|\alpha|}{l}\right| r^l = \sum_{l=0}^{\infty} (-1)^l \binom{-|\alpha|}{l} r^l. \tag{16.15.4}$$

Remember that these series converge, as in the previous section.

Of course,

$$\sum_{l=0}^{\infty} (-1)^l \binom{-|\alpha|}{l} r^l = b_{-|\alpha|}(-r) = (1 - r)^{-|\alpha|}, \tag{16.15.5}$$

using (16.14.8) in the second step. Thus

$$\sum_{l=0}^{\infty} \left|\binom{\alpha}{l}\right| r^l \le (1 - r)^{-|\alpha|}, \tag{16.15.6}$$

with equality when α is a real number with $\alpha \le 0$.

Let X be an indeterminate, so that $b_\alpha(X)$ may be defined as a formal power series in X with real or complex coefficients as in (16.13.1). In addition to (16.14.5), we get that

$$\|b_\alpha(X)\|_{1,r} \le (1 - r)^{-|\alpha|}, \tag{16.15.7}$$

in the notation of Section 15.7, because of (16.15.6). We also have equality when α is a real number with $\alpha \le 0$, as before.

Chapter 17

Norms, power series, and involutions

17.1 Another convergence property

Let α be a real number, so that $\binom{\alpha}{l} \in \mathbf{R}$ for every nonnegative integer l. Suppose that

$$0 < \alpha < 1. \tag{17.1.1}$$

It is easy to see that

$$(-1)^{l+1} \binom{\alpha}{l} > 0 \tag{17.1.2}$$

for every $l \geq 1$. This corresponds to part (a) of Exercise 1 on p125 of [8].

If $x \in \mathbf{R}$ and $|x| < 1$, then

$$b_\alpha(-x) = \sum_{j=0}^{\infty} \binom{\alpha}{j} (-x)^j = 1 - \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha}{j} x^j. \tag{17.1.3}$$

This means that

$$1 - \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha}{j} x^j = (1-x)^\alpha, \tag{17.1.4}$$

because of (16.14.8). This is closer to the formulation in [8].

We can use this to get that

$$\sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha}{j} = 1, \tag{17.1.5}$$

as in part (b) of Exercise 1 on p125 of [8]. Indeed,

$$\sum_{j=1}^{n} (-1)^{j+1} \binom{\alpha}{j} x^j \leq 1 - (1-x)^\alpha \leq 1 \tag{17.1.6}$$

for each positive integer n when $0 \le x < 1$, by (17.1.4). This implies that

$$\sum_{j=1}^{n} (-1)^{j+1} \binom{\alpha}{j} \le 1 \tag{17.1.7}$$

for every $n \ge 1$, by taking the limit as $x \to 1-$. This means that

$$\sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha}{j} \le 1, \tag{17.1.8}$$

which includes the convergence of the series on the left.

If $0 \le x < 1$, then

$$1 - (1-x)^{\alpha} = \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha}{j} x^j \le \sum_{j=1}^{\infty} (-1)^{j+1} \binom{\alpha}{j}, \tag{17.1.9}$$

using (17.1.4) in the first step. It follows that (17.1.5) holds, by taking the limit as $x \to 1-$. Of course, (17.1.5) implies that

$$\sum_{j=0}^{\infty} \left| \binom{\alpha}{j} \right| \tag{17.1.10}$$

converges. If X is an indeterminate, then we get that

$$b_{\alpha}(X) \in (\mathbf{R})_1^1[[X]], \tag{17.1.11}$$

in the notation of Section 15.7.

If α is a nonnegative integer, then (17.1.10) reduces to a finite sum, and $b_{\alpha}(X) \in \mathbf{R}[T]$, as before. If α is any nonnegative real number, then one can use (16.13.3) to get that (17.1.10) converges, so that (17.1.11) holds, using the remarks in Section 15.9. In this case, if x is a real or complex number, then one can define $b_{\alpha}(x)$ as a real or complex number as in (16.14.6). This defines a continuous real or complex-valued function on the set of $x \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, such that $|x| \le 1$, as in Section 9.12. It follows that (16.14.8) holds for all $x \in \mathbf{R}$ or $\mathbf{C}$ with $|x| \le 1$, because of the analogous statement when $|x| < 1$, and because both sides of the equation are continuous on the set where $|x| \le 1$.

17.2 Binomial series and Banach algebras

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\| = 1$, and let α be a real or complex number, appropriate. If $x \in \mathcal{A}$ and

$$\|x\|_{\mathcal{A}} < 1, \tag{17.2.1}$$

then

$$b_\alpha(x) = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j \tag{17.2.2}$$

defines an element of $\mathcal{A}$, as in Section 9.14, because of (16.14.5). If α is a nonnegative real number, then this works when

$$\|x\|_{\mathcal{A}} \le 1, \tag{17.2.3}$$

because of (17.1.11).

Let β be another real or complex number, as appropriate. If (17.2.1) holds, then

$$b_{\alpha+\beta}(x) = b_\alpha(x)\, b_\beta(x), \tag{17.2.4}$$

because of (16.13.3), as in Sections 9.13 and 15.11. Similarly, this works when α and β are nonnegative real numbers, and (17.2.3) holds.

Let us put

$$(e_{\mathcal{A}} + x)^\alpha = b_\alpha(x) \tag{17.2.5}$$

for every real or complex number α when (17.2.1) holds, and for every nonnegative real number α when (17.2.3) holds. If n is a positive integer, then

$$((e_{\mathcal{A}} + x)^\alpha)^n = (e_{\mathcal{A}} + x)^{n\,\alpha} \tag{17.2.6}$$

in both cases, because of (17.2.4). In particular,

$$((e_{\mathcal{A}} + x)^{1/n})^n = e_{\mathcal{A}} + x \tag{17.2.7}$$

when (17.2.3) holds. The $n = 2$ case corresponds to part (a) of Exercise 2 on p125 of [8]. This is also related to part (a) of Theorem 10.30 on p246 of [162].

Similarly, let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and suppose that $x \in \mathcal{A}$ is nilpotent. If $\alpha \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $b_\alpha(x)$ may be defined as an element of $\mathcal{A}$ as in (17.2.2), as in Section 15.11. If β is another real or complex number, then (17.2.4) holds, because of (16.13.3), as before. We may also express $b_\alpha(x)$ as in (17.2.5) in this case. Of course, (17.2.6) and (17.2.7) hold for all positive integers n, as before.

17.3 Some related formal power series

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let T be an indeterminate. If α is a real or complex number and $x \in \mathcal{A}$, then

$$b_{\alpha,x}(T) = b_\alpha(x\,T) = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j\, T^j \tag{17.3.1}$$

is a formal power series in T with coefficients in $\mathcal{A}$, as in Section 16.7. If β is another real or complex number, as appropriate, then

$$b_{\alpha+\beta,x}(T) = b_{\alpha,x}(T)\, b_{\beta,x}(T), \tag{17.3.2}$$

because of (16.13.3) and a remark in Section 16.7. If n is a positive integer, then it follows that

$$b_{\alpha,x}(T)^n = b_{n\,\alpha,x}(T). \tag{17.3.3}$$

In particular, we can take $\alpha = 1/n$, to get that

$$b_{1/n,x}(T)^n = 1 + x\,T. \tag{17.3.4}$$

Suppose that $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ is a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and that $x \in \mathcal{A}$ has the property that

$$\sum_{j=0}^{\infty} \left| \binom{\alpha}{j} \right| \|x^j\|_{\mathcal{A}} \tag{17.3.5}$$

converges as an infinite series of nonnegative real numbers. This implies that

$$b_{\alpha,x}(T) \in \mathcal{A}_1^1[[T]], \tag{17.3.6}$$

in the notation of Section 15.7, with

$$\|b_{\alpha,x}(T)\|_{1,1} = \|b_{\alpha,x}(T)\|_{\mathcal{A}_1^1[[T]]} = \sum_{j=0}^{\infty} \left| \binom{\alpha}{j} \right| \|x^j\|_{\mathcal{A}}. \tag{17.3.7}$$

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then

$$b_{\alpha,x}(t) = \sum_{j=0}^{\infty} \binom{\alpha}{j} t^j\, x^j \tag{17.3.8}$$

may be defined as an element of $\mathcal{A}$ as in Sections 15.8 and 16.7.

Suppose for the moment that x also has the property that

$$\sum_{j=0}^{\infty} \left| \binom{\beta}{j} \right| \|x^j\|_{\mathcal{A}} \tag{17.3.9}$$

converges as an infinite series of nonnegative real numbers, so that

$$b_{\beta,x}(T) \in \mathcal{A}_1^1[[T]], \tag{17.3.10}$$

with

$$\|b_{\beta,x}(T)\|_{1,1} = \sum_{j=0}^{\infty} \left| \binom{\beta}{j} \right| \|x^j\|_{\mathcal{A}}. \tag{17.3.11}$$

This implies that

$$b_{\alpha+\beta,x}(T) \in \mathcal{A}_1^1[[T]], \tag{17.3.12}$$

because of (17.3.2), as in Section 15.9. This means that

$$\sum_{j=0}^{\infty} \left| \binom{\alpha+\beta}{j} \right| \|x^j\|_{\mathcal{A}} \tag{17.3.13}$$

converges as an infinite series of nonnegative real numbers too, with

$$\sum_{j=0}^{\infty} \left| \binom{\alpha+\beta}{j} \right| \|x^j\|_{\mathcal{A}} = \|b_{\alpha+\beta,x}(T)\|_{1,1} \le \|b_{\alpha,x}(T)\|_{1,1} \, \|b_{\beta,x}(T)\|_{1,1}, \tag{17.3.14}$$

as before. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then $b_{\beta,x}(t)$ and $b_{\alpha+\beta,x}(t)$ may be defined as elements of $\mathcal{A}$ as well, and

$$b_{\alpha+\beta,x}(t) = b_{\alpha,x}(t) \, b_{\beta,x}(t), \tag{17.3.15}$$

as in Section 15.9 again.

If (17.3.5) converges, then it follows that

$$\sum_{j=0}^{\infty} \left| \binom{n\,\alpha}{j} \right| \|x^j\|_{\mathcal{A}} \tag{17.3.16}$$

converges as an infinite series of nonnegative real numbers for every positive integer n, so that

$$b_{n\,\alpha,x}(T) \in \mathcal{A}^1_1[[T]]. \tag{17.3.17}$$

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then $b_{n\,\alpha,x}(t)$ may be defined as an element of $\mathcal{A}$, and

$$b_{\alpha,x}(T)^n = b_{\alpha,x}(t)^n. \tag{17.3.18}$$

If $\alpha = 1/n$, then we get that

$$b_{1/n,x}(t)^n = e_{\mathcal{A}} + t\,x, \tag{17.3.19}$$

because of (17.3.4). This is related to part (a) of Theorem 10.30 on p246 of [162].

17.4 Logarithms and binomial series

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and suppose that $x \in \mathcal{A}$ is nilpotent. Remember that $\log(e_{\mathcal{A}} + x)$ may be defined as an element of $\mathcal{A}$ as in Section 16.5, and is nilpotent. If α is a real or complex number, as appropriate, then

$$\alpha \, \log(e_{\mathcal{A}} + x) \text{ is nilpotent in } \mathcal{A} \tag{17.4.1}$$

too, so that

$$\exp(\alpha \log(e_{\mathcal{A}} + x)) \tag{17.4.2}$$

may be defined as an element of $\mathcal{A}$, as in Sections 11.1 and 15.11.

Note that $b_\alpha(x)$ may be defined as an element of $\mathcal{A}$ as well, as in Section 15.11. Under these conditions,

$$b_\alpha(x) = \exp(\alpha \log(e_{\mathcal{A}} + x)), \tag{17.4.3}$$

because of the analogous statement for formal power series in Section 16.13, as in Section 15.13.

Let T be an indeterminate, and let r be a positive real number strictly less than 1. Remember that

$$a(T) = \log(1 + T) \in (\mathbf{R})^1_r[[T]], \tag{17.4.4}$$

in the notation of Section 15.7, with

$$\|a(T)\|_{1,r} = \|\log(1 + T)\|_{1,r} = \|\log(1 + t)\|_{(\mathbf{R})^1_r[[T]]} = -\log(1 - r), \tag{17.4.5}$$

as in Section 16.5. We also have that

$$f(T) = \exp(T) \in (\mathbf{R})^1_{r_1}[[T]] \tag{17.4.6}$$

for every $r_1 > 0$, with

$$\|f(T)\|_{1,r_1} = \|\exp(T)\|_{1,r_1} = \exp(r_1) \tag{17.4.7}$$

as in Section 16.5.

If α is a real or complex number, then

$$\alpha\, a(T) \in (\mathbf{R})^1_r[[T]] \text{ or } (\mathbf{C})^1_r[[T]], \tag{17.4.8}$$

as appropriate, with

$$\|\alpha\, a(T)\|_{1,r} = |\alpha|\, \|a(T)\|_{1,r} = -|\alpha| \log(1 - r). \tag{17.4.9}$$

It follows that

$$\begin{aligned} (f \circ (\alpha\, a))(T) &= f(\alpha\, a(T)) \\ &= \exp(\alpha \log(1 + T)) \in (\mathbf{R})^1_r[[T]] \text{ or } (\mathbf{C})^1_r[[T]], \end{aligned} \tag{17.4.10}$$

as appropriate, may be defined as in Section 15.15. More precisely, if $\alpha \neq 0$, then we can take

$$r_1 = -|\alpha| \log(1 - r), \tag{17.4.11}$$

to get that

$$\|f(\alpha\, a(T))\|_{1,r} \le \|f(T)\|_{1,r_1} = (1 - r)^{-|\alpha|}. \tag{17.4.12}$$

Note that the constant term in $\alpha\, a(T)$ is equal to 0, because of the analogous statement for $a(T)$, as in Section 16.4. This implies that

$$f(\alpha\, a(T)) = \exp(\alpha \log(1 + T)) \tag{17.4.13}$$

may be defined as a formal power series in T with coefficients in the real or complex numbers, as appropriate, as in Section 15.11. In fact,

$$\exp(\alpha \log(1+T)) = b_\alpha(T) \tag{17.4.14}$$

as a formal power series in T, as in Section 16.13. This is the same as in (17.4.10), as a formal power series in T, as in Section 15.15.

It follows that

$$b_\alpha(T) \in (\mathbf{R})^1_r[[T]] \text{ or } (\mathbf{C})^1_r[[T]], \tag{17.4.15}$$

as appropriate. Remember that this was obtained in Section 16.14 as well, because $r < 1$. If $\alpha \neq 0$, then (17.4.12) implies that

$$\|b_\alpha(T)\|_{1,r} \le (1-r)^{-|\alpha|}. \tag{17.4.16}$$

Of course, this is clear when $\alpha = 0$. Note that this is the same as in Section 16.15.

Now let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers, as appropriate, with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \le r$, then $\log(e_{\mathcal{A}} + x)$ may be defined as an element of $\mathcal{A}$ as in Section 16.5. This means that (17.4.2) may be defined as an element of $\mathcal{A}$ as in Section 10.4. Similarly, $b_\alpha(x)$ may be defined as an element of $\mathcal{A}$, as in Sections 9.14 and 17.2. We also have that (17.4.3) holds in this case, because of the remarks in Section 15.15.

17.5 More on logarithms, binomial series

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and suppose that $x \in \mathcal{A}$ has the property that

$$\sum_{l=1}^{\infty} (1/l)\, \|x^l\|_{\mathcal{A}} \tag{17.5.1}$$

converges as an infinite series of nonnegative real numbers. Let T be an indeterminate, and let us take $a(T) = \log(1+T)$ and

$$a_x(T) = \log(1 + x\,T) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l}\, x^l\, T^l, \tag{17.5.2}$$

as in Section 16.11. Remember that

$$a_x(T) \in \mathcal{A}^1_1[[T]], \tag{17.5.3}$$

as in Section 15.7. We also have that $f(T) = \exp(T)$ is an element of $(\mathbf{R})^1_{r_1}[[T]]$ for every $r_1 > 0$, as in Section 16.5.

If $\alpha \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $\alpha\, a_x(T)$ is an element of $\mathcal{A}^1_1[[T]]$, because of (17.5.3). This implies that

$$(f \circ (\alpha\, a_x))(T) = f(\alpha\, a_x(T)) = \exp(\alpha\, a_x(T)) \in \mathcal{A}^1_1[[T]] \tag{17.5.4}$$

may be defined as in Section 15.15.

Remember that the constant term in $a(T)$ is equal to 0, so that $\alpha\, a(T)$ has the same property. This means that

$$\exp(\alpha\, a(T)) = \exp(\alpha \log(1+T)) \tag{17.5.5}$$

may be defined as a formal power series in T with coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 15.11. This is equal to $b_\alpha(T)$, as in Section 16.13.

Similarly, the constant term in $\alpha\, a_x(T)$ is equal to 0, because of the analogous property of $a_x(T)$, so that $\exp(\alpha\, a_x(T))$ may be defined as a formal power series in T with coefficients in $\mathcal{A}$ as in Section 15.11. One can use this to get that

$$\exp(\alpha\, a_x(T)) = b_{\alpha,x}(T), \tag{17.5.6}$$

as in Section 16.8.

This definition of $\exp(\alpha\, a_x(T))$ as a formal power series in T with coefficients in $\mathcal{A}$ is equivalent to the one in Section 15.15 under these conditions, as before. It follows that

$$b_{\alpha,x}(T) \in \mathcal{A}_1^1[[T]], \tag{17.5.7}$$

because of (17.5.4). Another way to look at this will be mentioned in a moment.

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then $a_x(t)$ may be defined as an element of $\mathcal{A}$, as in Sections 15.8 and 16.11. This implies that

$$f(\alpha\, a_x(t)) = \exp(\alpha\, a_x(t)) \tag{17.5.8}$$

may be defined as an element of $\mathcal{A}$ as well, as in Sections 9.14 and 10.4. Similarly, $b_{\alpha,x}(t)$ may be defined as an element of $\mathcal{A}$ as in Section 15.8 and the previous section.

It is easy to see that

$$\exp(\alpha\, a_x(t)) = f(\alpha\, a_x(t)) = (f \circ (\alpha\, a_x))(t) = b_{\alpha,x}(t), \tag{17.5.9}$$

using some remarks in Section 15.15 in the second step, and (17.5.6) and the remark at the beginning of the paragraph that followed in the third step. This reduces to a statement in Section 16.11 when $\alpha = 1$.

Observe that the convergence of the series (17.5.1) implies that

$$\|x^{l_0}\|_{\mathcal{A}} < 1 \tag{17.5.10}$$

for some positive integer l_0. This means that

$$\|x^{l_0}\|_{\mathcal{A}}^{1/l_0} < 1, \tag{17.5.11}$$

so that

$$r_{\mathcal{A}}(x) < 1, \tag{17.5.12}$$

in the notation of Section 6.13. Remember that $\|x^l\|_{\mathcal{A}}^{1/l} \to r_{\mathcal{A}}(x)$ as $l \to \infty$, as in Section 6.14. It is easy to see that (17.5.12) implies (17.5.7), using this and the remarks in Section 16.14. Similarly, (17.5.12) implies that (17.5.1) converges.

17.6 Power series and involutions

Let T be an indeterminate, and let $f(T) = \sum_{j=0}^{\infty} f_j\, T^j$ be a formal power series in T with complex coefficients. Put

$$\overline{f}(T) = \sum_{j=0}^{\infty} \overline{f_j}\, T^j, \tag{17.6.1}$$

which is another element of $\mathbf{C}[[T]]$. It is easy to see that

$$f(T) \mapsto \overline{f}(T) \tag{17.6.2}$$

defines a conjugate-linear involution on $\mathbf{C}[[T]]$, as an algebra over the complex numbers, as in Section 6.4. Of course, $f(T)$ is a formal polynomial in T if and only if $\overline{f}(T) \in \mathbf{C}[T]$.

Similarly, if r is a positive real number, then

$$f(T) \in (\mathbf{C})^1_r[[T]] \tag{17.6.3}$$

if and only if

$$\overline{f}(T) \in (\mathbf{C})^1_r[[T]], \tag{17.6.4}$$

in the notation of Section 15.7. In this case,

$$\|\overline{f}(T)\|_{1,r} = \|f(T)\|_{1,r}. \tag{17.6.5}$$

If $0 < \rho \le \infty$, then

$$f(T) \in (\mathbf{C})_\rho[[T]] \tag{17.6.6}$$

if and only if

$$\overline{f}(T) \in (\mathbf{C})_\rho[[T]], \tag{17.6.7}$$

using the notation in Section 15.7 again.

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and an algebra involution $x \mapsto x^*$, which may be conjugate-linear in the complex case. If $x \in \mathcal{A}$ and $f(T)$ is a formal polynomial in T with real or complex coefficients, as appropriate, then

$$f(x)^* = f(x^*) \tag{17.6.8}$$

in the real case, and in the complex case when the involution is complex-linear, and

$$f(x)^* = \overline{f}(x^*) \tag{17.6.9}$$

in the complex case when the involution is conjugate-linear.

If $x \in \mathcal{A}$ is nilpotent, then it is easy to see that

$$x^* \text{ is nilpotent in } \mathcal{A} \tag{17.6.10}$$

too. If $f(T)$ is a formal power series in T with real or complex coefficients, as appropriate, then $f(x)$ may be defined as an element of $\mathcal{A}$ as in Section 15.11, and similarly for $f(x^*)$, and for $\overline{f}(x^*)$ in the complex case. We also have (17.6.8) or (17.6.9), as appropriate, as before.

Now let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and an algebra involution $x \mapsto x^*$, which may be conjugate-linear in the complex case. Suppose that there is a real number $C \geq 1$ such that

$$\|x^*\|_{\mathcal{A}} \leq C\,\|x\|_{\mathcal{A}} \tag{17.6.11}$$

for every $x \in \mathcal{A}$. Of course, this implies that

$$\|x\|_{\mathcal{A}} = \|(x^*)^*\|_{\mathcal{A}} \leq C\,\|x^*\|_{\mathcal{A}} \tag{17.6.12}$$

for every $x \in \mathcal{A}$. If $C = 1$, then we get that $\|x^*\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}$ for every $x \in \mathcal{A}$.

Suppose that $f(T)$ is an element of $(\mathbf{R})^1_r[[T]]$ or $(\mathbf{C})^1_r[[T]]$, as appropriate, for some $r > 0$. If $x \in \mathcal{A}$ and $\|x\|_{\mathcal{A}} \leq r$, then $f(x)$ may be defined as an element of $\mathcal{A}$, as in Sections 9.14 and 15.11. Similarly, if

$$\|x^*\|_{\mathcal{A}} \leq r, \tag{17.6.13}$$

then $f(x^*)$ may be defined as an element of $\mathcal{A}$, and $\overline{f}(x^*)$ may be defined as an element of $\mathcal{A}$ in the complex case. It is easy to see that (17.6.8) or (17.6.9) holds under these conditions, as appropriate.

Of course, (17.6.13) follows from the hypothesis that $\|x\|_{\mathcal{A}} \leq r$ when $C = 1$. We shall consider another version of this with a simpler role for C in Section 17.8.

17.7 Coefficients in $\mathcal{A}$

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with an involution $x \mapsto x^*$ that may be conjugate-linear in the complex case, and let T be an indeterminate. If $f(T) = \sum_{j=0}^{\infty} f_j\,T^j$ is a formal power series in T with coefficients in $\mathcal{A}$, then put

$$f^*(T) = \sum_{j=0}^{\infty} f_j^*\,T^j, \tag{17.7.1}$$

which is another element of $\mathcal{A}[[T]]$. One can check that

$$f(T) \mapsto f^*(T) \tag{17.7.2}$$

defines an algebra involution on $\mathcal{A}[[T]]$, which is conjugate-linear in the complex case when $x \mapsto x^*$ is conjugate-linear on $\mathcal{A}$. Note that (17.7.2) maps $\mathcal{A}[T]$ onto itself.

If $f(T)$ is a formal polynomial in T and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $f(t)$ may be defined as an element of $\mathcal{A}$, as in Section 15.4. It is easy to see that

$$f(t)^* = f^*(t) \tag{17.7.3}$$

in the real case, and in the complex case when the involution is complex-linear, and that

$$f(t)^* = f^*(\bar{t}) \tag{17.7.4}$$

in the complex case when the involution is conjugate-linear.

Let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$, and suppose that (17.6.11) holds on $\mathcal{A}$ for some $C \geq 1$. Also let r be a positive real number, and suppose that $f(T) \in \mathcal{A}^1_r[[T]]$, in the notation of Section 15.7. This implies that

$$f^*(T) \in \mathcal{A}^1_r[[T]] \tag{17.7.5}$$

too, with

$$\|f^*(T)\|_{1,r} \leq C\,\|f(T)\|_{1,r}. \tag{17.7.6}$$

Similarly, if $0 < \rho \leq \infty$ and $f(T) \in \mathcal{A}_\rho[[T]]$, then

$$f^*(T) \in \mathcal{A}_\rho[[T]] \tag{17.7.7}$$

as well.

Suppose that $\mathcal{A}$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$, and that $f(T) \in \mathcal{A}^1_r[[T]]$ for some $r > 0$, as before. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \leq r$, then $f(t)$ and $f^*(t)$ may be defined as elements of $\mathcal{A}$ as in Section 15.8, and $f^*(\bar{t})$ may be defined as an element of $\mathcal{A}$ in the complec case. We also have that (17.7.3) or (17.7.4) holds, as appropriate, under these conditions.

17.8 Convergence conditions and involutions

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, and let T be an indeterminate. If $a(T) = \sum_{l=0}^\infty a_l\,T^l$ is a formal power series in T with real or complex coefficients, as appropriate, and $x \in \mathcal{A}$, then put

$$a_x(T) = \sum_{l=0}^\infty a_l\,x^l\,T^l, \tag{17.8.1}$$

as in Section 16.7. In the complex case, $\overline{a}(T) \in \mathbf{C}[[T]]$ may be defined as in Section 17.6. Let $y \mapsto y^*$ be an involution on $\mathcal{A}$, which may be conjugate-linear in the complex case, so that $(a_x)^*(T) \in \mathcal{A}[[T]]$ may be defined as in the previous section. Observe that

$$(a_x)^*(T) = a_{x^*}(T) \tag{17.8.2}$$

in the real case, and in the complex case when the involution is complex-linear, and that

$$(a_x)^*(T) = \overline{a}_{x^*}(T) \tag{17.8.3}$$

in the complex case when the involution is conjugate-linear.

Now let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and an algebra involution $y \mapsto y^*$, which may be conjugate-linear in the complex case. Suppose that

$$\|y^*\|_{\mathcal{A}} \le C\, \|y\|_{\mathcal{A}} \tag{17.8.4}$$

for some $C \ge 1$ and every $y \in \mathcal{A}$, as before. Suppose also that $x \in \mathcal{A}$ has the property that

$$\sum_{l=0}^{\infty} |a_l|\, \|x^l\|_{\mathcal{A}} \tag{17.8.5}$$

converges as an infinite series of nonnegative real numbers. Note that

$$\|(x^*)^l\|_{\mathcal{A}} = \|(x^l)^*\|_{\mathcal{A}} \le C\, \|x^l\|_{\mathcal{A}} \tag{17.8.6}$$

for each l. It follows that

$$\sum_{l=0}^{\infty} |a_l|\, \|(x^*)^l\|_{\mathcal{A}} \tag{17.8.7}$$

converges too, by the comparison test, with

$$\sum_{l=0}^{\infty} |a_l|\, \|(x^*)^l\|_{\mathcal{A}} \le C \sum_{l=0}^{\infty} |a_l|\, \|x^l\|_{\mathcal{A}}. \tag{17.8.8}$$

This means that

$$a_x(T),\ a_{x^*}(T) \in \mathcal{A}^1_1[[T]], \tag{17.8.9}$$

in the notation of Section 15.7, with

$$\|a_{x^*}(T)\|_{1,1} \le C\, \|a_x(T)\|_{1,1}. \tag{17.8.10}$$

In the complex case, we also have that

$$\overline{a}_x(T),\ \overline{a}_{x^*}(T) \in \mathcal{A}^1_1[[T]], \tag{17.8.11}$$

with

$$\|\overline{a}_x(T)\|_{1,1} = \|a_x(T)\|_{1,1},\ \ \|\overline{a}_{x^*}(T)\|_{1,1} = \|a_{x^*}(T)\|_{1,1}. \tag{17.8.12}$$

If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then $a_x(t)$ and $a_{x^*}(t)$ may be defined as elements of $\mathcal{A}$, as in Sections 15.8 and 16.7. In the complex case, $\overline{a}_x(t)$ and $\overline{a}_{x^*}(t)$ may be defined as elements of $\mathcal{A}$ too, and of course these statements hold with t replaced by $\overline{t}$ as well. Under these conditions, one can check that

$$(a_x(t))^* = (a_x)^*(t) = a_{x^*}(t) \tag{17.8.13}$$

in the real case, and in the complex case when the involution is complex-linear, and that

$$(a_x(t))^* = (a_x)^*(\overline{t}) = \overline{a}_{x^*}(\overline{t}) \tag{17.8.14}$$

in the complex case when the involution is conjugate-linear.

17.9 Square roots from binomial series

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Also let T be an indeterminate, and for each $x \in \mathcal{A}$, put

$$b_{1/2,x}(T) = \sum_{j=0} \binom{1/2}{j} x^j \, T^j, \tag{17.9.1}$$

as in Section 17.3. This is a formal power series in T with coefficients in $\mathcal{A}$ such that

$$b_{1/2,x}(T)^2 = 1 + x\,T, \tag{17.9.2}$$

as before. Remember that

$$\sum_{j=0}^{\infty} \left| \binom{1/2}{j} \right| \tag{17.9.3}$$

converges as an infinite series of nonnegative real numbers, as in Section 17.1.

Suppose that

$$\sum_{j=0}^{\infty} \left| \binom{1/2}{j} \right| \|x^j\|_{\mathcal{A}} \tag{17.9.4}$$

converges as an infinite series of nonnegative real numbers, so that

$$b_{1/2,x}(T) \in \mathcal{A}^1_1[[T]], \tag{17.9.5}$$

in the notation of Section 15.7. If $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $|t| \le 1$, then

$$b_{1/2,x}(t) = \sum_{j=0}^{\infty} \binom{1/2}{j} t^j \, x^j \tag{17.9.6}$$

may be defined as an element of $\mathcal{A}$, as in Sections 15.8, 16.7, and 17.3. We also have that

$$b_{1/2,x}(t)^2 = e_{\mathcal{A}} + t\,x, \tag{17.9.7}$$

as in Section 17.3.

Let $y \mapsto y^*$ be an algebra involution on $\mathcal{A}$ that may be conjugate-linear in the complex case, and that satisfies (17.8.4) on $\mathcal{A}$ for some $C \ge 1$. This implies that

$$\sum_{j=0}^{\infty} \left| \binom{1/2}{j} \right| \|(x^*)^j\|_{\mathcal{A}} \tag{17.9.8}$$

converges as an infinite series of nonnegative real numbers, as in the previous section. This means that

$$b_{1/2,x^*}(T) \in \mathcal{A}^1_1[[T]], \tag{17.9.9}$$

so that $b_{1/2,x^*}(t)$ may be defined as an element of $\mathcal{A}$ when $|t| \le 1$, as before. In fact,

$$b_{1/2,x}(t)^* = b_{1/2,x^*}(t) \tag{17.9.10}$$

in the real case, and in the complex case when the involution is complex-linear, and

$$b_{1/2,x}(t)^* = b_{1/2,x^*}(\overline{t}) \tag{17.9.11}$$

in the complex case when the involution is conjugate-linear, as before.

In particular, if x is self-adjoint, so that $x^* = x$, then

$$b_{1/2,x}(t)^* = b_{1/2,x}(t) \tag{17.9.12}$$

in the real case, and in the complex case when the involution is complex-linear, and

$$b_{1/2,x}(t)^* = b_{1/2,x}(\overline{t}) \tag{17.9.13}$$

in the complex case when the involution is conjugate-linear. This implies that $b_{1/2,x}(t)$ is self-adjoint in $\mathcal{A}$ in the real case, in the complex case when the involution is complex-linear, and in the complex case when the involution is conjugate-linear and $t \in \mathbf{R}$. This is related to part (b) of Exercise 2 on p125 of [8].

17.10 Some more self-adjointness arguments

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a real or complex Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Suppose that $x \in \mathcal{A}$ satisfies

$$r_{\mathcal{A}}(x) < 1, \tag{17.10.1}$$

where $r_{\mathcal{A}}(x)$ is as in Section 6.13. This implies that (17.9.4) converges, as in Section 17.5. Thus

$$b_{1/2,x}(1) = \sum_{j=1}^{\infty} \binom{1/2}{j} x^j \tag{17.10.2}$$

may be defined as an element of $\mathcal{A}$ as before, and satisfies

$$b_{1/2,x}(1)^2 = e_{\mathcal{A}} + x. \tag{17.10.3}$$

We also have that $e_{\mathcal{A}} + x$ is invertible in $\mathcal{A}$, as in Section 6.13.

Let $y \mapsto y^*$ be an involution on $\mathcal{A}$ that may be conjugate-linear in the complex case. Suppose that x is also self-adjoint, so that the partial sums of the series on the right side of (17.10.2) are self-adjoint elements of $\mathcal{A}$ too. If the involution on $\mathcal{A}$ is continuous, then

$$b_{1/2,x}(1)^* = b_{1/2,x}(1), \tag{17.10.4}$$

as before. If $\mathcal{A}$ is complex and commutative, then this holds without asking that the involution be continuous, as in Section 14.4. This is basically a version of Theorem 11.20 on p278 of [162], with simpler hypotheses on x.

The same conclusion holds when $\mathcal{A}$ is complex and not necessarily commutative, as in Section 14.9. This uses the fact that the partial sums of the series on the right side of (17.10.2) commute with each other. This corresponds to Theorem 11.26 on p281 of [162].

17.11 Involutions and $r_{\mathcal{A}}(x)$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $x \in \mathcal{A}$, then $r_{\mathcal{A}}(x)$ may be defined as in Section 6.13, as mentioned at the beginning of Section 7.2. Let $y \mapsto y^*$ be an involution on $\mathcal{A}$, which may be conjugate-linear in the complex case.

It would sometimes be nice to have

$$r_{\mathcal{A}}(x^*) = r_{\mathcal{A}}(x) \tag{17.11.1}$$

for every $x \in \mathcal{A}$. It is easy to see that this holds when

$$r_{\mathcal{A}}(x^*) \le r_{\mathcal{A}}(x) \tag{17.11.2}$$

for every $x \in \mathcal{A}$. If the involution on $\mathcal{A}$ is continuous, so that (17.8.4) holds for some $C \ge 1$, then (17.11.2) follows from the fact that

$$\|(x^*)^l\|_{\mathcal{A}}^{1/l} = \|(x^l)^*\|_{\mathcal{A}}^{1/l} \le C^{1/l}\, \|x^l\|_{\mathcal{A}}^{1/l} \tag{17.11.3}$$

for every positive integer l, by taking the limit as $l \to \infty$.

If $\mathcal{A}$ is a complex Banach algebra with a multiplicative identity element, then (17.11.1) follows from the characterization of $r_{\mathcal{A}}(\cdot)$ in terms of the spectrum of an element of $\mathcal{A}$, as in Section 6.14. This also uses the characterization of the spectrum of x^* in terms of the spectrum of x, as in Section 14.1.

Suppose that $x \in \mathcal{A}$ is normal in the sense that it commutes with x^*, as in Section 7.5. This implies that

$$r_{\mathcal{A}}(x\,x^*) \le r_{\mathcal{A}}(x)\, r_{\mathcal{A}}(x^*), \tag{17.11.4}$$

as in Section 7.2. If (17.11.2) holds, then we get that

$$r_{\mathcal{A}}(x\,x^*) \le r_{\mathcal{A}}(x)^2. \tag{17.11.5}$$

17.12 Identity elements and $r_{\mathcal{A}}(\cdot)$

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If α is a positive real number, then

$$\|x\|_{\mathcal{A},\alpha} = \alpha\, \|x\|_{\mathcal{A}} \tag{17.12.1}$$

defines a norm on $\mathcal{A}$. If $\alpha \ge 1$, then it is easy to see that

$$\|\cdot\|_{\mathcal{A},\alpha} \text{ is submultiplicative on } \mathcal{A}. \tag{17.12.2}$$

Suppose for the moment that $\mathcal{A}$ is an associative algebra. If $x \in \mathcal{A}$, then $r_{\mathcal{A}}(x)$ may be defined as in Sections 6.13 and 7.2. If $\alpha \ge 1$, then we can define $r_{\mathcal{A},\alpha}(x)$ in the same way, with respect to the norm $\|\cdot\|_{\mathcal{A},\alpha}$ on $\mathcal{A}$. Observe that

$$r_{\mathcal{A},\alpha}(x) = \lim_{l\to\infty} \|x^l\|_{\mathcal{A},\alpha}^{1/l} = \lim_{l\to\infty} (\alpha^{1/l}\, \|x^l\|_{\mathcal{A}}^{1/l}) = r_{\mathcal{A}}(x), \tag{17.12.3}$$

because $\alpha^{1/l} \to 1$ as $l \to \infty$.

Let $\mathcal{A}_1 = \mathcal{A} \times \mathbf{R}$ or $\mathcal{A} \times \mathbf{C}$, as appropriate, considered as an algebra in the strict sense over the real or complex numbers, as appropriate, as in Section 6.15. Remember that

$$x \mapsto (x, 0) \tag{17.12.4}$$

is an injective algebra homomorphism from $\mathcal{A}$ into $\mathcal{A}_1$, and that $e_{\mathcal{A}_1} = (0, 1)$ is the multiplicative identity element in $\mathcal{A}_1$, by construction. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then $(e_{\mathcal{A}}, 0)$ is an idempotent element of $\mathcal{A}_1$, as in Section 7.6. If $\alpha \geq 1$, then

$$\|(x, t)\|_{\mathcal{A}_1, \alpha} = \|x\|_{\mathcal{A}, \alpha} + |t| = \alpha \, \|x\|_{\mathcal{A}} + |t| \tag{17.12.5}$$

defines a submultiplicative norm on $\mathcal{A}_1$, as before. Of course, this is the same as $\|\cdot\|_{\mathcal{A}_1}$ in Section 6.15 when $\alpha = 1$.

Suppose that $\mathcal{A}$ is an associative algebra again, so that $\mathcal{A}_1$ is associative as well, as in Section 6.15. If $(x, t) \in \mathcal{A}_1$ and $\alpha \geq 1$, then $r_{\mathcal{A}_1, \alpha}((x, t))$ may be defined in the usual way, with respect to the norm $\|\cdot\|_{\mathcal{A}_1, \alpha}$ on $\mathcal{A}_1$. Note that

$$r_{\mathcal{A}_1, \alpha}((x, 0)) = r_{\mathcal{A}, \alpha}(x) = r_{\mathcal{A}}(x) \tag{17.12.6}$$

for every $x \in \mathcal{A}$. Similarly,

$$r_{\mathcal{A}_1, \alpha}((0, t)) = |t| \tag{17.12.7}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate.

If $(x, t) \in \mathcal{A}_1$ and l is a positive integer, then one can check that

$$\|(x, t)^l\|_{\mathcal{A}_1, \alpha} \geq |t|^l. \tag{17.12.8}$$

This implies that

$$r_{\mathcal{A}_1, \alpha}((x, t)) \geq |t|. \tag{17.12.9}$$

It is easy to see that $(x, 0)$ and $(0, t)$ commute in $\mathcal{A}_1$. It follows that

$$r_{\mathcal{A}_1, \alpha}((x, t)) \leq r_{\mathcal{A}_1, \alpha}((x, 0)) + r_{\mathcal{A}_1, \alpha}((0, t)) = r_{\mathcal{A}}(x) + |t|, \tag{17.12.10}$$

where the first step is as in Section 9.7. We also have that

$$r_{\mathcal{A}_1, \alpha}((x, t)) = r_{\mathcal{A}_1, 1}((x, t)), \tag{17.12.11}$$

because

$$\|(x, t)^l\|_{\mathcal{A}_1, 1}^{1/l} \leq \|(x, t)^l\|_{\mathcal{A}_1, \alpha}^{1/l} \leq \alpha^{1/l} \, \|(x, t)^l\|_{\mathcal{A}_1, 1}^{1/l}. \tag{17.12.12}$$

Let $\mathcal{A}$ be an algebra in the strict sense again, and let $x \mapsto x^*$ be an involution on $\mathcal{A}$, which may be conjugate-linear in the complex case. If $(x, t) \in \mathcal{A}_1$, then put

$$(x, t)^* = (x^*, t) \tag{17.12.13}$$

in the real case, and in the complex case when the involution on $\mathcal{A}$ is complex-linear, and

$$(x, t)^* = (x^*, \overline{t}) \tag{17.12.14}$$

in the complex case when the involution on $\mathcal{A}$ is conjugate-linear. One can verify that this defines an involution on $\mathcal{A}_1$, which is conjugate-linear in the complex case when the involution on $\mathcal{A}$ is conjugate linear.

More precisely, this is the unique extension of the involution on $\mathcal{A}$ to an involution on $\mathcal{A}_1$, that is complex-linear or conjugate-linear in the complex case depending on whether the involution on $\mathcal{A}$ is complex-linear or conjugate-linear. This uses the fact that the multiplicative identity element in $\mathcal{A}_1$ is self-adjoint with respect to any involution on $\mathcal{A}_1$, as in Section 7.5.

Let $\|\cdot\|_{\mathcal{A}}$ be a submultiplicative norm on $\mathcal{A}$ again, and suppose that

$$\|x^*\|_{\mathcal{A}} \le C\,\|x\|_{\mathcal{A}} \tag{17.12.15}$$

for some $C \ge 1$ and all $x \in \mathcal{A}$. This implies that

$$\|x^*\|_{\mathcal{A},\alpha} \le C\,\|x\|_{\mathcal{A},\alpha} \tag{17.12.16}$$

for all $x \in \mathcal{A}$, so that

$$\|(x,t)^*\|_{\mathcal{A}_1,\alpha} \le C\,\|(x,t)\|_{\mathcal{A}_1,\alpha} \tag{17.12.17}$$

for all $(x,t) \in \mathcal{A}_1$. In particular, if the involution on $\mathcal{A}$ preserves $\|\cdot\|_{\mathcal{A}}$, then it preserves $\|\cdot\|_{\mathcal{A},\alpha}$, and the corresponding involution on $\mathcal{A}_1$ preserves $\|\cdot\|_{\mathcal{A}_1,\alpha}$. This corresponds to part of Proposition 2.5.4 on p58 of [8].

17.13 Limits at infinity

Let X be an infinite set, and remember that $c^{lai}(X,\mathbf{R})$, $c^{lai}(X,\mathbf{C})$ are the spaces of real or complex-valued functions f on X, as appropriate, with a limit at infinity, as in Section 3.3. This means that f may be expressed as

$$f = f_0 + t\,\mathbf{1}_X, \tag{17.13.1}$$

where f_0 is a real or complex-valued function on X that vanishes at infinity, as appropriate, $\mathbf{1}_X$ is the constant function on X equal to 1 at every point, and $t = t(f) \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, as before.

If f has a limit at infinity, then f is bounded on X, as before. Remember that $c^{lai}(X,\mathbf{R})$, $c^{lai}(X,\mathbf{C})$ are closed linear subspaces of $\ell^\infty(X,\mathbf{R})$, $\ell^\infty(X,\mathbf{C})$, respectively. It is easy to see that in fact

(17.13.2) $c^{lai}(X,\mathbf{R})$, $c^{lai}(X,\mathbf{C})$ are subalgebras of $\ell^\infty(X,\mathbf{R})$, $\ell^\infty(X,\mathbf{C})$,

respectively.

If f is as in (17.13.1), then t is called the limit of f at infinity on X, and it is uniquely determined by f. We have seen that

$$|t(f)| \le \|f\|_\infty, \tag{17.13.3}$$

and that

$$f \mapsto t(f) \tag{17.13.4}$$

is a linear functional on each of $c^{lai}(X, \mathbf{R})$, $c^{lai}(X, \mathbf{C})$. More precisely, one can check that this is an algebra homomorphism from $c^{lai}(X, \mathbf{R})$, $c^{lai}(X, \mathbf{C})$ into $\mathbf{R}$, $\mathbf{C}$, respectively.

Remember that $c_0(X, \mathbf{R})$, $c_0(X, \mathbf{C})$ are the spaces of real and complex-valued functions on X that vanish at infinity, respectively, as in Section 1.13. These are commutative and associative algebras over the real and complex numbers, as appropriate, without multiplicative identity elements, because X is an infinite set. Of course, $\mathbf{1}_X$ is the multiplicative identity element in each of $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, and in particular in each of $c^{lai}(X, \mathbf{R})$, $c^{lai}(X, \mathbf{C})$.

Let us take $\mathcal{A} = c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, and let $\mathcal{A}_1$ be as in Section 6.15 and the previous section. It is easy to see that

$$(f_0, t) \mapsto f_0 + t\,\mathbf{1}_X \tag{17.13.5}$$

defines an algebra isomorphism from $\mathcal{A}_1$ onto $c^{lai}(X, \mathbf{R})$ or $c^{lai}(X, \mathbf{C})$, as appropriate.

Similarly, let X be a metric space or a Hausdorff topological space that is locally compact, as in Section 5.1, and not compact. Let us say that a real or complex-valued function f on X has a *limit at infinity* on X if f can be expressed as in (17.13.1), where $t = t(f) \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and f_0 vanishes at infinity on X as in Section 5.2. In this case, one can verify that t is unique, and one may call it the limit of f at infinity on X.

This is the same as saying that f has a limit at the point at infinity in the one-point compactification of X, and that this limit is equal to t. The previous remarks for infinite sets correspond to using the discrete metric or topology on X.

One can check that the space of real or complex-valued functions f on X with a limit t at infinity is a subalgebra of the space of all real or complex-valued functions on X, as a commutative algebra over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with respect to pointwise multiplication of functions. The mapping from f to $t(f)$ is an algebra homomorphism from the space of real or complex-valued functions on X with a limit at infinity to $\mathbf{R}$ or $\mathbf{C}$, as appropriate. These statements are analogous to standard facts about the limit of a function at a point in a topological space, and one can reduce to that case using the one-point compactification of X, as in the preceding paragraph.

If f is bounded on X, then $|t|$ is less than or equal to the supremum norm of f on X, as before. Note that $\mathbf{1}_X$ does not vanish at infinity on X, because X is not compact.

Remember that $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$ are the spaces of continuous real and complex-valued functions on X that vanish at infinity, as in Section 5.2. These are closed linear subspaces of the spaces $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ of bounded continuous real and complex-valued functions on X, as appropriate, as before.

Let

$$C^{lai}(X, \mathbf{R}),\ C^{lai}(X, \mathbf{C}) \tag{17.13.6}$$

be the spaces of continuous real or complex-valued functions f on X with a limit at infinity, as appropriate. This means that f can be expressed as in (17.13.1), where f_0 is continuous on X and vanishes at infinity. One can check that

(17.13.7) $$C^{lai}(X, \mathbf{R}),\ C^{lai}(X, \mathbf{C}) \text{ are closed subalgebras of } C_b(X, \mathbf{R}),\ C_b(X, \mathbf{C}),$$

respectively.

A continuous real or complex-valued function f on X has a limit at infinity if and only if it extends to a continuous real or complex-valued function to the one-point compactification of X, as appropriate, whose value at the point at infinity is the same as the limit of f at infinity. In this case, the supremum norm of f on X is the same as the supremum norm of its continuous extension to the one-point compactification of X.

Let us now take $\mathcal{A} = C_0(X, \mathbf{R})$ or $C_0(X, \mathbf{C})$, which is a commutative associative algebra over the real or complex numbers, as appropriate, with respect to pointwise multiplication of functions. If $\mathcal{A}_1$ is as in Section 6.15 and the previous section again, then it is easy to see that (17.13.5) defines an algebra isomorphism from $\mathcal{A}_1$ onto $C^{lai}(X, \mathbf{R})$ or $C^{lai}(X, \mathbf{C})$, as appropriate.

17.14 Some weighted ℓ^∞ spaces

Let X be a nonempty set, and let w be a positive real-valued function on X. As in Section 15.3, we let $\ell^\infty_w(X, \mathbf{R})$ and $\ell^\infty_w(X, \mathbf{C})$ be the spaces of real and complex-valued functions f on X such that

(17.14.1) $$|f(x)|\, w(x)$$

is bounded on X, with the corresponding weighted supremum norm

(17.14.2) $$\|f\|_{\infty,w} = \|f\|_{sup,w} = \sup_{x \in X} (|f(x)|\, w(x)).$$

These are the same as the usual ℓ^∞ spaces and supremum norm when $w \equiv 1$ on X, as in Section 1.1.

If c is a positive real number, then $c\,w$ is another positive real-valued function on X, and $\ell^\infty_{c\,w}(X, \mathbf{R})$, $\ell^\infty_{c\,w}(X, \mathbf{C})$ are the same as $\ell^\infty_w(X, \mathbf{R})$, $\ell^\infty_w(X, \mathbf{C})$, respectively. If f is an element of one of these spaces, then

(17.14.3) $$\|f\|_{\infty,c\,w} = c\,\|f\|_{\infty,w}.$$

Let w_1, w_2 be positive real-valued functions on X, and suppose for the moment that

(17.14.4) $$c\,w_1 \le w_2$$

on X. This implies that

(17.14.5) $$\ell^\infty_{w_2}(X, \mathbf{R}) \subseteq \ell^\infty_{w_1}(X, \mathbf{R}),\ \ell^\infty_{w_2}(X, \mathbf{C}) \subseteq \ell^\infty_{w_1}(X, \mathbf{R}),$$

with

(17.14.6) $$c\,\|f\|_{\infty,w_1} = \|f\|_{\infty,c\,w_1} \le \|f\|_{\infty,w_2}$$

for all $f \in \ell^\infty_{w_2}(X, \mathbf{R})$ or $\ell^\infty_{w_2}(X, \mathbf{C})$.

Let w_1, w_2 be any positive real-valued functions on X again, and suppose for the moment that f, g are real or complex-valued functions on X such that

$$|f|\, w_1 \text{ and } |g|\, w_2 \text{ are bounded on } X. \tag{17.14.7}$$

Under these conditions, $w_1\, w_2$ is a positive real-valued function on X, and

$$f\, g \in \ell^\infty_{w_1\, w_2}(X, \mathbf{R}) \text{ or } \ell^\infty_{w_1\, w_2}(X, \mathbf{C}), \tag{17.14.8}$$

as appropriate. We also have that

$$\|f\, g\|_{w_1\, w_2} \le \|f\|_{w_1}\, \|g\|_{w_2}. \tag{17.14.9}$$

Suppose now that

$$w \ge c \tag{17.14.10}$$

on X, so that

$$c\, w \le w^2 \tag{17.14.11}$$

on X. If $f, g \in \ell^\infty_w(x, \mathbf{R})$ or $\ell^\infty_w(X, \mathbf{C})$, then

$$f\, g \in \ell^\infty_{w^2}(X, \mathbf{R}) \text{ or } \ell^\infty_{w^2}(X, \mathbf{C}), \tag{17.14.12}$$

as appropriate, as in (17.14.8). This implies that

$$f\, g \in \ell^\infty_w(X, \mathbf{R}) \text{ or } \ell^\infty_w(X, \mathbf{C}), \tag{17.14.13}$$

as appropriate, so that $\ell^\infty_w(X, \mathbf{R})$, $\ell^\infty_w(X, \mathbf{C})$ are subalgebras of the algebras of bounded real and complex-valued functions on X, respectively. We also get that

$$c\, \|f\, g\|_{\infty, w} = \|f\, g\|_{\infty, c\, w} \le \|f\, g\|_{\infty, w^2} \le \|f\|_{\infty, w}\, \|g\|_{\infty, w}. \tag{17.14.14}$$

Suppose that $c = 1$, so that

$$w \ge 1 \tag{17.14.15}$$

on X. In this case, $\ell^\infty_w(X, \mathbf{R})$ and $\ell^\infty_w(X, \mathbf{C})$ are Banach algebras over the real and complex numbers, respectively, with respect to $\|\cdot\|_{\infty, w}$. Of course, the constant function $\mathbf{1}_X$ on X equal to 1 at every point is contained in these spaces if and only if w is bounded on X. In fact,

$$\|\mathbf{1}_X\|_{\infty, w} = 1 \tag{17.14.16}$$

if and only if $w \equiv 1$ on X, because of (17.14.15).

Let $f \in \ell^\infty_w(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$ be given, and observe that for each integer $l \ge 2$,

$$\|f\|_\infty^l = \|f^l\|_\infty \le \|f^l\|_{\infty, w} \le \|f^{l-1}\|_\infty\, \|f\|_{\infty, w} = \|f\|_\infty^{l-1}\, \|f\|_{\infty, w}. \tag{17.14.17}$$

This implies that

$$\|f\|_\infty \le \|f^l\|_{\infty, w}^{1/l} \le \|f\|_\infty^{1-(1/l)}\, \|f\|_{\infty, w}^{1/l}. \tag{17.14.18}$$

Let us take $\mathcal{A} = \ell^\infty_w(X, \mathbf{R})$ or $\ell^\infty_w(X, \mathbf{C})$, as appropriate, so that $r_{\mathcal{A}}(f)$ may be defined as in Sections 6.13 and 7.2. Using (17.14.18), we obtain that

$$r_{\mathcal{A}}(f) = \lim_{l \to \infty} \|f^l\|_{\infty, w}^{1/l} = \|f\|_\infty. \tag{17.14.19}$$

17.15 Continuous weights

Let X be a nonempty metric or topological space, and let w be a continuous positive real-valued function on X. Consider the spaces

$$C_{b,w}(X, \mathbf{R}) = C(X, \mathbf{R}) \cap \ell^\infty_w(X, \mathbf{R}), \tag{17.15.1}$$

$$C_{b,w}(X, \mathbf{C}) = C(X, \mathbf{C}) \cap \ell^\infty_w(X, \mathbf{C}) \tag{17.15.2}$$

of continuous real and complex-valued functions on X, respectively, such that (17.14.1) is bounded on X. These are linear subspaces of $C(X, \mathbf{R})$, $C(X, \mathbf{C})$ and $\ell^\infty_w(X, \mathbf{R})$, $\ell^\infty_w(X, \mathbf{C})$, respectively. Of course, these are the same as the usual spaces $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$ of bounded continuous real and complex-valued functions on X when $w \equiv 1$ on X, as in Section 1.4. If X is equipped with the discrete metric or topology, then every function on X is continuous, and $C_{b,w}(X, \mathbf{R})$, $C_{b,w}(X, \mathbf{C})$ are the same as $\ell^\infty_w(X, \mathbf{R})$, $\ell^\infty_w(X, \mathbf{C})$, respectively.

Note that

$$f \in C_{b,w}(X, \mathbf{R}) \text{ or } C_{b,w}(X, \mathbf{C}) \tag{17.15.3}$$

if and only if

$$w\, f \in C_b(X, \mathbf{R}) \text{ or } C_{b,w}(X, \mathbf{C}), \tag{17.15.4}$$

as appropriate. Remember that

$$f \mapsto w\, f \tag{17.15.5}$$

is an isometric linear mapping from each of $\ell^\infty_w(X, \mathbf{R})$, $\ell^\infty_w(X, \mathbf{C})$ onto $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$, respectively, as in Section 15.3. This mapping sends $C_{b,w}(X, \mathbf{R})$, $C_{b,w}(X, \mathbf{C})$ onto $C_b(X, \mathbf{R})$, $C_b(X, \mathbf{C})$, respectively. It follows that

(17.15.6) $C_{b,w}(X, \mathbf{R})$, $C_{b,w}(X, \mathbf{C})$ are closed subsets of $\ell^\infty_w(X, \mathbf{R})$, $\ell^\infty_w(X, \mathbf{C})$,

respectively, because of the analogous statement when $w \equiv 1$ on X, as in Section 1.4. We also have that

(17.15.7) $C_{b,w}(X, \mathbf{R})$, $C_{b,w}(X, \mathbf{C})$ are Banach spaces

with respect to the weighted supremum norm $\|f\|_{sup,w}$ in (17.14.2).

Remember that $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ are the spaces of continuous real and complex-valued functions on X with compact support, as in Section 5.1. Observe that

$$C_{com}(X, \mathbf{R}) \subseteq C_{b,w}(X, \mathbf{R}),\ C_{com}(X, \mathbf{C}) \subseteq C_{b,w}(X, \mathbf{C}), \tag{17.15.8}$$

because w is bounded on compact subsets of X. We also have that (17.15.5) maps $C_{com}(X, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$ onto themselves.

Let

$$C_{0,w}(X, \mathbf{R}),\ C_{0,w}(X, \mathbf{C}) \tag{17.15.9}$$

be the spaces of continuous real and complex-valued functions on X, respectively, such that

(17.15.10) $w\, f$ vanishes at infinity on X, as in Section 5.2. These are the same as the usual spaces $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$ of continuous real and complex-valued functions on X that vanish at infinity when $w \equiv 1$ on X. If X is equipped with the discrete metric or topology, then these are the same as $c_{0,w}(X, \mathbf{R})$, $c_{0,w}(X, \mathbf{C})$ in Section 15.3. Of course, (17.15.5) maps $C_{0,w}(X, \mathbf{R})$, $C_{0,w}(X, \mathbf{C})$ onto $C_0(X, \mathbf{R})$, $C_0(X, \mathbf{C})$, respectively, This implies that

(17.15.11) $C_{0,w}(X, \mathbf{R})$, $C_{0,w}(X, \mathbf{C})$ are closed linear subspaces of $C_{b,w}(X, \mathbf{R})$, $C_b(X, \mathbf{C})$,

respectively, because of the analogous statement when $w \equiv 1$ on X.

In fact,

(17.15.12) $C_{0,w}(X, \mathbf{R})$, $C_{0,w}(X, \mathbf{C})$ are the same as the closures of $C_{com}(X, \mathbf{R})$, $C_{com}(X, \mathbf{C})$ in $C_{b,w}(X, \mathbf{R})$, $C_{b,w}(X, \mathbf{C})$,

respectively. This follows from the analogous statement when $w \equiv 1$ on X, as in Section 5.2.

Let w_1, w_2 be continuous positive real-valued functions on X such that (17.14.4) holds for some positive real number c. This implies that

$$C_{b,w_2}(X, \mathbf{R}) \subseteq C_{b,w_1}(X, \mathbf{R}),\ C_{b,w_2}(X, \mathbf{C}) \subseteq C_{b,w_1}(X, \mathbf{C}), \tag{17.15.13}$$

as in (17.14.5). Similarly,

$$C_{0,w_2}(X, \mathbf{R}) \subseteq C_{0,w_1}(X, \mathbf{R}),\ C_{0,w_2}(X, \mathbf{C}) \subseteq C_{0,w_1}(X, \mathbf{C}). \tag{17.15.14}$$

If we also have that

$$c'\, w_2 \le w_1 \tag{17.15.15}$$

on X for some positive real number c', then we get that

$$C_{b,w_1}(X, \mathbf{R}) = C_{b,w_2}(X, \mathbf{R}),\ C_{b,w_1}(X, \mathbf{C}) = C_{b,w_1}(X, \mathbf{C}) \tag{17.15.16}$$

and

$$C_{0,w_1}(X, \mathbf{R}) = C_{0,w_2}(X, \mathbf{C}),\ C_{0,w_1}(X, \mathbf{C}) = C_{0,w_2}(X, \mathbf{C}). \tag{17.15.17}$$

Let w_1, w_2 be any continuous positive real-valued functions on X, so that

(17.15.18) $w_1\, w_2$ is a continuous positive real-valued function on X

as well. If f, g are continuous real or complex-valued functions on X such that (17.14.7) holds, then

$$f\, g \in C_{b,w_1\, w_2}(X, \mathbf{R}) \text{ or } C_{b,w_1\, w_2}(X, \mathbf{C}), \tag{17.15.19}$$

as appropriate, as in (17.14.8). If we also have that

(17.15.20) $f\, w_1$ or $g\, w_2$ vanishes at infinity on X,

as in Section 5.2, then

$$f\, g \in C_{0,w_1\, w_2}(X, \mathbf{R}) \text{ or } C_{0,w_1\, w_2}(X, \mathbf{C}), \tag{17.15.21}$$

as appropriate.

Chapter 18

Algebras, involutions, and positivity

18.1 One-sided multiplicative identity elements

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers. An element e_L of $\mathcal{A}$ is said to be a *left multiplicative identity element* in $\mathcal{A}$ if

$$e_L\, x = x \tag{18.1.1}$$

for every $x \in \mathcal{A}$. Similarly, an element e_R of $\mathcal{A}$ is said to be a *right multiplicative identity element* in $\mathcal{A}$ if

$$x\, e_R = x \tag{18.1.2}$$

for every $x \in \mathcal{A}$. Thus a multiplicative identity element in $\mathcal{A}$ is the same as an element of $\mathcal{A}$ that is both a left and right multiplicative identity element. If $\mathcal{A}$ has a left multiplicative identity element e_L and a right multiplicative element e_R, then it is easy to see that

$$e_L = e_R, \tag{18.1.3}$$

so that this is a multiplicative identity element of $\mathcal{A}$.

If $a \in \mathcal{A}$, then let

$$L_a(x) = a\, x \tag{18.1.4}$$

be the corresponding left multiplication operator on $\mathcal{A}$, as in Section 6.3. Note that $e_L \in \mathcal{A}$ is a left multiplicative identity element in $\mathcal{A}$ if and only if

$$L_{e_L} = I_{\mathcal{A}}, \tag{18.1.5}$$

the identity mapping on $\mathcal{A}$. Similarly, $e_R \in \mathcal{A}$ is a right multiplicative identity element in $\mathcal{A}$ if and only if

$$L_a(e_R) = a \tag{18.1.6}$$

for every $a \in \mathcal{A}$. If $\mathcal{A}$ has a right multiplicative identity element, then it follows that

$$a \mapsto L_a \tag{18.1.7}$$

is a one-to-one mapping from $\mathcal{A}$ into the space $\mathcal{L}(\mathcal{A})$ of all linear mappings from $\mathcal{A}$ into itself.

Suppose for the moment that $x \mapsto x^*$ is an algebra involution on $\mathcal{A}$, which may be conjugate-linear in the complex case. Observe that $e_L \in \mathcal{A}$ is a left multiplicative identity element in $\mathcal{A}$ if and only if

$$e_L^* \text{ is a right multiplicative identity element in } \mathcal{A}. \tag{18.1.8}$$

Under these conditions, we get that

$$e_L^* = e_L, \tag{18.1.9}$$

as in (18.1.3), and that e_L is a multiplicative identity element in $\mathcal{A}$. Of course, there are analogous statements for right multiplicative identity elements in $\mathcal{A}$.

Of course, (18.1.7) is a linear mapping from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$, as in Section 6.3. The kernel of this mapping is equal to

$$\mathcal{A}_L = \{a \in \mathcal{A} : a\, y = 0 \text{ for every } y \in \mathcal{A}\}. \tag{18.1.10}$$

Thus (18.1.7) is one-to-one on $\mathcal{A}$ if and only if

$$\mathcal{A}_L = \{0\}. \tag{18.1.11}$$

Note that (18.1.10) is automatically a right ideal in $\mathcal{A}$.

Suppose now that $\mathcal{A}$ is an associative algebra. Remember that (18.1.7) is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$ in this case, as in Section 6.3. This implies that (18.1.10) is a two-sided ideal in $\mathcal{A}$, as in Section 6.11, which can also be verified directly.

If $e_L \in \mathcal{A}$ is a left multiplicative identity element in $\mathcal{A}$, then

$$(x\, e_L)\, y = x\, (e_L\, y) = x\, y \tag{18.1.12}$$

for every $x, y \in \mathcal{A}$. This means that

$$(x\, e_L - x)\, y = 0 \tag{18.1.13}$$

for every $x, y \in \mathcal{A}$, which is the same as saying that

$$x\, e_L - x \in \mathcal{A}_L \tag{18.1.14}$$

for every $x \in \mathcal{A}$. If (18.1.11) holds, then we get that

$$x\, e_L = x \tag{18.1.15}$$

for every $x \in \mathcal{A}$, so that e_L is a right multiplicative identity element in $\mathcal{A}$ too. There are analogous statements for right multiplicative identity elements in $\mathcal{A}$, as usual.

Suppose that $a_0 \in \mathcal{A}$ has the property that

$$L_{a_0} \text{ is a one-to-one mapping from } \mathcal{A} \text{ onto itself.} \tag{18.1.16}$$

This implies that there is an $x_0 \in \mathcal{A}$ such that

$$a_0\, x_0 = L_{a_0}(x_0) = a_0. \tag{18.1.17}$$

It follows that

$$a_0\,(x_0\, y) = (a_0\, x_0)\, y = a_0\, y \tag{18.1.18}$$

for every $y \in \mathcal{A}$. This means that

$$x_0\, y = y \tag{18.1.19}$$

for every $y \in \mathcal{A}$, because L_{a_0} is injective, so that x_0 is a left multiplicative identity element in $\mathcal{A}$.

Of course, (18.1.16) also implies that there is a $b_0 \in \mathcal{A}$ such that

$$a_0\, b_0 = L_{a_0}(b_0) = x_0. \tag{18.1.20}$$

This means that

$$L_{a_0} \circ L_{b_0} = L_{a_0\, b_0} = L_{x_0} = I_{\mathcal{A}}. \tag{18.1.21}$$

It follows that

$$L_{a_0}^{-1} = L_{b_0}, \tag{18.1.22}$$

which is to say that we also have that

$$L_{b_0} \circ L_{a_0} = I_{\mathcal{A}}. \tag{18.1.23}$$

If (18.1.11) holds, then x_0 is the multiplicative identity element in $\mathcal{A}$, and (18.1.23) implies that

$$b_0\, a_0 = x_0. \tag{18.1.24}$$

Of course, this means that a_0 is invertible in $\mathcal{A}$, with $a_0^{-1} = b_0$.

18.2 Some subalgebras of $\mathcal{L}(\mathcal{A})$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers again, and let us continue to use the notation in (18.1.4) and (18.1.10). Put

$$\mathcal{B} = \{L_a : a \in \mathcal{A}\}. \tag{18.2.1}$$

This is a subalgebra of $\mathcal{L}(\mathcal{A})$, as an algebra with respect to composition of linear mappings on $\mathcal{A}$, because (18.1.7) is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$.

Observe that

$$I_{\mathcal{A}} \in \mathcal{B} \tag{18.2.2}$$

if and only if $\mathcal{A}$ has a left multiplicative identity element. If (18.1.11) holds, then this is the same as saying that $\mathcal{A}$ has a multiplicative identity element, as in the previous section.

Let $\mathcal{B}_1$ be the subset of $\mathcal{L}(\mathcal{A})$ consisting of linear mappings of the form

$$L_a + t\, I_{\mathcal{A}}, \tag{18.2.3}$$

where $a \in \mathcal{A}$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. This is the same as the linear subspace of $\mathcal{L}(\mathcal{A})$ spanned by $\mathcal{B}$ and $I_{\mathcal{A}}$. Thus

$$\mathcal{B} = \mathcal{B}_1 \tag{18.2.4}$$

if and only if (18.2.2) holds. It is easy to see that

$$\mathcal{B}_1 \text{ is a subalgebra of } \mathcal{L}(\mathcal{A}) \tag{18.2.5}$$

too. In fact, $\mathcal{B}$ is a two-sided ideal in $\mathcal{B}_1$, as in Section 6.11.

If

$$L_a + t\, I_{\mathcal{A}} = 0 \tag{18.2.6}$$

for some $a \in \mathcal{A}$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $t \neq 0$, then (18.2.2) holds. If (18.2.2) does not hold, then

$$L_a + t\, I_{\mathcal{A}} \mapsto t \tag{18.2.7}$$

is a well-defined linear mapping from $\mathcal{B}_1$ onto $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with kernel $\mathcal{B}$. More precisely, this is an algebra homomorphism from $\mathcal{B}_1$ onto $\mathbf{R}$ or $\mathbf{C}$, as appropriate, in this case.

Let $\mathcal{A}_1 = \mathcal{A} \times \mathbf{R}$ or $\mathcal{A} \times \mathbf{C}$, as appropriate, be defined as an associative algebra over the real or complex numbers, as appropriate, as in Section 6.15. Consider the mapping from $\mathcal{A}_1$ into $\mathcal{L}(\mathcal{A})$ defined by

$$(a, t) \mapsto L_a + t\, I_{\mathcal{A}}. \tag{18.2.8}$$

One can check that this defines an algebra homomorphism from $\mathcal{A}_1$ into $\mathcal{L}(\mathcal{A})$. Of course, $\mathcal{B}_1$ is the same as the image of $\mathcal{A}_1$ under this homomorphism.

Suppose that (18.1.11) holds, so that (18.1.7) is one-to-one on $\mathcal{A}$. If the kernel of (18.2.8) is nontrivial, then (18.2.6) holds with $t \neq 0$, so that (18.2.2) holds, as before. Otherwise, if (18.2.2) does not hold, then it follows that (18.2.8) is an algebra isomorphism from $\mathcal{A}_1$ onto $\mathcal{B}_1$.

Suppose that $\mathcal{A} \neq \{0\}$, so that $I_{\mathcal{A}} \neq \{0\}$. If (18.2.2) holds, then it is easy to see that the kernel of (18.2.8) is nontrivial.

18.3 Another condition on L_a

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $a \in \mathcal{A}$, then let L_a be as in (18.1.4), which is a bounded linear mapping from $\mathcal{A}$ into itself, with

$$\|L_a\|_{op} \leq \|a\|_{\mathcal{A}}, \tag{18.3.1}$$

as in Section 6.3. Thus (18.2.1) defines a linear subspace of the space $\mathcal{BL}(\mathcal{A})$ of all bounded linear mappings from $\mathcal{A}$ into itself.

In this section, we ask that there be a nonnegative real number C such that

$$\|a\|_{\mathcal{A}} \le C\, \|L_a\|_{op} \tag{18.3.2}$$

for every $a \in \mathcal{A}$. Of course, this implies in particular that $a \mapsto L_a$ is a one-to-one linear mapping from $\mathcal{A}$ into $\mathcal{BL}(\mathcal{A})$, so that (18.1.11) holds, as before. If $\mathcal{A}$ has a right multiplicative identity element e_R, then

$$\|a\|_{\mathcal{A}} = \|L_a(e_R)\|_{\mathcal{A}} \le \|L_a\|_{op}\, \|e_R\|_{\mathcal{A}} \tag{18.3.3}$$

for every $a \in \mathcal{A}$, as in Section 6.3.

Suppose that $\mathcal{A}$ is complete with respect to the metric associated to $\|\cdot\|_{\mathcal{A}}$. This implies that that $\mathcal{B}$ is complete with respect to the restriction to $\mathcal{B}$ of the metric associated to the operator norm on $\mathcal{BL}(\mathcal{A})$, because (18.3.1) and (18.3.2). This means that

$$\mathcal{B} \text{ is a closed set in } \mathcal{BL}(\mathcal{A}), \tag{18.3.4}$$

as in Section 1.6.

Suppose that

$$I_{\mathcal{A}} \notin \mathcal{B}. \tag{18.3.5}$$

This implies that there is a positive real number c_1 such that

$$\|L_a + I_{\mathcal{A}}\|_{op} \ge c_1 \tag{18.3.6}$$

for every $a \in \mathcal{A}$, because of (18.3.4). It follows that

$$\|L_a + t\, I_{\mathcal{A}}\|_{op} \ge c_1\, |t| \tag{18.3.7}$$

for every $a \in \mathcal{A}$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate.

Let $\mathcal{B}_1$ be the linear subspace of $\mathcal{BL}(\mathcal{A})$ spanned by $\mathcal{B}$ and $I_{\mathcal{A}}$, which is equivalent to the definition of $\mathcal{B}_1$ in the previous section in this case. As before, (18.2.7) defines a linear functional on $\mathcal{B}_1$ with kernel equal to $\mathcal{B}$. Let us check that

$$\mathcal{B}_1 \text{ is a closed set in } \mathcal{BL}(\mathcal{A}) \tag{18.3.8}$$

under these conditions.

Let $\{a_j\}_{j=1}^{\infty}$ and $\{t_j\}_{j=1}^{\infty}$ be sequences of elements of $\mathcal{A}$ and $\mathbf{R}$ or $\mathbf{C}$, as appropriate, respectively. Suppose that the corresponding sequence

$$\{L_{a_j} + t_j\, I_{\mathcal{A}}\}_{j=1}^{\infty} \tag{18.3.9}$$

of elements of $\mathcal{B}_1$ converges to an element of $\mathcal{BL}(\mathcal{A})$. In particular, this implies that (18.3.9) is a Cauchy sequence in $\mathcal{BL}(\mathcal{A})$. This means that $\{t_j\}_{j=1}^{\infty}$ is a Cauchy sequence in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, because of (18.3.7). It follows that $\{t_j\}_{j=1}^{\infty}$ converges to $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, because $\mathbf{R}$ and $\mathbf{C}$ are complete with respect to their standard metrics.

This implies that $\{L_{a_j}\}_{j=1}^\infty$ converges in $\mathcal{BL}(\mathcal{A})$, because (18.3.9) converges in $\mathcal{BL}(\mathcal{A})$. In fact, $\{L_{a_j}\}_{j=1}^\infty$ converges to an element of $\mathcal{B}$, because of (18.3.4). This means that (18.3.9) converges to an element of $\mathcal{B}_1$.

Suppose now that $\mathcal{A}$ is associative as an algebra over the real or complex numbers, so that $\mathcal{B}$ and $\mathcal{B}_1$ are subalgebras of $\mathcal{BL}(\mathcal{A})$, as in the previous section. In this case, (18.3.6) and thus (18.3.7) hold with $c_1 = 1$. Otherwise, there would be an element of $\mathcal{B}$ that is invertible as an element of $\mathcal{B}_1$, as in Section 6.5. This is not possible, because (18.2.7) is an algebra homomorphism from $\mathcal{B}_1$ onto $\mathbf{R}$ or $\mathbf{C}$, as appropriate.

18.4 An involution on $\mathcal{B}_1$

Let $\mathcal{A}$ be an associative algebra over the real or complex with an involution $x \mapsto x^*$, which may be conjugate-linear in the complex case. If $a \in \mathcal{A}$, then let L_a be as in (18.1.4), and remember that $a \mapsto L_a$ is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$, as in Section 6.3. Suppose that

$$a \mapsto L_a \text{ is one-to-one on } \mathcal{A}, \tag{18.4.1}$$

which is the same as saying that (18.1.11) holds, as before. Let $\mathcal{B}$, $\mathcal{B}_1$ be the subalgebras of $\mathcal{L}(\mathcal{A})$ defined in Section 18.2.

Suppose from now on in this section that

$$\mathcal{A} \text{ does not have a multiplicative identity element.} \tag{18.4.2}$$

This implies that

$$\mathcal{A} \text{ does not have a left multiplicative identity element,} \tag{18.4.3}$$

as in Section 18.1. It follows that (18.3.5) holds, because of (18.4.3), as in Section 18.2. This means that (18.2.7) is an algebra homomorphism from $\mathcal{B}_1$ onto $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with kernel $\mathcal{B}$, as before.

If $a \in \mathcal{A}$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then put

$$(L_a + t\, I_{\mathcal{A}})^* = L_{a^*} + t\, I_{\mathcal{A}} \tag{18.4.4}$$

in the real case, and in the complex case when the involution on $\mathcal{A}$ is complex-linear, and

$$(L_a + t\, I_{\mathcal{A}})^* = L_{a^*} + \bar{t}\, I_{\mathcal{A}} \tag{18.4.5}$$

in the complex case when the involution on $\mathcal{A}$ is conjugate-linear. One can check that this defines an algebra involution on $\mathcal{B}_1$, which is conjugate-linear in the complex case when the involution on $\mathcal{A}$ is conjugate-linear.

Remember that $\mathcal{A}_1 = \mathcal{A} \times \mathbf{R}$ or $\mathcal{A} \times \mathbf{C}$, as appropriate, may be defined as an associative algebra over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 6.15, and that (18.2.8) defines an algebra isomorphism from $\mathcal{A}_1$ onto $\mathcal{B}_1$ under these conditions. The involution on $\mathcal{B}_1$ defined in the preceding paragraph corresponds exactly to the one defined on $\mathcal{A}_1$ as in Section 17.12 with respect to this isomorphism.

If $x \in \mathcal{A}$, then

$$(L_a + t\, I_{\mathcal{A}})(x) = a\, x + t\, x. \tag{18.4.6}$$

If $y \in \mathcal{A}$, then

$$x^* \, (L_a)^*(y) = x^* \, L_{a^*}(y) = x^* \, a^* \, y = (a\, x)^* \, y. \tag{18.4.7}$$

Using this, it is easy to see that

$$x^* \, ((L_a + t\, I_{\mathcal{A}})^*(y)) = (a\, x + t\, x)^* \, y. \tag{18.4.8}$$

This implies that

$$x^* \, ((L_a + t\, I_{\mathcal{A}})^*((L_a + t\, I_{\mathcal{A}})(x))) = (a\, x + t\, x)^* \, (a\, x + t\, x). \tag{18.4.9}$$

18.5 Multiplication operators on C^* algebras

Let $\mathcal{A}$ be an algebra in the strict sense over the real or complex numbers with a submultiplicative norm $\|\cdot\|_{\mathcal{A}}$. If $a \in \mathcal{A}$, then we let L_a be as in (18.1.4) again, which is a bounded linear mapping from $\mathcal{A}$ into itself, as before.

Let $x \mapsto x^*$ be an algebra involution on $\mathcal{A}$, which may be conjugate-linear in the complex case, and suppose that $\|\cdot\|_{\mathcal{A}}$ satisfies the C^* identity

$$\|x^* \, x\|_{\mathcal{A}} = \|x\|_{\mathcal{A}}^2 \tag{18.5.1}$$

for every $x \in \mathcal{A}$, as in Section 7.7. Remember that this implies that

$$\|x^*\|_{\mathcal{A}} = \|x\|_{\mathcal{A}} \tag{18.5.2}$$

for every $x \in \mathcal{A}$, so that

$$\|x\, x^*\|_{\mathcal{A}} = \|x^*\|_{\mathcal{A}}^2 = \|x\|_{\mathcal{A}} \, \|x^*\|_{\mathcal{A}} \tag{18.5.3}$$

for every $x \in \mathcal{A}$. If $a \in \mathcal{A}$, then

$$\|L_a(a^*)\|_{\mathcal{A}} = \|a\, a^*\|_{\mathcal{A}} = \|a\|_{\mathcal{A}} \, \|a^*\|_{\mathcal{A}}. \tag{18.5.4}$$

It follows that

$$\|L_a\|_{op} = \|a\|_{\mathcal{A}} \tag{18.5.5}$$

for every $a \in \mathcal{A}$ under these conditions, because of (18.3.1). Note that $\mathcal{A}$ is not asked to have a left or right multiplicative identity element here.

Of course, (18.5.5) implies that

(18.5.6) $\quad a \mapsto L_a$ is a one-to-one linear mapping from $\mathcal{A}$ into $\mathcal{BL}(\mathcal{A})$,

so that (18.1.11) holds, as in Section 18.3. Suppose now that $\mathcal{A}$ is associative as an algebra over the real or complex numbers, and let $\mathcal{B}$, $\mathcal{B}_1$ be as in Section 18.2. Suppose also that (18.4.2) holds, so that we can define an involution on $\mathcal{B}_1$ as in the previous section.

We would like to show that

$$\|(L_a + t\,I_{\mathcal{A}})^* \circ (L_a + t\,I_{\mathcal{A}})\|_{op} = \|L_a + t\,I_{\mathcal{A}}\|_{op}^2 \tag{18.5.7}$$

for every $a \in \mathcal{A}$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. To do this, it sufices to verify that

$$\|(L_a + t\,I_{\mathcal{A}})\|_{op}^2 \le \|(L_a + t\,I_{\mathcal{A}})^* \circ (L_a + t\,I_{\mathcal{A}})\|_{op}, \tag{18.5.8}$$

as in Section 7.7. Let $x \in \mathcal{A}$ be given, and let us check that

$$\|(L_a + t\,I_{\mathcal{A}})(x)\|_{\mathcal{A}}^2 \le \|(L_a + t\,I_{\mathcal{A}})^* \circ (L_a + t\,I_{\mathcal{A}})\|_{op}\,\|x\|_{\mathcal{A}}^2. \tag{18.5.9}$$

Observe that

$$\|(L_a + t\,I_{\mathcal{A}})(x)\|_{\mathcal{A}}^2 = \|a\,x + t\,x\|_{\mathcal{A}}^2 = \|(a\,x + t\,x)^*\,(a\,x + t\,x)\|_{\mathcal{A}}, \tag{18.5.10}$$

using (18.4.6) in the first step, and (18.5.1) in the second step. We can use (18.4.9) and (18.5.2) to get that

$$\begin{aligned} \|(a\,x + t\,x)^*\,(a\,x + t\,x)\|_{\mathcal{A}} &= \|x^*\,((L_a + t\,I_{\mathcal{A}})^*((L_a + t\,a)(x)))\|_{\mathcal{A}} \\ &\le \|x^*\|_{\mathcal{A}}\,\|(L_a + t\,I_{\mathcal{A}})^*((L_a + t\,I_{\mathcal{A}})(x))\|_{\mathcal{A}} \\ &\le \|(L_a + t\,I_{\mathcal{A}})^* \circ (L_a + t\,I_{\mathcal{A}})\|_{op}\,\|x\|_{\mathcal{A}}^2. \end{aligned} \tag{18.5.11}$$

This corresponds to Proposition 2.9.2 on p75 of [8], some remarks on p264 of [167], and Exercise 5 on p300 of [167].

18.6 Positive linear functionals

Let $\mathcal{A}$ be an associative algebra over the complex numbers with a conjugate-linear involution

$$y \mapsto y^*. \tag{18.6.1}$$

Let us say that a linear functional λ on $\mathcal{A}$, as a vector space over the complex numbers, is *positive in the strict sense* with respect to this involution if for every $x \in \mathcal{A}$, we have that

$$\lambda(x^*\,x) \in \mathbf{R} \tag{18.6.2}$$

and

$$\lambda(x^*\,x) \ge 0. \tag{18.6.3}$$

Of course, this is the same as saying that

$$\lambda(x\,x^*) \in \mathbf{R} \tag{18.6.4}$$

and

$$\lambda(x\,x^*) \ge 0 \tag{18.6.5}$$

for every $x \in \mathcal{A}$, by replacing x with x^*. One may also say that λ is a *positive functional in the strict sense* on $\mathcal{A}$ in this case. This corresponds to Definition 21.16 on p316 of [91], and some other versions of this are mentioned at the beginning of Section 4.7 on p122 of [8], in Definition 11.30 on p283 of [162], and at the beginning of Section 4 in Chapter VI on p270 of [167].

If λ is any linear functional on $\mathcal{A}$ that satisfies (18.6.2) for every $x \in \mathcal{A}$, then

$$\lambda(x^* y) = \overline{\lambda(y^* x)} \tag{18.6.6}$$

for every $x, y \in \mathcal{A}$, as in part (i) of Theorem 21.17 on p316 of [91]. Indeed, if $x, y \in \mathcal{A}$ and $t \in \mathbf{C}$, then

$$\lambda((x + t\, y)^* \, (x + t\, y)) = \lambda(x^* x) + t\, \lambda(x^* y) + \bar{t}\, \lambda(y^* x) + |t|^2 \lambda(y^* y). \tag{18.6.7}$$

This implies that

$$t\, \lambda(x^* y) + \bar{t}\, \lambda(y^* x) \in \mathbf{R}, \tag{18.6.8}$$

because of (18.6.2). One can get (18.6.6) from this by taking $t = 1, i$. Note that (18.6.6) implies (18.6.2), by taking $y = x$.

If λ is any linear functional on $\mathcal{A}$, then

$$b_\lambda(x, y) = \lambda(y^* x) \tag{18.6.9}$$

defines a sesquilinear form on $\mathcal{A}$, as in Section 5.15. The condition (18.6.6) is the same as saying that (18.6.9) is Hermitian symmetric on $\mathcal{A}$, as in Section 8.5. In this case, (18.6.3) is the same as saying that (18.6.9) is nonnegative as a Hermitian form on $\mathcal{A}$, as in Section 8.10. This implies that

$$|b_\lambda(x, y)| \le b_\lambda(x, x)^{1/2}\, b_\lambda(y, y)^{1/2} \tag{18.6.10}$$

for every $x, y \in \mathcal{A}$, as before. Equivalently, this means that

$$|\lambda(y^* x)| \le \lambda(x^* x)^{1/2}\, \lambda(y^* y)^{1/2} \tag{18.6.11}$$

for every $x, y \in \mathcal{A}$. This corresponds to part (ii) of Theorem 21.17 on p316 of [91]. This also corresponds to part of Proposition 4.7.1 on p123 of [8], part (b) of Theorem 11.31 on p284 of [162], and a remark in the proof of 4.1 on p271 of [167].

Now let $\mathcal{A}$ be an associative algebra over the real numbers with an involution as in (18.6.1). Let us say that a linear functional on $\mathcal{A}$, as a vector space over the real numbers, is *positive in the strict sense* with respect to the involution if

$$\lambda(x^* y) = \lambda(y^* x) \tag{18.6.12}$$

for every $x, y \in \mathcal{A}$, and if (18.6.3) holds for every $x \in \mathcal{A}$. As before, we may also say that λ is a *positive functional in the strict sense* on $\mathcal{A}$ under these conditions.

If λ is any linear functional on $\mathcal{A}$, then (18.6.9) defines a bilinear form on $\mathcal{A}$, as in Section 5.15. Of course, (18.6.12) is the same as saying that

$$b_\lambda \text{ is symmetric on } \mathcal{A}, \tag{18.6.13}$$

as in Section 8.5. If we have this, then (18.6.3) is the same as saying that b_λ is nonnegative as a symmetric bilinear form on $\mathcal{A}$, as in Section 8.10. This implies that (18.6.10) holds, which is the same as saying that (18.6.11) holds, as before.

18.7 Positive functionals and identity elements

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with an involution $y \mapsto y^*$ that is conjugate-linear in the complex case. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then $e_{\mathcal{A}}^* = e_{\mathcal{A}}$, as in Section 6.4. If λ is a linear functional on $\mathcal{A}$ that satisfies (18.6.6) or (18.6.12), as appropriate, then we can take $y = e_{\mathcal{A}}$ to get that

$$\lambda(x^*) = \overline{\lambda(x)} \tag{18.7.1}$$

or

$$\lambda(x^*) = \lambda(x), \tag{18.7.2}$$

as appropriate, for every $x \in \mathcal{A}$. Conversely, it is easy to see that each of these conditions implies (18.6.6) or (18.6.12), as appropriate.

If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, and if λ is a positive linear functional on $\mathcal{A}$ in the strict sense, then we can take $y = e_{\mathcal{A}}$ in (18.6.11) to get that

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})^{1/2}\, \lambda(x^*\, x)^{1/2} \tag{18.7.3}$$

for every $x \in \mathcal{A}$. Note that

$$\lambda(e_{\mathcal{A}}) \ge 0, \tag{18.7.4}$$

by taking $x = e_{\mathcal{A}}$ in (18.6.3).

Suppose that we do not ask that $\mathcal{A}$ have a multiplicative identity element, and that λ is a positive linear functional on $\mathcal{A}$ in the strict sense with respect to the involution. Let us say that λ is *positive in the strong sense* with respect to the involution if (18.7.1) or (18.7.2) holds for every $x \in \mathcal{A}$, as appropriate, and there is a nonnegative real number c such that

$$|\lambda(x)| \le c^{1/2}\, \lambda(x^*\, x)^{1/2} \tag{18.7.5}$$

for every $x \in \mathcal{A}$. We may also say that λ is a *positive functional in the strong sense* on $\mathcal{A}$ with respect to the involution in this case. If $\mathcal{A}$ has a multiplicative identity element, then positivity in the strict sense implies positivity in the strong sense, as before, and we may simply say that λ is a *positive linear functional* or *positive functional* on $\mathcal{A}$ with respect to the involution.

In [8, 162], positive linear functionals are considered on algebras with multiplicative identity elements. In [167], positivity is defined for bounded linear functionals on complex C^* algebras that may not have a multiplicative identity element, and (18.7.1) is basically included in the definition. If the algebra does not have a multiplicative identity element, then a positive linear functional on the algebra may be extended to a positive linear functional on the unitization of the algebra, as in Corollary 1 on p271 of [167].

In the complex case, it is easy to see that (18.7.1) holds if and only if

$$\lambda(x) \in \mathbf{R} \tag{18.7.6}$$

for every $x \in \mathcal{A}$ that is self-adjoint. In the real case, (18.7.2) holds if and only if

$$\lambda(x) = 0 \tag{18.7.7}$$

for every $x \in \mathcal{A}$ that is anti-self-adjoint, as in Section 7.5.

18.8 Positive linear functionals on $\mathcal{A}_1$

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with an involution $y \mapsto y^*$ that is conjugate-linear in the complex case again, and let

$$\mathcal{A}_1 = \mathcal{A} \times \mathbf{R} \text{ or } \mathcal{A} \times \mathbf{C}, \tag{18.8.1}$$

as appropriate, be defined as an associative algebra over the real or complex numbers, as appropriate, as in Section 6.15. Also let

$$(x,t) \mapsto (x,t)^* \tag{18.8.2}$$

be the involution on $\mathcal{A}_1$ defined in Section 17.12. This involution on $\mathcal{A}_1$ is conjugate-linear in the complex case, because of the correspoinding property of the involution on $\mathcal{A}$.

Remember that $e_{\mathcal{A}_1} = (0,1)$ is the multiplicative identity element in $\mathcal{A}_1$. If λ_1 is a positive linear functional on $\mathcal{A}_1$, then

$$\lambda(x) = \lambda_1((x,0)) \tag{18.8.3}$$

defines a positive linear functional on $\mathcal{A}$ in the strong sense. More precisely, (18.7.5) holds in this case with

$$c = \lambda_1(e_{\mathcal{A}_1}) = \lambda_1((0,1)). \tag{18.8.4}$$

This corresponds to the "only if" part of Theorem 21.18 on p317 of [91].

Conversely, suppose that λ is a positive linear functional on $\mathcal{A}$ in the strong sense, and let c be a nonnegative real number such that (18.7.5) holds for every $x \in \mathcal{A}$. Put

$$\lambda_1((x,t)) = \lambda(x) + t\,c \tag{18.8.5}$$

for every $(x,t) \in \mathcal{A}_1$, which defines a linear functional on $\mathcal{A}_1$. Clearly

$$\lambda_1((x,t)^*) = \overline{\lambda_1((x,t))} \tag{18.8.6}$$

or

$$\lambda_1((x,t)^*) = \lambda_1((x,t)), \tag{18.8.7}$$

as appropriate, for each $(x,t) \in \mathcal{A}_1$, because of (18.7.1) or (18.7.2), as appropriate.

If $(x,t) \in \mathcal{A}_1$, then

$$(x,t)^*\,(x,t) = (x^*\,x + t\,x^* + \overline{t}\,x, |t|^2) \tag{18.8.8}$$

in the complex case, and

$$(x,t)^*\,(x,t) = (x^*\,x + t\,x^* + t\,x, t^2) \tag{18.8.9}$$

in the real case. This implies that

$$\lambda_1((x,t)^*\,(x,t)) = \lambda(x^*\,x) + 2\,\mathrm{Re}(\overline{t}\,\lambda(x)) + |t|^2\,c \tag{18.8.10}$$

in the complex case, and

$$\lambda((x,t)^*\,(x,t)) = \lambda(x^*\,x) + 2\,t\,\lambda(x) + t^2 c \tag{18.8.11}$$

in the real case, using (18.7.1) and (18.7.2), as appropriate. It follows that

$$\begin{aligned} \lambda((x,t)^*\,(x,t)) &\geq \lambda(x^*\,x) - 2\,|t|\,|\lambda(x)| + |t|^2\,c \\ &\geq \lambda(x^*\,x) - 2\,|t|\,c^{1/2} + |t|^2\,c \\ &= (\lambda(x^*\,x)^{1/2} - |t|\,c^{1/2})^2 \geq 0 \end{aligned} \tag{18.8.12}$$

in both cases, using (18.7.5) in the second step. This corresponds to the "if" part of Theorem 21.18 on p317 of [91].

18.9 Another nonnegativity condition

Let X be a nonempty metric or topological space, and remember that

$$C_{com}(X,\mathbf{R}),\ C_{com}(X,\mathbf{C}) \tag{18.9.1}$$

are the spaces of continuous real and complex-valued functions on X with compact support, respectively, as in Section 5.1. A linear functional λ on either of these spaces is said to be *nonnegative* if

$$\lambda(f) \geq 0 \tag{18.9.2}$$

for every nonnegative real-valued continuous function f on X with compact support. In the complex case, this includes the condition that

$$\lambda(f) \in \mathbf{R} \tag{18.9.3}$$

for all such functions f.

Of course, $C_{com}(X,\mathbf{R})$ and $C_{com}(X,\mathbf{C})$ are commutative associative algebras over the real and complex numbers, respectively, with respect to pointwise multiplication of functions on X. We may also consider these algebras to be equipped with involutions, where we take the involution on $C_{com}(X,\mathbf{R})$ to be the identity mapping, and the involution on $C_{com}(X,\mathbf{C})$ to be defined by taking the complex conjugate

$$\overline{f} \tag{18.9.4}$$

of $f \in C_{com}(X,\mathbf{C})$.

In both cases, a nonnegative linear functional in the sense considered here is a positive linear functional in the strict sense as in Section 18.6, with respect to the corresponding involution. Conversely, every positive linear functional in the strict sense on these algebras is nonnegative in the sense considered here. This uses the fact that if f is a nonnegative real-valued continuous function on X with compact support, then its square root

$$f^{1/2} \tag{18.9.5}$$

is a nonnegative real-valued continuous function on X with the same support as f.

Note that (18.7.2) holds trivially for any linear functional on $C_{com}(X, \mathbf{R})$, because the involution is taken to be the identity mapping on $C_{com}(X, \mathbf{R})$. In the complex case, (18.7.1) is equivalent to the condition that (18.9.3) hold when

(18.9.6) f is a continuous real-valued function on X with compact support,

as in (18.7.6).

If f is any continuous real-valued function on X, then

$$f_+ = \max(f, 0) \tag{18.9.7}$$

and

$$f_- = \max(-f, 0) \tag{18.9.8}$$

are nonnegative continuous real-valued functions on X. Note that

$$f = f_+ - f_- \tag{18.9.9}$$

on X. If f has compact support in X, then

(18.9.10) f_+, f_- have compact supports as well.

This uses the well-known fact that a closed set in X that is contained in a compact set is compact too. If λ is a linear functional on $C_{com}(X, \mathbf{C})$ that satisfies (18.9.3) when f is a nonnegative real-valued continuous function on X with compact support, then it follows that (18.9.3) holds for every continuous real-valued function f on X with compact support.

18.10 Nonnegativity and local compactness

Let X be a nonempty metric or topological space again, and let λ be a nonnegative linear functional on $C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$. Thus λ is a positive linear functional in the strict sense on $C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$, as appropriate, as a commutative associative algebra over the real or complex numbers, as appropriate, with respect to the involution mentioned in the previous section. We have also seen that λ satisfies one of the conditions needed to be a positive linear functional in the strong sense, as in Section 18.7. The other condition may be restated as saying that there is a nonnegative real number c such that

$$|\lambda(f)| \le c^{1/2}\, \lambda(|f|^2)^{1/2} \tag{18.10.1}$$

for every $f \in C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$, as appropriate.

Suppose for the moment that this condition holds, and let a be a continuous real-valued function on X with compact support such that

$$0 \le a \le 1 \tag{18.10.2}$$

on X. This implies that

$$a^2 \le a \tag{18.10.3}$$

on X, so that $a - a^2 \ge 0$ on X, and thus

$$\lambda(a - a^2) \ge 0. \tag{18.10.4}$$

This means that $\lambda(a^2) \le \lambda(a)$, which we can use in (18.10.1) to get that

$$\lambda(a) \le c. \tag{18.10.5}$$

Suppose now that X is a nonempty metric space or Hausdorff topological space that is locally compact, as in Section 5.1. If f is a continuous real or complex-valued function on X with compact support, then the version of Urysohn's lemma mentioned in Section 5.1 implies that there is a continuous real-valued function b on X with compact support such that

$$0 \le b \le 1 \tag{18.10.6}$$

on X and

$$b = 1 \text{ on } \operatorname{supp} f. \tag{18.10.7}$$

This means that

$$f = b\, f \tag{18.10.8}$$

on X, so that

$$|\lambda(f)| = |\lambda(b\, f)| \le \lambda(b^2)^{1/2}\, \lambda(|f|^2)^{1/2}, \tag{18.10.9}$$

where the second step is as in Section 18.6. If (18.10.5) holds for some $c \ge 0$ and all continuous real-valued functions a on X with compact support that satisfy (18.10.2) on X, then we can take $a = b^2$ to get that (18.10.1) holds. This implies that λ is positive in the strong sense on $C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$, as appropriate, as in Section 18.7.

Let μ be a nonnegative Borel measure on X such that

$$\mu(K) < +\infty \tag{18.10.10}$$

for all compact subsets K of X. This implies that continuous real and complex-valued functions on X with compact support are integrable on X with respect to μ, so that

$$\lambda_\mu(f) = \int_X f\, d\mu \tag{18.10.11}$$

defines a nonnegative linear functional on $C_{com}(X, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$. If

$$\mu(X) < \infty, \tag{18.10.12}$$

then

$$\left|\int_X f\, d\mu\right| \le \int_X |f|\, d\mu \le \mu(X)^{1/2} \Big(\int_X |f|^2\, d\mu\Big)^{1/2} \tag{18.10.13}$$

for all square-integrable functions f on X with respect to μ, by the Cauchy–Schwarz inequality. In this case, we get that λ_μ is positive in the strong sense on $C_{com}(X, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$, as in Section 18.7. Of course, if X is compact, then (18.10.12) follows from (18.10.10). If λ is any nonnegative linear functional on $C_{com}(X, \mathbf{R})$ or $C_{com}(X, \mathbf{C})$, then a version of another type of *Riesz representation theorem* states that there is a unique nonnegative Borel measure μ on X with suitable regularity properties such that (18.10.10) holds, and λ can be expressed as in (18.10.11). If there is a nonnegative real number c such that (18.10.5) holds for all continuous real-valued functions a on X that satisfy (18.10.2), then

$$\mu(X) \leq c. \tag{18.10.14}$$

This is all much simpler when X is equipped with the discrete metric or topology, so that $C_{com}(C, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$ are the same as $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$, respectively, as mentioned in Section 5.1. If ρ is a nonnegative real-valued function on X, then

$$\lambda_\rho(f) = \sum_{x \in X} f(x) \rho(x) \tag{18.10.15}$$

defines a nonnegative linear functional on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$. If ρ is summable on X, then one can use the Cauchy–Schwarz inequality for sums to get that λ_ρ is positive in the strong sense on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$, as before. Conversely, if λ_ρ is positive in the strong sense on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, then ρ is summable on X, with sum less than or equal to the corresponding constant c. Any nonnegative linear functional on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ can be expressed as (18.10.15) for a unique nonnegative real-valued function ρ on X.

18.11 Positivity on Banach algebras

Let $(\mathcal{A}, \| \cdot \|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and an involution $y \mapsto y^*$ that is conjugate linear in the complex case. Also let λ be a positive linear functional on $\mathcal{A}$, as in Section 18.7. Suppose that $w \in \mathcal{A}$ has the property that

$$e_{\mathcal{A}} - w = v^* \, v \tag{18.11.1}$$

for some $x \in \mathcal{A}$. This implies that

$$\lambda(e_{\mathcal{A}}) - \lambda(w) = \lambda(e_{\mathcal{A}} - w) = \lambda(v^* \, v) \geq 0, \tag{18.11.2}$$

so that

$$\lambda(w) \leq \lambda(e_{\mathcal{A}}). \tag{18.11.3}$$

Note that

$$w^* = w \tag{18.11.4}$$

in this case, because $e_{\mathcal{A}}$ and $v \, v^*$ are self-adjoint.

If

$$\|w\|_{\mathcal{A}} \le 1, \tag{18.11.5}$$

then $b_{1/2}(-w) \in \mathcal{A}$ may be defined as in Section 17.2, and

$$b_{1/2}(-w)^2 = e_{\mathcal{A}} - w. \tag{18.11.6}$$

Suppose that

$$\|y^*\|_{\mathcal{A}} \le C\,\|y\|_{\mathcal{A}} \tag{18.11.7}$$

for some $C \ge 1$ and all $y \in \mathcal{A}$, so that the involution is continuous on $\mathcal{A}$. If w satisfies (18.11.4) too, then

$$b_{1/2}(-w)^* = b_{1/2}(-w), \tag{18.11.8}$$

so that (18.11.1) holds, with $v = b_{1/2}(-w)$. It follows that (18.11.3) holds, as before. If w is any self-adjoint element of $\mathcal{A}$, then we get that

$$|\lambda(w)| \le \lambda(e_{\mathcal{A}})\,\|w\|_{\mathcal{A}}, \tag{18.11.9}$$

by multiplying w by $\pm 1/\|w\|_{\mathcal{A}}$ when $w \ne 0$.

If x is any element of $\mathcal{A}$, then

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})^{1/2}\,\lambda(x^*\,x)^{1/2} \le \lambda(e_{\mathcal{A}})\,\|x^*\,x\|_{\mathcal{A}}^{1/2}, \tag{18.11.10}$$

using (18.7.3) in the first step, and (18.11.9) in the second step. This implies that

$$|\lambda(x)| \le C^{1/2}\,\lambda(e_{\mathcal{A}})\,\|x\|_{\mathcal{A}}, \tag{18.11.11}$$

because of (18.11.7). This corresponds to part of Proposition 4.7.1 on p123 of[8], and to Theorem 21.19 on p317 of [91], when $C = 1$. This also corresponds to part of part (e) of Theorem 11.31 on p284 of [162], and it is related to 4.1 and Corollary 1 on p271 of [167].

If $z \in \mathcal{A}$ is self-adjoint, then

$$|\lambda(z)| \le \lambda(e_{\mathcal{A}})^{1/2}\,\lambda(z^2)^{1/2}, \tag{18.11.12}$$

because of (18.7.3). If $x \in \mathcal{A}$ and l is a positive integer, then we can take $z = (x^*\,x)^l$, to get that

$$\lambda((x^*\,x)^l) \le \lambda(e_{\mathcal{A}})^{1/2}\,\lambda((x^*\,x)^{2\,l})^{1/2}. \tag{18.11.13}$$

It is well known that

$$\sum_{j=1}^{n} 2^{-j} = 1 - 2^{-n} \tag{18.11.14}$$

for every positive integer n. One can use (18.11.13) repeatedly to get that

$$\lambda(x^*\,x) \le \lambda(e_{\mathcal{A}})^{1-2^{-n}}\,\lambda((x^*\,x)^{2^n})^{2^{-n}} \tag{18.11.15}$$

for every nonnegative integer n. Note that

$$\lambda((x^* x)^{2^n}) \le C^{1/2} \lambda(e_{\mathcal{A}}) \|(x^* x)^{2^n}\|_{\mathcal{A}}, \tag{18.11.16}$$

as in (18.11.11).

It follows that

$$\lambda(x^* x) \le C^{2^{-n-1}} \lambda(e_{\mathcal{A}}) \|(x^* x)^{2^n}\|_{\mathcal{A}}^{2^{-n}} \tag{18.11.17}$$

for every nonnegative integer n. Taking the limit as $n \to \infty$, we get that

$$\lambda(x^* x) \le \lambda(e_{\mathcal{A}}) \, r_{\mathcal{A}}(x^* x), \tag{18.11.18}$$

where $r_{\mathcal{A}}(\cdot)$ is as in Section 6.13. This implies that

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})^{1/2} \lambda(x^* x)^{1/2} \le \lambda(e_{\mathcal{A}}) \, r_{\mathcal{A}}(x^* x)^{1/2}, \tag{18.11.19}$$

using (18.7.3) in the first step. This corresponds to Corollary 21.21 on p319 of [91]. This also corresponds to part (c) of Theorem 11.31 on p284 of [162], and we shall say more about this in the next section.

18.12 Another approach to estimating λ

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a real or complex Banach algebra with a multiplicative identity element $e_{\mathcal{A}}$ and $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$. Suppose that $w \in \mathcal{A}$ satisfies

$$r_{\mathcal{A}}(w) < 1, \tag{18.12.1}$$

where $r_{\mathcal{A}}(w)$ is as in Section 6.13. Remember that

$$r_{\mathcal{A}}(t\, w) = |t| \, r_{\mathcal{A}}(w) \tag{18.12.2}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 10.14. In particular, $r_{\mathcal{A}}(-w) < 1$, so that $b_{1/2,-w}(1) \in \mathcal{A}$ may be defined as in Section 17.10, and satisfies

$$b_{1/2,-w}(1)^2 = e_{\mathcal{A}} - w. \tag{18.12.3}$$

Let $y \mapsto y^*$ be an involution on $\mathcal{A}$ that is conjugate-linear in the complex case, and suppose that w is self-adjoint. If this involution is continuous on $\mathcal{A}$, then

$$b_{1/2,-w}(1)^* = b_{1/2,-w}(1), \tag{18.12.4}$$

as in Section 17.10. This also holds when $\mathcal{A}$ is complex, without asking that the involution be continuous, as before.

Let λ be a positive linear functional on $\mathcal{A}$ again. Under the conditions mentioned in the previous paragraphs, (18.11.3) holds, as before. If w is any self-adjoint element of $\mathcal{A}$, $t \in \mathbf{R}$, and

$$|t| \, r_{\mathcal{A}}(w) < 1, \tag{18.12.5}$$

then $t\,w$ is self-adjoint in $\mathcal{A}$, and

$$\lambda(t\,w) \le \lambda(e_{\mathcal{A}}). \tag{18.12.6}$$

This implies that

$$|\lambda(w)| \le \lambda(e_{\mathcal{A}})\, r_{\mathcal{A}}(w) \tag{18.12.7}$$

for every self-adjoint $w \in \mathcal{A}$.

If x is any element of $\mathcal{A}$, then x^* is self-adjoint, and we can take $w = x^*\,x$ in (18.12.7) to get (18.11.18), and thus (18.11.19). This corresponds to part (c) of Theorem 11.31 on p284 of [162].

18.13 Some related properties of λ

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach algebra over the real or complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and an involution $y \mapsto y^*$ which is conjugate-linear in the complex case. Also let λ be a positive linear functional on $\mathcal{A}$, and suppose that

(18.13.1) the involution on $\mathcal{A}$ is continuous, or $\mathcal{A}$ is complex.

Remember that (18.11.19) holds for every $x \in \mathcal{A}$ under these conditions, as in the previous two sections. If x is normal, then we can combine this with (17.11.5) to get that

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})\, r_{\mathcal{A}}(x). \tag{18.13.2}$$

This corresponds to part (d) of Theorem 11.31 on p284 of [162].

Note that (18.13.2) implies that

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})\, \|x\|_{\mathcal{A}}. \tag{18.13.3}$$

This also follows from (18.11.11) when (18.11.7) holds with $C = 1$. If $\mathcal{A}$ is commutative, then every element of $\mathcal{A}$ is normal. This means that (18.13.3) holds for every $x \in \mathcal{A}$ when $\mathcal{A}$ is commutative, and (18.13.1) holds. This corresponds to part of part (e) of Theorem 11.31 on p284 of [162].

If $\mathcal{A}$ is complex, then the first part of part (e) of Theorem 11.31 on p284 of [162] states that

(18.13.4) λ is a bounded linear functional on $\mathcal{A}$.

In this statement, $\mathcal{A}$ is not asked to be commutative, nor is the involution on $\mathcal{A}$ asked to be continuous.

18.14 Positivity and homomorphisms

Let $\mathcal{A}$ be an associative algebra over the real or complex numbers with an involution $y \mapsto y^*$ that is conjugate-linear in the complex case. Also let h be an algebra homomorphism from $\mathcal{A}$ into $\mathbf{R}$ or $\mathbf{C}$, as appropriate, as in Section 6.9. Suppose that for every $x \in \mathcal{A}$,

$$h(x^*) = \overline{h(x)} \tag{18.14.1}$$

in the complex case, or

$$h(x^*) = h(x) \tag{18.14.2}$$

in the real case. In both cases, we get that

$$h(x^*\, x) = h(x^*)\, h(x) = |h(x)|^2, \tag{18.14.3}$$

so that h is positive in the strong sense with respect to the involution on $\mathcal{A}$, as in Section 18.7.

Let U be a nonempty open subset of the complex plane, and suppose that

$$\overline{z} \in U \tag{18.14.4}$$

for every $z \in U$. Remember that the space $\mathcal{H}(U)$ of holomorphic functions on U is a commutative associative algebra over the complex numbers with respect to pointwise multiplication of functions, as in Section 13.12. If $f \in \mathcal{H}(U)$, then let f^* be the complex-valued function defined on U by

$$f^*(z) = \overline{f(\overline{z})}. \tag{18.14.5}$$

It is well known that f^* is holomorphic on U as well. This defines a conjugate-linear involution on $\mathcal{H}(U)$.

If $w \in U$, then

$$h_w(f) = f(w) \tag{18.14.6}$$

defines an algebra homomorphism from $\mathcal{H}(U)$ onto $\mathbf{C}$, as in Section 13.12. Observe that

$$h_w(f^*) = \overline{f(\overline{w})} = \overline{h_{\overline{w}}(f)}. \tag{18.14.7}$$

Thus h_w satisfies (18.14.1) if and only if

$$\overline{w} = w, \tag{18.14.8}$$

which is to say that $w \in \mathbf{R}$.

Let U_1 be the open unit disk in the complex plane, and let $A(U_1)$ be the disk algebra, consisting of continuous complex-valued functions on the closed unit disk that are holomorphic on U_1, as in Section 13.13. This is a Banach algebra with respect to the supremum norm, as before, and (18.14.5) defines an involution on $A(U_1)$ that preserves the supremum norm. Note that (18.14.6) defines an algebra homomorphism from $A(U_1)$ onto $\mathbf{C}$ for every w in the closed unit disk, as before. This example is mentioned in Exercise 3 on p126 of [8], and in Exercise 10 on p289 of [162].

18.15 Positivity and commutativity

Let $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a commutative Banach algebra over the complex numbers with a multiplicative identity element $e_{\mathcal{A}}$, $\|e_{\mathcal{A}}\|_{\mathcal{A}} = 1$, and a conjugate-linear involution $y \mapsto y^*$, and let λ be a positive linear functional on $\mathcal{A}$. If $x \in \mathcal{A}$, then

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})\, r_{\mathcal{A}}(x) \tag{18.15.1}$$

as in Section 18.13. Remember that $\mathrm{Sp}(\mathcal{A})$ is the space of all nonzero algebra homomorphisms from $\mathcal{A}$ into $\mathbf{C}$, as in Section 12.10, and that

$$\widehat{x}(h) = h(x) \tag{18.15.2}$$

is called the Gelfand transform of x. We also have that

$$\|\widehat{x}\|_{sup} = \|\widehat{x}\|_{sup,\mathrm{Sp}(\mathcal{A})} = r_{\mathcal{A}}(x), \tag{18.15.3}$$

where the left side is the supremum norm of $\widehat{x}$ on $\mathrm{Sp}(\mathcal{A})$, as before. Combining this with (18.15.1), we get that

$$|\lambda(x)| \le \lambda(e_{\mathcal{A}})\, \|\widehat{x}\|_{sup}. \tag{18.15.4}$$

Remember too that

$$\widehat{\mathcal{A}} = \{\widehat{x} : x \in \mathcal{A}\} \tag{18.15.5}$$

is a subalgebra of the algebra $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$ of continuous complex-valued functions on $\mathrm{Sp}(\mathcal{A})$ that contains the constant functions and separates points in $\mathrm{Sp}(\mathcal{A})$, as in Section 12.11. We would like to define a linear functional $\widehat{\lambda}$ on $\widehat{\mathcal{A}}$ by

$$\widehat{\lambda}(\widehat{x}) = \lambda(x) \tag{18.15.6}$$

for every $x \in \mathcal{A}$. Note that $\lambda(x) = 0$ when $\widehat{x} = 0$, because of (18.15.4). One can use this to check that $\widehat{\lambda}$ is well defined on $\widehat{\mathcal{A}}$ in this way. Of course,

$$|\widehat{\lambda}(\widehat{x})| \le \lambda(e_{\mathcal{A}})\, \|\widehat{x}\|_{sup} \tag{18.15.7}$$

for every $x \in \mathcal{A}$, by (18.15.4).

Suppose that $y \mapsto y^*$ is symmetric on $\mathcal{A}$, in the sense that

$$h(y^*) = \overline{h(y)} \tag{18.15.8}$$

for every $y \in \mathcal{A}$, as on p285 of [162]. This implies that $\widehat{\mathcal{A}}$ is invariant under complex conjugation, so that

$$\widehat{\mathcal{A}} \text{ is dense in } C(\mathrm{Sp}(\mathcal{A}), \mathbf{C}) \tag{18.15.9}$$

with respect to the supremum metric, by the Stone–Weierstrass theorem. It follows that

$$\begin{gathered}\widehat{\lambda} \text{ has a unique extension to a bounded} \\ \text{linear functional on } C(\mathrm{Sp}(\mathcal{A}), \mathbf{C}),\end{gathered} \tag{18.15.10}$$

with respect to the supremum norm, and it is easy to see that this extension is nonnegative on $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$. This corresponds to part of the proof of Theorem 11.32 on p285 of [162]. This is related to a famous theorem of Bochner, as in [162].

In fact, the argument in [162] shows more than this. One can use the Hahn–Banach theorem to get an extension of $\widehat{\lambda}$ to a bounded linear functional on $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, with dual norm equal to $\lambda(e_{\mathcal{A}})$ with respect to the supremum norm on $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$, without using (18.15.8). A version of the Riesz representation theorem implies that this extension corresponds to a regular complex Borel measure μ on $\mathrm{Sp}(\mathcal{A})$, with total variation norm equal to $\lambda(e_{\mathcal{A}})$. In particular,

$$\mu(\mathrm{Sp}(\mathcal{A})) = \widehat{\lambda}(\widehat{e_{\mathcal{A}}}) = \lambda(e_{\mathcal{A}}), \tag{18.15.11}$$

and one can use this to get that μ is a nonnegative real measure on $\mathrm{Sp}(\mathcal{A})$. If (18.15.8) holds, then μ is uniquely determined by λ, because of (18.15.9).

Conversely, if (18.15.8) holds, then any nonnegative linear functional on $C(\mathrm{Sp}(\mathcal{A}), \mathbf{C})$ leads to a positive linear functional on $\mathcal{A}$, using composition with the Gelfand transform. This corresponds to another part of the proof of Theorem 11.32 on p285 of [162].

Chapter 19

Algebras over fields

19.1 Algebras and homomorphisms

Let k be a field, such as the real or complex numbers, and let V, W, and Z be vector spaces over k. A mapping B from $V \times W$ into Z is said to be *bilinear* if $B(v, w)$ is linear as a function of $v \in V$ for each $w \in W$, and linear as a function of $w \in W$ for each $v \in V$, as in Section 5.13.

We can define the notion of an *algebra in the strict sense* over k in the same way as in Section 6.1, for algebras over $\mathbf{R}$ or $\mathbf{C}$. Thus $\mathcal{A}$ is an algebra in the strict sense over k if $\mathcal{A}$ is a vector space over k equipped with a mapping from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$ that is bilinear over k. The properties of associativity and commutativity may be defined as before, as well as multiplicative identity elements.

If X is a nonempty set, then the space of all k-valued functions on X is a commutative associative algebra over k, with respect to pointwise multiplication of functions. If $1 = 1_k$ is the multiplicative identity element in k, then the function $\mathbf{1}_X$ equal to 1_k at every point in X is the multiplicative identity in this algebra.

If V is a vector space over k, then the space $\mathcal{L}(V)$ of linear mappings from V into itself is an associative algebra over k with respect to composition of mappings. The identity mapping $I = I_V$ on V is the multiplicative identity element in $\mathcal{L}(V)$, as before.

If $\mathcal{A}$ is an algebra in the strict sense over k and $\mathcal{A}_0$ is a linear subspace of $\mathcal{A}$ such that

$$a\, b \in \mathcal{A}_0 \tag{19.1.1}$$

for every $a, b \in \mathcal{A}_0$, then $\mathcal{A}_0$ is an algebra in the strict sense over k with respect to the restriction of multiplication on $\mathcal{A}$ to $\mathcal{A}_0$. In this case, $\mathcal{A}_0$ is said to be a *subalgebra* of $\mathcal{A}$, as in Section 6.1. If $\mathcal{A}$ is associative or commutative, then $\mathcal{A}_0$ has the same property, as before.

Let $\mathcal{A}$, $\mathcal{B}$ be algebras in the strict sense over k. A linear mapping ϕ from $\mathcal{A}$ into $\mathcal{B}$ is said to be an *algebra homomorphism* if

$$\phi(x\, y) = \phi(x)\, \phi(y) \tag{19.1.2}$$

for every $x, y \in \mathcal{A}$, as in Section 6.3. If $\mathcal{A}$ and $\mathcal{B}$ have multiplicative identity elements $e_{\mathcal{A}}$ and $e_{\mathcal{B}}$, respectively, then one may also wish to ask that

$$\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}, \tag{19.1.3}$$

as before. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$ and $\phi(\mathcal{A}) = \mathcal{B}$, then (19.1.2) implies that $\phi(e_{\mathcal{A}})$ is the multiplicative identity element in $\mathcal{B}$. A one-to-one homomorphism ϕ from $\mathcal{A}$ onto $\mathcal{B}$ is called as *algebra isomorphism*, in which case ϕ^{-1} is an isomorphism from $\mathcal{B}$ onto $\mathcal{A}$.

A linear mapping ϕ from $\mathcal{A}$ into $\mathcal{B}$ is said to be an *opposite algebra homomorphism* if

$$\phi(x\,y) = \phi(y)\,\phi(x) \tag{19.1.4}$$

for every $x, y \in \mathcal{A}$, as in Section 6.4. If $\mathcal{A}$ and $\mathcal{B}$ have multiplicative identity elements $e_{\mathcal{A}}$ and $e_{\mathcal{B}}$, respectively, then one may wish to ask that (19.1.3) hold, as usual. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$ and $\phi(\mathcal{A}) = \mathcal{B}$, then (19.1.4) implies that $\phi(e_{\mathcal{A}})$ is the multiplicative identity element in $\mathcal{B}$. A one-to-one opposite algebra homomorphism ϕ from $\mathcal{A}$ onto $\mathcal{B}$ is called an *opposite algebra isomorphism*, which means that ϕ^{-1} is an opposite algebra isomorphism from $\mathcal{B}$ onto $\mathcal{A}$. If ϕ is an opposite algebra homomorphism from $\mathcal{A}$ onto itself, and

$$\phi \circ \phi = I_{\mathcal{A}}, \tag{19.1.5}$$

then ϕ is said to be an *algebra involution* on $\mathcal{A}$, as before.

Let $\mathcal{C}$ be another algebra in the strict sense over k. If ϕ is a homomorphism from $\mathcal{A}$ into $\mathcal{B}$, and if ψ is a homomorphism from $\mathcal{B}$ into $\mathcal{C}$, then $\psi \circ \phi$ is a homomorphism from $\mathcal{A}$ into $\mathcal{C}$. If ϕ and ψ are opposite algebra homomorphisms, then $\psi \circ \phi$ is a homomorphism. If one of ϕ and ψ is a homomorphism and the other is an opposite algebra homomorphism, then $\psi \circ \phi$ is an opposite algebra homomorphism.

19.2 Ideals and identity elements

Let k be a field, let $\mathcal{A}$ be an algebra in the strict sense over k, and consider

$$\mathcal{A}_1 = \mathcal{A} \times k \tag{19.2.1}$$

initially as a vector space over k, with respect to coordinatewise addition and scalar multiplication. We can use multiplication on $\mathcal{A}$ to define multiplication on $\mathcal{A}_1$ by

$$(a, t)\,(b, z) = (a\,b + z\,a + t\,b, t\,z) \tag{19.2.2}$$

for every $a, b \in \mathcal{A}$ and $t, z \in k$, as in Section 6.15. This makes $\mathcal{A}_1$ into an algebra in the strict sense over k, for which $a \mapsto (a, 0)$ is an injective algebra homomorphism from $\mathcal{A}$ into $\mathcal{A}_1$. By construction,

$$e_{\mathcal{A}_1} = (0, 1_k) \tag{19.2.3}$$

is the multiplicative identity element in $\mathcal{A}_1$. If $\mathcal{A}$ is associative or commutative, then $\mathcal{A}_1$ has the same property, as before.

Let $\mathcal{A}$ be an algebra over k in the struct sense again, and let $\mathcal{I}$ be a linear subspace of $\mathcal{A}$. If

$$a\,x \in \mathcal{I} \tag{19.2.4}$$

for every $a \in \mathcal{A}$ and $x \in \mathcal{I}$, then $\mathcal{I}$ is said to be a *left ideal* in $\mathcal{A}$. Similarly, if

$$x\,a \in \mathcal{I} \tag{19.2.5}$$

for every $a \in \mathcal{A}$ and $x \in \mathcal{I}$, then $\mathcal{I}$ is said to be a *right ideal* in $\mathcal{A}$. Note that left and right ideals in $\mathcal{A}$ are both subalgebras of $\mathcal{A}$. If $\mathcal{A}$ is commutative, then left and right ideals in $\mathcal{A}$ are the same.

If $\mathcal{I}$ is both a left and right ideal in $\mathcal{A}$, then $\mathcal{I}$ is said to be a *two-sided ideal* in $\mathcal{A}$, as in Section 6.11. If $\mathcal{B}$ is another algebra in the strict sense over k, then the kernel of any homomorphism from $\mathcal{A}$ into $\mathcal{B}$ is a two-sided ideal in $\mathcal{A}$. The kernel of an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ is a two-sided ideal in $\mathcal{A}$ as well.

If $\mathcal{I}$ is a linear subspace of $\mathcal{A}$, then the quotient space $\mathcal{A}/\mathcal{I}$ may be defined as a vector space over k in a standard way. Let q be the natural quotient mapping from $\mathcal{A}$ onto $\mathcal{A}/\mathcal{I}$, which is a linear mapping with kernel equal to $\mathcal{I}$. If $\mathcal{I}$ is a two-sided ideal in $\mathcal{A}$, then one can define multiplication on $\mathcal{A}/\mathcal{I}$ in such a way that

$$q(a)\,q(b) = q(a\,b) \tag{19.2.6}$$

for every $a, b \in \mathcal{A}$, as mentioned in Section 6.11. In this case, $\mathcal{A}/\mathcal{I}$ may be considered as an algebra in the strict sense over k, and q is an algebra homomorphism from $\mathcal{A}$ onto $\mathcal{A}/\mathcal{I}$. Of course, if $\mathcal{A}$ is commutative or associative, then $\mathcal{A}/\mathcal{I}$ has the same property, as before.

Let V be a vector space over k. If $v \in V$, then

$$\{T \in \mathcal{L}(V) : T(v) = 0\} \tag{19.2.7}$$

is a left ideal in $\mathcal{L}(V)$. If λ is a linear functional on V, then

$$\{T \in \mathcal{L}(V) : \lambda \circ T = 0\} \tag{19.2.8}$$

is a right ideal in $\mathcal{L}(V)$.

19.3 Invertible elements and homomorphisms

Let k be a field, and let $\mathcal{A}$ be an associative algebra over k with a multiplicative identity element $e_{\mathcal{A}}$. An element x of $\mathcal{A}$ is said to be *invertible* in $\mathcal{A}$ if there is an element x^{-1} of $\mathcal{A}$ such that

$$x\,x^{-1} = x^{-1}\,x = e_{\mathcal{A}}, \tag{19.3.1}$$

as in Section 6.5. The collection $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$ is a group with respect to multiplication, as before. If $x \in \mathcal{A}$ is *nilpotent*, so that $x^l = 0$ for some positive integer l, then

$$e_{\mathcal{A}} - x \in G(\mathcal{A}), \tag{19.3.2}$$

as in Section 6.13. If x is any element of $\mathcal{A}$, then the *spectrum* of x with respect to $\mathcal{A}$ may be defined by

$$\sigma_{\mathcal{A}}(x) = \{\lambda \in k : \lambda\, e_{\mathcal{A}} - x \not\in G(\mathcal{A})\}, \tag{19.3.3}$$

as in Section 6.8.

Let h be a homomorphism from $\mathcal{A}$ into k, as an algebra into itself, and let 0_k and 1_k be the additive and multiplicative identity elements in k. It is easy to see that $h(e_{\mathcal{A}})$ is either 0_k or 1_k, as in Section 6.9, and that $h \equiv 0_k$ on $\mathcal{A}$ when $h(e_{\mathcal{A}}) = 0_k$. Suppose that $h(e_{\mathcal{A}}) = 1_k$, and observe that

$$h(a) \neq 0_k \text{ when } a \in G(\mathcal{A}), \tag{19.3.4}$$

as before. If a is any element of $\mathcal{A}$, then it follows that

$$h(a) \in \sigma_{\mathcal{A}}(a), \tag{19.3.5}$$

as before. We may use the notation $\mathrm{Sp}(\mathcal{A}) = \mathrm{Sp}_k(\mathcal{A})$ for the set of all nonzero algebra homomorphisms from $\mathcal{A}$ into k, as in Sections 12.10 and 12.14, particularly when $\mathcal{A}$ is commutative.

Let $\mathcal{B}$ be another associative algebra over k with a multiplicative identity element $e_{\mathcal{B}}$, and let ϕ be a homomorphism or an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ such that $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. If $x \in \mathcal{A}$ is invertible, then $\phi(x)$ is invertible in $\mathcal{B}$, with

$$\phi(x^{-1}) = \phi(x)^{-1}, \tag{19.3.6}$$

as in Section 10.11. This implies that

$$\phi(G(\mathcal{A})) \subseteq G(\mathcal{B}), \tag{19.3.7}$$

as before. It follows that

$$\sigma_{\mathcal{B}}(\phi(y)) \subseteq \sigma_{\mathcal{A}}(y) \tag{19.3.8}$$

for every $y \in \mathcal{A}$. If $\beta \in \mathrm{Sp}(\mathcal{B})$, then

$$\widehat{\phi}(\beta) = \beta \circ \phi \tag{19.3.9}$$

is an element of $\mathrm{Sp}(\mathcal{A})$, as in Section 12.14.

Let $V \neq \{0\}$ be a vector space over k, and let T be a linear mapping from V into itself. Remember that $\lambda \in k$ is said to be an *eigenvalue* of T if there is a nonzero *eigenvector* $v \in V$ corresponding to λ, so that

$$T(v) = \lambda\, v. \tag{19.3.10}$$

The set $\sigma_p(T)$ of eigenvalues of T is called the *point spectrum* of T, as in Section 9.1. It is easy to see that

$$\sigma_p(T) \subseteq \sigma_{\mathcal{L}(V)}(T), \tag{19.3.11}$$

and that

$$\sigma_p(T) = \sigma_{\mathcal{L}(V)}(T) \tag{19.3.12}$$

when V has finite dimension, as before. If k is algebraically closed, then

$$\sigma_{\mathcal{L}(V)}(T) \neq \emptyset \tag{19.3.13}$$

when V has finite dimension.

19.4 Left and right multiplication operators

Let k be a field, and let $\mathcal{A}$ be an algebra over k in the strict sense. An element e_L of $\mathcal{A}$ is said to be a *left multiplicative identity element* in $\mathcal{A}$ if

$$e_L\, x = x \tag{19.4.1}$$

for every $x \in \mathcal{A}$, as in Section 18.1. Similarly, $e_R \in \mathcal{A}$ is said to be a *right multiplicative identity element* in $\mathcal{A}$ if

$$x\, e_R = x \tag{19.4.2}$$

for every $x \in \mathcal{A}$. Of course, a multiplicative identity element in $\mathcal{A}$ is the same as an element of $\mathcal{A}$ that is both a left and right multiplicative identity element. If $\mathcal{A}$ has a left multiplicative identity element e_L and a right multiplicative identity element e_R, then $e_L = e_R$, so that this is a multiplicative identity element in $\mathcal{A}$, as before.

If $a \in \mathcal{A}$, then let

$$L_a(x) = a\, x \tag{19.4.3}$$

and

$$R_a(x) = x\, a \tag{19.4.4}$$

be the corresponding left and right multiplication operators on $\mathcal{A}$, as in Sections 6.3 and 6.4. Of course,

$$a \mapsto L_a \tag{19.4.5}$$

and

$$a \mapsto R_a \tag{19.4.6}$$

are linear mappings from $\mathcal{A}$ into the space $\mathcal{L}(\mathcal{A})$ of linear mappings from $\mathcal{A}$ into itself, as before. Remember that $e_L \in \mathcal{A}$ is a left multiplicative identity element in $\mathcal{A}$ if and only if

$$L_{e_L} = I_{\mathcal{A}}, \tag{19.4.7}$$

as in Section 18.1. This happens if and only if

$$R_a(e_L) = a \tag{19.4.8}$$

for every $a \in \mathcal{A}$. Similarly, $e_R \in \mathcal{A}$ is a right multiplicative identity element in $\mathcal{A}$ if and only if

$$R_{e_R} = I_{\mathcal{A}}, \tag{19.4.9}$$

which is the same as saying that

$$L_a(e_R) = a \tag{19.4.10}$$

for every $a \in \mathcal{A}$.

Suppose now that $\mathcal{A}$ is an associative algebra over k, so that (19.4.5) is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$, as in Section 6.3. Similarly, (19.4.6) is an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$, as in Section 6.4. Suppose also that $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, so that (19.4.5) and (19.4.6) are injective, as in (19.4.8) and (19.4.10).

If $\mathcal{A}$ has finite dimension as a vector space over k, and k is algebraically closed, then

$$\sigma_{\mathcal{L}(\mathcal{A})}(L_a) \neq \emptyset \tag{19.4.11}$$

for every $a \in \mathcal{A}$, as in (19.3.13). This implies that

$$\sigma_{\mathcal{A}}(a) \neq \emptyset \tag{19.4.12}$$

for every $a \in \mathcal{A}$, as in (19.3.8).

19.5 Centralizers and invertibility

Let k be a field, and let $\mathcal{A}$ be an associative algebra over k. If E is a nonempty subset of $\mathcal{A}$, then the *centralizer* of E in $\mathcal{A}$ is defined by

$$C(E) = C_{\mathcal{A}}(E) = \{a \in \mathcal{A} : a\,x = x\,a \text{ for every } x \in E\}, \tag{19.5.1}$$

as in Section 14.5. One can check that

$$C_{\mathcal{A}}(E) \text{ is a subalgebra of } \mathcal{A}, \tag{19.5.2}$$

as before. Of course, if $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then

$$e_{\mathcal{A}} \in C_{\mathcal{A}}(E). \tag{19.5.3}$$

Note that

$$E \subseteq C_{\mathcal{A}}(C_{\mathcal{A}}(E)), \tag{19.5.4}$$

as before.

If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$ and $a \in C_{\mathcal{A}}(E)$ is invertible in $\mathcal{A}$, then one can check that

$$a^{-1} \in C_{\mathcal{A}}(E), \tag{19.5.5}$$

as in Section 14.5. This implies that

$$\sigma_{\mathcal{A}}(y) = \sigma_{C_{\mathcal{A}}(E)}(y) \tag{19.5.6}$$

for every $y \in C_{\mathcal{A}}(E)$, as before.

If E_1, E_2 are nonempty subsets of $\mathcal{A}$ and

$$E_1 \subseteq E_2, \tag{19.5.7}$$

then

$$C_{\mathcal{A}}(E_2) \subseteq C_{\mathcal{A}}(E_1), \tag{19.5.8}$$

as in Section 14.5. The elements of E commute with each other if and only if

$$E \subseteq C_{\mathcal{A}}(E), \tag{19.5.9}$$

in which case

$$C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \subseteq C_{\mathcal{A}}(E), \tag{19.5.10}$$

by (19.5.8). Observe that

$$C_{\mathcal{A}}(E_1) \text{ is a commutative subalgebra of } \mathcal{A} \tag{19.5.11}$$

when

$$C_{\mathcal{A}}(E_1) \subseteq E_1. \tag{19.5.12}$$

This implies that

$$C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \text{ is a commutative subalgebra of } \mathcal{A} \tag{19.5.13}$$

when the elements of E commute with each other, as in Section 14.5.

If E is any nonempty subset of $\mathcal{A}$, then

$$\mathcal{B} = C_{\mathcal{A}}(C_{\mathcal{A}}(E)) \tag{19.5.14}$$

is a subalgebra of $\mathcal{A}$ that contains E. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then $e_{\mathcal{A}} \in \mathcal{B}$, and

$$\sigma_{\mathcal{A}}(b) = \sigma_{\mathcal{B}}(b) \tag{19.5.15}$$

for every $b \in \mathcal{B}$, as in (19.5.6). This holds in particular when $b \in E$, as in Section 14.5.

If $a \in \mathcal{A}$, then let L_a and R_a be as in (19.4.3) and (19.4.4), respectively, and put

$$L_{\mathcal{A}} = \{L_a : a \in \mathcal{A}\}, \ R_{\mathcal{A}} = \{R_a : a \in \mathcal{A}\}, \tag{19.5.16}$$

as in Section 14.7. These are subalgebras of $\mathcal{L}(\mathcal{A})$, and

$$L_a \circ R_b = R_b \circ L_a \tag{19.5.17}$$

for every $a, b \in \mathcal{A}$, because $\mathcal{A}$ is associative, as before. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then

$$C_{\mathcal{L}(\mathcal{A})}(L_{\mathcal{A}}) = R_{\mathcal{A}}, \ C_{\mathcal{L}(\mathcal{A})}(R_{\mathcal{A}}) = L_{\mathcal{A}}, \tag{19.5.18}$$

as in Section 14.7. One can use this to get that L_a is invertible in $\mathcal{L}(\mathcal{A})$ exactly when a is invertible in $\mathcal{A}$, and similarly for R_a, as before. One can use this to get that

$$\sigma_{\mathcal{L}(\mathcal{A})}(L_b), \sigma_{\mathcal{L}(\mathcal{A})}(R_b) = \sigma_{\mathcal{A}}(b) \tag{19.5.19}$$

for every $b \in \mathcal{A}$, as in Section 14.7 again.

19.6 Some algebras of polynomials

Let k be a field, let n be a positive integer, and let $X_1, \ldots, X_n$ be n commuting indeterminates. One can define *formal polynomials* in $X_1, \ldots, X_n$ with coefficients in k in the same way as in Section 13.3. The space $k[X_1, \ldots, X_n]$ of these formal polynomials is a commutative associative algebra over k with a multiplicative identity element, as before.

Let $a = (a_1, \ldots, a_n)$ be an n-tuple of commuting elements of $\mathcal{A}$. If $p(X) = p(X_1, \ldots, X_n)$ is an element of $k[X_1, \ldots, X_n]$, then one can define

$$p(a) = p(a_1, \ldots, a_n) \in \mathcal{A} \tag{19.6.1}$$

as in Section 13.4. One may also use the notation $p_{\mathcal{A}}(a) = p_{\mathcal{A}}(a_1, \ldots, a_n)$ to indicate the role of $\mathcal{A}$ here, as before. Note that

$$p(X) \mapsto p(a) \tag{19.6.2}$$

defines an algebra homomorphism from $k[X_1, \ldots, X_n]$ into $\mathcal{A}$, as before.

Let k^n be the space of n-tuples of elements of k, as usual. If $w \in k^n$, then

$$h_w(p(X)) = p(w) \tag{19.6.3}$$

defines an algebra homomorphism from $k[X_1, \ldots, X_n]$ onto k, as in Section 13.5. Of course,

$$h_w(X_j) = w_j \tag{19.6.4}$$

for each $j = 1, \ldots, n$, by construction. In fact,

$$w \mapsto h_w \tag{19.6.5}$$

is a one-to-one mapping from k^n onto $\mathrm{Sp}(k[X_1, \ldots, X_n])$, as before.

Let us now take $n = 1$, and let X be a single indeterminate. If $p(X)$ is an element of $k[X]$ and $a \in \mathcal{A}$, then $p(a) = p_{\mathcal{A}}(a) \in \mathcal{A}$ may be defined as before, which is essentially the same as in Section 8.13. In particular, $p(X)$ determines a polynomial function p_k on k. One can check that

$$p_k(\sigma_{\mathcal{A}}(a)) \subseteq \sigma_{\mathcal{A}}(p_{\mathcal{A}}(a)), \tag{19.6.6}$$

as in Section 8.13. This may be considered as at least part of a version of the *spectral mapping theorem*, as before.

One might also like to have that

$$\sigma_{\mathcal{A}}(p_{\mathcal{A}}(a)) \subseteq p_k(\sigma_{\mathcal{A}}(a)) \tag{19.6.7}$$

under suitable conditions, as before. If $p(X)$ is a constant polynomial, then this holds when $\sigma_{\mathcal{A}}(a) \neq \emptyset$. If $p(X)$ is not a constant polynomial, and k is algebraically closed, then one can verify that (19.6.7) holds, as in Section 8.13.

19.7 Subalgebras of finite dimension

Let k be a field, let $\mathcal{B}$ be an associative algebra over k with a multiplicative identity element $e = e_{\mathcal{B}}$, and let $\mathcal{A}$ be a subalgebra of $\mathcal{B}$ that contains e. Note that

(19.7.1) $$G(\mathcal{A}) \subseteq G(\mathcal{B}),$$

as in Section 7.3. If $x \in \mathcal{A}$, then

(19.7.2) $$\sigma_{\mathcal{B}}(x) \subseteq \sigma_{\mathcal{A}}(x),$$

as before. This corresponds to (19.3.8) as well, with ϕ taken to be the obvious inclusion mapping from $\mathcal{A}$ into $\mathcal{B}$.

Suppose that $\mathcal{A}$ has finite dimension as a vector space over k, and that $e \neq 0$, to avoid trivialities. If $a \in \mathcal{A}$, then one can get nontrivial formal polynomial $p(X)$ with coefficients in k such that

(19.7.3) $$p_{\mathcal{A}}(a) = 0.$$

Note that this implies that

(19.7.4) $$\sigma_{\mathcal{A}}(a) \subseteq \{t \in k : p_k(t) = 0\},$$

as in (19.6.6).

If a is invertible in $\mathcal{B}$, then one can take $p(X)$ to have degree which is as small as possible, to get that the constant term is not zero. One can use this to get a formal polynomial $q(X)$ with coefficients in k such that

(19.7.5) $$a\, q_{\mathcal{A}}(a) = e.$$

This means that

(19.7.6) $$a^{-1} = q_{\mathcal{A}}(a) \in \mathcal{A}.$$

It follows that

(19.7.7) $$G(\mathcal{A}) = \mathcal{A} \cap G(\mathcal{B})$$

in this case. This implies that

(19.7.8) $$\sigma_{\mathcal{A}}(x) = \sigma_{\mathcal{B}}(x)$$

for every $x \in \mathcal{A}$.

19.8 An invariant linear subspace

Let k be a field, let V be a vector space over k, and let W be a linear subspace of V. Put

(19.8.1) $$\mathcal{L}_W(V) = \{T \in \mathcal{L}(V) : T(W) \subseteq W\},$$

which is a subalgebra of $\mathcal{L}(V)$ that contains I_V.

If $T \in \mathcal{L}_W(V)$, then let T_W be the restriction of T to W. Note that

$$T \mapsto T_W \tag{19.8.2}$$

defines a homomorphism from $\mathcal{L}_W(V)$ into $\mathcal{L}(W)$, as algebras over k. The kernel of this homomorphism is equal to

$$\{T \in \mathcal{L}(V) : T = 0 \text{ on } W\}. \tag{19.8.3}$$

Thus this a two-sided ideal in $\mathcal{L}_W(V)$, which is a left ideal in $\mathcal{L}(V)$ too.

The quotient space V/W may be defined as a vector space over k in a natural way. Let $q_{V/W}$ be the natural quotient mapping from V onto V/W, which is a linear mapping with kernel equal to W. If T is any linear mapping from V into itself, then

$$q_{V/W} \circ T \tag{19.8.4}$$

is a linear mapping from V into V/W. Clearly $T \in \mathcal{L}_W(V)$ if and only if

$$q_{V/W} \circ T = 0 \text{ on } W. \tag{19.8.5}$$

In this case, there is a unique linear mapping $T^{V/W}$ from V/W into itself such that

$$T^{V/W} \circ q_{V/W} = q_{V/W} \circ T, \tag{19.8.6}$$

by standard arguments.

One can check that

$$T \mapsto T^{V/W} \tag{19.8.7}$$

is a homomorphism from $\mathcal{L}_W(V)$ into $\mathcal{L}(V/W)$, as algebras over k. The kernel of this homomorphism is equal to

$$\{T \in \mathcal{L}(V) : T(V) \subseteq W\}. \tag{19.8.8}$$

This is another two-sided ideal in $\mathcal{L}_W(V)$, which is also a right ideal in $\mathcal{L}(V)$.

Of course, the group $G(\mathcal{L}(V))$ of invertible elements of $\mathcal{L}(V)$ consists of the one-to-one linear mappings T from V onto itself. This is also known as the *general linear group* of V, which may be denoted $GL(V)$ as well.

It is easy to see that

$$G(\mathcal{L}_W(V)) = \{T \in GL(V) : T(W) = W\}. \tag{19.8.9}$$

If W has finite dimension, as a vector space over k, then

$$G(\mathcal{L}_W(V)) = GL(V) \cap \mathcal{L}_W(V). \tag{19.8.10}$$

Indeed, if W has finite dimention, $T \in \mathcal{L}_W(V)$, and T is injective on W, then $T(W) = W$.

The *codimension* of W as a linear subspace of V may be defined as the dimension of V/W, as a vector space over k. If W has finite codimension in V, then (19.8.10) holds. More precisely, if $T \in \mathcal{L}_W(V)$ and $T(V) = V$, then

$$T^{V/W}(V/W) = V/W, \tag{19.8.11}$$

by (19.8.6). If $\dim(V/W) < \infty$, then it follows that

(19.8.12) $$\ker T^{V/W} = \{0\}.$$

This means that

(19.8.13) $$T^{-1}(W) = W.$$

Suppose that V has finite dimension, so that W and V/W have finite dimension too, with

(19.8.14) $$\dim V = \dim W + \dim(V/W).$$

In this case, $\mathcal{L}(V)$ has finite dimension as well, with

(19.8.15) $$\dim \mathcal{L}(V) = (\dim V)^2,$$

by standard arguments. One can use a basis for V that includes a basis for W to get that

(19.8.16) $$\dim \mathcal{L}_W(V) = (\dim W)^2 + (\dim V - \dim W)\,(\dim V).$$

19.9 Algebraic duals

Let k be a field, and let V be a vector space over k. As in Section 3.1, a *linear functional* on V is a linear mapping from V into k, considered as a one-dimensional vector space over itself. The space V^{alg} of all linear functionals on V may be called the *algebraic dual* of V, as before. This is a linear subspace of the space of all k-valued functions on V, as usual.

Let W be another vector space over k, and let T be a linear mapping from V into W. If μ is a linear functional on W, then

(19.9.1) $$T^{\mathrm{alg}}(\mu) = \mu \circ T$$

is a linear functional on V, as in Section 3.13. This defines a linear mapping from W^{alg} into V^{alg}, which is the *dual linear mapping* associated to T, as before.

The space $\mathcal{L}(V, W)$ of all linear mappings from V into W is a linear subspace of the space of all W-valued functions on V, as in Section 2.2. Note that

(19.9.2) $$T \mapsto T^{\mathrm{alg}}$$

is a linear mapping from $\mathcal{L}(V, W)$ into $\mathcal{L}(W^{\mathrm{alg}}, V^{\mathrm{alg}})$, as in Section 3.13.

If $T \neq 0$, then $T(v) \neq 0$ for some $v \in V$, which implies that $\mu(T(v)) \neq 0$ for some $\mu \in W^{\mathrm{alg}}$. This means that

(19.9.3) $$(T^{\mathrm{alg}}(\mu))(v) = (\mu \circ T)(v) = \mu(T(v)) \neq 0,$$

so that $T^{\mathrm{alg}} \neq 0$. Thus (19.9.2) is one-to-one on $\mathcal{L}(V, W)$.

Let Z be another vector space over k. If T_1 is a linear mapping from V into W, and T_2 is a linear mapping from W into Z, then

(19.9.4) $$(T_2 \circ T_1)^{\mathrm{alg}} = T_1^{\mathrm{alg}} \circ T_2^{\mathrm{alg}},$$

as linear mappings from Z^{alg} into W^{alg}, as in Section 3.13.

In particular, we may consider linear mappings from V into itself, and the corresponding dual linear mappings on V^{alg}. The remarks in the previous paragraphs imply that (19.9.2) defines an injective opposite algebra homomorphism from $\mathcal{L}(V)$ into $\mathcal{L}(V^{\mathrm{alg}})$. Note that

$$(I_V)^{\mathrm{alg}} = I_{V^{\mathrm{alg}}}. \tag{19.9.5}$$

Suppose for the moment that V has finite dimension, so that (19.8.15) holds. In this case, we also have that

$$\dim V^{\mathrm{alg}} = \dim V, \tag{19.9.6}$$

and

$$\dim \mathcal{L}(V^{\mathrm{alg}}) = (\dim V^{\mathrm{alg}})^2 = (\dim V)^2. \tag{19.9.7}$$

It follows that (19.9.2) maps $\mathcal{L}(V)$ onto $\mathcal{L}(V^{\mathrm{alg}})$ under these conditions.

If $v \in V$, then

$$\widehat{L}_v(\lambda) = \widehat{L}_v^V(\lambda) = \lambda(v) \tag{19.9.8}$$

defines an element of the algebraic dual $(V^{\mathrm{alg}})^{\mathrm{alg}}$ of V^{alg}, as in Section 3.14. More precisely,

$$v \mapsto \widehat{L}_v \tag{19.9.9}$$

defines a one-to-one linear mapping from V into $(V^{\mathrm{alg}})^{\mathrm{alg}}$, as before.

If V has finite dimension, then

$$\dim V = \dim V^{\mathrm{alg}} = \dim (V^{\mathrm{alg}})^{\mathrm{alg}}, \tag{19.9.10}$$

as before. This implies that (19.9.9) maps V onto $(V^{\mathrm{alg}})^{\mathrm{alg}}$ in this case.

If T is a linear mappng from V into W, then the second dual mapping $(T^{\mathrm{alg}})^{\mathrm{alg}}$ may be defined as a linear mapping from $(V^{\mathrm{alg}})^{\mathrm{alg}}$ into $(W^{\mathrm{alg}})^{\mathrm{alg}}$ by taking the dual linear mapping associated to T^{alg}, as in Section 3.15. If $w \in W$, then let $\widehat{L}_w^W \in (W^{\mathrm{alg}})^{\mathrm{alg}}$ be as in (19.9.8). One can check that

$$(T^{\mathrm{alg}})^{\mathrm{alg}}(\widehat{L}_v^V) = \widehat{L}_{T(v)}^W \tag{19.9.11}$$

for every $v \in V$, as in Section 3.15.

19.10 Algebraic duals and linear subspaces

Let k be a field, let V be a vector space over k, and let V_0 be a linear subspace over V. Also let T_0 be the natural inclusion mapping from V_0 into V, which sends every element of V_0 to itself, considered as an element of V. If μ is a linear functional on V, then

$$(T_0)^{\mathrm{alg}}(\mu) = \mu \circ T_0 \tag{19.10.1}$$

is the same as the restriction of μ to V_0, considered as an element of $(V_0)^{\mathrm{alg}}$. This defines a linear mapping from V^{alg} into $(V_0)^{\mathrm{alg}}$, as in the previous section.

In fact,

$$(T_0)^{\mathrm{alg}}(V^{\mathrm{alg}}) = (V_0)^{\mathrm{alg}}. \tag{19.10.2}$$

This is the same as saying that every linear functional on V_0 can be extended to a linear functional on V. This can be seen using elementary arguments when V_0 has finite codimension in V, and in particular when V has finite dimension. Otherwise, there are standard arguments based on the axiom of choice using Zorn's lemma or Hausdorff's maximality principle.

Put

$$(V_0)^{\perp_{\mathrm{alg}}} = \ker (T_0)^{\mathrm{alg}} = \{\mu \in V^{\mathrm{alg}} : \mu \equiv 0 \text{ on } V_0\}, \tag{19.10.3}$$

which is a linear subspace of V^{alg}. If V_0 has finite dimension, then the codimension of $(V_0)^{\perp_{\mathrm{alg}}}$ in V^{alg} is equal to

$$\dim (V_0)^{\mathrm{alg}} = \dim V_0, \tag{19.10.4}$$

because of (19.10.2). If V has finite dimension, then we get that

$$\begin{aligned} \dim V = \dim V^{\mathrm{alg}} &= \dim (V_0)^{\perp_{\mathrm{alg}}} + \dim (V_0)^{\mathrm{alg}} \\ &= \dim (V_0)^{\perp_{\mathrm{alg}}} + \dim V_0. \end{aligned} \tag{19.10.5}$$

If V_1 is a linear subspace of V^{alg}, then put

$$^{\perp_{\mathrm{alg}}}V_1 = \bigcap_{\lambda \in V_1} \ker \lambda = \{v \in V : \lambda(v) = 0 \text{ for every } \lambda \in V_1\}, \tag{19.10.6}$$

which is a linear subspace of V. Clearly

$$V_0 \subseteq {}^{\perp_{\mathrm{alg}}}((V_0)^{\perp_{\mathrm{alg}}}). \tag{19.10.7}$$

In fact,

$$V_0 = {}^{\perp_{\mathrm{alg}}}((V_0)^{\perp_{\mathrm{alg}}}). \tag{19.10.8}$$

This means that if $v \in V$ and v is not an element of V_0, then there is a linear functional μ on V such that $\mu \equiv 0$ on V_0, and

$$\mu(v) \neq 0. \tag{19.10.9}$$

One can first get such a linear functional on the linear subspace of V spanned by V_0 and v, and then use an extension of that to a linear functional on V.

Let T be a linear mapping from V into itself. If

$$T(V_0) \subseteq V_0, \tag{19.10.10}$$

then it is easy to see that

$$T^{\mathrm{alg}}(\mu) = \mu \circ T \equiv 0 \text{ on } V_0 \text{ for every } \mu \in (V_0)^{\perp_{\mathrm{alg}}}. \tag{19.10.11}$$

Equivalently, this means that

$$T^{\mathrm{alg}}((V_0)^{\perp_{\mathrm{alg}}}) \subseteq (V_0)^{\perp_{\mathrm{alg}}}. \tag{19.10.12}$$

This is also the same as saying that

$$T(V_0) \subseteq \ker \mu \text{ for every } \mu \in (V_0)^{\perp_{\mathrm{alg}}}, \tag{19.10.13}$$

which means that

$$T(V_0) \subseteq {}^{\perp_{\mathrm{alg}}}((V_0)^{\perp_{\mathrm{alg}}}). \tag{19.10.14}$$

This shows that (19.10.12) is equivalent to (19.10.10), because of (19.10.8).

19.11 Bilinear functionals

Let k be a field, and let V, W be vector spaces over k. A bilinear mapping b from $V \times W$ into k may be called a *bilinear functional* on $V \times W$, as in Section 5.15. This is called a *bilinear function* on p251 of [30], and a *bilinear form* on p337 of [124] and p338 of [138]. The term "bilinear form" is defined another way on p281 of [30], as a type of formal polynomial with coefficients in k. This corresponds to a bilinear functional on the Cartesian product of two finite-dimensional vector spaces over k with the appropriate dimensions, using bases for these vector spaces, as on p282 of [30].

Let b be a bilinear functional on $V \times W$, and put

$$b_{1,w}(v) = b(v, w) \tag{19.11.1}$$

and

$$b_{2,v}(w) = b(v, w) \tag{19.11.2}$$

for all $v \in V$ and $w \in W$, as in Sections 5.13 and 5.15. Note that $b_{1,w}$ is a linear functional on V for each $w \in W$, and that $w \mapsto b_{1,w}$ is a linear mapping from W into V^{alg}, as before. Similarly, $b_{2,v}$ is a linear functional on W for every $v \in V$, and $v \mapsto b_{2,v}$ is a linear mapping from V into W^{alg}. Any linear mapping from W into V^{alg} or from V into W^{alg} corresponds to a bilinear functional on $V \times W$ in this way, as before.

Let 0_k and 1_k be the additive and multiplicative identity elements in k, as before. If $1_k + 1_k = 0_k$, then k is said to have characteristic two, as usual. This is the same as saying that $-1_k = 1_k$, and otherwise $1_k + 1_k$ has a multiplicative inverse in k.

Suppose now that $V = W$. In this case, one may refer to b as a bilinear form on V, as in Sections 5.15 and 8.5. If

$$b(v, w) = b(w, v) \tag{19.11.3}$$

for all $v, w \in V$, then we say that b is *symmetric* on V, as in Section 8.5. If

$$b(v, w) = -b(w, v) \tag{19.11.4}$$

for all $v, w \in V$, then b is said to be *antisymmetric* on V, as before, or *skew-symmetric*, as on p341 of [138]. If

$$b(v, v) = 0 \tag{19.11.5}$$

for all $v \in V$, then b is said to be *alternating* on V, as on p329, 354 of [124], and p341 of [138].

It is easy to see that b is alternating when b is antisymmetric and the characteristic of k is not equal to 2. If the characteristic of k is equal to 2, then (19.11.3) and (19.11.4) are the same.

Note that

$$b(v+w, v+w) = b(v,v) + b(v,w) + b(w,v) + b(w,w) \tag{19.11.6}$$

for all $v, w \in V$, as in Section 8.5. If b is alternating on V, then it follows that b is antisymmetic on V, as before. If b is symmetric on V, and if the characteristic of k is not equal to 2, then one can use (19.11.6) to get that b is determined by the values of $b(u,u)$, $u \in V$, as in Section 8.5 again.

If b is any bilinear form on V, then

$$b(v,w) + b(w,v) \tag{19.11.7}$$

is a symmetric bilinear form on V. If the characteristic of k is equal to 2, then (19.11.7) is alternating on V.

19.12 Nondegenerate bilinear forms

Let k be a field, and let V be a finite-dimensional vector space over k. A bilinear form b on V is said to be *nondegenerate* if for every $v \in V$ with $v \neq 0$ there is a $w \in V$ such that

$$b(v,w) \neq 0. \tag{19.12.1}$$

If $b_{2,v}$ is as in (19.11.2), then b is nondegenerate on V if and only if

$$v \mapsto b_{2,v} \text{ is one-to-one on } V, \tag{19.12.2}$$

as a linear mapping from V into V^{alg}. Equivalently, this means that

$$v \mapsto b_{2,v} \text{ maps } V \text{ onto } V^{\mathrm{alg}}, \tag{19.12.3}$$

because $\dim V = \dim V^{\mathrm{alg}}$.

Similarly, if $b_{1,w}$ is as in (19.11.1), then

$$w \mapsto b_{1,w} \text{ is one-to-one on } V, \tag{19.12.4}$$

as a linear mapping from V into V^{alg}, if and only if for every $w \in V$ with $w \neq 0$ there is a $v \in V$ such that (19.12.1) holds. As before, (19.12.4) holds if and only if

$$w \mapsto b_{1,w} \text{ maps } V \text{ onto } V^{\mathrm{alg}}, \tag{19.12.5}$$

because $\dim V = \dim V^{\mathrm{alg}}$. It is easy to see that (19.12.5) implies that b is nondegenerate on V, in the formulation used earlier. Conversely, (19.12.3) implies the version of nondegeneracy that is equivalent to (19.12.4).

Suppose that b is nondegenerate on V, and let T be a linear mapping from V into itself. If $w \in V$, then put

$$\mu_w(v) = \mu_{T,w}(v) = b(T(v), w) \tag{19.12.6}$$

for each $v \in V$, which defines a linear functional on V. It follows that there is a unique element $T^{*_b}(w)$ of V such that

$$b(T(v), w) = \mu_w(v) = b(v, T^{*_b}(w)) \tag{19.12.7}$$

for every $v \in V$, as in the preceding paragraph. It is easy to see that T^{*_b} is a linear mapping from V into itself, which may be called the *adjoint* of T with respect to b. If $k = \mathbf{R}$ and b is an inner product on V, then this corresponds to the adjoint of T in the sense of Section 3.5.

We also have that

$$T \mapsto T^{*_b} \tag{19.12.8}$$

is a linear mapping from $\mathcal{L}(V)$ into itself. If $T \neq 0$, then there is a $v \in V$ such that $T(v) \neq 0$, which implies that

$$b(T(v), w) \neq 0 \tag{19.12.9}$$

for some $w \in V$. This means that

$$b(v, T^{*_b}(w)) \neq 0, \tag{19.12.10}$$

so that $T^{*_b} \neq 0$. Thus (19.12.8) is injective on $\mathcal{L}(V)$. It follows that (19.12.8) maps $\mathcal{L}(V)$ onto itself, because V has finite dimension, by hypothesis.

If T_1, T_2 are linear mappings from V into itself, then

$$b(T_2(T_1(v)), w) = b(T_1(v), T_2^{*_b}(w)) = b(v, T_1^{*_b}(T_2^{*}(w))) \tag{19.12.11}$$

for all $v, w \in V$. This implies that

$$(T_2 \circ T_1)^{*_b} = T_1^{*_b} \circ T_2^{*_b}, \tag{19.12.12}$$

so that (19.12.8) is an opposite algebra isomorphism from $\mathcal{L}(V)$ onto itself.

19.13 More on bilinear forms

Let us continue with the same notation and hypotheses as in the previous section. If $r \in k$ and $r \neq 0$, then it is easy to see that

(19.13.1) $\qquad r\,b$ is nondegenerate as a bilinear form on V.

If T is a linear mapping from V into itself, then T^{*_b} and $T^{*_{r\,b}}$ may be defined as linear mappings from V into itself, as in the previous section, and one can check that

$$T^{*_{r\,b}} = T^{*_b}. \tag{19.13.2}$$

Note that

$$c(v, w) = b(w, v) \tag{19.13.3}$$

is another bilinear form on V, and that the nondegeneracy of b on V is equivalent to the nondegeneracy of c on V, as before. If T is a linear mapping from V into itself, then there is a unique linear mapping T^{*_c} from V into itself such that

$$c(T(v), w) = c(v, T^{*_c}(w)) \tag{19.13.4}$$

for every $v, w \in V$, as before. Equivalently, this means that

$$b(w, T(v)) = b(T^{*_c}(w), v) \tag{19.13.5}$$

for every $v, w \in V$. Of course, this is the same as saying that

$$b(v, T(w)) = b(T^{*_c}(v), w) \tag{19.13.6}$$

for every $v, w \in V$, so that

$$T = (T^{*_c})^{*_b}. \tag{19.13.7}$$

Similarly,

$$T = (T^{*_b})^{*_c}. \tag{19.13.8}$$

Thus

$$T \mapsto T^{*_c} \tag{19.13.9}$$

is the inverse of (19.12.8).

Of course, $b = c$ if and only if b is symmetric on V, and $b = -c$ if and only if b is antisymmetric on V. In both cases, we get that

$$T^{*_b} = T^{*_c} \tag{19.13.10}$$

for every $T \in \mathcal{L}(V)$, as in (19.13.2).

It follows that

$$(T^{*_b})^{*_b} = T \tag{19.13.11}$$

for every $T \in \mathcal{L}(V)$ when b is symmetric or antisymmetric on V, because of (19.13.7). This means that (19.12.8) defines an involution on $\mathcal{L}(V)$ in these two cases.

Note that $T \in \mathcal{L}(V)$ satisfies

$$T^{*_b} = T \tag{19.13.12}$$

if and only if

$$b(T(v), w) = b(v, T(w)) \tag{19.13.13}$$

for every $v, w \in V$. Similarly,

$$T^{*_b} = -T \tag{19.13.14}$$

if and only if

$$b(T(v), w) = -b(v, T(w)) \tag{19.13.15}$$

for every $v, w \in V$.

If A is a linear mapping from V into itself, then

$$b_A(v, w) = b(A(v), w) \tag{19.13.16}$$

defines another bilinear form on V. One can check that every bilinear form on V corresponds to a unique $A \in \mathcal{L}(V)$ in this way, because b is nondegenerate on V. We also have that

$$b_A \text{ is nondegenerate on } V \tag{19.13.17}$$

exactly when A is injective on V, which is the same as saying that A is invertible on V, because V has finite dimension. In this case, if $T \in \mathcal{L}(V)$, then

$$\begin{aligned} b_A(T(v), w) &= b(A(T(v)), w) = b(A(v), (A^{-1})^{*_b}(T^{*_b}(A^{*_b})(w))) \\ &= b_A(v, (A^{-1})^{*_b}(T^{*_b}(A^{*_b}(w)))) \end{aligned} \tag{19.13.18}$$

for every $v, w \in V$. This means that if $T^{*_{b_A}}$ is defined on V using b_A as in the previous section, then

$$T^{*_{b_A}} = (A^{-1})^{*_b} \circ T^{*_b} \circ A^{*_b}. \tag{19.13.19}$$

Suppose for the moment that b is symmetric on V, so that

$$b_A(w, v) = b(A(w), v) = b(v, A(w)) \tag{19.13.20}$$

for every $v, w \in V$. Observe that

$$b_A(v, w) = b_A(w, v) \tag{19.13.21}$$

for every $v, w \in V$ if and only if $A^{*_b} = A$, and that

$$b_A(v, w) = -b_A(w, v) \tag{19.13.22}$$

for all $v, w \in V$ if and only if $A^{*_b} = -A$. Similarly, suppose for the moment that b is antisymmetric on V, so that

$$b_A(w, v) = b(A(w), v) = -b(v, A(w)) \tag{19.13.23}$$

for all $v, w \in V$. In this case, (19.13.21) holds if and only if $A^{*_b} = -A$, and (19.13.22) holds if and only if $A^{*_b} = A$.

19.14 Involutions and conjugations

Let k be a field, and let $\mathcal{A}$ be an associative algebra over k with a multiplicative identity element $e_{\mathcal{A}}$. Also let $x \mapsto x^*$ be an opposite algebra isomorphism from $\mathcal{A}$ onto itself, and note that $e_{\mathcal{A}}^* = e_{\mathcal{A}}$. Of course, if x is an invertible element of $\mathcal{A}$, then x^* is invertible too, with

$$(x^*)^{-1} = (x^{-1})^*. \tag{19.14.1}$$

Put

$$U(\mathcal{A}) = \{x \in \mathcal{A} : x\, x^* = x^*\, x = e_{\mathcal{A}}\}, \tag{19.14.2}$$

as in Section 7.5. One can check that this is a subgroup of the group $G(\mathcal{A})$ of invertible elements of $\mathcal{A}$, as before. Equivalently,

$$U(\mathcal{A}) = \{x \in G(\mathcal{A}) : x^{-1} = x^*\}. \tag{19.14.3}$$

If $x \in U(\mathcal{A})$, then

$$(x^*)^* = (x^{-1})^* = (x^*)^{-1} = (x^{-1})^{-1} = x. \tag{19.14.4}$$

Let us say that $y \in \mathcal{A}$ is *self-adjoint* with respect to $x \mapsto x^*$ if

$$y^* = y, \tag{19.14.5}$$

and that y is *anti-self-adjoint* if

$$y^* = -y, \tag{19.14.6}$$

as in Section 7.5. If $x \mapsto x^*$ is an involution on $\mathcal{A}$, then

$$y + y^* \tag{19.14.7}$$

is self-adjoint, and

$$y - y^* \tag{19.14.8}$$

is anti-self-adjoint. If k has characteristic 2, then (19.14.5) and (19.14.6) are the same, and (19.14.7) is equal to (19.14.8). If k does not have characteristic 2, and if $x \mapsto x^*$ is an involution on $\mathcal{A}$, then every element of $\mathcal{A}$ can be expressed in a unique way as a sum of self-adjoint and anti-self-adjoint elements, as in Section 7.5.

Let a be an invertible element of $\mathcal{A}$, so that

$$x \mapsto a\, x\, a^{-1} \tag{19.14.9}$$

is an automorphism of $\mathcal{A}$. This implies that

$$x^{*_a} = a\, x^*\, a^{-1} \tag{19.14.10}$$

also defines an opposite algebra isomorphism from $\mathcal{A}$ onto itself. If $x \in \mathcal{A}$, then

$$(x^{*_a})^{*_a} = a\, (a\, x^*\, a^{-1})^*\, a^{-1} = a\, (a^{-1})^*\, (x^*)^*\, a^*\, a^{-1}. \tag{19.14.11}$$

If a is self-adjoint, then a^{-1} is self-adjoint, and

$$(x^{*_a})^{*_a} = (x^*)^* \tag{19.14.12}$$

for all $x \in \mathcal{A}$. Similarly, if a is anti-self-adjoint, then a^{-1} is anti-self-adjoint, and (19.14.12) holds. If $x \mapsto x^*$ is an involution on $\mathcal{A}$, then it follows that (19.14.10) defines an involution on $\mathcal{A}$ in both of these cases.

Let V be a finite-dimensional vector space over k, and let b be a nondegenerate bilinear form on V. Thus $T \mapsto T^{*_b}$ defines an opposite algebra isomorphism from $\mathcal{L}(V)$ onto itself, as in Section 19.12. If $T \in \mathcal{L}(V)$, then

$$b(T(v), T(w)) = b(v, T^{*_b}(T(w))) \tag{19.14.13}$$

for every $v, w \in V$. This means that

$$b(T(v), T(w)) = b(v, w) \tag{19.14.14}$$

for all $v, w \in V$ if and only if

$$T^{*_b} \circ T = I \tag{19.14.15}$$

on V.

Note that (19.14.15) implies that is injective on V. This means that T is invertible on V, because V has finite dimension. Thus (19.14.15) is the same as saying that T is invertible on V, with

$$T^{-1} = T^{*_b}. \tag{19.14.16}$$

This is related to remarks in Sections 2.10, 3.5, and 7.5 when $k = \mathbf{R}$ and b is an inner product on V.

19.15 Nondegenerate sesquilinear forms

Let V be a finite-dimensional vector space over the complex numbers, and let b be a sesquilinear form on V, as in Section 5.15. We say that b is *nondegenerate* on V if for every $v \in V$ with $v \neq 0$ there is a $w \in V$ such that

$$b(v, w) \neq 0. \tag{19.15.1}$$

Remember that $V_{\mathbf{R}}$ is the same as V, considered as a vector space over the real numbers, as in Section 1.1. The dimension of $V_{\mathbf{R}}$ is equal to twice the dimension of V. It is easy to see that the real part of b defines a bilinear form on $V_{\mathbf{R}}$. One can check that b is nondegenerate on V if and only if

$$\operatorname{Re} b \text{ is nondegenerate on } V_{\mathbf{R}}. \tag{19.15.2}$$

As in Section 19.12, (19.15.2) holds if and only if for every $w \in V$ with $w \neq 0$ there is a $v \in V$ such that

$$\operatorname{Re} b(v, w) \neq 0. \tag{19.15.3}$$

One can use this to get that b is nondegenerate on V if and only if for every $w \in V$ with $w \neq 0$ there is a $v \in V$ such that (19.15.1) holds.

If $w \in V$, then $b_{1,w}(v) = b(v, w)$ defines a linear functional on V, as before. In this case, $w \mapsto b_{1,w}$ defines a conjugate-linear mapping from V into V^{alg}. Conversely, any conjugate-linear mapping from V into V^{alg} corresponds to a sesquilinear form on V in this way.

Similarly, if $v \in V$, then

$$\widetilde{b}_{2,v}(w) = \overline{b(v,w)} \tag{19.15.4}$$

is complex-linear in w, as in Section 5.15. We also have that

$$v \mapsto \widetilde{b}_{2,v} \tag{19.15.5}$$

defines a conjugate-linear mapping from V into V^{alg}, and that every such mapping corresponds to a sesquilinear form on V in this way.

Observe that b is nondegenerate on V if and only if

$$v \mapsto \widetilde{b}_{2,v} \text{ is one-to-one on } V. \tag{19.15.6}$$

This is equivalent to the condition that

$$v \mapsto \widetilde{b}_{2,v} \text{ maps } V \text{ onto } V^{\mathrm{alg}}, \tag{19.15.7}$$

because V and V^{alg} have the same dimension. One can consider (19.15.5) as a real-linear mapping here, in order to use standard results from linear algebra.

Similarly,

$$w \mapsto b_{1,w} \text{ is one-to-one on } V, \tag{19.15.8}$$

as a conjugate-linear mapping from V into V^{alg}, if and only if for every $w \in V$ with $w \neq 0$ there is a $v \in V$ such that (19.15.1) holds. This condition holds if and only if

$$w \mapsto b_{1,w} \text{ maps } V \text{ onto } V^{\mathrm{alg}}, \tag{19.15.9}$$

as in the preceding paragraph. It is easy to see that (19.15.9) implies the earlier formulation of nondegeneracy of b on V, as in Section 19.12. Conversely, (19.15.7) implies the version of nondegeneracy that is equivalent to (19.15.8), as before. This is another way to look at the equivalence of these two formuations of nondegeneracy of b on V.

Suppose that b is nondegenerate on V, let T be a linear mapping from V into itself, and put

$$\mu_w(v) = \mu_{T,w}(v) = b(T(v), w) \tag{19.15.10}$$

for every $v, w \in V$, as in Section 19.12. If $w \in V$, then (19.15.10) defines a linear functional on V, and thus there is a unique element $T^{*_b}(w)$ ov V such that

$$b(T(v), w) = \mu_w(v) = b(v, T^{*_b}(w)) \tag{19.15.11}$$

for all $v \in V$, as in the previous paragraph. This defines a linear mapping T^{*_b} from V into itself, as in Section 19.12, which may be called the *adjoint* of T with respect to b. This corresponds to the adjoint of T in the sense of Section 3.5 when b is an inner product on V.

In this case,

$$T \mapsto T^{*_b} \tag{19.15.12}$$

is a conjugate-linear mapping from $\mathcal{L}(V)$ into itself. If T_1, T_2 are linear mappings from V into itself, then

$$(T_2 \circ T_1)^{*_b} = T_1^{*_b} \circ T_2^{*_b}, \tag{19.15.13}$$

as in Section 19.12.

Observe that

$$c(w, v) = \overline{b(v, w)} \tag{19.15.14}$$

is another sesquilinear form on V, and that the nondegeneracy of b on V is equivalent to the nondegeneracy of c on V, by the earlier remarks. If T is a linear mapping from V into itself, then there is a unique linear mapping T^{*_c} from V into itself such that

$$c(T(v), w) = c(v, T^{*_c}(w)) \tag{19.15.15}$$

for every $v, w \in V$, as before, which is the same as saying that

$$b(w, T(v)) = b(T^{*_c}(w), v) \tag{19.15.16}$$

for all $v, w \in V$. This implies that

$$T = (T^{*_c})^{*_b}, \tag{19.15.17}$$

as in Section 19.13. Similarly,

$$T = (T^{*_b})^{*_c}, \tag{19.15.18}$$

so that $T \mapsto T^{*_c}$ is the inverse of (19.15.12), as before.

Note that $b = c$ exactly when b is Hermitian symmetric on V, as in Section 8.5. In this case,

$$(T^{*_b})^{*_b} = T \tag{19.15.19}$$

for every $T \in \mathcal{L}(V)$, so that (19.15.12) defines a conjugate-linear algebra involution on $\mathcal{L}(V)$.

If A is a linear mapping from V into itself, then

$$b_A(v, w) = b(A(v), w) \tag{19.15.20}$$

defines another sesquilinear form on V, and one can check that every sesquilinear form on V corresponds to a unique $A \in \mathcal{L}(V)$ in this way, because b is nondegenerate on V, as in Section 19.13. Note that

$$b_A \text{ is nondegenerate on } V \tag{19.15.21}$$

exactly when A is injective on V, which means that A is invertible on V, because V has finite dimension, by hypothesis, as before. In this case, if $T \in \mathcal{L}(V)$, and $T^{*_{b_a}}$ is defined using b_A in the same way as before, then one can check that

$$T^{*_{b_A}} = (A^{-1})^{*_b} \circ T^{*_b} \circ A^{*_b}, \tag{19.15.22}$$

as in Section 19.13.

If b is Hermitian symmetric, then one can verify that b_A is Hermitian symmetric on V exactly when $A^{*_b} = A$.

If $T \in \mathcal{L}(V)$, then

$$b(T(v), T(w)) = b(v, w) \tag{19.15.23}$$

for every $v, w \in V$ if and only if

$$T^{*_b} \circ T = I \tag{19.15.24}$$

on V, as in the previous section. This happens exactly when T is invertible on V, with

$$T^{-1} = T^{*_b}, \tag{19.15.25}$$

as before. This is related to remarks in Sections 2.10, 3.5, and 7.5 when b is an inner product on V, as before.

Chapter 20

Representations of algebras

20.1 Representations and modules

Let k be a field, and let $\mathcal{A}$ be an associative algebra over k. A *representation* $\rho = \rho^V$ of $\mathcal{A}$ on a vector space V over k is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(V)$. If $x \in \mathcal{A}$, then we may use $\rho_x = \rho_x^V$ for the corresponding linear mapping on V. Thus

$$\rho_{x\,y} = \rho_x \circ \rho_y \tag{20.1.1}$$

for every $x, y \in \mathcal{A}$, as in (19.1.2). If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then we may wish to ask that

$$\rho_{e_{\mathcal{A}}} = I_V, \tag{20.1.2}$$

as in (19.1.3).

This corresponds to Definition 21.4 on p312 of [91], and to the definition on p384 of [124]. More precisely, if $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then one can look at the behavior of $\rho_{e_{\mathcal{A}}}$ on $\mathcal{A}$, as in Note 21.5 on p313 of [91], and we shall return to this in Section 20.9. After those remarks, it is stated in [91] that representations of an associative algebra with a multiplicative identity element will be asked to satisfy (20.1.2). In [124], $\mathcal{A}$ is asked to have a multiplicative identity element, and (20.1.2) is included in the definition of a representation, because the analogous condition is included in the definition of a ring homomorphism.

If $x \in \mathcal{A}$ and $v \in V$, then we may put

$$x \cdot v = \rho_x(v), \tag{20.1.3}$$

to indicate the action of ρ_x on V. Using this notation, (20.1.1) is the same as saying that

$$(x\,y) \cdot v = x \cdot (y \cdot v) \tag{20.1.4}$$

for every $x, y \in \mathcal{A}$ and $v \in V$. Similarly, if $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then (20.1.2) is the same as saying that

$$e_{\mathcal{A}} \cdot v = v \tag{20.1.5}$$

for every $v \in V$. We may also refer to the representation of $\mathcal{A}$ on V by saying that V is a *left module* over $\mathcal{A}$.

A vector space V over k is said to be a *right module* over $\mathcal{A}$ if it is equipped with an action of $\mathcal{A}$ on the right that corresponds to an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(V)$. This means that if $x \in \mathcal{A}$ and $v \in V$, then

$$v \cdot x \tag{20.1.6}$$

is defined as an element of V, with the following properties. If $x \in \mathcal{A}$, then

$$v \mapsto v \cdot x \tag{20.1.7}$$

is a linear mapping from V into itself. If $v \in V$, then

$$x \mapsto v \cdot x \tag{20.1.8}$$

is a linear mapping from $\mathcal{A}$ into V. This means that the mapping from $x \in \mathcal{A}$ to (20.1.7) is linear as a mapping from $\mathcal{A}$ into $\mathcal{L}(V)$. The condition that this mapping be an opposite algebra homomorphism is the same as saying that

$$v \cdot (x\, y) = (v \cdot x) \cdot y \tag{20.1.9}$$

for every $x, y \in \mathcal{A}$ and $v \in V$. If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then we may wish to ask that

$$v \cdot e_{\mathcal{A}} = v \tag{20.1.10}$$

for every $v \in V$, as before.

A left module over $\mathcal{A}$ is sometimes simply called a *module* over $\mathcal{A}$. Of course, if $\mathcal{A}$ is commutative, then left and right modules over $\mathcal{A}$ are the same.

20.2 Some examples and remarks

Let k be a field, and let V be a vector space over k. If $\mathcal{A}$ is a subalgebra of $\mathcal{L}(V)$, then the obvious inclusion mapping from $\mathcal{A}$ into $\mathcal{L}(V)$ defines a representation of $\mathcal{A}$ on V.

Let $\mathcal{A}$ be any associative algebra over k, and if $a, x \in \mathcal{A}$, then put

$$L_a(x) = a\, x, \tag{20.2.1}$$

as in Sections 6.3 and 19.4. This defines L_a as a linear mapping from $\mathcal{A}$ into itself, which is the *left multiplication operator* associated to a. In fact,

$$a \mapsto L_a \tag{20.2.2}$$

defines a homomorphism from $\mathcal{A}$ into the algebra $\mathcal{L}(\mathcal{A})$ of linear mappings on $\mathcal{A}$, as before, and thus a representation of $\mathcal{A}$ on itself, as a vector space over k. Equivalently, $\mathcal{A}$ may be considered as a left module over itself, as a vector space over k, where the action of $a \in \mathcal{A}$ on $\mathcal{A}$ on the left is defined using multiplication on $\mathcal{A}$.

Similarly, $\mathcal{A}$ may be considered as a right module over itself, as a vector space over k, where the action of $a \in \mathcal{A}$ on $\mathcal{A}$ on the right is defined using multiplication in $\mathcal{A}$ again. Equivalently, if $a, x \in \mathcal{A}$, then put

$$R_a(x) = x\,a, \tag{20.2.3}$$

as in Sections 6.4 and 19.4. This defines a linear mapping from $\mathcal{A}$ into itself, which is the *right multiplication operator* associated to a. We also have that

$$a \mapsto R_a \tag{20.2.4}$$

defines an opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{L}(\mathcal{A})$, as before.

Let V be a vector space over k again. If V is a left module over $\mathcal{A}$, then

$$(a, v) \mapsto a \cdot v \tag{20.2.5}$$

is a bilinear mapping from $\mathcal{A} \times V$ into V. Similarly, if V is a right module over $\mathcal{A}$, then

$$(v, a) \mapsto v \cdot a \tag{20.2.6}$$

is a bilinear mapping from $V \times \mathcal{A}$ into V.

Suppose now that $k = \mathbf{R}$ or $\mathbf{C}$, and that V is a vector space over k with a norm $\|\cdot\|_V$. In this case, we may be interested in representations of associative algebras $\mathcal{A}$ over k corresponding to homomorphisms from $\mathcal{A}$ into the algebra $\mathcal{BL}(V)$ of bounded linear mappings on V. Similarly, we may be interested in actions of $\mathcal{A}$ on V on the right, corresponding to opposite algebra homomorphisms from $\mathcal{A}$ into $\mathcal{BL}(V)$.

If $\|\cdot\|_{\mathcal{A}}$ is a norm on $\mathcal{A}$, then we may be interested in representations of $\mathcal{A}$ on V corresponding to homomorphisms from $\mathcal{A}$ into $\mathcal{BL}(V)$ that are bounded linear mappings, with respect to the operator norm on $\mathcal{BL}(\mathcal{A})$ associated to $\|\cdot\|_V$. This means that there is a nonnegative real number C such that

$$\|a \cdot v\|_V \le C \|a\|_{\mathcal{A}} \, \|v\|_V \tag{20.2.7}$$

for every $a \in \mathcal{A}$ and $v \in V$. This is the same as saying that (20.2.5) is bounded as a bilinear mapping from $\mathcal{A} \times V$ into V, as in Section 5.13.

Similarly, we may be interested in actions of $\mathcal{A}$ on V on the right that correspond to opposite algebra homomorphisms from $\mathcal{A}$ into $\mathcal{BL}(V)$ that are bounded as linear mappings, so that

$$\|v \cdot a\|_V \le C \, \|a\|_{\mathcal{A}} \, \|v\|_V \tag{20.2.8}$$

for some $C \ge 0$ and all $a \in \mathcal{A}$, $v \in V$. This is the same as saying that (20.2.6) is bounded as a bilinear mapping from $V \times \mathcal{A}$ into V. If $\|\cdot\|_{\mathcal{A}}$ is submultiplicative on $\mathcal{A}$, then the actions of $\mathcal{A}$ on itself on the left and right satisfy these conditions with $C = 1$.

20.3 Opposite algebras

Let k be a field, and let $\mathcal{A}$ be an algebra in the strict sense over k, as in Section 19.1. The corresponding *opposite algebra* $\mathcal{A}^{op}$ is defined to be the same as $\mathcal{A}$ as a vector space over k, with the order of multiplication exchanged. More precisely, if $x \in \mathcal{A}$, then we may use x^{op} to refer to x as an element of $\mathcal{A}^{op}$. With this notation, multiplication in $\mathcal{A}^{op}$ is defined by

$$x^{op}\, y^{op} = (y\, x)^{op}. \tag{20.3.1}$$

Of course, $\mathcal{A}^{op}$ is an algebra in the strict sense over k too.

Thus multiplication in $\mathcal{A}^{op}$ is the same as in $\mathcal{A}$ exactly when $\mathcal{A}$ is commutative. If $\mathcal{A}$ is associative, then $\mathcal{A}^{op}$ is associative as well. This corresponds to the notion of an opposite ring, as on p109 of [43], and on p185 of [138].

If $\mathcal{A}$ has a multiplicative identity element $e_{\mathcal{A}}$, then $e_{\mathcal{A}}^{op}$ is the multiplicative identity element in $\mathcal{A}^{op}$. Clearly $\mathcal{A}$ and $\mathcal{A}^{op}$ have the same subalgebras, and the same two-sided ideals. Left ideals in $\mathcal{A}$ correspond to right ideals in $\mathcal{A}^{op}$, and right ideals in $\mathcal{A}$ correspond to left ideals in $\mathcal{A}^{op}$.

Let $\mathcal{B}$ be another algebra over k in the strict sense. An opposite algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ is the same as an algebra homomorphism from $\mathcal{A}^{op}$ into $\mathcal{B}$. This is also the same as an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}^{op}$. Similarly, an opposite algebra isomorphism from $\mathcal{A}$ onto $\mathcal{B}$ corresponds exactly to an algebra isomorphism from $\mathcal{A}^{op}$ onto $\mathcal{B}$, and to an algebra isomorphism from $\mathcal{A}$ onto $\mathcal{B}^{op}$.

Suppose that $\mathcal{A}$ is an associative algebra over k, and let V be a vector space over k. If V is a left module over $\mathcal{A}$, then V may be considered as a right module over $\mathcal{A}^{op}$, with

$$v \cdot x^{op} = x \cdot v \tag{20.3.2}$$

for every $x \in \mathcal{A}$ and $v \in V$. Similarly, if V is a right module over $\mathcal{A}$, then V may be considered as a left module over $\mathcal{A}^{op}$, with

$$x^{op} \cdot v = v \cdot x \tag{20.3.3}$$

for every $x \in \mathcal{A}$ and $v \in V$. This corresponds to Proposition 16 on p185 of [138], for modules over rings, and it is also mentioned on p109 of [43].

20.4 Invariant linear subspaces

Let k be a field, let $\mathcal{A}$ be an associative algebra over k, and let ρ^V be a representation of $\mathcal{A}$ on a vector space V over k. A linear subspace W of V is said to be *invariant under* ρ^V if

$$\rho_x^V(W) \subseteq W \tag{20.4.1}$$

for every $x \in \mathcal{A}$. In this case, for each $x \in \mathcal{A}$,

$$\text{let } \rho_x^W \text{ be the restriction of } \rho_x^V \text{ to } W. \tag{20.4.2}$$

It is easy to see that this defines a representation of $\mathcal{A}$ on W. This may be called a *subrepresentation* of ρ^V.

Equivalently, if V is a left module over $\mathcal{A}$, then a linear subspace W of V is said to be a *submodule* of V if

$$x \cdot w \in W \tag{20.4.3}$$

for every $x \in \mathcal{A}$ and $w \in W$. This means that W may be considered as a left module over $\mathcal{A}$ as well, using the restriction of the action of $\mathcal{A}$ on V to W.

Similarly, if V is a right module over $\mathcal{A}$, then a linear subspace W of V is said to be a *submodule* of V if

$$w \cdot x \in W \tag{20.4.4}$$

for every $x \in \mathcal{A}$ and $w \in W$. This implies that W may be considered as a right module over $\mathcal{A}$, with respect to the restriction of the action of $\mathcal{A}$ on V to W.

Remember that $\mathcal{A}$ may be considered as a left and right module over itself, as in Section 20.2. A linear subspace $\mathcal{I}$ of $\mathcal{A}$ is a submodule of $\mathcal{A}$, as a left or right module over itself, exactly when $\mathcal{I}$ is a left or right ideal in $\mathcal{A}$, as appropriate, as in Section 19.2.

Suppose for the moment that $k = \mathbf{R}$ or $\mathbf{C}$, and that V is a vector space over k with a norm $\|\cdot\|_V$. Suppose also that V is a left or right module over $\mathcal{A}$, and that the action of each element of $\mathcal{A}$ on V is a bounded linear mapping on V. If W is a submodule of V, as a module over $\mathcal{A}$, then it is easy to see that

$$\overline{W} \text{ is a submodule of } V, \tag{20.4.5}$$

as a module over $\mathcal{A}$. Here $\overline{W}$ is the closure of W in V with respect to the metric associated to $\|\cdot\|_V$, as usual.

Let k be any field again, let $X_1, \ldots, X_n$ be commuting indeterminates, and let $k[X_1, \ldots, X_n]$ be the algebra of formal polynomials in $X_1, \ldots, X_n$ with coefficients in k, as in Section 19.6. Also let V be a vector space over k, and let $T_1, \ldots, T_n$ be n commuting linear mappings from V into itself, so that

$$T_j \circ T_l = T_l \circ T_j \tag{20.4.6}$$

for all $j, l = 1, \ldots, n$. This leads to a representation of $k[X_1, \ldots, X_n]$ on V, as in Section 19.6 again. This representation is characterized by the properties that

$$1_k \cdot v = v \tag{20.4.7}$$

for every $v \in V$, and

$$X_j \cdot v = T_j(v) \tag{20.4.8}$$

for each $j = 1, \ldots, n$ and $v \in V$. If W is a linear subspace of V such that

$$T_j(W) \subseteq W \tag{20.4.9}$$

for each $j = 1, \ldots, n$, then W is invariant under this representation.

In particular, we can take $n = 1$ here. Equivalently, if X is an indeterminate, then we let $k[X]$ be the algebra of formal polynomials in X with coefficients in k, as before. If T is a linear mapping from V into itself, then there is a unique representation of $k[X]$ on V that satisfies (20.4.7) and

$$X \cdot v = T(v) \tag{20.4.10}$$

for every $v \in V$. If W is a linear subspace of V such that

$$T(W) \subseteq W, \tag{20.4.11}$$

then W is invariant under this representation.

20.5 More on invariant linear subspaces

Let k be a field, let V be a vector space over k, and let W be a linear subspace of V. Remember that $\mathcal{L}_W(V)$ be the subalgebra of $\mathcal{L}(V)$ of linear mappings T from V into itself that map W into itself, as in Section 19.8. One may consider V as a left module over $\mathcal{L}_W(V)$, as mentioned at the beginning of Section 20.2. Of course, W is a submodule of V, as a left module over $\mathcal{L}_W(V)$, by construction.

Suppose now that $k = \mathbf{R}$ or $\mathbf{C}$, and that $\|\cdot\|_V$ is a norm on V. Consider the space

$$\mathcal{BL}_W(V) = \mathcal{BL}(V) \cap \mathcal{L}_W(V) \tag{20.5.1}$$

of bounded linear mappings T from V into itself such that $T(W) \subseteq W$. This is a subalgebra of $\mathcal{BL}(V)$ that contains the identity mapping on V. If W is also a closed set in V, with respect to the metric associated to $\|\cdot\|_V$, then it is easy to see that

$$\mathcal{BL}_W(V) \text{ is a closed set in } \mathcal{BL}(V), \tag{20.5.2}$$

with respect to the metric associated to the operator norm. Note that

$$\mathcal{BL}_W(V) \subseteq \mathcal{BL}_{\overline{W}}(V). \tag{20.5.3}$$

If $T \in \mathcal{L}_W(V)$, then let T_W be the restriction of T to W, as in Section 19.8. Remember that $T \mapsto T_W$ defines a homomorphism from $\mathcal{L}_W(V)$ into $\mathcal{L}(W)$, as algebras over k, as before. The restriction of this homomorphism to $\mathcal{BL}_W(V)$ defines a homomorphism into $\mathcal{BL}(W)$. More precisely, if $T \in \mathcal{BL}_W(V)$, then $T_W \in \mathcal{BL}(W)$, and the operator norm of T_W on W is less than or equal to the operator norm of T on V.

The kernel of $T \mapsto T_W$ on $\mathcal{BL}_W(V)$ is equal to

$$\{T \in \mathcal{BL}(V) : T = 0 \text{ on } W\}. \tag{20.5.4}$$

This is a two-sided ideal in $\mathcal{BL}_W(V)$, and a left ideal in $\mathcal{BL}(V)$. Note that (20.5.4) is a closed set in $\mathcal{BL}(V)$, with respect to the metric associated to the operator norm. It is easy to see that (20.5.4) is the same as

$$\{T \in \mathcal{BL}(V) : T = 0 \text{ on } \overline{W}\}. \tag{20.5.5}$$

Let V/W be the usual quotient space, and let $q_{V/W}$ be the corresponding quotient mapping from V onto V/W. Suppose from now on in this section that W is a closed linear subspace of V, and let $\|\cdot\|_{V/W}$ be the corresponding quotient norm on V/W, as in Section 6.10. If $T \in \mathcal{BL}_W(V)$, then there is a unique linear mapping $T^{V/W}$ from V/W into itself such that

$$T^{V/W} \circ q_{V/W} = q_{V/W} \circ T, \tag{20.5.6}$$

as in Section 19.8. Let us check that $T^{V/W}$ is bounded, with respect to the quotient norm.

If $v \in V$ and $w \in W$, then

$$\begin{aligned} \|T^{V/W}(q_{V/W}(v))\|_{V/W} &= \|T^{V/W}(q_{V/W}(v-w))\|_{V/W} \\ &= \|q_{V/W}(T(v-w))\|_{V/W} \\ &\le \|T(v-w)\|_V \le \|T\|_{op,VV}\,\|v-w\|_V, \end{aligned} \tag{20.5.7}$$

using the definition of the quotient norm in the third step. This implies that

$$\|T^{V/W}(q_{V/W}(v))\|_{V/W} \le \|T\|_{op,VV}\,\|q_{V/W}(v)\|_{V/W}, \tag{20.5.8}$$

by the definition of the quotient norm. This shows that $T^{V/W}$ is bounded on V/W, with operator norm less than or equal to the operator norm of T on V.

As in Section 19.8,

$$T \mapsto T^{V/W} \tag{20.5.9}$$

is a homomorphism from $\mathcal{BL}_W(V)$ into $\mathcal{BL}(V/W)$, as algebras over k. The kernel of this homomorphism is equal to

$$\{T \in \mathcal{BL}(V) : T(V) \subseteq W\}, \tag{20.5.10}$$

as before. This is another two-sided ideal in $\mathcal{BL}_W(V)$, and a right ideal in $\mathcal{BL}(V)$. We also have that (20.5.10) is a closed set in $\mathcal{BL}(V)$, with respect to the metric associated to the operator norm, because W is a closed set in V, by hypothesis.

20.6 Quotient representations

Let k be a field, let $\mathcal{A}$ be an associative algebra over k, and let ρ^V be a representation of $\mathcal{A}$ on a vector space V over k. Also let W be a linear subspace of V that is invariant under ρ, as in Section 20.4. If $x \in \mathcal{A}$, and $q_{V/W}$ is the natural quotient mapping from V onto V/W, then there is a unique linear mapping $\rho_x^{V/W}$ from V/W into itself such that

$$\rho_x^{V/W} \circ q_{V/W} = q_{V/W} \circ \rho_x^V, \tag{20.6.1}$$

as in Section 19.8. One can check that

$$\rho^{V/W} \text{ is a representation of } \mathcal{A} \text{ on } V/W. \tag{20.6.2}$$

This may be called a *quotient representation* of ρ^V.

Equivalently, if V is a left module over $\mathcal{A}$, and W is a submodule of V, then there is a unique action of $\mathcal{A}$ on V/W such that

$$x \cdot q_{V/W}(v) = q_{V/W}(x \cdot v) \tag{20.6.3}$$

for every $x \in \mathcal{A}$ and $v \in V$. This makes V/W a left module over $\mathcal{A}$, which may be called a *quotient module*. Similarly, if V is a right module over $\mathcal{A}$, and W is a submodule of V, then there is a unique action of $\mathcal{A}$ on V/W such that

$$q_{V/W}(v) \cdot x = q_{V/W}(v \cdot x) \tag{20.6.4}$$

for every $x \in \mathcal{A}$ and $v \in V$. This makes V/W a right module over $\mathcal{A}$, which may also be called a quotient module.

In particular, if $\mathcal{I}$ is a left or right ideal in $\mathcal{A}$, then

(20.6.5) $\mathcal{A}/\mathcal{I}$ may be considered as a left or right module over $\mathcal{A}$,

as appropriate. If $\mathcal{I}$ is a two-sided ideal in $\mathcal{A}$, then $\mathcal{A}$ may be considered as both a left and right module over $\mathcal{A}$, and as an associative algebra over k, as in Section 19.2. If $q_{\mathcal{A}/\mathcal{I}}$ is the natural quotient homomorphism from $\mathcal{A}$ onto $\mathcal{A}/\mathcal{I}$, as in Section 19.2, then the action of $x \in \mathcal{A}$ on $\mathcal{A}/\mathcal{I}$ on the left and the right corresponds to multiplication by $q_{\mathcal{A}/\mathcal{I}}(x)$ on the left and the right in $\mathcal{A}/\mathcal{I}$.

Suppose now that $k = \mathbf{R}$ or $\mathbf{C}$, and that V is a vector space over k with a norm $\|\cdot\|_V$. Let W be a closed linear subspace of V, and let $\|\cdot\|_{V/W}$ be the corresponding quotient norm on V/W, as in Section 6.10. Suppose that V is a left or right module over $\mathcal{A}$, where the action of elements of $\mathcal{A}$ on V are bounded linear mappings on V. In this case, the action of elements of $\mathcal{A}$ on V/W are bounded linear mappings on V/W, as in the previous section. More precisely, the operator norm of the action of $a \in \mathcal{A}$ on V/W is less than or equal to the operator norm of the action of a on V, as before.

Let $\|\cdot\|_{\mathcal{A}}$ be a norm on $\mathcal{A}$, and suppose that there is a nonnegative real number C such that the operator norm of the action of $a \in \mathcal{A}$ on V is less than or equal $C\,\|a\|_{\mathcal{A}}$. This implies that the operator norm of the action of a on V/W is less than or equal to $C\,\|a\|_{\mathcal{A}}$ as well.

20.7 Homomorphisms between representations

Let k be a field, let $\mathcal{A}$ be an associative algebra over k, and let V_1, V_2 be vector spaces over k. Also let ρ^{V_1}, ρ^{V_2} be representations of $\mathcal{A}$ on V_1, V_2, respectively. A linear mapping T from V_1 into V_2 is said to *intertwine* these two representations if

$$T \circ \rho_x^{V_1} = \rho_x^{V_2} \circ T, \tag{20.7.1}$$

as linear mappings from V_1 into V_2, for every $x \in \mathcal{A}$. We may also say that T is a *homomorphism* between these two representations in this case.

Equivalently, if V_1 and V_2 are left modules over $\mathcal{A}$, then a linear mapping T from V_1 into V_2 is said to be a *homomorphism* from V_1 into V_2, as modules over $\mathcal{A}$, if

$$T(x \cdot v_1) = x \cdot T(v_1) \tag{20.7.2}$$

for every $x \in \mathcal{A}$ and $v_1 \in V_1$. Similarly, if V_1 and V_2 are right modules over $\mathcal{A}$, then a linear mapping T from V_1 into V_2 is said to be a *homomorphism* from V_1 into V_2, as modules over $\mathcal{A}$, if

$$T(v_1 \cdot x) = T(v_1) \cdot x \tag{20.7.3}$$

for every $x \in \mathcal{A}$ and $v_1 \in V_!$.

Let V be a left or right module over $\mathcal{A}$, and let W be a submodule of V. The obvious inclusion mapping from W into V may be considered as a homomorphism from W into V, as modules over $\mathcal{A}$. The corresponding quotient mapping $q_{V/W}$ from V onto V/W is a module homomorphism too, as in the previous section.

Let V_1, V_2 be both left or both right modules over $\mathcal{A}$ again, and let T be a homomorphism from V_1 into V_2, as modules over $\mathcal{A}$. It is easy to see that

$$\ker T \text{ is a submodule of } V_1 \tag{20.7.4}$$

and

$$T(V_1) \text{ is a submodule of } V_2. \tag{20.7.5}$$

If T is also a one-to-one mapping from V_1 onto V_2, then T^{-1} is a homomorphism from V_2 onto V_1, as modules over $\mathcal{A}$. In this case, we may say that T is an *isomorphism* from V_1 onto V_2, as modules over $\mathcal{A}$. If V_1 and V_2 are left modules over $\mathcal{A}$, then we may say that T is an *isomorphism* between these two representations, or that these two representations of $\mathcal{A}$ are *equivalent*.

Suppose now that $k = \mathbf{R}$ or $\mathbf{C}$, and that $\| \cdot \|_{V_1}$, $\| \cdot \|_{V_2}$ are norms on V_1, V_2, respectively. If V_1, V_2 are both left or both right modules over $\mathcal{A}$, then one may be interested in module homomorphisms that are also bounded as linear mappings from V_1 into V_2.

Similarly, one may be interested in isomorphisms T from V_1 onto V_2 such that T and T^{-1} are both bounded linear mappings. One may also be interested in such an isomorphism T that is an isometric linear mapping from V_1 onto V_2.

In particular, if V_1 and V_2 are Hilbert spaces, then this means that T is a unitary mapping, as in Section 2.10. One may say that V_1 and V_2 are *unitarily equivalent* in this case, or that the corresponding representations are unitarily equivalent when V_1 and V_2 are left modules over $\mathcal{A}$. This corresponds to a definition on p58 of [8], and part of Definition 21.8 on p314 of [91], at least for a suitable class of representations.

20.8 Direct sums of representations

Let k be a field, and let $\mathcal{A}$ be an associative algebra over k. Also let V_1 and V_2 be vector spaces over k that are both left or both right modules over k. Remember that the Cartesian product $V_1 \times V_2$ of V_1 and V_2 may be considered as a vector space over k with respect to coordinatewise addition and scalar multiplication, which is the direct sum of V_1 and V_2, as in Section 5.12. We may consider $V = V_1 \times V_2$ as a left or right module over $\mathcal{A}$, as appropriate, with

$$x \cdot (v_1, v_2) = (x \cdot v_1, x \cdot v_2) \tag{20.8.1}$$

or

$$(v_1, v_2) \cdot x = (v_1 \cdot x, v_2 \cdot x) \tag{20.8.2}$$

for all $x \in \mathcal{A}$, $v_1 \in V_1$, and $v_2 \in V_2$, as appropriate. This is the *direct sum* of V_1 and V_2, as modules over $\mathcal{A}$.

If V_1 and V_2 are left modules over $\mathcal{A}$, then we can look at this in terms of the corresponding representations ρ^{V_1} and ρ^{V_2} of $\mathcal{A}$ as well. Namely,

$$\rho_x^V((v_1, v_2)) = (\rho_x^{V_1}(v_1), \rho_x^{V_2}(v_2)) \tag{20.8.3}$$

defines a representation ρ^V of $\mathcal{A}$ on V, which is the *direct sum* of ρ^{V_1} and ρ^{V_2}.

Note that

$$V_1 \times \{0\} \text{ and } \{0\} \times V_2 \text{ are submodules of } V. \tag{20.8.4}$$

Of course, $v_1 \mapsto (v_1, 0)$ and $v_2 \mapsto (0, v_2)$ are isomorphisms from V_1 and V_2 onto these submodules of V, respectively. One could also consider these mappings as injective homomorphisms from V_1 and V_2 into V, respectively, as modules over $\mathcal{A}$.

Similarly, the coordinate mappings $(v_1, v_2) \mapsto v_1$ and $(v_1, v_2) \mapsto v_2$ are homomorphisms from V onto V_1 and V_2, respectively, as modules over $\mathcal{A}$. The kernels of these homomorphisms are $\{0\} \times V_2$ and $V_1 \times \{0\}$, respectively.

Suppose now that $k = \mathbf{R}$ or $\mathbf{C}$. If $\langle \cdot, \cdot \rangle_{V_1}$ and $\langle \cdot, \cdot \rangle_{V_2}$ are inner products on V_1 and V_2, respectively, then we can get an inner product $\langle \cdot, \cdot \rangle_V$ on V as in Section 5.12. This is used in the definition of a direct sum on p57 of [8], under suitable conditions.

If $\| \cdot \|_{V_1}$, $\| \cdot \|_{V_2}$ are norms on V_1, V_2, respectively, then we can define a norm $\| \cdot \|_{V,p}$ on V for $1 \le p \le \infty$ as in Section 5.12. If T_1, T_2 are bounded linear mappings on V_1, V_2, respectively, then it is easy to see that

$$T((v_1, v_2)) = (T_1(v_1), T_2(v_2)) \tag{20.8.5}$$

defines a bounded linear mapping on V with respect to $\| \cdot \|_{V,p}$ for each p. More precisely, the operator norm of T with respect to $\| \cdot \|_{V,p}$ is equal to

$$\max(\|T_1\|_{op,V_1 V_1}, \|T_2\|_{op,V_2 V_2}) \tag{20.8.6}$$

for each p.

If the actions of $x \in \mathcal{A}$ on V_1 and V_2 are bounded linear mappings, then the action of x on V is a bounded linear mapping with respect to $\| \cdot \|_{V,p}$ for each p. Let $\| \cdot \|_{\mathcal{A}}$ be a norm on $\mathcal{A}$, and suppose that there are nonnegative real numbers C_1, C_2 such that the operator norms of the actions of x on V_1 and V_2 are less than or equal to $C_1 \|x\|_{\mathcal{A}}$ and $C_2 \|x\|_{\mathcal{A}}$, respectively. This implies that the operator norm of the action of x on V with respect to $\| \cdot \|_{V,p}$ is less than or equal to

$$\max(C_1, C_2) \|x\|_{\mathcal{A}} \tag{20.8.7}$$

for each p.

20.9 Nondegenerate representations

Let k be a field, let $\mathcal{A}$ be an associative algebra over k, and let V be a vector space over k that is a left or right module over $\mathcal{A}$. If V is a left module over $\mathcal{A}$, then it is easy to see that

$$\{v \in V : x \cdot v = 0 \text{ for every } x \in \mathcal{A}\} \tag{20.9.1}$$

is a submodule of V. Equivalently, if ρ^V is the corresponding representation on V, then this is the same as

$$\bigcap_{x \in \mathcal{A}} \ker \rho^V_x . \tag{20.9.2}$$

Similarly, if V is a right module over $\mathcal{A}$, then

$$\{v \in V : v \cdot x = 0 \text{ for every } x \in \mathcal{A}\} \tag{20.9.3}$$

is a submodule of V.

Let us say that V is *nondegenerate* as a left or right module over $\mathcal{A}$ if (20.9.1) or (20.9.3) is equal to $\{0\}$, as appropriate. If V is a left module over $\mathcal{A}$, corresponding to the representation ρ^V, then this is the same as saying that (20.9.2) is equal to $\{0\}$. In this case, we may say that ρ^V is *nondegenerate*, as a representation on V. This corresponds to the definition of nondegeneracy mentioned on p57 of [8], for a certain type of representation.

20.10

Let k be any field again, and let V be a vector space over k that is a left or right module over $\mathcal{A}$. Also let x be an idempotent element of $\mathcal{A}$, so that $x^2 = x$, as in Section 7.6. If V is a left module over $\mathcal{A}$, then it follows that

$$v \mapsto x \cdot v \tag{20.10.1}$$

is idempotent as an element of $\mathcal{L}(V)$. This means that (20.10.1) is a projection on V, as in Section 8.2. Similarly, if V is a right module over $\mathcal{A}$, then

$$v \mapsto v \cdot x \tag{20.10.2}$$

is a projection on V.

In particular, if $e_{\mathcal{A}}$ is a multiplicative identity element of $\mathcal{A}$, then

$$v \mapsto e_{\mathcal{A}} \cdot v \tag{20.10.3}$$

is a projection on V when V is a left module over $\mathcal{A}$, and

$$v \mapsto v \cdot e_{\mathcal{A}} \tag{20.10.4}$$

is a projection on V when V is a right module over $\mathcal{A}$. If P is any projection on V, then V corresponds to the direct sum of the kernel of P and $P(V)$, as a vector space over k, as in Section 8.2.

Part V

Algebras, norms, and operators, 4

Chapter 21

Group and semigroup algebras

Appendix A

Absolute values on fields

A.1 Metrics and ultrametrics

Let X be a set, and let $d(x, y)$ be a nonnegative real-valued function defined for $x, y \in X$. As usual, $d(x, y)$ is said to be a *metric* on X if it satisfies the following three conditions. First,

$$d(x, y) = 0 \text{ if and only if } x = y. \tag{A.1.1}$$

Second,

$$d(x, y) = d(y, x) \quad \text{for every } x, y \in X. \tag{A.1.2}$$

Third,

$$d(x, z) \leq d(x, y) + d(y, z) \quad \text{for every } x, y, z \in X, \tag{A.1.3}$$

which is known as the *triangle inequality*.

The condition

$$d(x, z) \leq \max(d(x, y), d(y, z)) \quad \text{for every } x, y, z \in X \tag{A.1.4}$$

is the *ultrametric* version of the triangle inequality. If $d(x, y)$ satisfies (A.1.1), (A.1.2), and (A.1.4) on X, then $d(x, y)$ is said to be an *ultrametric* on X. Of course, (A.1.4) implies (A.1.3), so that an ultrametric on X is a metric in particular.

The *discrete metric* on X is defined by putting $d(x, y)$ equal to 0 when $x = y$, and to 1 otherwise. It is easy to see that this is an ultrametric on X.

If $d(\cdot, \cdot)$ is any metric on X, then the *open ball* in X centered at $x \in X$ with radius $r > 0$ with respect to $d(\cdot, \cdot)$ is defined as usual by

$$B(x, r) = B_d(x, r) = \{y \in X : d(x, y) < r\}. \tag{A.1.5}$$

Similarly, the *closed ball* in X centered at x with radius $r \geq 0$ with respect to $d(\cdot, \cdot)$ is defined by

$$\overline{B}(x, r) = \overline{B}_d(x, r) = \{y \in X : d(x, y) \leq r\}. \tag{A.1.6}$$

It is well known that open balls are open sets in X, and that closed balls are closed sets. If $d(\cdot,\cdot)$ is an ultrametric on X, then one can check that open balls in X are closed sets. One can also verify that closed balls in X of positive radius are open sets in this case.

Let a be a positive real number. If $a \le 1$, then it is well known that

$$(r+t)^a \le r^a + t^a \tag{A.1.7}$$

for all nonnegative real numbers r, t, as in Section 1.12. If $d(\cdot,\cdot)$ is a metric on X, then it follows that

$$d(x,y)^a \tag{A.1.8}$$

is a metric on X as well. If $d(\cdot,\cdot)$ is an ultrametric on X, then it is easy to see that (A.1.8) is an ultrametric on X for every $a > 0$. In both cases, we have that

$$B_{d^a}(x, r^a) = B_d(x,r) \tag{A.1.9}$$

for every $x \in X$ and $r > 0$, and

$$\overline{B}_{d^a}(x, r^a) = \overline{B}_d(x,r) \tag{A.1.10}$$

for every $x \in X$ and $r \ge 0$.

In particular, one can use (A.1.9) to get that $d(\cdot,\cdot)$ and (A.1.8) determine the same topology on X. This means that convergence of sequences in X is the same for $d(\cdot,\cdot)$ and (A.1.8), which is easy to see directly anyway. These metrics also determine the same Cauchy sequences in X, for instance. Using this, one can check that X is complete as a metric space with respect to $d(\cdot,\cdot)$ if and only if X is complete with respect to (A.1.8).

If $d(\cdot,\cdot)$ is a metric on X and Y is a subset of X, then the restriction of $d(x,y)$ to $x,y \in Y$ defines a metric on Y. Similarly, if $d(\cdot,\cdot)$ is an ultrametric on X, then the restriction of $d(x,y)$ to $x,y \in Y$ is an ultrametric on Y.

See [42, 95, 195] for some topics related to ultrametrics.

A.2 Absolute value functions

Let k be a field. A nonnegative real-valued function $|\cdot|$ on k is said to be an *absolute value function* on k if it satisfies the following three conditions. First,

$$|x| = 0 \text{ if and only if } x = 0. \tag{A.2.1}$$

Second,

$$|x\,y| = |x|\,|y| \quad \text{for every } x,y \in k. \tag{A.2.2}$$

Third,

$$|x+y| \le |x| + |y| \quad \text{for every } x,y \in k. \tag{A.2.3}$$

Let us say that $|\cdot|$ is an *ultrametric absolute value function* on k if it satisfies (A.2.1), (A.2.2), and

$$|x+y| \le \max(|x|,|y|) \quad \text{for every } x,y \in k. \tag{A.2.4}$$

Clearly (A.2.4) implies (A.2.3), so that an ultrametric absolute value function on k is an absolute value function in the previous sense.

It is well known that the standard absolute values on $\mathbf{R}$ and $\mathbf{C}$ are absolute value functions in this sense. The *trivial absolute value function* may be defined on any field k by putting $|x|$ equal to 0 when $x = 0$, and to 1 otherwise. It is easy to see that this defines an ultrametric absolute value function on k.

If $|\cdot|$ is an absolute value function on a field k and k_0 is a subfield of k, then the restriction of $|x|$ to $x \in k_0$ defines an absolute value function on k_0. If $|\cdot|$ is an ultrametric absolute value function on k, then its restriction to k_0 is an ultrametric absolute value function as well.

If $|\cdot|$ is an absolute value function on a field k, then

$$d(x, y) = |x - y| \tag{A.2.5}$$

defines a metric on k. More precisely, in order to check that this is symmetric in x and y, one needs to verify that

$$|-1_k| = 1, \tag{A.2.6}$$

where 1_k is the multiplicative identity element in k, and $1 = 1_{\mathbf{R}}$ is the multiplicative identity element in $\mathbf{R}$. To do this, one should first get that

$$|1_k| = 1. \tag{A.2.7}$$

One can obtain (A.2.7) using the facts that $1_k^2 = 1_k$ and $1_k \neq 0$ in k, and one can obtain (A.2.6) from this and the fact that $(-1_k)^2 = 1_k$. If $|\cdot|$ is an ultrametric absolute value function on k, then (A.2.5) is an ultrametric on k.

Let p be a prime number. The *p-adic absolute value* $|x|_p$ of a rational number x is defined as follows. Of course, we put $|0|_p = 0$, and so we may suppose that $x \neq 0$. In this case,

$$x = p^j \, (a/b) \tag{A.2.8}$$

for some integers a, b, and j, where $a, b \neq 0$ and neither a nor b is a multiple of p, and we put

$$|x|_p = p^{-j}. \tag{A.2.9}$$

It is easy to see that this is well defined, even though the expression (A.2.8) for x is not unique.

One can check that this defines an ultrametric absolute value function on the field $\mathbf{Q}$ of rational numbers. This implies that

$$d_p(x, y) = |x - y|_p \tag{A.2.10}$$

defines an ultrametric on $\mathbf{Q}$, as before. This is known as the *p-adic metric* on $\mathbf{Q}$.

The field $\mathbf{Q}_p$ of *p-adic numbers* can be obtained by completing $\mathbf{Q}$ with respect to the p-adic metric, in much the same way that the real numbers can be obtained by completing $\mathbf{Q}$ with respect to the standard Euclidean metric.

Similarly, if k is a field with an absolute value function $|\cdot|$, and if k is not complete with respect to the associated metric (A.2.5), then one can pass to a completion. It is well known that the completion is also a field, and that $|\cdot|$ extends to an absolute value function on the completion in a natural way. If $|\cdot|$ is an ultrametric absolute value function on k, then its extension to the completion of k is an ultrametric absolute value function on the completion.

A.3 The archimedean property

Let k be a field. If $x \in k$ and n is a positive integer, then let $n \cdot x$ be the sum of n x's in k. It is easy to see that

$$m \cdot (n \cdot x) = (m\,n) \cdot x \tag{A.3.1}$$

for all $m, n \in \mathbf{Z}_+$ and $x \in k$. Note that

$$n \cdot x = (n \cdot 1_k)\, x \tag{A.3.2}$$

for every $n \in \mathbf{Z}_+$ and $x \in k$. We also have that

$$(m\,n) \cdot 1_k = m \cdot (n \cdot 1_k) = (m \cdot 1_k)\,(n \cdot 1_k) \tag{A.3.3}$$

for every $m, n \in \mathbf{Z}_+$.

Let $|\cdot|$ be an absolute value function on k. We say that $|\cdot|$ is *archimedean* on k if there are positive integers n such that

$$|n \cdot 1_k| \text{ is arbitrarily large.} \tag{A.3.4}$$

Thus $|\cdot|$ is *non-archimedean* on k if there is a positive real number C such that

$$|n \cdot 1_k| \le C \tag{A.3.5}$$

for all $n \in \mathbf{Z}_+$.

If there is a positive integer n_0 such that

$$|n_0 \cdot 1_k| > 1, \tag{A.3.6}$$

then $|\cdot|$ is archimedean on k. Indeed, in this case,

$$|n_0^j \cdot 1_k| = |(n_0 \cdot 1_k)^j| = |n_0 \cdot 1_k|^j \to \infty \quad \text{as } j \to \infty. \tag{A.3.7}$$

If $|\cdot|$ is non-archimedean on k, then it follows that (A.3.5) holds with $C = 1$.

If $|\cdot|$ is an ultrametric absolute value function on k, then it is easy to see that (A.3.5) holds with $C = 1$, so that $|\cdot|$ is non-archimedean on k. Conversely, if $|\cdot|$ is non-archimedean on k, then it is well known that $|\cdot|$ is an ultrametric absolute value function on k. This corresponds to Lemma 1.5 on p16 of [44], and Theorem 2.2.2 on p28 of [80]. Note that the term "non-archimedean" is sometimes used directly for absolute value functions that satisfy the ultrametric version of the triangle inequality, or something equivalent to it in a straightforward way, as in [44, 80]. This is equivalent to the terminology being used here, because of the result that was just mentioned.

A.4 Equivalent absolute value functions

Let k be a field, and let $|\cdot|$ be an absolute value function on k. If a is a positive real number, then

$$|x|^a \tag{A.4.1}$$

satisfies the first two conditions in the definition of an absolute value function. If $a \le 1$, then (A.4.1) satisfies the triangle inequality on k, because of (A.1.7). and thus defines an absolute value function on k. If $|\cdot|$ is an ultrametric absolute value function on k, then it is easy to see that (A.4.1) is an ultrametric absolute value function on k for every $a > 0$.

A pair $|\cdot|_1$, $|\cdot|_2$ of absolute value functions on k are said to be *equivalent* on k if there is a positive real number a such that

$$|x|_2 = |x|_1^a \tag{A.4.2}$$

for every $x \in k$. Of course, this implies that

$$|x-y|_2 = |x-y|_1^a \tag{A.4.3}$$

for every $x, y \in k$. In particular, this means that the metrics on k associated to $|\cdot|_1$ and $|\cdot|_2$ determine the same topology on k, as in Section A.1.

Conversely, if $|\cdot|_1$ and $|\cdot|_2$ are absolute value functions on k whose associated metrics determine the same topology on k, then it is well known that $|\cdot|_1$ and $|\cdot|_2$ are equivalent in the sense of (A.4.2). This corresponds to Lemma 3.2 on p20 of [44], and to Lemma 3.1.2 on p42 of [80].

A famous theorem of Ostrowski states that any absolute value function on $\mathbf{Q}$ is either equivalent to the standard Euclidean absolute value function, trivial, or equivalent to the p-adic absolute value function for some p. This corresponds to Theorem 2.1 on p16 of [44], and to Theorem 3.1.3 on p44 of [80].

Suppose that k is a field with an archimedean absolute value function $|\cdot|$, and that k is complete with respect to the metric associated to $|\cdot|$. Another famous theorem of Ostrowski states that k is isomorphic to $\mathbf{R}$ or $\mathbf{C}$, in such a way that $|\cdot|$ corresponds to an absolute value function that is equivalent to the standard Euclidean absolute value function. See Theorem 1.1 on p33 of [44].

A.5 Norms and seminorms

Let k be a field with an absolute value function $|\cdot|$, and let V be a vector space over k. A nonnegative real-valued function N on V is said to be a *seminorm* or *pseudonorm* with respect to $|\cdot|$ on k if

$$N(t\,v) = |t|\,N(v) \tag{A.5.1}$$

for every $t \in k$ and $v \in V$, and

$$N(v+w) \le N(v) + N(w) \tag{A.5.2}$$

for every $v, w \in V$. Of course, if $k = \mathbf{R}$ or $\mathbf{C}$, equipped with the standard Euclidean absolute value function, then this is the same as in Section 1.2.

If N is a nonnegative real-valued function on V that satisfies (A.5.1) and

$$N(v + w) \le \max(N(v), N(w)) \tag{A.5.3}$$

for every $v, w \in V$, then we say that N is a *semi-ultranorm* or *pseudo-ultranorm* on V with respect to $|\cdot|$ on k. As usual, (A.5.3) implies (A.5.2), so that a semi-ultranorm on V is a seminorm in particular. If N is a semi-ultranorm on V, and $N(v) > 0$ for some $v \in V$, then one can check that $|\cdot|$ is an ultrametric absolute value function on k.

If N is a seminorm on V with respect to $|\cdot|$ on k, and if

$$N(v) > 0 \text{ for every } v \in V \text{ with } v \neq 0, \tag{A.5.4}$$

then we say that N is a *norm* on V with respect to $|\cdot|$ on k. This is the same as in Section 1.1 when $k = \mathbf{R}$ or $\mathbf{C}$ with the standard absolute value function, as before. If N is a semi-ultranorm on V with respect to $|\cdot|$ on k that satisfies (A.5.3), then we say that N is an *ultranorm* on V with respect to $|\cdot|$ on k. Note that $|\cdot|$ is a norm on k, as a one-dimensional vector space over itself, and an ultranorm on k when $|\cdot|$ is ultrametric absolute value function on k.

If N is a norm on V, then it is easy to see that

$$d_N(v, w) = N(v - w) \tag{A.5.5}$$

is a metric on V. This uses (A.2.7) to get that (A.5.5) is symmetric in v and w. If N is an ultranorm on V, then (A.5.5) is an ultrametric on V.

Suppose for the moment that $|\cdot|$ is the trivial absolute value function on k. The *trivial ultranorm* on V is defined by taking $N(v) = 1$ when $v \in V$ and $v \neq 0$, and $N(0) = 0$. It is easy to see that this defines an ultranorm on V, for which the corresponding metric is the discrete metric.

Let a be a positive real number, and suppose that $|\cdot|^a$ is an absolute value function on k too. Suppose that N is a nonnegative real-valued function on V that satisfies (A.5.1), so that

$$N(t\, v)^a = |t|^a\, N(v)^a \tag{A.5.6}$$

for every $t \in k$ and $v \in V$. If we also have that

$$N(v + w)^a \le N(v)^a + N(w)^a, \tag{A.5.7}$$

then N^a is a seminorm on V with respect to $|\cdot|^a$ on k. If N is a seminorm on V with respect to $|\cdot|$ on k and $a \le 1$, then (A.5.7) follows from (A.5.2), because of (A.1.7). If N satisfies (A.5.4) and (A.5.7), then N^a is a norm on V with respect to $|\cdot|^a$ on k.

Similarly, if $|\cdot|$ is an ultrametric absolute value function on k, and N is a semi-ultranorm on V with respect to $|\cdot|$ on k, then N^a is a semi-ultranorm on V with respect to $|\cdot|^a$ on k for every $a > 0$. In this case, if N is an ultranorm on V with respect to $|\cdot|$ on k, then N^a is an ultranorm on V with respect to $|\cdot|^a$ on k for every $a > 0$.

A.6 Some norms on k^n

Let k be a field with an absolute value function $|\cdot|$, and let n be a positive integer. Of course, the space k^n of n-tuples of elements of k is a vector space over k, with respect to coordinatewise addition and scalar multiplication. If $v \in k^n$, then put

$$\|v\|_\infty = \max_{1 \le j \le n} |v_j|, \tag{A.6.1}$$

as in Section 1.3. One can check that this defines a norm on k^n with respect to $|\cdot|$ on k, as before. If $|\cdot|$ is an ultrametric absolute value function on k, then one can verify that (A.6.1) defines an ultranorm on k^n with respect to $|\cdot|$ on k.

Let r be a positive real number, and put

$$\|v\|_r = \Big(\sum_{j=1}^n |v_j|^r\Big)^{1/r} \tag{A.6.2}$$

for every $v \in k^n$. It is easy to see that this defines a norm on k^n with respect to $|\cdot|$ on k when $r = 1$. If $1 < r < \infty$, then one can check that this defines a norm on k^n with respect to $|\cdot|$ on k, using Minkowski's inequality for finite sums, as in Section 1.3. We also have that

$$\|v\|_\infty \le \|v\|_r \le n^{1/r}\, \|v\|_r \tag{A.6.3}$$

for every $v \in k^n$ and $0 < r < \infty$, as before.

Let $e_1, \ldots, e_n$ be the standard basis vectors for k^n, so that the lth coordinate of e_j is equal to 1 when $j = l$, and to 0 otherwise. If N is a seminorm on k^n with respect to $|\cdot|$ on k, then it is easy to see that

$$N(v) \le \sum_{j=1}^n N(e_j)\, |v_j| \tag{A.6.4}$$

for every $v \in k^n$, as in Section 1.11. This implies that

$$N(v) \le \Big(\max_{1 \le j \le n} N(e_j)\Big)\, \|v\|_1 \tag{A.6.5}$$

for every $v \in k^n$. Similarly,

$$N(v) \le \Big(\sum_{j=1}^n N(e_j)\Big)\, \|v\|_\infty \tag{A.6.6}$$

for every $v \in k^n$. If N is a semi-ultranorm on k^n with respect to $|\cdot|$ on k, then we get that

$$N(v) \le \max_{1 \le j \le n} (N(e_j) |v_j|) \le \Big(\max_{1 \le j \le n} N(e_j)\Big)\, \|v\|_\infty \tag{A.6.7}$$

for every $v \in k^n$.

Suppose that k is complete with respect to the metric associated to $|\cdot|$. This implies that k^n is complete with respect to the metric associated to (A.6.1), by standard arguments. Similarly, k^n is complete with respect to the metric associated to (A.6.2) when $r \geq 1$.

Let N be a norm on k^n with respect to $|\cdot|$ on k. Suppose that there is a positive real number c such that

$$c\,\|v\|_\infty \leq N(v) \tag{A.6.8}$$

for every $v \in k^n$. Under these conditions, one can check that k^n is also complete with respect to the metric associated to N.

In fact, if k is complete with respect to the metric associated to $|\cdot|$, and N is any norm on k^n with respect to $|\cdot|$ on k, then it is well known that there is a $c > 0$ such that (A.6.8) holds for every $v \in k^n$. This corresponds to Lemma 2.1 on p116 of [44], and to Proposition 5.2.3 on p138 of [80].

If V is a vector space over k of positive finite dimension n, then V is isomorphis to k^n, as a vector space over k. If N_V is a norm on V with respect to $|\cdot|$ on k, and if k is complete with respect to the metric associatived to $|\cdot|$, then V is complete with respect to the metric associated to N_V, because of the analogous statement for k^n in the preceding paragraphs.

Let W be a vector space over k with a norm N_W with respect to $|\cdot|$ on k. Of course, if V is a linear subspace of W, then the restriction of N_W to V defines a norm on V. If V has finite dimension as a vector space over k, and if k is complete with respect to the metric associated to $|\cdot|$, then V is complete with respect to the metric associated to the restriction of N_W to V. This is the same as the restriction to V of the metric on W associated to N_W. This implies that V is a closed set in W with respect to the metric associated to N_W, as in Section 1.6.

A.7 Sequences and series

Let k be a field with an absolute value function $|\cdot|$, let V be a vector space over k, and let N be a norm on V with respect to $|\cdot|$ on k. If $\{v_j\}_{j=1}^\infty$ and $\{w_j\}_{j=1}^\infty$ are sequences of elements of V that converge to $v, w \in V$, respectively, with respect to the metric d_N associated to N, then it is easy to see that

$$\{v_j + w_j\}_{j=1}^\infty \text{ converges to } v + w \tag{A.7.1}$$

with respect to d_N. Similarly, if $\{t_j\}_{j=1}^\infty$ is a sequence of elements of k that converges to $t \in k$ with respect to the metric associated to $|\cdot|$, then

$$\{t_j\, v_j\}_{j=1}^\infty \text{ converges to } t\, v \tag{A.7.2}$$

in V with respect to N, as in Section 1.5.

An infinite series $\sum_{j=1}^\infty v_j$ with terms in V is said to *converge* in V with respect to N if the sequence of partial sums $\sum_{j=1}^n v_j$ converges in V with respect to d_N, in which case we put

$$\sum_{j=1}^{\infty} v_j = \lim_{n\to\infty} \sum_{j=1}^{n} v_j, \tag{A.7.3}$$

as usual. Under these conditions, if $t \in k$, then $\sum_{j=1}^{\infty} t\, v_j$ converges in V as well, with

$$\sum_{j=1}^{\infty} (t\, v_j) = t \sum_{j=1}^{\infty} v_j. \tag{A.7.4}$$

If $\sum_{j=1}^{\infty} w_j$ is another convergent series of elements of V, then $\sum_{j=1}^{\infty}(v_j + w_j)$ converges, with

$$\sum_{j=1}^{\infty} (v_j + w_j) = \sum_{j=1}^{\infty} v_j + \sum_{j=1}^{\infty} w_j, \tag{A.7.5}$$

as before.

If $\sum_{j=1}^{\infty} v_j$ is any infinite series with terms in V, then the corresponding sequence of partial sums is a Cauchy sequence with respect to d_N if and only if for every $\epsilon > 0$ there is a positive integer $L(\epsilon)$ such that

$$N\Big(\sum_{j=l}^{n} v_j\Big) < \epsilon \tag{A.7.6}$$

for all integers $n \geq l \geq L(\epsilon)$, as in Section 1.5 again. This implies that

$$\lim_{j\to\infty} N(v_j) = 0, \tag{A.7.7}$$

as before. Note that

$$N\Big(\sum_{j=l}^{n} v_j\Big) \leq \sum_{j=l}^{n} N(v_j) \tag{A.7.8}$$

for all positive integers $l \leq n$, as in Section 1.7. If N is an ultranorm on V, then

$$N\Big(\sum_{j=l}^{n} v_j\Big) \leq \max_{l\leq j\leq n} N(v_j) \tag{A.7.9}$$

for all $l \leq n$.

We say that $\sum_{j=1}^{\infty} v_j$ *converges absolutely* with respect to N if

$$\sum_{j=1}^{\infty} N(v_j) \tag{A.7.10}$$

converges as an infinite series of nonnegative real numbers, as in Section 1.7. This implies that the corresponding sequence of partial sums is a Cauchy sequence with respect to d_N, as before.

If V is complete with respect to d_N, then V is said to be a *Banach space* with respect to N, as in Section 1.5. In this case, if $\sum_{j=1}^{\infty} v_j$ converges absolutely with respect to N, then $\sum_{j=1}^{\infty} v_j$ converges in V, with

$$N\Big(\sum_{j=1}^{\infty} v_j\Big) \le \sum_{j=1}^{\infty} N(v_j), \tag{A.7.11}$$

as in Section 1.7. We also have that the completeness of V with respect to d_N is characterized by the condition that absolute convergence of an infinite series in V with respect to N implies that the series converges, as before.

If N is an ultranorm on V, then (A.7.7) implies that the sequence of partial sums $\sum_{j=1}^{n} v_j$ is a Cauchy sequence in V with respect to d_N, because of (A.7.9). If V is a Banach space with respect to N too, then it follows that $\sum_{j=1}^{\infty} v_j$ converges in V. One can check that

$$N\Big(\sum_{j=1}^{\infty} v_j\Big) \le \max_{j\ge 1} N(v_j) \tag{A.7.12}$$

in this case. More precisely, one can verify that the maximum on the right side is attained, by reducing to the maximum of finitely many terms when at least one of the terms is not zero.

If V is not already complete with respect to d_N, then V has a completion that is a Banach space over k, as in Section 1.15. If N is an ultranorm on V, then its extension to the completion of V is an ultranorm on the completion of V.

A.8 Discrete absolute value functions

Let k be a field with an absolute value function $|\cdot|$ again. Observe that

$$\{|t| : t \in k,\, t \ne 0\} \tag{A.8.1}$$

is a subgroup of the multiplicative group $\mathbf{R}_+$ of positive real numbers. This may be called the *valuation group* associatied to $|\cdot|$ on k, as on p42 of [44]. Of course, this is the trivial subgroup $\{1\}$ of $\mathbf{R}_+$ if and only if $|\cdot|$ is the trivial absolute value function on k.

Put

$$\rho_1 = \sup\{|t| : t \in k,\, |t| < 1\}, \tag{A.8.2}$$

so that

$$0 \le \rho_1 \le 1, \tag{A.8.3}$$

by construction. It is easy to see that

$$\rho_1 = 0 \tag{A.8.4}$$

if and only if $|\cdot|$ is the trivial absolute value function on k.

If

(A.8.5) $$\rho_1 = 1,$$

then

(A.8.6) $$1 \text{ is a limit point of (A.8.1)},$$

with respect to the standard Eulidean metric on $\mathbf{R}$. Conversely, one can check that (A.8.6) implies (A.8.5), because the valuation group (A.8.2) contains the multiplicative inverses of its elements. In this case, one can verify that (A.8.1) is dense in $\mathbf{R}_+$, with respect to the standard Euclidean metric.

If

(A.8.7) $$\rho_1 < 1,$$

then $|\cdot|$ is said to be *discrete* on k, as on p42 of [44]. If $|\cdot|$ is nontrivial on k too, so that $\rho_1 > 0$, then it is not too difficult to show that (A.8.1) consists exactly of the integer powers of ρ_1. In particular, this means that the supremum on the right side of (A.8.2) is attained. This corresponds to Lemma 1.2 on p42 of [44], and to part of Problem 63 on p39 of [80]. More precisely, in these statements in [44, 80], $|\cdot|$ is asked to be an ultrametric absolute value function on k, and this is used to consider other versions of these properties. However, this is not really needed for the versions of these properties mentioned here. If $|\cdot|$ is discrete on k, then one can show that $|\cdot|$ is an ultrametric absolute value function on k, as in the next paragraph.

If $|\cdot|$ is archimedean on k, then it is easy to see that k has characteristic 0, so that k contains a copy of $\mathbf{Q}$. This implies that $|\cdot|$ on k induces an archimedean absolute value function on $\mathbf{Q}$. This induced absolute value function on $\mathbf{Q}$ is equivalent to the standard Euclidean absolute value function on $\mathbf{Q}$, by Ostrowski's theorem, as in Section A.4. One can use this to get that $|\cdot|$ is not discrete on k, because the standard Euclidean absolute value function on $\mathbf{Q}$ is not discrete. If $|\cdot|$ is a discrete absolute value function on k, then it follows that $|\cdot|$ is non-archimedean on k, so that $|\cdot|$ is an ultrametric absolute value function on k, as in Section A.3.

A.9 Supremum seminorms and c_0 spaces

Let X be a nonempty set, let k be a field, and let W be a vector space over k. The space of all functions on X with values in W is a vector space over k with respect to pointwise addition and scalar multiplication, as in Section 2.1.

Let $|\cdot|$ be an absolute value function on k, and let N_W be a seminorm on W with respect to $|\cdot|$ on k. Let us say that a W-valued function f on X is *bounded* with respect to N_W on W if

(A.9.1) $$N_W(f(x)) \text{ is bounded as a real-valued function on } X.$$

Let

(A.9.2) $$\ell^\infty(X, W) = \ell^\infty_{N_W}(X, W)$$

be the space of all W-valued functions on X that are bounded with respect to N_W. This is a linear subspace of the space of all W-valued functions on X, as in Section 2.3.

If $f \in \ell^\infty_{N_W}(X, W)$, then put

$$\|f\|_\infty = \|f\|_{sup} = \|f\|_{\ell^\infty(X,W)} = \|f\|_{\ell^\infty_{N_W}(X,W)} = \sup_{x \in X} N_W(f(x)), \tag{A.9.3}$$

as before. One can check that this defines a seminorm on $\ell^\infty_{N_W}(X, W)$, which is the *supremum seminorm* associated to N_W. If N_W is a semi-ultranorm on W, then (A.9.3) is a semi-ultranorm on $\ell^\infty_{N_W}(X, W)$.

If N_W is a norm on W, then (A.9.3) is a norm on $\ell^\infty_{N_W}(X, W)$, which is the *supremum norm* associated to N_W on W. The metric on $\ell^\infty_{N_W}(X, W)$ corresponding to the supremum norm is the same as the *supremum metric* corresponding to the metric on W associated to N_W, as in Section 2.3. If N_W is an ultranorm on W, then (A.9.3) is an ultranorm on $\ell^\infty_{N_W}(X, W)$.

If N_W is a norm on W, and W is a Banach space with respect to N_W, then

$$\ell^\infty_{N_W}(X, W) \text{ is a Banach space} \tag{A.9.4}$$

with respect to (A.9.3). This is analogous to the statements in Sections 1.6 and 2.3 for vector spaces over the real and complex numbers. In particular, if k is complete with respect to the metric associated to $|\cdot|$, then one can take $W = k$.

If f is a function on X with values in W, then the *support* of f may be defined as the set $\operatorname{supp} f$ of $x \in X$ such that $f(x) \neq 0$, as in Sections 1.12 and 2.3. Let

$$c_{00}(X, W) \tag{A.9.5}$$

be the space of W-valued functions f on X such that $\operatorname{supp} f$ has only finitely many elements, as before. This is a linear subspace of the space of all W-valued functions on X, as a vector space over k.

Let N_W be a seminorm on W with respect to $|\cdot|$ on k again. A W-valued function f on X is said to *vanish at infinity* on X with respect to N_W if

$$N_W(f(x)) \text{ vanishes at infinity on } X, \tag{A.9.6}$$

as a nonnegative real-valued function on X, as in Section 2.6. Let

$$c_0(X, W) = c_{0,N_W}(X, W) \tag{A.9.7}$$

be the space of W-valued functions on X that vanish at infinity with respect to N_W. It is easy to see that this is a linear subspace of $\ell^\infty_{N_W}(X, W)$, as before. If N_W is a norm on W, then $c_{0,N_W}(X, W)$ is the same as the closure of $c_{00}(X, W)$ in $\ell^\infty_{N_W}(X, W)$ with respect to the supremum metric, as in Sections 1.13 and 2.6.

A.10 Lipschitz and operator seminorms

Let (X, d_X) be a nonempty metric space, let k be a field with an absolute value function $|\cdot|$, and let W be a vector space over k with a norm N_W with respect to $|\cdot|$ on k. Thus W may also be considered as a metric space with respect to the corresponding metric d_{N_W}, so that the space $\mathrm{Lip}(X, W)$ of all Lipschitz mappings from X into W may be defined as in Section 2.1. One can check that

(A.10.1) $\mathrm{Lip}(X, W)$ is a linear subspace of the space of all W-valued functions on X,

as a vector space over k with respect to pointwise addition and scalar multiplication on X, as before.

If $f \in \mathrm{Lip}(X, W)$, then the corresponding Lipschitz constant $\mathrm{Lip}(f) = \mathrm{Lip}_{X,W}(f)$ may be defined as in Section 2.1. One can verify that

(A.10.2) $\mathrm{Lip}(f)$ is a seminorm on $\mathrm{Lip}(X, W)$

with respect to $|\cdot|$, on k, as before. Remember that $\mathrm{Lip}(f) = 0$ if and only if f is constant on X. Similarly, if N_W is an ultranorm on W, then one can check that

(A.10.3) $\mathrm{Lip}(f)$ is a semi-ultranorm on $\mathrm{Lip}(X, W)$.

Let V be another vector space over k, and let us denote the space of all linear mappings from V into W as $\mathcal{L}(V, W)$, as in Section 2.2. This is a linear subspace of the space of all W-valued functions on V, as a vector space over k.

Let N_V, N_W be seminorms on V, W, respectively, with respect to $|\cdot|$ on k. A linear mapping T from V into W is said to be *bounded* with respect to these seminorms if there is a nonnegative real number C such that

$$N_W(T(v)) \le C\, N_V(v) \tag{A.10.4}$$

for every $v \in V$, as in Section 2.2. This implies that

$$N_W(T(u) - T(v)) = N_W(T(u - v)) \le C\, N_V(u - v) \tag{A.10.5}$$

for every $u, v \in V$, as before. If N_V, N_W are norms on V, W, respectively, then (A.10.5) says that T is Lipschitz with constant C with respect to the associated metrics d_{N_V}, d_{N_W}. Otherwise, d_{N_V} and d_{N_W} may be considered as semimetrics on V and W, respectively, as in Section 1.2.

Let $\mathcal{BL}(V, W)$ be the space of all bounded linear mappings from V into W with respect to N_V, N_W, respectively. One can check that this is a linear subspace of $\mathcal{L}(V, W)$, as in Section 2.2.

If $T \in \mathcal{BL}(V, W)$, then put

$$\|T\|_{op} = \|T\|_{op,VW} = \inf\{C \ge 0 : \text{(A.10.4) holds}\}. \tag{A.10.6}$$

One can verify that

(A.10.7) $\|\cdot\|_{op}$ defines a seminorm on $\mathcal{BL}(V, W)$

with respect to $|\cdot|$ on k, which is the *operator seminorm* associated to N_V, N_W. It is easy to see that the infimum on the right side of (A.10.6) is automatically attained, so that

$$N_W(T(v)) \le \|T\|_{op,VW}\, N_V(v) \tag{A.10.8}$$

for every $v \in V$. If N_W is a norm on W, then

$$\|\cdot\|_{op} \text{ is a norm on } \mathcal{BL}(V,W), \tag{A.10.9}$$

which is the *operator norm* associated to N_V, N_W.

Similarly, if N_W is a semi-ultranorm on W, then one can check that

$$\|\cdot\|_{op} \text{ is a semi-ultranorm on } \mathcal{BL}(V,W). \tag{A.10.10}$$

If N_W is an ultranorm on W, then

$$\|\cdot\|_{op} \text{ is an ultranorm on } \mathcal{BL}(V,W). \tag{A.10.11}$$

Let Z be another vector space over k, with a seminorm N_Z with respect to $|\cdot|$ on k. If T_1 is a bounded linear mapping from V into W, and T_2 is a bounded linear mapping from W into Z, then $T_2 \circ T_1$ is a bounded linear mapping from V into Z, with

$$\|T_2 \circ T_1\|_{op,VZ} \le \|T_1\|_{op,VW}\, \|T_2\|_{op,WZ}, \tag{A.10.12}$$

as in Section 2.2.

If N_W is a norm on W, and W is complete with respect to the metric associated to N_W, then

$$\mathcal{BL}(V,W) \text{ is complete} \tag{A.10.13}$$

with respect to the metric associated to the operator norm. This follows from the same type of argument as in Section 2.2.

Let us suppose from now on in this section that N_V and N_W are norms on V and W, respectively. Note that (A.10.6) is the same as the Lipschitz constant $\mathrm{Lip}(T)$ of T with respect to the metrics associated to N_V, N_W, as in Section 2.2.

Let T be a linear mapping from V into W, and suppose for the moment that $N_W(T(v))$ is bounded on a ball of positive radius in V centered at 0 with respect to the metric associated to N_V. Of course, this condition holds in particular when T is continuous at 0 with respect to the metrics associated to N_V, N_W. If $|\cdot|$ is not the trivial absolute value function on k, then one can check that

$$T \text{ is a bounded linear mapping from } V \text{ into } W. \tag{A.10.14}$$

Let V_0 be a linear subspace of V that is dense in V with respect to the metric associated to N_V, and let T_0 be a bounded linear mapping from V_0 into W, with respect to the restriction of N_V to V_0. If W is complete with respect to the metric associated to N_W, then

$$\begin{aligned}&\text{there is a unique extension of } T_0 \text{ to a}\\ &\text{bounded linear mapping } T \text{ from } V \text{ into } W,\end{aligned} \tag{A.10.15}$$

as in Section 2.2. One can check that the operator norm of T on V is equal to the operator norm of T_0 on V_0, as before.

A.11 ℓ^r Spaces

Let X be a nonempty set, let r be a positive real number, and let k be a field with an absolute value function $|\cdot|$. Also let W be a vector space over k with a seminorm N_W with respect to $|\cdot|$ on k. Consider the space

$$\ell^r(X, W) = \ell^r_{N_W}(X, W) \tag{A.11.1}$$

of W-valued functions f on X such that

$$N_W(f(x)) \text{ is } r\text{-summable as a nonnegative real-valued function on } X, \tag{A.11.2}$$

as in Section 2.5. This was discussed in Section 2.6 for vector spaces over the real or complex numbers with a norm. As before, one can take $W = k$, with $|\cdot|$ as the norm.

One can check that $\ell^r_{N_W}(X, W)$ is a linear subspace of the space of all W-valued functions on X, as in Section 2.6. If $f \in \ell^r_{N_W}(X, W)$, then put

$$\|f\|_r = \|f\|_{\ell^r(X,W)} = \|f\|_{\ell^r_{N_W}(X,W)} = \Big(\sum_{x \in X} N_W(f(x))^r\Big)^{1/r}, \tag{A.11.3}$$

as before. If $r \geq 1$, then this defines a seminorm on $\ell^r_{N_w}(X, W)$ with respect to $|\cdot|$, and a norm when N_W is a norm on W. If $0 < r \leq 1$, then this satisfies the usual homogeneity property of a seminorm, and

$$\|f + g\|_r^r \leq \|f\|_r^r + \|g\|_r^r \tag{A.11.4}$$

for all $f, g \in \ell^r_{N_W}(X, W)$. If N_W is a norm on W, then we get that

$$\|f - g\|_r^r \tag{A.11.5}$$

is a metric on $\ell^r_{N_W}(X, W)$ when $r \leq 1$, as usual.

If $0 < r_1 \leq r_2 \leq +\infty$, then

$$\ell^{r_1}_{N_W}(X, W) \subseteq \ell^{r_2}_{N_W}(X, W), \tag{A.11.6}$$

as in Section 2.6. If $f \in \ell^{r_1}_{N_W}(X, W)$, then

$$\|f\|_{r_2} \leq \|f\|_{r_1}, \tag{A.11.7}$$

as before. If $0 < r < \infty$, then

$$c_{00}(X, W) \subseteq \ell^r_{N_W}(X, W) \subseteq c_{0,N_W}(X, W), \tag{A.11.8}$$

as before, where $c_{0,N_W}(X, W)$ is as in Section A.9. If N_W is a norm on W, then one can check that

$$c_{00}(X, W) \text{ is dense in } \ell^r_{N_W}(X, W) \tag{A.11.9}$$

when $r < \infty$, as in Section 2.6. This uses the metric associated to (A.11.3) when $r \geq 1$, and the metric (A.11.5) when $r \leq 1$.

If N_W is a norm on W and W is complete with respect to the associated metric, then

(A.11.10) $$\ell^r_{N_W}(X, W) \text{ is complete}$$

with respect to the metric associated to (A.11.3) when $r \geq 1$, and with respect to the metric (A.11.5) when $r \leq 1$, as in Section 2.6.

A.12 Some bounded linear mappings

Let X be a nonempty set, let k be a field, and let W be a vector space over k. If $f \in c_{00}(X, W)$, then

(A.12.1) $$\sum_{x \in X} f(x)$$

may be defined as an element of W as in Section 2.3. This defines a linear mapping from $c_{00}(X, W)$ into W, as before.

If a is any W-valued function on X and $f \in c_{00}(X, k)$, then

(A.12.2) $$a\, f \in c_{00}(X, W).$$

This means that

(A.12.3) $$T_a(f) = \sum_{x \in X} a(x)\, f(x)$$

defines an element of W, and that T_a defines a linear mapping from $c_{00}(X, k)$ into W, as in Section 2.3.

If $y \in X$, then let δ_y be the k-valued function on X equal to 1 at y at to 0 at every other point in X, as in Section 1.12. It is easy to see that

(A.12.4) the collection of δ_y's, $y \in X$, is a basis for $c_{00}(X, k)$,

as a vector space over k, as before. If $0 < r \leq \infty$, then

(A.12.5) $$\|\delta_y\|_{\ell^r(X,k)} = 1$$

for every $y \in X$, using $|\cdot|$ as the norm on k on the left side.

Of course,

(A.12.6) $$T_a(\delta_y) = a(y)$$

for every $y \in X$, as in Section 2.3. If T is any linear mapping from $c_{00}(X, k)$ into W, then

(A.12.7) $$a(y) = a_T(y) = T(\delta_y)$$

defines a W-valued function on X. It is easy to see that this is the unique W-valued function on X such that

(A.12.8) $$T = T_a,$$

as before.

Let N_W be a seminorm on W with respect to $|\cdot|$ on k, and suppose that $1 \le r \le \infty$, so that the restriction of the ℓ^r norm to $c_{00}(X,k)$ defines a norm on $c_{00}(X,k)$ with respect to $|\cdot|$ on k. Also let a be a W-valued function on X again, and suppose that

(A.12.9) T_a is bounded as a linear mapping from $c_{00}(X,k)$ into W, with respect to the ℓ^r norm on $c_{00}(X,k)$ and N_W on W.

Under these conditions, we get that

$$N(a(y)) = N_W(T_a(\delta_y)) \le \|T_a\|_{op,r}\,\|\delta_y\|_{\ell^r(X,k)} = \|T_a\|_{op,r} \tag{A.12.10}$$

for every $y \in X$, where $\|T_a\|_{op,r}$ is the operator seminorm of T_a corresponding to the restriction of the ℓ^r norm to $c_{00}(X,k)$ and N_W on W. This means that a is bounded on X with respect to N_W on W, with

$$\|a\|_{\ell^\infty_{N_W}(X,W)} \le \|T_a\|_{op,r}. \tag{A.12.11}$$

If a is any element of $\ell^\infty_{N_W}(X,W)$ and $f \in c_{00}(X,k)$, then

$$N_W(T_a(f)) \le \sum_{x\in X} N_W(a(x))\,|f(x)| \le \|a\|_{\ell^\infty_{N_w}(X,W)}\,\|f\|_{\ell^1(X,k)}, \tag{A.12.12}$$

as in Section 2.3. This implies that

(A.12.13) T_a is a bounded linear mapping from $c_{00}(X,k)$ into W, with respect to the ℓ^1 norm on $c_{00}(X,k)$ and N_W on W,

with

$$\|T_a\|_{op,1} \le \|a\|_{\ell^\infty_{N_W}(X,W)}. \tag{A.12.14}$$

Thus

$$\|T_a\|_{op,1} = \|a\|_{\ell^\infty_{N_W}(X,W)}, \tag{A.12.15}$$

because of (A.12.11). If N_W is a norm on W, and W is complete with respect to the associated metric, then

(A.12.16) T_a has a unique extension to a bounded linear mapping from $\ell^1(X,k)$ into W, with the same operator norm,

as in Section A.10.

If N_W is a semi-ultranorm on W with respect to $|\cdot|$ on k, then

$$N_W(T_a(f)) \le \max_{x\in X}(N_W(a(x))\,|f(x)|) \le \|a\|_{\ell^\infty_{N_W}(X,W)}\,\|f\|_{\ell^\infty(X,k)} \tag{A.12.17}$$

for every $f \in c_{00}(X,k)$. In this case,

(A.12.18) T_a is a bounded linear mapping from $c_{00}(X,k)$ into W, with respect to the ℓ^∞ norm on $c_{00}(X,k)$ and N_W on W, with

$$\|T_a\|_{op,\infty} \le \|a\|_{\ell^\infty_{N_W}(X,W)}. \tag{A.12.19}$$

This means that

$$\|T_a\|_{op,\infty} = \|a\|_{\ell^\infty_{N_W}(X,W)}, \tag{A.12.20}$$

by (A.12.11). If N_W is an ultranorm on W, and W is complete with respect to the associated ultrametric, then

$$\begin{aligned} &T_a \text{ has a unique extension to a bounded linear mapping} \\ &\text{from } c_0(X,k) \text{ into } W, \text{ with the same operator norm,} \end{aligned} \tag{A.12.21}$$

as in Section A.10 again.

A.13 Sums of vectors

Let X be a nonempty set, and let k be a field with an absolute value function $|\cdot|$. Also let W be a vector space over k with a norm N_W with respect to $|\cdot|$ on k, and let f be a function on X with values in W. As in Section 2.7, we say that

$$\sum_{x \in X} f(x) \textit{ converges in the generalized sense} \tag{A.13.1}$$

with respect to the metric associated to N_W if there is a $w \in W$ with the following property: for every $\epsilon > 0$ there is a finite set $A(\epsilon) \subseteq X$ such that for every finite set $A \subseteq X$ with

$$A(\epsilon) \subseteq A, \tag{A.13.2}$$

we have that

$$N_W\Big(\sum_{x \in A} f(x) - w\Big) < \epsilon. \tag{A.13.3}$$

One can check that such a $w \in W$ is unique when it exists, as before, in which case it is considered to be the value of the sum. This is equivalent to the convergence of the corresponding net of sums over finite subsets of X, as before.

Let us say that

$$\sum_{x \in X} f(x) \text{ satisfies } \textit{the generalized Cauchy condition} \tag{A.13.4}$$

with respect to N_W if for every $\epsilon > 0$ there is a finite subset $A_0(\epsilon)$ of X with the following property, as before: if A, B are finite subsets of X such that

$$A_0(\epsilon) \subseteq A, B, \tag{A.13.5}$$

then

$$N_W\Big(\sum_{x \in A} f(x) - \sum_{x \in B} f(x)\Big) < \epsilon. \tag{A.13.6}$$

This is the same as saying that the corresponding net of sums over finite subsets of X is a Cauchy net with respect to the metric associated to N_W, as before. If the sum converges in the generalized sense, then it satisfies the generalized Cauchy condition with

$$A_0(\epsilon) = A(\epsilon/2) \tag{A.13.7}$$

for every $\epsilon > 0$, as before. If N_W is an ultranorm on W, then one can take

$$A_0(\epsilon) = A(\epsilon) \tag{A.13.8}$$

for each $\epsilon > 0$.

As in Section 2.7, the generalized Cauchy condition is the same as saying that for every $\epsilon > 0$ there is a finite subset $A_1(\epsilon)$ of X with the following property: if C is a finite subset of X such that

$$A_1(\epsilon) \cap C = \emptyset, \tag{A.13.9}$$

then

$$N_W\Big(\sum_{x \in C} f(x)\Big) < \epsilon. \tag{A.13.10}$$

The previous version implies this one, with

$$A_1(\epsilon) = A_0(\epsilon) \tag{A.13.11}$$

for each $\epsilon > 0$, as before. This version also implies the previous one, with

$$A_0(\epsilon) = A_1(\epsilon/2) \tag{A.13.12}$$

for every $\epsilon > 0$, as before. If N_W is an ultranorm on W, then one can take

$$A_0(\epsilon) = A_1(\epsilon) \tag{A.13.13}$$

for every $\epsilon > 0$.

If C is any nonempty finite subset of X, then

$$N_W\Big(\sum_{x \in C} f(x)\Big) \le \sum_{x \in C} N_W(f(x)), \tag{A.13.14}$$

as in Section 2.7. If $f \in \ell^1_{N_W}(X, W)$, then one can use this to get that the sum satisfies the generalized Cauchy condition, as before. If the sum converges in the generalized sense, then one can verify that

$$N_W\Big(\sum_{x \in X} f(x)\Big) \le \sum_{x \in X} N_W(f(x)), \tag{A.13.15}$$

as before.

The generalized Cauchy condition implies that

$$f \text{ vanishes at infinity on } X \text{ with respect to } N_W, \tag{A.13.16}$$

as before. If N_W is an ultranorm on W, then

$$N_W\Big(\sum_{x \in C} f(x)\Big) \le \max_{x \in C} N_W(f(x)) \tag{A.13.17}$$

for every nonempty finite subset C of X. In this case, if f vanishes at infinity on X with respect to N_W, then it is easy to see that the sum satisfies the generalized Cauchy condition, using the remarks in the preceding paragraph. If the sum converges in the generalized sense, then one can check that

$$N_W\Big(\sum_{x \in X} f(x)\Big) \le \max_{x \in X} N_W(f(x)). \tag{A.13.18}$$

Note that the maximum on the right is attained on X, because $N_W(f(x))$ vanishes at infinity on X, by hypothesis.

Let $\{x_j\}_{j=1}^{\infty}$ be a sequence of distinct elements of X such that

$$\operatorname{supp} f \subseteq \{x_j : j \in \mathbf{Z}_+\}. \tag{A.13.19}$$

If $\sum_{x \in X} f(x)$ converges in the generalized sense, then it is easy to see that

$$\sum_{j=1}^{\infty} f(x_j) = \sum_{x \in X} f(x), \tag{A.13.20}$$

including the convergese of the series on the left, as in Section 2.7. If $\sum_{x \in X} f(x)$ satisfies the generalized Cauchy condition, and if the sum on the left side of (A.13.20) converges, then one can check that $\sum_{x \in X} f(x)$ converges in the generalized sense, with sum as in (A.13.20), as before.

If $\sum_{x \in X} f(x)$ satisfies the generalized Cauchy condition, then

$$\Big\{\sum_{j=1}^{n} f(x_j)\Big\}_{n=1}^{\infty} \text{ is a Cauchy sequence in } W \tag{A.13.21}$$

with respect to the metric associated to N_W, as in Section 2.7. If W is a Banach space with respect to N_W, then it follows that the series on the left side of (A.13.20) converges in W. This implies that $\sum_{x \in W} f(x)$ converges in the generalized sense, with sum as in (A.13.20), as in the preceding paragraph.

A.14 Some related spaces of functions

Let us continue with the same notation and hypotheses as at the beginning of the previous section. Let

$$Sum_{N_W}(X, W) \tag{A.14.1}$$

be the space of W-valued functions f on X such that $\sum_{x \in X} f(x)$ converges in the generalized sense, as in Section 2.8. This is a linear subspace of the space of all W-valued functions on X, and the sum defines a linear mapping from $Sum_{N_W}(X, W)$ into W, as before.

Let

$$GCC_{N_W}(X, W) \tag{A.14.2}$$

be the space of all W-valued functions f on X such that $\sum_{x \in X} f(x)$ satisfies the generalized Cauchy condition, as in Section 2.8 again. This is a linear subspace of the space of all W-valued functions on X too, with

$$c_{00}(X, W) \subseteq Sum_{N_W}(X, W) \subseteq GCC_{N_W}(X, W) \subseteq c_{0,N_W}(X, W), \tag{A.14.3}$$

as before. If W is a Banach space, then

$$Sum_{N_W}(X, W) = GCC_{N_W}(X, W), \tag{A.14.4}$$

as in the previous section.

We have seen that

$$\ell^1_{N_W}(X, W) \subseteq GCC_{N_W}(X, W) \tag{A.14.5}$$

so that

$$\ell^1_{N_W}(X, W) \subseteq Sum_{N_W}(X, W) \tag{A.14.6}$$

when W is a Banach space. If N_W is an ultranorm on W, then

$$GCC_{N_W}(X, W) = c_{0,N_W}(X, W), \tag{A.14.7}$$

as in the previous section. If W is also a Banach space, then it follows that

$$Sum_{N_W}(X, W) = c_{0,N_W}(X, W) \tag{A.14.8}$$

in this case.

As in Section 2.9, we say that a W-valued function f on X has *bounded finite sums* on X with respect to N_W if the norms

$$N_W\Big(\sum_{x \in A} f(x)\Big) \tag{A.14.9}$$

of the sums of f over all finite subsets A of X are bounded. The space

$$BFS_{N_W}(X, W) \tag{A.14.10}$$

of all W-valued functions on X with bounded finite sums is a linear subspace of the space of all W-valued functions on X, as before. One can check that

$$GCC_{N_W}(X, W) \subseteq BFS_{N_W}(X, W) \subseteq \ell^\infty_{N_W}(X, W), \tag{A.14.11}$$

as before.

If $f \in BFS_{N_W}(X, W)$, then we put

$$\begin{aligned} \|f\|_{BFS} &= \|f\|_{BFS_{N_W}(X,W)} \\ &= \sup\Big\{N_W\Big(\sum_{x \in A} f(x)\Big) : A \text{ is a finite subset of } X\Big\}. \end{aligned} \tag{A.14.12}$$

This defines a norm on $BFS_{N_W}(X, W)$ with respect to $|\cdot|$ on k, as before. Clearly

$$\|f\|_{\ell^\infty_{N_W}(X,W)} \le \|f\|_{BFS_{N_W}(X,W)} \tag{A.14.13}$$

for every $f \in BFS_{N_W}(X, W)$. If $f \in \ell^1_{N_W}(X, W)$, then $f \in BFS_{N_W}(X, W)$, with

$$\|f\|_{BFS_{N_W}(X,W)} \le \|f\|_{\ell^1_{N_W}(X,W)}, \tag{A.14.14}$$

as before. If $f \in Sum_{N_W}(X, W)$, then $f \in BFS_{N_W}(X, W)$, by (A.14.3) and (A.14.11), and

$$N_W\Big(\sum_{x \in X} f(x)\Big) \le \|f\|_{BFS_{N_W}(X,W)}, \tag{A.14.15}$$

as before.

Suppose for the moment that N_W is an ultranorm on W. If $f \in \ell^\infty_{N_W}(X, W)$, then it is easy to see that $f \in BFS_{N_W}(X, W)$, with

$$\|f\|_{BFS_{N_W}(X,W)} \le \|f\|_{\ell^\infty_{N_W}(X,W)}. \tag{A.14.16}$$

This implies that

$$BFS_{N_W}(X, W) = \ell^\infty_{N_W}(X, W), \tag{A.14.17}$$

with

$$\|f\|_{BFS_{N_W}(X,W)} = \|f\|_{\ell^\infty_{N_W}(X,W)} \tag{A.14.18}$$

for every $f \in \ell^\infty_{N_W}(X, W)$.

If N_W is any norm on W, then

$$GCC_{N_W}(X, W) \text{ is the same as the closure of } c_{00}(X, W) \text{ in } BFS_{N_W}(X, W), \tag{A.14.19}$$

as in Section 2.9. If W is a Banach space, then

$$BFS_{N_W}(X, W) \text{ is complete} \tag{A.14.20}$$

with respect to the metric associated to (A.14.12), as before.

Appendix B

More on metrics and norms

B.1 q-Metrics and q-semimetrics

Let X be a set, and let $d(x, y)$ be a nonnegative real-valued function defined for $x, y \in X$ such that

$$d(x, x) = 0 \tag{B.1.1}$$

for every $x \in X$, and

$$d(x, y) = d(y, x) \tag{B.1.2}$$

for every $x, y \in X$. Also let q be a positive real number, and suppose that

$$d(x, z)^q \leq d(x, y)^q + d(y, z)^q \tag{B.1.3}$$

for every $x, y, z \in X$. Under these conditions, we say that $d(\cdot, \cdot)$ is a *q-semimetric* or *q-pseudometric* on X. If we also have that

$$d(x, y) > 0 \tag{B.1.4}$$

when $x \neq y$, then we say that $d(\cdot, \cdot)$ is a *q-metric* on X. Thus q-semimetrics and q-metrics are the same as ordinary semimetrics and metrics when $q = 1$, as in Sections 1.2 and A.1.

Of course, (B.1.3) is the same as saying that

$$d(x, z) \leq (d(x, y)^q + d(y, z)^q)^{1/q}. \tag{B.1.5}$$

The right side is monotonically decreasing in q, as in Section 1.12. This implies that the property of being a q-semimetric or q-metric is more restrictive as q increases.

If $d(\cdot, \cdot)$ satisfies (B.1.1), (B.1.2), and

$$d(x, z) \leq \max(d(x, y), d(y, z)) \tag{B.1.6}$$

for every $x, y, z \in X$, then we say that $d(\cdot, \cdot)$ is a *semi-ultrametric* or *pseudo-ultrametric* on X. If $d(\cdot, \cdot)$ satisfies (B.1.4) as well, then $d(\cdot, \cdot)$ is an ultrametric on X, as in Section A.1. It is easy to see that a semi-ultrametric on X is a q-semimetric on X for every $q > 0$, and similarly an ultrametric on X is a q-metric for every $q > 0$.

Note that

$$\lim_{q \to \infty} (d(x,y)^q + d(y,z)^q)^{1/q} = \max(d(x,y), d(y,z)) \tag{B.1.7}$$

for every $x, y, z \in X$, as in Section 1.3. Because of this, we shall consider semi-ultrametrics and ultrametrics as being q-semimetrics and q-metrics, respectively, with $q = \infty$.

Of course, if $d(\cdot,\cdot)$ is a q-semimetric on X and Y is a subset of X, then the restriction of $d(x,y)$ to $x, y \in Y$ is a q-semimetric on Y. Similarly, if $d(\cdot,\cdot)$ is a q-metric on X, then the restriction of $d(x,y)$ to $x, y \in Y$ is a q-metric on Y.

B.2 More on q-metrics and q-semimetrics

Let X be a set, and let a be a positive real number. If $d(\cdot,\cdot)$ is a q-semimetric on X for some $q > 0$, then it is easy to see that

$$d(x,y)^a \tag{B.2.1}$$

is a (q/a)-semimetric on X. Similarly, if $d(\cdot,\cdot)$ is a q-metric on X, then (B.2.1) is a (q/a) metric on X. If $q = \infty$, then this means that if $d(\cdot,\cdot)$ is a semi-ultrametric or ultrametric on X, then (B.2.1) has the same property, as in Section A.1.

If $d(\cdot,\cdot)$ is a q-semimetric or q-metric on X and $a \le 1$, then it follows that (B.2.1) has the same property, as in Section A.1. This uses the fact that the property of being a q-semimetric of q-metric is more restrictive as q increases, as in the previous section.

If $d(\cdot,\cdot)$ is a q-semimetric on X, then one can define open and closed balls in X with respect to $d(\cdot,\cdot)$ in the usual way, as in Section A.1. The open and closed balls in X with radius r with respect to $d(\cdot,\cdot)$ are the same as the open and closed balls in X with radius r^a with respect to (B.2.1), respectively, and with the same center, as before.

One can use open balls with respect to $d(\cdot,\cdot)$ to define what it means for a subset of X to be an open set with respect to $d(\cdot,\cdot)$ in the usual way. The open subsets of X with respect to $d(\cdot,\cdot)$ are the same as the open sets with respect to (B.2.1), because of the remark about the corresponding open balls in the preceding paragraph.

It is easy to see that this defines a topology on X, as usual. If $q \ge 1$, then $d(\cdot,\cdot)$ is a semimetric on X, and otherwise

$$d(x,y)^q \tag{B.2.2}$$

is a semimetric on X that determines the same topology as $d(\cdot,\cdot)$. The topology determined by a semimetric has many of the same properties as the topology determined by a metric. In particular, one can check that open balls with respect to a semimetric are open sets. This implies that open balls with respect to $d(\cdot,\cdot)$ are open sets, which could also be verified more directly when $q < 1$.

Similarly, closed balls with respect to a semimetric are closed sets, by standard arguments. This implies that closed balls with respect to $d(\cdot,\cdot)$ are closed sets, which could be verified more directly when $q < 1$ as well.

If $d(\cdot,\cdot)$ is a semi-ultrametric on X, then open balls in X with respect to $d(\cdot,\cdot)$ are closed sets too, and closed balls of positive radius are open sets, as in Section A.1.

It is well known and easy to see that the topology determined by a metric is Hausdorff. This implies that the topology determined by a q-metric is Hausdorff, which could be verified more directly, as usual. If $d(\cdot,\cdot)$ is a q-semimetric on X, and the topology determined on X by $d(\cdot,\cdot)$ satisfies the first or zeroth separation condition, then $d(\cdot,\cdot)$ is a q-metric on X.

It is well known that the topology determined by a metric is normal in the strong sense, and regular in the strong sense in particular, as mentioned in Section 5.5. This implies that the topology determined by a q-metric has the same properties, as before. It is not too difficult to show that the topology determined by a semimetric is regular and normal in the strict sense, which implies the analogous statements for q-semimetrics.

If $d(\cdot,\cdot)$ is a q-metric on X, then one can define the notions of Cauchy sequences and completeness with respect to $d(\cdot,\cdot)$ in the usual way. It is easy to see that $d(\cdot,\cdot)$ and (B.2.1) determine the same Cauchy sequences in X, as in Section A.1. This implies that X is complete with respect to $d(\cdot,\cdot)$ if and only if X is complete with respect to (B.2.1), as before.

B.3 q-Absolute value functions

Let k be a field, and let $|\cdot|$ be a nonnegative real-valued function on k such that

(B.3.1) $$|x| = 0 \text{ if and only if } x = 0$$

and

(B.3.2) $$|x\,y| = |x|\,|y|$$

for every $x, y \in k$. This implies that

(B.3.3) $$|1_k| = |-1_k| = 1,$$

as in Section A.2.

Let q be a positive real number, and suppose that

(B.3.4) $$|x+y|^q \le |x|^q + |y|^q$$

for every $x, y \in k$. In this case, we say that $|\cdot|$ is a *q-absolute value function* on k. This is the same as an ordinary absolute value function, as in Section A.2, when $q = 1$.

Note that (B.3.4) is the same as saying that

$$|x + y| \leq (|x|^q + |y|^q)^{1/q}. \tag{B.3.5}$$

The right side is monotonically decreasing in q, as in Section 1.12, so that this condition becomes more restrictive as q increases.

It is easy to see that an ultrametric absolute value function on k is a q-absolute value function for every $q > 0$. We also have that

$$\lim_{q \to \infty} (|x|^q + |y|^q)^{1/q} \tag{B.3.6}$$

for every $x, y \in k$, as in Section 1.3. Ultrametric absolute value functions on k will be considered as q-absolute value functions with $q = \infty$.

If $|\cdot|$ is a q-absolute value function on k for some $q > 0$ and a is a positive real number, then

$$|x|^a \tag{B.3.7}$$

is a (q/a)-absolute value function on k. If $q = \infty$, then this means that (B.3.7) is an ultrametric absolute value function on k when $|\cdot|$ is an ultrametric absolute value function on k, as in Section A.4.

If $|\cdot|$ is a q-absolute value function on k, then $|\cdot|$ is an absolute value function on k when $q \geq 1$, and otherwise

$$|x|^q \tag{B.3.8}$$

is an absolute value function on k. This means that q-absolute value functions have the same types of properties as ordinary absolute value functions, as in Sections A.2, A.3, A.4, and A.8.

In particular, if $|\cdot|$ is a q-absolute value function on k, then

$$d(x, y) = |x - y| \tag{B.3.9}$$

is a q-metric on k.

B.4 Quasimetric absolute value functions

Let k be a field again, and let $|\cdot|$ be a nonnegative real-valued function on k that satisfies (B.3.1) and (B.3.2). Suppose that there is a real number $C \geq 1$ such that

$$|1_k + z| \leq C \tag{B.4.1}$$

for every $z \in k$ with

$$|z| \leq 1, \tag{B.4.2}$$

where 1_k is the multiplicative identity element in k. In this case, $|\cdot|$ may be called a *valuation* on k, as in Definition 1.1 on p12 of [44]. This term is also sometimes used in other but related ways.

Suppose that $x, y \in k$ satisfy

$$|x| \leq |y|. \tag{B.4.3}$$

If $y \neq 0$, then $z = x/y$ satisfies (B.4.2), and (B.4.1) implies that

$$|x + y| = |1_k + z|\,|y| \le C\,|y|. \tag{B.4.4}$$

Let us say that $|\cdot|$ is a *quasimetric absolute value function* on k if it satisfies (B.3.1) and (B.3.2), as before, and if there is a real number $C_1 \geq 1$ such that

$$|x + y| \le C_1\,(|x| + |y|) \tag{B.4.5}$$

for every $x, y \in k$. This is the same as saying that $|\cdot|$ is an absolute value function on k when $C_1 = 1$.

Alternatively, one may ask that there be a real number $C_0 \geq 1$ such that

$$|x + y| \le C_0\,\max(|x|, |y|) \tag{B.4.6}$$

for every $x, y \in k$. This means that $|\cdot|$ is an ultrametric absolute value function on k when $C_0 = 1$.

Clearly (B.4.6) implies (B.4.5), with $C_1 = C_0$. Conversely, (B.4.5) implies (B.4.6), with $C_0 = 2\,C_1$.

Similarly, (B.4.6) implies (B.4.1), with $C = C_0$. Conversely, (B.4.1) implies (B.4.6), with $C_0 = C$, as in (B.4.4). Thus a quasimetric absolute value function on k is the same as a valuation on k in the sense of [44].

If a is a positive real number, then (B.3.7) also satisfies (B.3.1) and (B.3.2). If $|\cdot|$ is a valuation on k in the sense of [44] with constant C, then

$$|1_k + z|^a \le C^a \tag{B.4.7}$$

for every $z \in k$ with

$$|z|^a \le 1, \tag{B.4.8}$$

so that (B.3.7) is a valuation on k in the sense of [44] with constant C^a. This corresponds to Lemma 1.1 on p13 of [44]. Equivalently, if $|\cdot|$ is a quasimetric absolute value function on k satisfying (B.4.6) for some C_0, then

$$|x + y|^a \le C_0^a\,\max(|x|^a, |y|^a) \tag{B.4.9}$$

for every $x, y \in k$.

Lemma 1.2 on p13 of [44] says that $|\cdot|$ is an absolute value function on k if and only if $|\cdot|$ is a valuation on k in the sense of [44] with constant $C = 2$. Of course, the "only if" part of this statement is very easy to see. This means that $|\cdot|$ is an absolute value function on k if and only if $|\cdot|$ is a quasimetric absolute value function on k that satisfies (B.4.6) with $C_0 = 2$.

If $|\cdot|$ is a valuation on k in the sense of [44], or equivalently a quasimetric absolute value function, then it follows that (B.3.7) is an absolute value function on k when $a > 0$ is sufficiently small, as in the corollary on p14 of [44].

Let q be a positive real number, and remember that

$$|\cdot| \text{ is a } q\text{-absolute value function on } k \tag{B.4.10}$$

if and only if $|x|^q$ is an absolute value function on k. Lemma 1.2 on p13 of [44] implies that this happens if and only if $|\cdot|$ is a valuation on k in the sense of [44] with constant $C = 2^{1/q}$, which is the same as saying that $|\cdot|$ is a quasimetric absolute value function on k that satisfies (B.4.6) with $C_0 = 2^{1/q}$. If $|\cdot|$ is a valuation on k in the sense of [44], or equivalently a quasimetric absolute value function, then we get that (B.4.10) holds when $q > 0$ is sufficiently small.

B.5 q-Norms and q-seminorms

Let k be a field, and let $|\cdot|$ be a q_k-absolute value function on k for some $q_k > 0$. Also let V be a vector space over k, let N be a nonnegative real-valued function on V, and let q_N be a positive real number. We say that N is a *q_N-seminorm* or *q_N-pseudonorm* on V with respect to $|\cdot|$ on k if it satisfies the following two conditions. The first condition is the usual homogeneity property

$$N(t\,v) = |t|\,N(v) \tag{B.5.1}$$

for every $t \in k$ and $v \in V$. The second condition is the q_N-seminorm version of the *triangle inequality*, which is that

$$N(v + w)^{q_N} \le N(v)^{q_N} + N(w)^{q_N} \tag{B.5.2}$$

for every $v, w \in V$.

If we also have that $N(v) > 0$ when $v \neq 0$, then N is said to be a *q_N-norm* on V with respect to $|\cdot|$ on k. If $q_k = q_N = 1$, then q_N-seminorms and q_N-norms on V with respect to $|\cdot|$ on k are the same as ordinary seminorms and norms, as in Section A.5. Of course, $|\cdot|$ is a q_k-norm on k, as a one-dimensional vector space over itself, as before.

As usual, (B.5.2) is the same as saying that

$$N(v + w) \le (N(v)^{q_N} + N(w)^{q_N})^{1/q_N} \tag{B.5.3}$$

for every $v, w \in V$. The right side is monotonically decreasing in q_N, as in Section 1.12, so that this condition is more restrictive as q_N increases, as before.

If N satisfies (B.5.1) and

$$N(v + w) \le \max(N(v), N(w)) \tag{B.5.4}$$

for every $v, w \in V$, then we say that N is a *semi-ultranorm* or *pseudo-ultranorm* on V with respect to $|\cdot|$ on k, as in Section A.5. If $N(v) > 0$ when $v \neq 0$ too, then N is said to be an *ultranorm* on V with respect to $|\cdot|$ on k, as before. A semi-ultranorm on V is a q_N-seminorm for every $q_N > 0$, as usual, and similarly an ultranorm is a q_N-norm for every $q_N > 0$.

As in Section 1.3,

$$\lim_{q_N \to \infty} (N(v)^{q_N} + N(w)^{q_N})^{1/q_N} = \max(N(v), N(w)) \tag{B.5.5}$$

for every $v, w \in V$. Semi-ultranorms and ultranorms on V may be considered as q_N-seminorms and q_N-norms, respectively, with $q_N = \infty$, as before.

If N is a q_N-seminorm on V with respect to $|\cdot|$ on k, and if $N(v) > 0$ for some $v \in V$, then it is easy to see that $|\cdot|$ is a q_N-absolute value function on k. In particular, if N is a q_N-norm on V, and $V \neq \{0\}$, then $|\cdot|$ is a q_N-absolute value function on k.

If N is a q_N-seminorm on V with respect to $|\cdot|$ on k, then

$$d_N(v, w) = N(v - w) \tag{B.5.6}$$

is a q_N-semimetric on V. If N is a q_N-norm on V, then this is a q_N-metic on V.

Let a be a positive real number, and remember that $|\cdot|^a$ defines a (q_k/a)-absolute value function on k, as in Section B.3. If N is a q_N-seminorm on V with respect to $|\cdot|$ on k, then it is easy to see that

$$N(v)^a \tag{B.5.7}$$

is a (q_N/a)-semimetric on V with respect to $|\cdot|^a$ on k. Similarly, if N is a q_N-norm on V with respect to $|\cdot|$ on k, then this is a (q_N/a)-norm on V with respect to $|\cdot|^a$ on k.

B.6 ℓ^r Spaces and q-seminorms

Let X be a nonempty set, and let k be a field with a q_k-absolute value function $|\cdot|$ for some $q_k > 0$. Also let W be a vector space over k, and let N_W be a q_W-seminorm on W with respect to $|\cdot|$ on k for some $q_W > 0$. Consider the space

$$\ell^\infty(X, W) = \ell^\infty_{N_W}(X, W) \tag{B.6.1}$$

of all W-valued functions f on X that are bounded with respect to N_W, so that $N_W(f(x))$ is bounded as a real-valued function on X, as in Section A.9. In this case, we put

$$\|f\|_\infty = \|f\|_{sup} = \|f\|_{\ell^\infty(X,W)} = \|f\|_{\ell^\infty_{N_W}(X,W)} = \sup_{x \in X} N(f(x)), \tag{B.6.2}$$

as before.

If r is a positive real number, then we let

$$\ell^r(X, W) = \ell^r_{N_W}(X, W) \tag{B.6.3}$$

be the space of W-valued functions f on X such that $N_W(f(x))$ is r-summable as a nonnegative real-valued function on X, as in Section A.11. Under these conditions, we put

$$\|f\|_r = \|f\|_{\ell^r(X,W)} = \|f\|_{\ell^r_{N_W}(X,W)} = \Big(\sum_{x \in X} N_W(f(x))^r\Big)^{1/r}, \tag{B.6.4}$$

as before.

One can check that $\ell^r_{N_W}(X, W)$ is a linear subspace of the space of all W-valued functions on X for each $0 < r \le \infty$. Note that (B.6.2) and (B.6.4) satisfy the usual homogeneity condition (B.5.1) with respect to $|\cdot|$ on k.

If

$$r \le q_W, \tag{B.6.5}$$

then one can verify that

$$\|\cdot\|_{\ell^r_{N_W}(X,W)} \text{ is an } r\text{-seminorm on } \ell^r_{N_W}(X, W), \tag{B.6.6}$$

with respect to $|\cdot|$ on k. This uses the fact that

$$N_W \text{ is an } r\text{-seminorm on } W \tag{B.6.7}$$

with respect to $|\cdot|$ on k, because of (B.6.5), as in the previous section. If N_W is a q_W-norm on W, then

$$\|\cdot\|_{\ell^r_{N_W}(X,W)} \text{ is an } r\text{-norm on } \ell^r_{N_W}(X, W). \tag{B.6.8}$$

If

$$q_W \le r, \tag{B.6.9}$$

then

$$\|\cdot\|_{\ell^r_{N_W}(X,W)} \text{ is a } q_W\text{-seminorm on } \ell^r_{N_W}(X, W) \tag{B.6.10}$$

with respect to $|\cdot|$ on k. This can be seen fairly directly when $r = \infty$, and otherwise it can be obtained using Minkowski's inequality for sums. If N_W is a q_W-norm on W, then we get that

$$\|\cdot\|_{\ell^r_{N_W}(X,W)} \text{ is a } q_W\text{-norm on } \ell^r_{N_W}(X, W), \tag{B.6.11}$$

as before.

B.7 More on ℓ^r spaces

Let us continue with the same notation and hypotheses as in the previous section. We may refer to (B.6.2) as the *supremum q_W-seminorm* on $\ell^\infty_{N_W}(X, W)$ associated to N_W, or as the *supremum q_W-norm* when N_W is a q_W-norm on W. This leads to the corresponding *supremum q_W-semimetric* or *supremum q_W-metric* on $\ell^\infty_{N_W}(X, W)$ in the usual way.

If $0 < r_1 \le r_2 \le +\infty$, then

$$\ell^{r_1}_{N_W}(X, W) \subseteq \ell^{r_2}_{N_W}(X, W), \tag{B.7.1}$$

as in Sections 2.6 and A.11. In this case,

$$\|f\|_{\ell^{r_2}_{N_W}(X,W)} \le \|f\|_{\ell^{r_1}_{N_W}(X,W)} \tag{B.7.2}$$

for every $f \in \ell^{r_1}_{N_W}(X, W)$, as before.

Let us say that a W-valued function f on X *vanishes at infinity* on X with respect to N_W if $N_W(f(x))$ vanishes at infinity as a nonnegative real-valued function on X, as in Section A.9. The space

$$c_0(X, W) = c_{0,N_W}(X, W) \tag{B.7.3}$$

of these functions is a linear subspace of $\ell^\infty_{N_W}(X, W)$, as before. Remember that $c_{00}(X, W)$ is the space of Wvalued functions on X with finite support, as in Section A.9. One can check that

$$c_{0,N_W}(X, W) \text{ is the closure of } c_{00}(X, W) \text{ in } \ell^\infty_{N_W}(X, W) \tag{B.7.4}$$

with respect to the supremum q_W-semimetric, as before.

If $0 < r < \infty$, then

$$c_{00}(X, W) \subseteq \ell^r_{N_W}(X, W) \subseteq c_{0,N_W}(X, W), \tag{B.7.5}$$

as in Section A.11. In this case, one can verify that

$$c_{00}(X, W) \text{ is dense in } \ell^r_{N_W}(X, W) \tag{B.7.6}$$

with respect to the q_W or r-semimetric, as appropriate, associated to (B.6.4), as before.

Suppose that N_W is a q_W-norm on W, and that W is complete with respect to the corresponding q_W-metric. One can check that

$$\ell^r_{N_W}(X, W) \text{ is complete} \tag{B.7.7}$$

with respect to the q_W or r-metric associated to (B.6.2) or (B.6.4), as appropriate, as in Sections A.9 and A.11.

B.8 q-Absolute convergence

Let k be a field with a q_k-absolute value function for some $q_k > 0$, let V be a vector space over k, and let N be a q_N-norm on V with respect to $|\cdot|$ on k for some $q_N > 0$. Of course, one can define convergence of sequences in V with respect to the q_N-metric d_N associated to N as in Section B.5 in the usual way, which is the same as convergence of sequences with respect to the topology determined on V by d_N as in Section B.2. One can check that convergent sequences in V with respect to d_N have the usual properties in terms of sums and scalar multiplication, as in Section A.7.

Convergence of an infinite series $\sum_{j=1}^\infty v_j$ with terms in V with respect to N means that the corresponding sequence of partial sums $\sum_{j=1}^n v_j$ converges in V with respect to d_N, as before. This also has the usual properties in terms of sums and scalar multiplication, because of the analogous statements for sequences.

The sequence of partial sums is a Cauchy sequence with respect to d_N if and only if for every $\epsilon > 0$ there is a positive integer $L(\epsilon)$ such that

(B.8.1) $$N\Big(\sum_{j=l}^{n} v_j\Big) < \epsilon$$

for all integers $n \geq l \geq L(\epsilon)$, as in Section A.7. This implies that $\{v_j\}_{j=1}^\infty$ converges to 0 with respect to d_N, as usual.

Suppose that $q_N < \infty$, and observe that

(B.8.2) $$N\Big(\sum_{j=l}^{n} v_j\Big)^{q_N} \leq \sum_{j=l}^{n} N(v_j)^{q_N}$$

for all positive integers $l \leq n$. We say that $\sum_{j=1}^\infty$ *converges q_N-absolutely* with respect to N if

(B.8.3) $$\sum_{j=1}^{\infty} N(v_j)^{q_N}$$

converges as an infinite series of nonnegative real numbers. This is the same as ordinary absolute convergence when $q_N = 1$, as in Section A.7. This implies that the corresponding sequence of partial sums is a Cauchy sequence with respect to d_N, because of (B.8.2), as before.

If V is complete with respect to d_N, as in Section B.2, then we say that V is a *q_N-Banach space*. This is the same as a Banach space when $q_N = 1$, as in Section A.7. If V is a q_N-Banach space with respect to N and $\sum_{j=1}^\infty v_j$ converges q_N-absolutely with respect to N, then $\sum_{j=1}^\infty v_j$ converges in V, and one can check that

(B.8.4) $$N\Big(\sum_{j=1}^{\infty} v_j\Big)^{q_N} \leq \sum_{j=1}^{\infty} N(v_j)^{q_N}.$$

One can also verify that the completeness of V with respect to d_N is characterized by the condition that q_N-absolutely convergent series in V with respect to N converge in V, as before. If V is not already complete with respect to d_N, then V has a completion that is a q_N-Banach space over k, as usual.

B.9 Some more Lipschitz conditions

Let X, Y be sets with q_X, q_Y-semimetrics d_X, d_Y for some $q_X, q_Y > 0$, and let α be a positive real number. A mapping f from X into Y is said to be *Lipschitz of order α* with respect to d_X, d_Y if there is a nonnegative real number C such that

(B.9.1) $$d_Y(f(x), f(w)) \leq C\, d_X(x, w)^\alpha$$

for every $x, w \in X$. This is the same as in Section 2.1 when $\alpha = 1$. As before, we may say that f is Lipschitz of order α with constant C, to be more precise.

One can define *uniform continuity* of mappings from X into Y with respect to d_X, d_Y in the same way as for metric spaces. In particular,

(B.9.2) Lipschitz mappings of order α are uniformly continuous,

as in Section 2.1. Constant mappings from X into Y are Lipschitz of order α with $C = 0$. If f is Lipschitz of order α with constant $C = 0$, and if d_Y is a q_Y-metric on Y, then f is constant on X.

Let a_X, a_Y be positive real numbers, so that $d_X^{a_X}$, $d_Y^{a_Y}$ are (q_X/a_X), (q_Y/a_Y)-semimetrics on X, Y, respectively, as in Section B.2. It is easy to see that (B.9.1) is the same as saying that

$$d_Y(f(x), f(w))^{a_Y} \le C^{a_Y}\,(d_X(x,w)^{a_X})^{\alpha\, a_Y/a_X} \tag{B.9.3}$$

for every $x, w \in X$. This means that f is Lipschitz of order α with constant C with respect to d_X, d_Y if and only if

(B.9.4) f is Lipschitz of order $\alpha\, a_Y/a_X$ with constant C^{a_Y}

with respect to $d_X^{a_X}$, $d_Y^{a_Y}$.

The space of all Lipschitz mappings from X into Y of order α with respect to d_X, d_Y may be denoted

$$\mathrm{Lip}_\alpha(X, Y). \tag{B.9.5}$$

If $\alpha = 1$, then this corresponds to $\mathrm{Lip}(X, Y)$, as in Section 2.1.

Suppose that $X \neq \emptyset$, and that f is a Lipschitz mapping of order α from X into Y. We would like to put

$$\begin{aligned} \text{(B.9.6)}\quad \mathrm{Lip}_\alpha(f) &= \mathrm{Lip}_{\alpha,X,Y}(f) \\ &= \sup\left\{\frac{d_Y(f(x), f(w))}{d_X(x,w)^\alpha} : x, w \in X,\ d_X(x,w) > 0\right\}, \end{aligned}$$

at least if there are $x, w \in X$ such that $d_X(x, w) > 0$. Otherwise, this may be interpreted as being equal to 0. This corresponds to $\mathrm{Lip}(f)$ as in Section 2.1 when $\alpha = 1$, and we may also use this notation here in this case. One can check that $\mathrm{Lip}_\alpha(f)$ is the smallest nonnegative real number C such that f is Lipschitz of order α with constant C, as before.

Let k be a field with a q_k-absolute value function $|\cdot|$ for some $q_k > 0$, and suppose now that Y is a vector space over k with a q_Y-seminorm N_Y with respect to $|\cdot|$ on k. In this case, we can take d_Y to be the q_Y-semimetric d_{N_Y} on Y associated to N_Y. One can check that

(B.9.7) $\mathrm{Lip}_\alpha(X, Y)$ is a linear subspace of the space of all functions on X with values in Y,

as in Sections 2.1 and A.10. More precisely,

(B.9.8) $\mathrm{Lip}_\alpha(f)$ is a q_Y-seminorm on $\mathrm{Lip}_\alpha(X, Y)$

with respect to $|\cdot|$ on k, as before.

Note that the space of uniformly continuous mappings from X into Y is a linear subspace of the space of all mappings from X into Y under these conditions.

B.10 Operator q-seminorms

Let k be a field with a q_k-absolute value function for some $q_k > 0$, and let V, W be vector spaces over k with q_V, q_W-seminorms N_V, N_W, respectively, with respect to $|\cdot|$ on k, for some $q_V, q_W > 0$. A linear mapping T from V into W is said to be *bounded* with respect to N_V, N_W if there is a nonnegative real number C such that

$$N_W(T(v)) \le C\, N_V(v) \tag{B.10.1}$$

for every $v \in V$, as in Sections 2.2 and A.10. This means that

$$N_W(T(u) - T(v)) \le C\, N_V(u - v) \tag{B.10.2}$$

for every $u, v \in V$, as before. This says that T is Lipschitz of order 1 with constant C with respect to the q_V, q_W-semimetrics d_{N_V}, d_{N_W} associated to N_V, N_W, respectively, as in the previous section.

Let $\mathcal{BL}(V, W)$ be the space of all bounded linear mappings from V into W with respect to N_V, N_W, respectively, as before. One can verify that this is a linear subspace of the space $\mathcal{L}(V, W)$ of all linear mappings from V into W, as before. If $T \in \mathcal{BL}(V, W)$, then put

$$\|T\|_{op} = \|T\|_{op,VW} = \inf\{C \ge 0 : (B.10.1) \text{ holds}\}, \tag{B.10.3}$$

as in Section A.10. This is the same as the Lipschitz constant $\mathrm{Lip}(T) = \mathrm{Lip}_1(T)$ of T with respect to the q_V, q_W-metrics associated to N_V, N_W, respectively, as before.

In particular,

$$\|\cdot\|_{op} \text{ is a } q_W\text{-seminorm on } \mathcal{BL}(V, W) \tag{B.10.4}$$

with respect to $|\cdot|$ on k, as in the previous section. This is the *operator q_W-seminorm* on $\mathcal{BL}(V, W)$ associated to N_V, N_W. Note that the infimum on the right side of (B.10.3) is automatically attained, so that

$$N_W(T(v)) \le \|T\|_{op,VW}\, N_V(v) \tag{B.10.5}$$

for every $v \in V$. If N_W is a q_W-norm on W, then

$$\|\cdot\|_{op} \text{ is a } q_W\text{-norm on } \mathcal{BL}(V, W), \tag{B.10.6}$$

as in Section A.10. In this case, $\|\cdot\|_{op}$ may be called the *operator q_W-norm* on $\mathcal{BL}(V, W)$ associated to N_V, N_W.

Let Z be another vector space over k, with a q_Z-seminorm N_Z with respect to $|\cdot|$ on k for some $q_Z > 0$. If T_1, T_2 are bounded linear mappings from V, W into W, Z, respectively, then $T_2 \circ T_1$ is a bounded linear mapping from V into Z, with

$$\|T_2 \circ T_1\|_{op,VZ} \le \|T_1\|_{op,VW}\, \|T_2\|_{op,WZ}, \tag{B.10.7}$$

as in Sections 2.2 and A.10.

Let T be a linear mapping from V into W again, and suppose for the moment that $N_W(T(v))$ is bounded on a ball of positive radius in V centered at 0 with respect to the q_V-semimetric associated to N_V. If $|\cdot|$ is not the trivial absolute value function on k, then one can verify that

(B.10.8) T is a bounded linear mapping from V into W

with respect to N_V, N_W, as in Section A.10. In particular, the hypothesis on T holds when T is continuous at 0 with respect to the q_V, q_W-semimetrics associated to N_V, N_W, respectively, as before.

If N_W is a q_W-norm on W, and W is complete with respect to the q_W-metric associated to N_W, then

(B.10.9) $\mathcal{BL}(V, W)$ is complete

with respect to the q_W-metric associated to $\|\cdot\|_{op}$, as in Sections 2.2 and A.10.

Let V_0 be a linear subspace of V that is dense in V with respect to the q_V-semimetric associated to N_V, and let T_0 be a bounded linear mapping from V_0 into W, with respect to the restriction of N_V to V_0. If N_W is a q_W-norm on W, and W is complete with respect to the q_W-metric associated to N_W, then

(B.10.10) there is a unique extension of T_0 to a bounded linear mapping T from V into W,

as in Sections 2.2 and A.10. One can verify that the operator q_W-norm of T on V is equal to the operator q_W-norm of T_0 on V_0, as before.

B.11 Some more bounded linear mappings

Let X be a nonempty set, let k be a field with a q_k-absolute value function for some $q_k > 0$, and let W be a vector space over k with a q_W-seminorm N_W with respect to $|\cdot|$ on k for some $q_W > 0$. If a is a W-valued function on X and f is a k-valued function on X with finite support, then put

$$T_a(f) = \sum_{x \in X} a(x)\, f(x), \tag{B.11.1}$$

as in Section A.12. This defines a linear mapping from $c_{00}(X, k)$ into W, as before.

Let δ_y be the k-valued function on X equal to 1 at $y \in X$ and to 0 at every other point in X, as before. Remember that $\|\delta_y\|_{\ell^r(X,k)} = 1$ for every $r > 0$, using $|\cdot|$ as a q_k-norm on k, as a one-dimensional vector space over itself. We also have that

$$T_a(\delta_y) = a(y) \tag{B.11.2}$$

for every $y \in X$, as before.

Suppose that

(B.11.3) T_a is bounded as a linear mapping from $c_{00}(X, k)$ into W, with respect to $\|\cdot\|_{\ell^r(X,k)}$ on $c_{00}(X, k)$ and N_W on W

for some $r > 0$. This implies that

$$N(a(y)) \leq \|T\|_{op,r} \tag{B.11.4}$$

for every $y \in X$, as in Section A.12, where $\|T_a\|_{op,r}$ is the operator q_W-seminorm of T_a with respect to $\|\cdot\|_{\ell^r(X,k)}$ on $c_{00}(X,k)$ and N_W on W. Equivalently, a is bounded on X with respect to N_W on W, with

$$\|a\|_{\ell^\infty_{N_W}(X,W)} \leq \|T_a\|_{op,r}. \tag{B.11.5}$$

If a is any element of $\ell^\infty_{N_W}(X,W)$, $f \in c_{00}(X,W)$, and $q_W < \infty$, then

$$\begin{aligned} N_W(T_a(f))^{q_W} &\leq \sum_{x \in X} N_W(a(x))^{q_W}\,|f(x)|^{q_W} \\ &\leq \|a\|^{q_W}_{\ell^\infty_{N_W}(X,W)}\,\|f\|^{q_W}_{\ell^{q_W}(X,k)}, \end{aligned} \tag{B.11.6}$$

as before. This means that

(B.11.7) T_a is a bounded linear mapping from $c_{00}(X,k)$ into W, with respect to $\|\cdot\|_{\ell^{q_W}(X,k)}$ on $c_{00}(X,k)$ and N_W on W,

with

$$\|T_a\|_{op,q_W} \leq \|a\|_{\ell^\infty_{N_W}(X,W)}. \tag{B.11.8}$$

More precisely, we have that

$$\|T_a\|_{op,q_W} = \|a\|_{\ell^\infty_{N_W}(X,W)}, \tag{B.11.9}$$

because of (B.11.5). If N_W is a q_W-norm on W, and W is complete with respect to the associated q_W-metric, then it follows that

(B.11.10) T_a has a unique extension to a bounded linear mapping from $\ell^{q_W}(X,k)$ into W, with the same operator q_W-norm,

as in the previous section.

If $q_W = \infty$, then

$$N_W(T_a(f)) \leq \|a\|_{\ell^\infty_{N_W}(X,W)}\,\|f\|_{\ell^\infty(X,k)} \tag{B.11.11}$$

for every $f \in c_{00}(X,k)$, so that T_a is a bounded linear mapping from $c_{00}(X,k)$ into W, with respect to $\|\cdot\|_{\ell^\infty(X,k)}$ on $c_{00}(X,k)$ and N_W on W, as in Section A.12. We also have that

$$\|T_a\|_{op,\infty} = \|a\|_{\ell^\infty_{N_W}(X,W)} \tag{B.11.12}$$

in this case, because of (B.11.5), as before. If N_W is an ultranorm on W, and W is complete with respect to the associated ultrametric, then T_a has a unique extension to a bounded linear mapping from $c_0(X,k)$ into W, with the same operator norm, as before.

B.12 More on sums of vectors

Let X be a nonempty set, let k be a field with a q_k-absolute value function $|\cdot|$ for some $q_k > 0$ again, and let W be a vector space over k with a q_W-norm N_W with respect to $|\cdot|$ on k for some $q_W > 0$. If f is a W-valued function on X, then the *convergence* of the sum $\sum_{x \in X} f(x)$ *in the generalized sense* with respect to the q_W-metric associated to N_W can be defined in the same way as in Section A.13. Similarly, the *generalized Cauchy condition* for the sum can be defined in the same way as before. These have essentially the same properties as before, with $\epsilon/2$ replaced with

$$\epsilon/2^{q_W} \tag{B.12.1}$$

in a couple of places when $q_W < \infty$.

If $q_W < \infty$, then

$$N_W\Big(\sum_{x \in C} f(x)\Big)^{q_W} \le \sum_{x \in C} N_W(f(x))^{q_W} \tag{B.12.2}$$

for every nonempty finite subset C of X. This implies that the sum satisfies the generalized Cauchy condition when

$$f \in \ell^{q_W}_{N_W}(X, W), \tag{B.12.3}$$

as before. If the sum converges in the generalized sense, then one can check that

$$N_W\Big(\sum_{x \in X} f(x)\Big)^{q_W} \le \sum_{x \in X} N_W(f(x))^{q_W}, \tag{B.12.4}$$

as before.

One can define $Sum_{N_W}(X, W)$ and $GCC_{N_W}(X, W)$ in the same was as in Section A.14, and with basically the same properties as before. If $q_W < \infty$, then

$$\ell^{q_W}_{N_W}(X, W) \subseteq GCC_{N_W}(X, W), \tag{B.12.5}$$

as in the preceding paragraph. This implies that

$$\ell^{q_W}_{N_W}(X, W) \subseteq Sum_{N_W}(X, W) \tag{B.12.6}$$

when W is a q_W-Banach space, as before.

One can also define $BFS_{N_W}(X, W)$ and $\|f\|_{BFS_{N_W}(X,W)}$ in the same way as before, and with essentially the same properties as before. More precisely,

$$\|f\|_{BFS_{N_W}(X,W)} \text{ is a } q_W\text{-norm on } BFS_{N_W}(X, W), \tag{B.12.7}$$

with respect to $|\cdot|$ on k. It is easy to see that

$$\ell^{q_W}_{N_W}(X, W) \subseteq BFS_{N_W}(X, W), \tag{B.12.8}$$

with

$$\|f\|_{BFS_{N_W}(X,W)} \le \|f\|_{\ell^{q_W}_{N_W}(X,W)} \tag{B.12.9}$$

for every $f \in \ell^{q_W}_{N_W}(X, W)$.

Bibliography

[1] F. Acker, *A note on the spectral theorem in the finite-dimensional real case*, American Mathematical Monthly **121** (2014), 942–945.

[2] F. Albiac and N. Kalton, *A characterization of real $\mathcal{C}(K)$-spaces*, American Mathematical Monthly **114** (2007), 737–743.

[3] F. Albiac and N. Kalton, *Topics in Banach Space Theory*, 2nd edition, Springer, 2016.

[4] H. Alexander and J. Wermer, *Several Complex Variabkes and Banach Algebras*, 3rd edition, Springer-Verlag, 1998.

[5] T. Alfred-Jones, *A short proof of the uniform smoothness of certain Lebesgue spaces*, American Mathematical Monthly **127** (2020), 643–648.

[6] S. Artstein-Avidan, A. Giannopoulos, and V. Milman, *Asymptotic Geometric Analysis, Part I*, American Mathematical Society, 2015.

[7] S. Artstein-Avidan, A. Giannopoulos, and V. Milman, *Asymptotic Geometric Analysis, Part II*, American Mathematical Society, 2021.

[8] W. Arveson, *A Short Course on Spectral Theory*, Springer-Verlag, 2002.

[9] M. Atiyah, *K-Theory*, Notes by D. Anderson, 2nd edition, Addison-Wesley, 1989.

[10] M. Atiyah and I. Macdonald, *Introduction to Commutative Algebra*, Addison-Wesley, 1969.

[11] G. Bachman, *Introduction to p-Adic Numbers and Valuation Theory*, Academic Press, 1964.

[12] G. Bachman and L. Narici, *Functional Analysis*, Dover, 2000.

[13] J. Baker, *Isometries in normed spaces*, American Mathematical Monthly **78** (1971), 655–658.

[14] R. Bartle, *Nets and filters in topology*, American Mathematical Monthly **62** (1955), 551–557; correction, ibid., **70** (1963), 52–53.

[15] R. Beals, *Topics in Operator Theory*, University of Chicago Press, 1971.

[16] R. Beals, *Advanced Mathematical Analysis: Periodic Functions and Distributions, Complex Analysis, Laplace Transform and Applications*, Springer-Verlag, 1973.

[17] R. Beals, *Analysis: An Introduction*, Cambridge University Press, 2004.

[18] B. Beauzamy, *Introduction to Banach Spaces and their Geometry*, 2nd edition, North-Holland, 1985.

[19] J. Benedetto and W. Czaja, *Integration and Modern Analysis*, Birkhäuser, 2009.

[20] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis*, American Mathematical Society, 2000.

[21] S. Berberian, *The spectral mapping theorem for a Hermitian operator*, American Mathematical Monthly **70** (1963), 1049–1051.

[22] S. Berberian, *Notes on Spectral Theory*, Van Nostrand, 1966.

[23] S. Berberian, *Baer $*$-Rings*, Springer-Verlag, 1972.

[24] S. Berberian, *Lectures in Functional Analysis and Operator Theory*, Springer-Verlag, 1974.

[25] S. Berberian, *Introduction to Hilbert Space*, Chelsea, 1976.

[26] S. Berberian, *A First Course in Real Analysis*, Springer-Verlag, 1994.

[27] S. Berberian, *Fundamentals of Real Analysis*, Springer-Verlag, 1999.

[28] S. Berberian, *Measure and Integration*, AMS Chelsea, 2011.

[29] M. Berger, *Convexity*, American Mathematical Monthly **97** (1990), 650–678.

[30] G. Birkhoff and S. Mac Lane, *A Survey of Modern Algebra*, 4th edition, Macmillan, 1977.

[31] R. Boas, *A Primer of Real Functions*, 4th edition, revised and updated by H. Boas, Mathematical Association of America, 1996.

[32] R. Bott and L. Tu, *Differential Forms in Algebraic Topology*, Springer-Verlag, 1982.

[33] A. Bowers and N. Kalton, *An Introductory Course in Functional Analysis*, Springer, 2014.

[34] J. Brace, *Compactness in the weak topology*, Mathematics Magazine **28** (1955), 125–134.

[35] A. Browder, *Introduction to Function Algebras*, Benjamin, 1969.

[36] A. Browder, *Mathematical Analysis: An Introduction*, Springer-Verlag, 1996.

[37] A. Browder, *On Bernstein's inequality and the norm of Hermitian operators*, American Mathematical Monthly **78** (1971), 871–873.

[38] A. Browder, *Topology in the complex plane*, American Mathematical Monthly **107** (2000), 393–401.

[39] A. Browder, *Complex numbers and the ham sandwich theorem*, American Mathematical Monthly **113** (2006), 935–937.

[40] P. Bullen, *An inequality for variations*, American Mathematical Monthly **90** (1983), 560.

[41] E. Burger and T. Struppeck, *Does* $\sum_{n=0}^{\infty} 1/n!$ *really converges? Infinite series and p-adic analysis or "You can sum some of the series some of the time and some of the series none of the time... but can you sum some of the series all of the time?"*, American Mathematical Monthly **103** (1996), 565–577.

[42] D. Burton and J. Coleman, *Quasi-Cauchy sequences*, American Mathematical Monthly **117** (2010), 328–333.

[43] H. Cartan and S. Eilenberg, *Homological Algebra*, with an appendix by D. Buchsbaum, Princeton University Press, 1999.

[44] J. Cassels, *Local Fields*, Cambridge University Press, 1986.

[45] R. Coifman and G. Weiss, *Analyse Harmonique Non-Commutative sur Certains Espaces Homogènes*, Lecture Notes in Mathematics **242**, Springer-Verlag, 1971.

[46] R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis*, Bulletin of the American Mathematical Society **83** (1977), 569–645.

[47] A. Connes, *Noncommutative Geometry*, Academic Press, 1994.

[48] J. Conway, *A Course in Functional Analysis*, 2nd edition, Springer-Verlag, 1990.

[49] J. Conway, *Mathematical Connections: A Capstone Course*, American Mathematical Society, 2010.

[50] J. Conway, *A Course in Abstract Analysis*, American Mathematical Society, 2012.

[51] J. Conway, *A Course in Point Set Topology*, Springer, 2014.

[52] J. Conway, *A First Course in Analysis*, Cambridge University Press, 2018.

[53] A. Cuoco, *Visualizing the p-adic integers*, American Mathematical Monthly **98** (1991), 355–364.

[54] H. Derksen, *The fundamental theorem of algebra and linear algebra*, American Mathemtical Monthly **110** (2003), 620–623.

[55] R. Dressler and K. Stromberg, *The Tonelli integral*, American Mathematical Monthly **81** (1974), 67–68.

[56] P. Duren, *Theory of H^p Spaces*, Academic Press, 1970.

[57] P. Duren and A. Schuster, *Bergman Spaces*, American Mathematical Society, 2004.

[58] Y. Eidelman, V. Milman, and A. Tsolomitis, *Functional Analysis: An Introduction*, American Mathematical Society, 2004.

[59] M. Eisen and H. Gindler, *On the spectrum of a Laurent form*, American Mathematical Monthly **73** (1966), no. 4, part II, 79–90.

[60] A. Escassut, *Ultrametric Banach Algebras*, World Scientific, 2003.

[61] A. Escassut, *Banach Algebras of Ultrametric Functions*, World Scientific, 2022.

[62] H. Federer, *Geometric Measure Theory*, Springer-Verlag, 1969.

[63] F. Ficken, *Some uses of linear spaces in analysis*, American Mathematical Monthly **66** (1959), 259–275.

[64] G. Folland, *Real Analysis: Modern Techniques and their Applications*, 2nd edition, Wiley, 1999.

[65] G. Folland, *A Guide to Advanced Real Analysis*, Mathematical Association of America, 2009.

[66] G. Folland, *A Course in Abstract Harmonic Analysis*, 2nd edition, CRC Press, 2016.

[67] W. Fulks, *Bounded linear functionals on L_q*, American Mathematical Monthly **72** (1965), 750–753.

[68] T. Gamelin, *Uniform Algebras*, Prentice-Hall, 1969.

[69] T. Gamelin, *Uniform Algebras and Jensen Measures*, Cambridge University Press, 1978.

[70] T. Gamelin and R. Greene, *Introduction to Topology*, 2nd edition, Dover, 1999.

[71] J. Garnett, *Bounded Analytic Functions*, revised first edition, Springer, 2007.

[72] B. Gelbaum, *Banach algebras and their applications*, American Mathematical Monthly **71** (1964), 248–256.

[73] B. Gelbaum, *Modern Real and Complex Analysis*, Wiley, 1995.

[74] L. Gillman and M. Jerison, *Rings of Continuous Functions*, Springer-Verlag, 1976.

[75] H. Gindler, *Extensions of linear transformations*, American Mathematical Monthly **t1** (1964), 525–529.

[76] H. Gindler, *A Spectral Mapping Theorem for Polynomials*, American Mathematical Monthly **72** (1965), 528–530.

[77] C. Goffman, *Real Functions*, Rinehart & Co., 1953.

[78] C. Goffman and G. Pedrick, *First Course in Functional Analysis*, Prentice-Hall, 1965.

[79] P. Górka, *Separability of a metric space is equivalent to the existence of a Borel measure*, American Mathematical Monthly **128** (2021), 84–86.

[80] F. Gouvêa, *p-Adic Numbers: An Introduction*, 2nd edition, Springer-Verlag, 1997. A third edition was published in 2020.

[81] S. Grabiner, *A short proof of Runge's theorem*, American Mathematical Monthly **83** (1976), 807–808.

[82] S. Grabiner, *The Tietze extension theorem and the open mapping theorem*, American Mathematical Monthly **93** (1986), 190-191.

[83] S. Gudder, *Inner product spaces*, American Mathematical Monthly **81** (1974), 29–36. A correction was published in ibid., **82** (1975), 251–252. A second correction, with S. Holland, was published in ibid., **82** (1975), 818.

[84] O. Hájek, *Metric completion simplified*, American Mathematical Monthly **75** (1968), 62-63.

[85] P. Halmos, *What does the spectral theorem say?*, American Mathematical Monthly **70** (1963), 241–247.

[86] P. Halmos, *Finite-Dimensional Hilbert Spaces*, American Mathematical Monthly **77** (1970), 457–464.

[87] P. Halmos, *Finite-Dimensional Vector Spaces*, 2nd edition, Springer-Verlag, 1974.

[88] P. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, 2nd edition, AMS Chelsea, 1998.

[89] I. Halperin, *The spectral theorem*, American Mathematical Monthly **71** (1964), 408–410.

[90] G. Helmberg, *Curiosities concerning weak topology in Hilbert space*, American Mathematical Monthly **113** (2006), 447–452.

[91] E. Hewitt and K. Ross, *Abstract Harmonic Analysis*, Volume I, 2nd edition, Springer-Verlag, 1979.

[92] E. Hewitt and K. Ross, *Abstract Harmonic Analysis*, Volume II, Springer-Verlag, 1970.

[93] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, 1975.

[94] K. Hoffman, *Analysis in Euclidean Space*, Prentice-Hall, 1975.

[95] J. Holly, *Pictures of ultrametric spaces, the p-adic numbers, and valued fields*, American Mathematical Monthly **108** (2001), 721–728.

[96] K. Hoffman, *Banach Spaces of Analytic Functions*, Dover, 1988.

[97] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, 2nd printing, revised, Springer-Verlag, 1978.

[98] R. James, *Bases in Banach spaces*, American Mathematical Monthly **89** (1982), 625-640.

[99] J. Johnson, *An elementary characterization of weak convergence of measures*, American Mathematical Monthly **92** (1985), 136–137.

[100] F. Jones, *Lebesgue Integration on Euclidean Space*, Jones and Bartlett, 1993.

[101] N. Kalton, N. Peck, and J. Roberts, *An F-Space Sampler*, Cambridge University Press, 1984.

[102] I. Kaplansky, *Set Theory and Metric Spaces*, 2nd edition, Chelsea, 1977.

[103] Y. Katznelson, *An Introduction to Harmonic Analysis*, 3rd edition, Cambridge University Press, 2004.

[104] J. Kelley, *General Topology*, Springer-Verlag, 1975.

[105] J. Kelley, I. Namioka, et al., *Linear Topological Spaces*, Springer-Verlag, 1976.

[106] J. Kelley and T. Srinivasan, *Measure and Integral*, Springer-Verlag, 1988.

[107] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.

[108] A. Knapp, *Advanced Real Analysis*, Birkhäuser, 2005.

[109] V. Komornik, *Lectures on Functional Analysis and the Lebesgue Integral*, translated from the 2002 French original, Springer-Verlag, 2016.

[110] P. Koosis, *Introduction to H_p Spaces*, 2nd edition, with two appendices by V. Havin, Cambridge University Press, 1998.

[111] A. Korányi, *Around the finite-dimensional spectral theorem*, American Mathematical Monthly **108** (2001), 120–125.

[112] S. Krantz, *What is several complex variables?*, American Mathematical Monthly **94** (1987), 236–256.

[113] S. Krantz, *Function Theory of Several Complex Variables*, AMS Chelsea, 2001.

[114] S. Krantz, *A Guide to Complex Variables*, Mathematical Association of America, 2008.

[115] S. Krantz, *A Guide to Real Variables*, Mathematical Association of America, 2009.

[116] S. Krantz, *A Guide to Topology*, Mathematical Association of America, 2009.

[117] S. Krantz, *Essentials of Topology with Applications*, CRC Press, 2010.

[118] S. Krantz, *A Guide to Functional Analysis*, Mathematical Association of America, 2013.

[119] S. Krantz, *Real Analysis and Foundations*, 4th edition, CRC Press, 2017.

[120] S. Krantz, *Elementary Introduction to the Lebesgue Integral*, CRC Press, 2018.

[121] S. Krantz and H. Parks, *A Primer of Real-Analytic Functions*, 2nd edition, Birkhäuser, 2002.

[122] S. Krantz and H. Parks, *The Implicit Function Theorem: History, Theory, and Applications*, Birkhäuser / Springer, 2013.

[123] B. Kripke, *One more reason why sequences are not enough*, American Mathematical Monthly **74** (1967), 563–565.

[124] S. Lang, *Algebra*, Addison-Wesley, 1970. A second edition was published by Addison-Wesley in 1984, and a revised third edition was published by Springer-Verlag in 2002.

[125] S. Lang, *Real and Functional Analysis*, 3rd edition, Springer-Verlag, 1993.

[126] S. Lang, *Undergraduate Analysis*, 2nd edition, Springer-Verlag, 1997.

[127] P. Lax, *Functional Analysis*, Wiley, 2002.

[128] P. Lax, *Rethinking the Lebesgue integral*, American Mathematical Monthly **116** (2009), 863–881.

[129] R. Leslie, *The l_2 matrix norm and the spectral radius*, American Mathematical Monthly **97** (1990), 339–341.

[130] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces*, Lecture Notes in Mathematics **338**, Springer-Verlag, 1973.

[131] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces I: Sequence Spaces*, Springer-Verlag, 1977.

[132] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces II: Function Spaces*, Springer-Verlag, 1979.

[133] E. Lorch, *Spectral Theory*, Oxford University Press, 1962.

[134] E. Lorch, *Szeged in 1934*, American Mathematical Monthly **100** (1993), 219–230.

[135] D. Luecking and L. Rubel, *Complex Analysis: A Functional Analsysis Apporach*, Springer-Verlag, 1984.

[136] C. MacDuffee, *The p-Adic Numbers of Hensel*, American Mathematical Monthly **45** (1938), 500–508.

[137] R. Macías and C. Segovia, *Lipschitz functions on spaces of homogeneous type*, Advances in Mathematics **33** (1979), 257–270.

[138] S. Mac Lane and G. Birkhoff, *Algebra*, 3rd edition, Chelsea, 1988.

[139] I. Maddox, *Elements of Functional Analysis*, 2nd edition, Cambridge University Press, 1988.

[140] A. Martin, *On the spectral theorem for compact symmetric operators*, American Mathematical Monthly **72** (1965), 892–894.

[141] W. Martin, *Functions of several complex variables*, American Mathematical Monthly **52** (1945), 17–27.

[142] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press, 1995.

[143] E. McShane, *Partial orderings and Moore–Smith limits*, American Mathematical Monthly **59** (1952), 1–11.

[144] B. Mendelson, *Introduction to Topology*, 3rd edition, Dover, 1990.

[145] A. Monna, *Analyse Non-Archimédienne*, Springer-Verlag, 1970.

[146] A. Monna, *Functional Analysis in Historical Perspective*, Wiley, 1973.

[147] L. Narici and E. Beckenstein, *Strange terrain — non-Archimedean spaces*, American Mathematical Monthly **88** (1981), 667–676.

[148] L. Narici and E. Beckenstein, *Topological Vector Spaces*, 2nd edition, CRC Press, 2011.

[149] L. Narici, E. Beckenstein, and G. Bachman, *Functional Analysis and Valuation Theory*, Dekker, 1971.

[150] A. Nijenhuis, *Strong derivatives and inverse mappings*, American Mathematical Monthly **81** (1974), 969–980; addendum, **83** (1975), 22.

[151] S. Ovchinnikov, *Functional Analysis: An Introductory Course*, Springer, 2018.

[152] D. Promislow, *A First Course in Functional Analysis*, Wiley, 2008.

[153] M. Range, *Complex analysis: a brief tour into higher dimensions*, American Mathematical Monthly **110** (2003), 89–108.

[154] C. Rickart, *General Theory of Banach Algebras*, Van Nostrand, 1960.

[155] C. Rickart, *Natural Function Algebras*, Springer-Verlag, 1979.

[156] N. Roy, *Extreme points of convex sets in infinite-dimensional spaces*, American Mathematical Monthly **94** (1987), 409–422.

[157] W. Rudin, *Function Theory in Polydisks*, Benjamin, 1969.

[158] W. Rudin, *Lectures on the Edge-of-the-Wedge Theorem*, American Mathematical Society, 1971.

[159] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.

[160] W. Rudin, *Real and Complex Analysis*, 2nd edition, McGraw-Hill, 1974. A third edition was published in 1987.

[161] W. Rudin, *Fourier Analysis on Groups*, Wiley, 1990.

[162] W. Rudin, *Functional Analysis*, McGraw-Hill, 1973. A second edition was published in 1991.

[163] W. Rudin, *Function Theory in the Unit Ball of* $\mathbf{C}^n$, Springer-Verlag, 2008.

[164] A. Russell, *A commutative Banach algebra of functions of bounded variation*, American Mathematical Monthly **87** (1980), 39–40.

[165] B. Rynne and M. Youngson, *Linear Functional Analysis*, 2nd edition, Springer-Verlag, 2008.

[166] K. Saxe, *Beginning Functional Analysis*, Springer-Verlag, 2002.

[167] H. Schaefer and M. Wolff, *Topological Vector Spaces*, 2nd edition, Springer-Verlag, 1999.

[168] R. Schatten, *"Closing-up" of sequence spaces*, American Mathematical Monthly **57** (1950), 603–616.

[169] R. Schatten, *A Non-Approximation Theorem*, American Mathematical Monthly **69** (1962), 745–750.

[170] H. Seid, *A corollary to the Gelfand–Mazur theorem*, American Mathematical Monthly **77** (1970), 282–283.

[171] J.-P. Serre, *Linear Representations of Finite Groups*, translated from the 2nd French edition by L. Scott, Springer-Verlag, 1977.

[172] J.-P. Serre, *Complex Semisimple Lie Algebras*, translated from the French by G. Jones, Springer-Verlag, 2001.

[173] J.-P. Serre, *Lie Algebras and Lie Groups*, Correctd fifth printing of the second (1992) edition, Lecture Notes in Mathematics **1500**, Springer-Verlag, 2006.

[174] R. Shakarchi, *Problems and Solutions for* Undergraduate Analysis [126], Springer-Verlag, 1998.

[175] N. Shioji, *Simple proofs of the uniform convexity of L^p and the Riesz representation theorem for L^p*, American Mathematical Monthly **125** (2018), 733–738.

[176] D. Singh, *The spectrum in a Banach algebra*, American Mathematical Monthly **113** (2006), 756–758.

[177] L. Steen, *Highlights in the history of spectral theory*, American Mathematical Monthly **80** (1973), 359–381.

[178] L. Steen and J. Seebach, Jr., *Counterexamples in Topology*, 2nd edition, Dover, 1995.

[179] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970.

[180] E. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.

[181] E. Stein and R. Shakarchi, *Functional Analysis: Introduction to Further Topics in Analysis*, Princeton University Press, 2011.

[182] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[183] E. Stout, *The Theory of Uniform Algebras*, Bogden & Quigley, 1971.

[184] K. Stromberg, *An Introduction to Classical Real Analysis*, AMS Chelsea, 1981.

[185] M. Taibleson, *Fourier Analysis on Local Fields*, Princeton University Press, 1975.

[186] A. Taylor, *New proofs of some theorems of Hardy by Banach spaces methods*, Mathematics Magazine **23** (1950), 115–124.

[187] A. Taylor, *Notes in the history of the uses of analyticity in operator theory*, American Mathematical Monthly **78** (1971), 331–342; addendum, 1104.

[188] A. Taylor, *General Theory of Functions and Integration*, reprint of the 1966 2nd edition, Dover, 1985.

[189] A. Taylor and D. Lay, *Introduction to Functional Analysis*, 2nd edition, Wiley, 1980.

[190] A. Thompson, *Minkowski Geometry*, Cambridge University Press, 1996.

[191] F. Trèves, *Topological Vector Spaces, Distributions and Kernels*, Dover, 2005.

[192] J. Väisälä, *A proof of the Mazur–Ulam theorem*, American Mathematical Monthly **110** (2003), 633–635.

[193] A. van Rooij, *Non-Archimedean Functional Analysis*, Dekker, 1978.

[194] J. Van Schaftingen, *A direct proof of the existence of eigenvalues and eigenvectors by Weierstrass's theorem*, American Mathematical Monthly **120** (2013), 741–746.

[195] J. Vaughan, *Examples of ultrametrics*, American Mathematical Monthly **82** (1975), 749–752.

[196] G. Weiss, *Complex methods in harmonic analysis*, American Mathematical Monthly **77** (1970), 465–474.

[197] R. Whitley, *The spectral theorem for a normal operator*, American Mathematical Monthly **75** (1968), 856–861.

[198] P. Wojtaszczyk, *Banach Spaces for Analysts*, Cambridge University Press, 1991.

[199] A. Zaanen, Some examples in weak sequential convergence, American Mathematical Monthly **69** (1962), 85–93.

[200] A. Zygmund, *Trigonometric Series*, Volumes I, II, 3rd edition, Cambridge University Press, 2002.

Index