An introduction to some aspects of functional analysis, 2: Bounded linear operators

Stephen Semmes
Rice University

Abstract

These notes are largely concerned with the strong and weak operator topologies on spaces of bounded linear operators, especially on Hilbert spaces, and related matters.

Contents

Part I
Basic notions

1 Norms and seminorms

Let V be a vector space over the real numbers $\mathbf{R}$ or the complex numbers $\mathbf{C}$. A nonnegative real-valued function N on V is a *seminorm* if

$$N(t\,v) = |t|\,N(v) \tag{1.1}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $v \in V$, and

$$N(v + w) \le N(v) + N(w) \tag{1.2}$$

for every $v, w \in V$. Here $|t|$ denotes the absolute value of $t \in \mathbf{R}$ or the modulus of $t \in \mathbf{C}$. If also $N(v) > 0$ when $v \neq 0$, then N is a *norm* on V.

If N is a norm on V, then

$$d(v, w) = N(v - w) \tag{1.3}$$

is a metric on V. A collection $\mathcal{N}$ of seminorms on V will be called *nice* if for every $v \in V$ with $v \neq 0$ there is an $N \in \mathcal{N}$ such that $N(v) > 0$. In this case, the topology on V associated to $\mathcal{N}$ is defined by saying that $U \subseteq V$ is an open set if for every $u \in U$ there are finitely many seminorms $N_1, \ldots, N_l \in \mathcal{N}$ and finitely many positive real numbers $r_1, \ldots, r_l$ such that

$$\{v \in V : N_j(u - v) < r_j \text{ for } j = 1, \ldots, l\} \subseteq U. \tag{1.4}$$

It is easy to see that V is Hausdorff with respect to this topology, and that the vector space operations of addition and scalar multiplication are continuous.

2 ℓ^p spaces

Let p be a real number, $p \ge 1$, and let ℓ^p be the space of sequences $a = \{a_j\}_{j=1}^\infty$ of real or complex numbers such that the infinite series

$$\sum_{j=1}^\infty |a_j|^p \tag{2.1}$$

converges. It is well known that this is a vector space with respect to termwise addition and scalar multiplication, and that

$$\|a\|_p = \Big(\sum_{j=1}^{\infty} |a_j|^p\Big)^{1/p} \tag{2.2}$$

defines a norm on this space. Similarly, the space ℓ^∞ of bounded sequences of real or complex numbers is a vector space, and

$$\|a\|_\infty = \sup\{|a_j| : j \ge 1\} \tag{2.3}$$

is a norm on ℓ^∞.

If $a_j = 0$ for all but finitely many j, then a is contained in ℓ^p for each p. Moreover, the space of these sequences is a dense linear subspace of ℓ^p when $p < \infty$. The closure of this space of sequences in ℓ^∞ is the space c_0 of sequences that converge to 0.

These spaces may also be denoted $\ell^p(\mathbf{Z}_+)$, $c_0(\mathbf{Z}_+)$, where $\mathbf{Z}_+$ is the set of positive integers. There are versions of these spaces for doubly-infinite sequences $a = \{a_j\}_{j=-\infty}^{\infty}$ as well, denoted $\ell^p(\mathbf{Z})$, $c_0(\mathbf{Z})$, where $\mathbf{Z}$ is the set of all integers. More precisely, if $p < \infty$, then $\ell^p(\mathbf{Z})$ is the space of doubly-infinite sequences such that

$$\sum_{j=-\infty}^{\infty} |a_j|^p \tag{2.4}$$

converges, which is the same as saying that

$$\sum_{j=1}^{\infty} |a_j|^p, \ \sum_{j=0}^{\infty} |a_{-j}|^p \tag{2.5}$$

both converge, equipped with the norm

$$\|a\|_p = \Big(\sum_{j=-\infty}^{\infty} |a_j|^p\Big)^{1/p}. \tag{2.6}$$

The space $\ell^\infty(\mathbf{Z})$ consists of bounded doubly-infinite sequences, with the norm

$$\|a\|_\infty = \sup\{|a_j| : j \in \mathbf{Z}\}. \tag{2.7}$$

As before, the space of doubly-infinite sequences a such that $a_j \to 0$ as $j \to \pm\infty$ defines a closed linear subspace $c_0(\mathbf{Z})$ of $\ell^\infty(\mathbf{Z})$.

3 Bounded linear mappings

Let V, W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$, $\|w\|_W$, respectively. A linear mapping $T : V \to W$ is said to be *bounded* if there is an $L \ge 0$ such that

$$\|T(v)\|_W \le L\, \|v\|_V \tag{3.1}$$

for every $v \in V$. It is easy to see that bounded linear mappings are continuous and even uniformly continuous with respect to the metrics on V, W associated to their norms. Conversely, a linear mapping is bounded if it is continuous at 0. The operator norm of a bounded linear mapping $T : V \to W$ is defined by

$$\|T\|_{op} = \sup\{\|T(v)\|_W : v \in V,\ \|v\|_V \le 1\}. \tag{3.2}$$

Equivalently, $L = \|T\|_{op}$ is the smallest nonnegative real number that satisfies the previous condition. The space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W is a vector space with respect to pointwise addition and scalar multiplication, and it is easy to see that $\|T\|_{op}$ defines a norm on this space.

Suppose that V_1, V_2, V_3 are vector spaces, all real or all complex, and equipped with norms $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_3$. If $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ are bounded linear mappings, then their composition $T_2 \circ T_1 : V_1 \to V_3$ is also a bounded linear mapping, with

$$\|T_2 \circ T_1\|_{op,13} \le \|T_1\|_{op,12}\, \|T_2\|_{op,23}, \tag{3.3}$$

where the subscripts indicate the spaces and norms involved. In particular, the space $\mathcal{BL}(V) = \mathcal{BL}(V, V)$ of bounded linear operators on a real or complex vector space V with a norm $\|v\|_V$ is an algebra with respect to composition. Of course, the identity operator I on V has operator norm 1.

4 Dual spaces

Let V be a real or complex vector space, equipped with a norm $\|v\|_V$. A *bounded linear functional* on V is a bounded linear mapping from V into $\mathbf{R}$ or $\mathbf{C}$, using the standard absolute value or modulus as the norm on the latter. The vector space of bounded linear functionals on V is the same as $\mathcal{BL}(V, \mathbf{R})$ or $\mathcal{BL}(V, \mathbf{C})$, and will be denoted V'. The *dual norm* of $\lambda \in V'$ is the same as its operator norm as a bounded linear mapping from V into $\mathbf{R}$ or $\mathbf{C}$, and may be denoted $\|\lambda\|_{V'}$.

Suppose that $1 \le p, q \le \infty$ are conjugate exponents, in the sense that

$$\frac{1}{p} + \frac{1}{q} = 1, \tag{4.1}$$

where $1/\infty = 0$. If $a = \{a_j\}_{j=1}^\infty$, $b = \{b_j\}_{j=1}^\infty$ are sequences of real or complex numbers, then let ab the sequence $\{a_j\, b_j\}_{j=1}^\infty$ of products of the terms of a, b. *Hölder's inequality* implies that $ab \in \ell^1$ when $a \in \ell^p$, $b \in \ell^q$, and that

$$\|ab\|_1 \le \|a\|_p\, \|b\|_q. \tag{4.2}$$

The analogous statements for doubly-infinite sequences also hold.

It follows that

$$\lambda_b(a) = \sum_{j=1}^\infty a_j\, b_j \tag{4.3}$$

defines a bounded linear functional on ℓ^p for each $b \in \ell^q$. Furthermore,

$$\|\lambda_b\|_{(\ell^p)'} = \|b\|_q, \tag{4.4}$$

because the dual norm of λ_b on ℓ^p is less than or equal to $\|b\|_q$ by Hölder's inequality, while the opposite inequality can be derived from specific choices of $a \in \ell^p$ for a given b. It is well known that every bounded linear functional on ℓ^p is of this form when $p < \infty$. If $b \in \ell^1$, then the restriction of λ_b to c_0 defines a bounded linear functional on c_0 with respect to the ℓ^∞ norm, the dual norm of this linear functional is also equal to $\|b\|_1$, and every bounded linear functional on c_0 is of this form. Again, there are analogous statements for doubly-infinite sequences.

5 Shift operators

If $a = \{a_j\}_{j=-\infty}^{\infty}$ is a doubly-infinite sequence of real or complex numbers, then let $b = T(a)$ be defined by

$$b_j = a_{j-1}. \tag{5.1}$$

This is the (forward) shift operator on the vector space of all such sequences, which is a one-to-one linear mapping of this vector space onto itself, whose inverse is known as the backward shift operator. Observe that T sends $\ell^p(\mathbf{Z})$, $1 \le p \le \infty$, and $c_0(\mathbf{Z})$ onto themselves, and satisfies

$$\|T(a)\|_p = \|a\|_p \tag{5.2}$$

for every $a \in \ell^p(\mathbf{Z})$. Thus T is a bounded linear operator on $\ell^p(\mathbf{Z})$ with operator norm 1, and its inverse has the same property.

Now let $a = \{a_j\}_{j=1}^{\infty}$ be an ordinary sequence of real or complex numbers, and let $b = T(a)$ be defined by

$$b_j = a_{j-1} \text{ when } j \ge 1,\ b_1 = 0. \tag{5.3}$$

This is the forward shift operator on the vector space of all such sequences, which is a one-to-one linear mapping of this vector space into itself. Clearly T sends $\ell^p(\mathbf{Z}_+)$, $1 \le p \le \infty$, and $c_0(\mathbf{Z}_+)$ into themselves, and satisfies (5.2) for every $a \in \ell^p(\mathbf{Z}_+)$, so that T is a bounded linear operator on $\ell^p(\mathbf{Z}_+)$ with operator norm 1. If we identify ordinary sequences $a = \{a_j\}_{j=1}^{\infty}$ with doubly-infinite sequences $\{a_j\}_{j=-\infty}^{\infty}$ such that $a_j = 0$ when $j \le 0$, then this shift operator is the same as the previous one restricted to this linear subspace.

The backward shift operator R on the space of ordinary sequences $a = \{a_j\}_{j=1}^{\infty}$ is defined by $c = R(a)$, where

$$c_j = a_{j+1}. \tag{5.4}$$

This is the same identifying a with a doubly-infinite sequence as in the previous paragraph, applying the earlier backward shift operator on doubly-infinite sequences, and projecting the result to an ordinary sequence by restricting the indices to positive integers. In particular, R sends $\ell^p(\mathbf{Z}_+)$, $1 \le p \le \infty$, and $c_0(\mathbf{Z}_+)$ onto themselves, and satisfies

$$\|R(a)\|_p \le \|a\|_p \tag{5.5}$$

when $a \in \ell^p(\mathbf{Z}_+)$. This means that R has operator norm less than or equal to 1 on $\ell^p(\mathbf{Z}_+)$ for each p, and one can check that the operator norm is actually equal to 1. Similarly, R has operator norm 1 on $c_0(\mathbf{Z})$.

6 Arbitrary sequences

Let V be the vector space of doubly-infinite sequences $a = \{a_j\}_{j=-\infty}^{\infty}$ of real or complex numbers, with respect to termwise addition and scalar multiplication. For each $j \in \mathbf{Z}$,

$$N_j(a) = |a_j| \tag{6.1}$$

defines a seminorm on V. The topology on V associated to this family of seminorms is the same as the product topology, where V is identified with the Cartesian product of infinitely many copies of $\mathbf{R}$ or $\mathbf{C}$. The shift operator on V described in the previous section is a homeomorphism with respect to this topology. The vector space W of ordinary sequences $a = \{a_j\}_{j=1}^{\infty}$ of real or complex numbers can be identified with the linear subspace of V consisting of doubly-infinite sequences $\{a_j\}_{j=-\infty}^{\infty}$ such that $a_j = 0$ when $j \le 0$. On W, we may as well consider only the seminorms N_j corresponding to $j \in \mathbf{Z}_+$. The topology on W determined by these seminorms is the same as the topology induced by the one on V. This is also the same as the product topology on W as the Cartesian product of infinitely many copies of $\mathbf{R}$ or $\mathbf{C}$. Note that the linear subspace of V identified with W in this way is closed. The forward shift operator is a linear homeomorphism of W onto a proper closed linear subspace, while the backward shift operator determines a continuous linear mapping of W onto itself.

7 Metrizability

Let V be a real or complex vector space, and let $\mathcal{N}$ be a nice collection of seminorms on V. If $N_1, \ldots, N_l$ are finitely many seminorms on V, then it is easy to see that

$$N(v) = \max(N_1(v), \ldots, N_l(v)) \tag{7.1}$$

is also a seminorm on V. If $\mathcal{N}$ has only finitely many elements, then it follows that their maximum is a norm on V that determines the same topology. If $\mathcal{N}$ is countably infinite, and $N_1, N_2, \ldots$ is an enumeration of its elements, then one can check that

$$\rho(v, w) = \max_{j \ge 1} \big(\min(N_j(v - w), 1/j) \big) \tag{7.2}$$

is a metric on V that determines the same topology as $\mathcal{N}$. Conversely, if $\mathcal{N}$ is a nice collection of seminorms on V for which there is a countable local base for the corresponding topology at 0, then the same topology is determined by only finitely or countably many elements of $\mathcal{N}$. It also happens sometimes that a countable subcollection of $\mathcal{N}$ suffices to determine the same topology on some interesting subsets of V. Hence the induced topology on these subsets can also be described by a metric.

8 Completeness

Remember that a metric space M is said to be *complete* if every sequence of elements of M converges to an element of M. For example, the real line and the complex plane are complete with respect to their standard metrics. A real or complex vector space V with a norm $\|v\|$ is said to be a *Banach space* if it is complete with respect to the metric associated to the norm. It is well known that the ℓ^p spaces described in Section 2 are complete. A closed subset Y of a complete metric space M is automatically complete, with respect to the restriction of the metric to Y, and thus c_0 is also a Banach space with respect to the ℓ^∞ norm.

Let V be a real or complex vector space equipped with a nice collection $\mathcal{N}$ of finitely or countably many seminorms. As in the previous section, there is a metric $\rho(v, w)$ on V that determines the same topology on V as $\mathcal{N}$ and is translation-invariant in the sense that

$$\rho(v + u, w + u) = \rho(v, w) \tag{8.1}$$

for every $u, v, w \in V$. It is easy to check that any other translation-invariant metric on V that determines the same topology also determines the same class of Cauchy sequences, which can be characterized directly in terms of differences of terms in the sequence and the topology of V at 0 as well. In this case, V is said to be a *Fréchet space* if it is complete. Banach spaces are automatically Fréchet spaces, and the spaces of arbitrary sequences of real or complex numbers give examples of Frćhet spaces for which the topology is not determined by a single norm.

9 Continuous extensions

Let M and N be metric spaces, and suppose that N is complete. If E is a dense set in M and $f : E \to N$ is uniformly continuous, then it is well known that there is a unique extension of f to a uniformly continuous mapping from M into N. For if x is any element of M, then there is a sequence $\{x_j\}_{j=1}^\infty$ of elements of E that converges to x. Thus $\{x_j\}_{j=1}^\infty$ is a Cauchy sequence as a sequence of elements of E, and uniform continuity implies that $\{f(x_j)\}_{j=1}^\infty$ is a Cauchy sequence in N that converges by completeness. If $\{y_j\}_{j=1}^\infty$ is another sequence of elements of E that converges to x, then $\{f(y_j)\}_{j=1}^\infty$ converges to the same element of N as $\{f(x_j)\}_{j=1}^{\infty}$. The extension of f to M at x can be defined as this limit, which does not depend on the choice of the sequence of elements of E that converges to x. Uniform continuity of this extension is easily inherited from uniform continuity of f. Uniqueness holds because two continuous functions on M that agree on a dense set are the same.

Now let V, W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$, $\|w\|_W$. If U is a dense linear subspace of V, T is a bounded linear mapping from U into W, and W is complete, then there is a unique extension of T to a bounded linear mapping from V into W. It follows from the statement in the previous paragraph that T has a unique extension to a uniformly continuous mapping from V into V, since bounded linear mappings are automatically uniformly continuous. One can also check that the extension is linear, because T is linear, and that the operator norm of the extension is the same as the operator norm of T.

10 Uniform boundedness

Let V, W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$, $\|w\|_W$. Suppose that $\mathcal{T}$ is a collection of bounded linear mappings from V into W which is uniformly bounded pointwise on V, in the sense that

$$\|T(v)\|_W,\ T \in \mathcal{T}, \tag{10.1}$$

are bounded for every $v \in V$. If V is complete, then the Banach–Steinhaus theorem implies that $\mathcal{T}$ is uniformly bounded, in the sense that the elements of $\mathcal{T}$ have uniformly bounded operator norm. To see this, consider

$$A_n = \{v \in V : \|T(v)\|_W \le n \text{ for every } T \in \mathcal{T}\} \tag{10.2}$$

for each positive integer n. The hypothesis of pointwise boundedness implies that

$$\bigcup_{n=1}^{\infty} A_n = V. \tag{10.3}$$

Each A_n is also a closed set in V, because every $T \in \mathcal{T}$ is bounded and hence continuous. Since V is complete, the Baire category theorem implies that A_n contains a nonempty open set for some n. Using linearity, one can show that the elements of $\mathcal{T}$ are uniformly bounded on the unit ball in V, which means exactly that they have uniformly bounded operator norms.

11 Bounded linear mappings, 2

Let V, W be vector spaces, both real or both complex, and equipped with norms. If W is complete, then the space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W is also complete, with respect to the operator norm. For suppose that $\{T_j\}_{j=1}^{\infty}$ is a Cauchy sequence of bounded linear mappings from V into W

in the operator norm. This implies that $\{T_j(v)\}_{j=1}^\infty$ is a Cauchy sequence of elements of W for each $v \in V$, which converges to an element $T(v)$ of W by completeness. Thus $\{T_j\}_{j=1}^\infty$ converges pointwise on V to $T : V \to W$, and it is easy to see that T is linear, since T_j is linear for each j. Using the fact that $\{T_j\}_{j=1}^\infty$ is a Cauchy sequence in the operator norm, one can show that T is bounded, and that $\{T_j\}_{j=1}^\infty$ converges to T in the operator norm, as desired. In particular, the dual space V' is complete, with respect to the dual norm.

12 Separability

Remember that a metric space is said to be *separable* if it contains a dense set with only finitely or countably many elements. For example, the set of rational numbers is a countable dense set in the real line, which is therefore separable. The complex plane is separable as well, since the set of complex numbers with rational real and imaginary parts is countable and dense. Let V be a real or complex vector space equipped with a norm. In order for V to be separable with respect to the metric associated to the norm, it suffices that there be a set $E \subseteq V$ with only finitely or countably many elements whose linear span $\operatorname{span} E$ is dense in V. By definition, $\operatorname{span} E$ consists of all finite linear combinations of elements of E, and is the smallest linear subspace of V that contains E. It is easy to see that the set of finite linear combinations of elements of E with rational coefficients is dense in $\operatorname{span} E$ in the real case, and that the set of finite linear combinations of elements of E whose coefficients have rational real and imaginary parts is dense in $\operatorname{span} E$ in the complex case. If E has only finitely or countable many elements, then these subsets of $\operatorname{span} E$ have only finitely or countably many elements too, and are dense in V when $\operatorname{span} E$ is dense in V. It follows from this criterion that finite-dimensional spaces are automatically separable. The ℓ^p spaces in Section 2 are separable when $p < \infty$, and c_0 is separable with respect to the ℓ^∞ norm.

13 Inner product spaces

Let V be a real or complex vector space. An *inner product* on V is a function $\langle v, w \rangle$ defined for $v, w \in V$ and with values in the real or complex numbers, as appropriate, with the following three properties. First, $\langle v, w \rangle$ is linear in v for each $w \in V$. Second,

$$\langle w, v \rangle = \langle v, w \rangle \tag{13.1}$$

for every $v, w \in V$ in the real case, and

$$\langle w, v \rangle = \overline{\langle v, w \rangle} \tag{13.2}$$

for every $v, w \in V$ in the complex case. Here $\overline{z}$ denotes the complex conjugate of a complex number z. Note that this symmetry condition implies that $\langle v, w \rangle$ is also linear in w in the real case, and is conjugate-linear in the complex case.

This also implies that $\langle v, v \rangle$ is automatically a real number for every $v \in V$ in the complex case. The third condition asks that

$$\langle v, v \rangle > 0 \tag{13.3}$$

for every $v \in V$ such that $v \neq 0$, in both the real and complex cases.

Put

$$\|v\| = \langle v, v \rangle^{1/2}. \tag{13.4}$$

The *Cauchy–Schwarz inequality*

$$|\langle v, w \rangle| \le \|v\| \, \|w\| \tag{13.5}$$

can be derived from the fact that

$$\langle v + t\, w, v + t\, w \rangle \ge 0 \tag{13.6}$$

for every $v, w \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. By expanding $\|v + w\|^2$ into a sum of inner products and applying the Cauchy–Schwarz inequality, one can show that $\|v\|$ satisfies the triangle inequality, and hence defines a norm on V. The *parallelogram law* states that

$$\|v + w\|^2 + \|v - w\|^2 = 2\,\|v\|^2 + 2\,\|w\|^2 \tag{13.7}$$

for every $v, w \in V$, which follows easily by expanding the terms on the left side into sums of inner products.

The standard inner products on $\mathbf{R}^n$, $\mathbf{C}^n$ are given by

$$\langle v, w \rangle = \sum_{j=1}^{n} v_j \, w_j \tag{13.8}$$

in the real case and

$$\langle v, w \rangle = \sum_{j=1}^{n} v_j \, \overline{w_j} \tag{13.9}$$

in the complex case, where $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$. The associated norm

$$\|v\| = \Big(\sum_{j=1}^{n} |v_j|^2 \Big)^{1/2} \tag{13.10}$$

is the same as the standard Euclidean norm. It is well known that a finite-dimensional inner product space V of dimension n is equivalent to $\mathbf{R}^n$ or $\mathbf{C}^n$ with the standard inner product, using the Gram–Schmidt process to convert an ordinary basis for V into an orthonormal basis.

Similarly, the standard inner products on the spaces of sequences of real or complex numbers in ℓ^2 are defined by

$$\langle a, b \rangle = \sum_{j=1}^{\infty} a_j \, b_j \tag{13.11}$$

in the real case and

$$\langle a, b \rangle = \sum_{j=1}^{\infty} a_j \, \overline{b_j} \tag{13.12}$$

in the complex case, where $a = \{a_j\}_{j=1}^{\infty}$, $b = \{b_j\}_{j=1}^{\infty}$. More precisely, the absolute convergence of these series can be derived from the Cauchy–Schwarz inequality for the standard inner products on $\mathbf{R}^n$, $\mathbf{C}^n$, or from the $p = q = 2$ case of Hölder's inequality, which is basically the same. The associated norms are equal to $\|a\|_2$ in Section 2. The standard inner products on the spaces of doubly-infinite sequences of real or complex numbers in $\ell^2(\mathbf{Z})$ are defined analogously, and the associated norms are also equal to $\|a\|_2$ in Section 2.

14 Hilbert spaces

A real or complex inner product space $(V, \langle v, w \rangle)$ is said to be a *Hilbert space* if it is complete with respect to the associated norm $\|v\|$. Finite-dimensional inner product spaces are complete, as are the ℓ^2 spaces in Section 2.

Suppose that V is a Hilbert space, and that $v_1, v_2, \ldots$ is an orthonormal sequence of vectors in V. This means that

$$\langle v_j, v_l \rangle = 0 \tag{14.1}$$

when $j \neq l$, and $\|v_j\| = 1$ for every j. If $a = \{a_j\}_{j=1}^{\infty}$ is a square-summable sequence of real or complex numbers, as appropriate, then

$$\sum_{j=1}^{\infty} a_j \, v_j \tag{14.2}$$

converges in V. To see this, it suffices to show that the corresponding sequence of partial sums $\sum_{j=1}^{n} a_j \, v_j$ is a Cauchy sequence, which then converges by completeness. Because of orthonormality,

$$\Big\| \sum_{j=k}^{n} a_j \, v_j \Big\|^2 = \sum_{j=k}^{n} |a_j|^2 \tag{14.3}$$

when $n \geq k \geq 1$, and this tends to 0 as $k \to \infty$ since $a \in \ell^2$, as desired.

Suppose now that V is infinite-dimensional and separable, so that there is a sequence $v_1, v_1, \ldots$ of elements of V whose linear span is dense in V. Using the Gram–Schmidt process, we may ask also that the v_j's be orthonormal. In this case,

$$a \mapsto \sum_{j=1}^{\infty} a_j \, v_j \tag{14.4}$$

defines a linear mapping from ℓ^2 into V. It is not difficult to show that this mapping is one-to-one, and that the standard inner product on ℓ^2 corresponds exactly to the given inner product on V under this mapping. One can also show that this mapping sends ℓ^2 onto V, because the linear span of the v_j's is dense in V.

15 Orthogonal projections

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let $\|v\|$ be the associated norm. Suppose that A is a nonempty closed convex set in V, and let v be any element of V. If

$$r = \inf\{\|v - w\| : w \in A\}, \tag{15.1}$$

then there is a $w \in A$ such that $\|v - w\| = r$. Of course, there is a sequence $\{w_j\}_{j=1}^{\infty}$ of elements of A such that

$$\lim_{j \to \infty} \|v - w_j\| = r. \tag{15.2}$$

It suffices to show that this sequence converges, because its limit would then have the required properties. By completeness, it is enough to show that $\{w_j\}_{j=1}^{\infty}$ is a Cauchy sequence. This can be done using the parallelogram law, and the fact that

$$\frac{w_j + w_l}{2} \in A \tag{15.3}$$

for each $j, l \geq 1$, by convexity.

Suppose now that A is a closed linear subspace of V. If v, w are as in the previous paragraph, and $u \in A$, then $w - u \in A$ too, and hence

$$\begin{aligned} \|v - w\|^2 = r^2 &\leq \|v - w + u\|^2 \\ &= \|v - w\|^2 + \langle v - w, u\rangle + \langle u, v - w\rangle + \|u\|^2. \end{aligned} \tag{15.4}$$

This can be used to show that

$$\langle v - w, u\rangle = 0. \tag{15.5}$$

Conversely, if $v \in V$, $w \in A$, and (15.5) holds for each $u \in A$, then

$$\|v - w + u\|^2 = \|v - w\|^2 + \|u\|^2 \tag{15.6}$$

for every $u \in A$, which implies that

$$\|v - w\| \leq \|v - w + u\|, \tag{15.7}$$

and $\|v - w\| = r$. Applying (15.6) to $u = w$, we also get that

$$\|v\|^2 = \|v - w\|^2 + \|w\|^2. \tag{15.8}$$

Suppose that $v \in V$, $w, w' \in A$, and w, w' both satisfy (15.5) for each $u \in A$. This implies that $w - w' \in A$,

$$\langle w - w', w - w'\rangle = \langle w - v, w - w'\rangle + \langle v - w', w - w'\rangle = 0, \tag{15.9}$$

and hence $w = w'$. Thus $w \in A$ is uniquely determined by the condition that (15.5) hold for every $u \in A$, and the *orthogonal projection* P_A of V onto A is defined by $P_A(v) = w$. It is easy to see from this characterization that P_A is a linear mapping on V, and that

$$\|P_A(v)\| \leq \|v\|. \tag{15.10}$$

Of course, $P_A(v) = v$ when $v \in A$, so that P_A has operator norm equal to 1, unless $A = \{0\}$, in which case $P_A = 0$.

16 Orthogonal complements

Let $(V, \langle \cdot, \cdot \rangle)$ be a real or complex Hilbert space, and let W be a linear subspace of V. The *orthogonal complement* $W^\perp$ of W is defined by

$$W^\perp = \{v \in V : \langle v, w \rangle = 0 \text{ for every } w \in W\}. \tag{16.1}$$

It is easy to see that this is a closed linear subspace of V, and that

$$W \cap W^\perp = \{0\}. \tag{16.2}$$

If W is a closed linear subspace of V, then it follows from the previous section that every element of V can be expressed in a unique way as a sum of elements of W and $W^\perp$, and that $W^\perp$ is the same as the kernel of the orthogonal projection of V onto W. Note that the orthogonal complement of W is equal to the orthogonal complement of its closure in V, and so we may as well restrict our attention to closed linear subspaces of V.

Clearly

$$W \subseteq (W^\perp)^\perp. \tag{16.3}$$

If W is a closed linear subspace of V, then

$$W = (W^\perp)^\perp. \tag{16.4}$$

For suppose that $v \in (W^\perp)^\perp$, and let us check that $v \in W$. As in the previous section, there is a $w \in W$ such that $v - w \in W^\perp$. However, $v - w \in (W^\perp)^\perp$, since $v \in (W^\perp)^\perp$ and $w \in W \subseteq (W^\perp)^\perp$, and hence $v - w = 0$, as desired.

17 Dual spaces, 2

Let $(V, \langle \cdot, \cdot \rangle)$ be a real or complex inner product space. For each $w \in V$,

$$\lambda_w(v) = \langle v, w \rangle \tag{17.1}$$

defines a linear functional on V. The Cauchy–Schwarz inequality implies that λ_w is a bounded linear functional on V, with dual norm less than or equal to the norm of w. The dual norm of λ_w is actually equal to the norm of w, because $\lambda_w(w) = \|w\|^2$. Suppose now that λ is an arbitrary bounded linear functional on V, and that V is a Hilbert space. The kernel Z of λ is a closed linear subspace of V, since λ is continuous. Of course, $Z = V$ if and only if $\lambda = 0$. Otherwise, Z has codimension 1 in V, and one can check that $\lambda = \lambda_w$ for some $w \in Z^\perp$.

18 The Hahn–Banach theorem

Let V be a real or complex vector space equipped with a norm $\|\cdot\|$. The *Hahn–Banach theorem* states that a bounded linear functional λ on a linear subspace W of V can be extended to a bounded linear functional on V with the same norm. More precisely, the norm of λ on W is defined using the restriction of the norm $\|\cdot\|$ to W. In the proof, one first extends λ to a linear functional on a linear subspace of V spanned by W and a single vector, while keeping the norm of the linear functional constant. If W has finite codimension in V, then it suffices to do this finitely many times. If V is separable, then a sequence of these extensions can be used to get an extension of λ to a dense linear subspace of V, and then to V itself, as in Section 9. If V is a Hilbert space, then the Hahn–Banach theorem can be derived from the characterization of bounded linear functions in the previous section.

As a corollary, for each $v \in V$ there is a bounded linear functional on V with norm 1 such that

$$\lambda(v) = \|v\|. \tag{18.1}$$

Indeed, one can start by defining λ on the subspace W of V spanned by v, and then use the Hahn-Banach theorem to extend λ to all of V. This corollary can be verified directly for the ℓ^p spaces in Section 2, as well as any inner product space.

Note that every $v \in V$ determines a bounded linear functional L_v on the dual space V' of bounded linear functionals on V, given by

$$L_v(\lambda) = \lambda(v) \tag{18.2}$$

for each $\lambda \in V'$. It is easy to see that the dual norm of L_v on V' is less than or equal to the norm of v, just by the definition of the dual norm. The corollary of the Hahn–Banach theorem mentioned in the previous paragraph implies that the dual norm of L_v is equal to the norm of v. Thus $v \mapsto L_v$ is an isometric linear embedding of V into the dual V'' of V'.

A Banach space V is said to be *reflexive* if every bounded linear functional on V' is of the form L_v for some $v \in V$. Hilbert spaces are reflexive, by the characterization of their bounded linear functionals in the previous section. If $1 < p < \infty$, then the ℓ^p spaces in Section 2 are reflexive, by the description of their duals in Section 4.

19 The weak topology

Let V be a real or complex vector space equipped with a norm. For each bounded linear functional λ on V,

$$N_\lambda(v) = |\lambda(v)| \tag{19.1}$$

defines a seminorm on V. The collection of all of these seminorms N_λ, $\lambda \in V'$, is a nice collection of seminorms on V in the sense of Section 1, because of the Hahn–Banach theorem. The topology on V that they determine is known as the *weak topology*. It is easy to see that every open set in V in the weak topology is also an open set in the topology determined by the norm, since bounded linear functionals are continuous. The converse holds only when V has finite dimension. For example, an open ball in V cannot be an open set in the weak topology when V is infinite-dimensional, because it does not contain an affine subspace of finite codimension.

Similarly, every closed set in V with respect to the weak topology is a closed set in the topology determined by the norm. The Hahn–Banach theorem implies that the converse holds for closed linear subspaces of V. For if W is a closed linear subspace of V with respect to the norm and $v \in V$ is not in W, then one can show that there is a bounded linear functional λ on the linear span of W and v such that $\lambda(w) = 0$ for every $w \in W$ and $\lambda(v) \neq 0$. The Hahn–Banach theorem implies that λ has an extension to a bounded linear functional on V, and this implies that there is a neighborhood of v in the weak topology on V that is disjoint from W.

Suppose that Λ is a collection of bounded linear functionals on V whose linear span is dense in V' with respect to the dual norm. In this case, N_λ, $\lambda \in \Lambda$, is also a nice collection of seminorms on V. If $E \subseteq V$ is bounded, which means that E is contained in a ball with respect to the norm, then the topology on E induced by the weak topology on V is the same as the topology induced by the seminorms N_λ, $\lambda \in \Lambda$. If V' is separable with respect to the dual norm, then it follows that the topology induced on bounded subsets of V by the weak topology is metrizable, as in Section 7.

20 The weak* topology

Let V be a real or complex vector space equipped with a norm. For each $v \in V$,

$$N_v(\lambda) = |\lambda(v)| \tag{20.1}$$

defines a seminorm on the dual space V' of bounded linear functionals on V. The collection of all of these seminorms is a nice collection of seminorms on V' in the sense of Section 1, simply because a linear functional λ on V is equal to 0 when $\lambda(v) = 0$ for every $v \in V$. The topology on V' determined by this collection of seminorms is known as the *weak* topology*. Every open set in V' with respect to the weak* topology is also an open set with respect to the topology determined by the dual norm, because $L_v(\lambda) = \lambda(v)$ is a bounded linear functional on V' for each $v \in V$.

Suppose that the linear span of $A \subseteq V$ is dense in V. This implies that N_v, $v \in A$, is also a nice collection of seminorms on V'. If $E \subseteq V'$ is bounded with respect to the dual norm, then the topology on E induced by the weak* topology on V' is the same as the topology determined by the seminorms N_v, $v \in A$. If V is separable, then it follows that the topology induced on bounded subsets of V' with respect to the dual norm by the weak* topology is metrizable, as in Section 7.

The *Banach–Alaoglu theorem* states that the closed unit ball B' in V' with respect to the dual norm is compact in the weak* topology. If V is separable, then the topology induced on B' by the weak* topology is metrizable, as in the previous paragraph. It is well known that compactness is equivalent to sequential compactness in this case.

21 Convergence of sequences

Let V be a real or complex vector space equipped with a norm. A sequence of elements $\{v_j\}_{j=1}^{\infty}$ of V converges to $v \in V$ in the weak topology if and only if

$$\lim_{j \to \infty} \lambda(v_j) = \lambda(v) \tag{21.1}$$

for every $\lambda \in V'$. Similarly, a sequence $\{\lambda_j\}_{j=1}^{\infty}$ of bounded linear functionals on V converges to $\lambda \in V'$ in the weak* topology if

$$\lim_{j \to \infty} \lambda_j(v) = \lambda(v) \tag{21.2}$$

for every $v \in V$. If $\{\lambda_j(v)\}_{j=1}^{\infty}$ converges as a sequence of real or complex numbers for every $v \in V$, then the limit automatically defines a linear functional λ on V. If the dual norms of the λ_j's are uniformly bounded, then it is easy to see that λ is also bounded.

If $\{\lambda_j\}_{j=1}^{\infty}$ is a sequence of bounded linear functionals on V which converges in the weak* topology, then $\{\lambda_j(v)\}_{j=1}^{\infty}$ is a bounded sequence of real or complex numbers for every $v \in V$. If V is complete, then it follows from the Banach–Steinhaus theorem in Section 10 that the λ_j's have uniformly bounded dual norms. If $\{v_j\}_{j=1}^{\infty}$ is a sequence of elements of V that converges in the weak topology, then one can use the Banach–Steinhaus theorem to show that the v_j's have bounded norm in V. More precisely, this uses the completeness of V', as in Section 11, instead of the completeness of V. This also uses the Hahn–Banach theorem, to get uniform boundedness of the v_j's from the uniform boundedness of $\lambda(v_j)$ over $\lambda \in V'$ with dual norm less than or equal to 1.

For any sequence $\{\lambda_j\}_{j=1}^{\infty}$ of linear functionals on V, the set of $v \in V$ such that $\{\lambda_j(v)\}_{j=1}^{\infty}$ converges in $\mathbf{R}$ of $\mathbf{C}$ is a linear subspace of V. If the λ_j's are bounded linear functionals on V with uniformly bounded dual norms, then one can show that this set is closed. It follows that a sequence $\{\lambda_j\}_{j=1}^{\infty}$ of bounded linear functionals on V with uniformly bounded dual norms converges in the weak* topology when $\{\lambda_j(v)\}_{j=1}^{\infty}$ converges in $\mathbf{R}$ or $\mathbf{C}$ for a set of v's with dense linear span in V.

Remember from Section 4 that ℓ^q can be identified with the dual of ℓ^p when $1 \le p < \infty$ and $1/p + 1/q = 1$, and that ℓ^1 can be identified with the dual of c_0 equipped with the ℓ^∞ norm. Suppose that

$$a(1) = \{a_l(1)\}_{l=1}^{\infty}, a(2) = \{a_l(2)\}_{l=1}^{\infty}, \ldots \tag{21.3}$$

is a sequence of elements of ℓ^q with uniformly bounded ℓ^q norms such that $\{a_l(j)\}_{j=1}^{\infty}$ converges as a sequence of real or complex numbers for each l. If

$$\lim_{j \to \infty} a_l(j) = a_l \tag{21.4}$$

and $a = \{a_l\}_{l=1}^{\infty}$, then $a \in \ell^q$ and $a(1), a(2), \ldots$ converges to a in the weak* topology, as bounded linear functionals on ℓ^p or c_0 when $q = 1$. The analogous statement for $\ell^q(\mathbf{Z})$ holds as well.

22 The strong operator topology

Let V, W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$, $\|w\|_W$, respectively. For each $v \in V$,

$$N_v(T) = \|T(v)\|_W \tag{22.1}$$

defines a seminorm on the vector space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W. The collection of these seminorms is a nice collection of seminorms in the sense of Section 1, since $T = 0$ as a linear mapping exactly when $T(v) = 0$ for every $v \in V$. The topology determined on $\mathcal{BL}(V, W)$ by this collection of seminorms is known as the *strong operator topology*.

If $W = \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $\mathcal{BL}(V, W)$ is the same as the dual V' of V, and the strong operator topology reduces to the weak* topology. If instead $V = \mathbf{R}$ or $\mathbf{C}$, then there is a natural identification between $\mathcal{BL}(V, W)$ and W. In this case, the strong operator topology reduces to the topology on W determined by the norm $\|w\|_W$.

Of course, open subsets of $\mathcal{BL}(V, W)$ in the strong operator topology are also open with respect to the topology determined by the operator norm. It is easy to see that closed balls in $\mathcal{BL}(V, W)$ with respect to the operator norm are closed sets in the strong operator topology.

If the linear span of $A \subseteq V$ is dense in V, then $N_v(T)$, $v \in A$, is also a nice collection of seminorms on $\mathcal{BL}(V, W)$, because $T(v) = 0$ for every $v \in V$ when $T : V \to W$ is a bounded linear transformation and $T(v) = 0$ for every $v \in A$. If $E \subseteq \mathcal{BL}(V, W)$ is bounded with respect to the operator norm, then the topology induced on E by the strong operator topology is the same as the topology determined by the seminorms $N_v(T)$, $v \in A$. Hence the topology induced on E by the strong operator topology is metrizable when V is separable, as in Section 7.

23 Convergence of sequences, 2

Let V, W be vector spaces, both real or both complex, and equipped with norms. A sequence $\{T_j\}_{j=1}^\infty$ of bounded linear mappings from V into W converges to another bounded linear mapping $T : V \to W$ in the strong operator topology when

$$\lim_{j \to \infty} T_j(v) = T(v) \tag{23.1}$$

for every $v \in V$. If $\{T_j(v)\}_{j=1}^\infty$ converges in W for every $v \in V$, then the limit defines a linear mapping $T : V \to W$. If the operator norms of the T_j's are uniformly bounded, then it is easy to see that T is bounded too.

If $\{T_j\}_{j=1}^\infty$ is a sequence of bounded linear mappings from V into W such that $\{T_j(v)\}_{j=1}^\infty$ converges in W for every $v \in V$, then $\{T_j(v)\}_{j=1}^\infty$ is bounded in W for every $v \in V$. If V is complete, then the Banach–Steinhaus theorem in Section 10 implies that the operator norms of the T_j's are uniformly bounded. Hence the limit is bounded as well, as in the previous paragraph.

Suppose that $\{T_j\}_{j=1}^\infty$ is a sequence of linear mappings from V into W, and that $T : V \to W$ is another linear mapping. The set of $v \in V$ such that $\{T_j(v)\}_{j=1}^\infty$ converges to $T(v)$ in W is a linear subspace of V. If the T_j's and T are bounded linear mappings, and if the T_j's have bounded operator norms, then the set of $v \in V$ such that $\{T_j(v)\}_{j=1}^\infty$ converges to $T(v)$ in W is closed. Under these conditions, it follows that $\{T_j\}_{j=1}^\infty$ converges to T in the strong operator topology when $\{T_j(v)\}_{j=1}^\infty$ converges to $T(v)$ in W for a set of v's whose linear span is dense in V.

For any sequence $\{T_j\}_{j=1}^\infty$ of linear mappings from V into W, the set of $v \in V$ such that $\{T_j\}_{j=1}^\infty$ converges in W is a linear subspace of V. Similarly, the set of $v \in V$ such that $\{T_j(v)\}_{j=1}^\infty$ is a Cauchy sequence in W is a linear subspace of V. If the T_j's are bounded linear mappings with uniformly bounded operator norms, then the set of $v \in V$ such that $\{T_j(v)\}_{j=1}^\infty$ is a Cauchy sequence in W is closed in V. If W is complete, then it follows that the set of $v \in V$ such that $\{T_j(v)\}_{j=1}^\infty$ converges in W is closed in V when the T_j's are uniformly bounded. Thus $\{T_j\}_{j=1}^\infty$ converges in the strong operator topology when W is complete, the T_j's are uniformly bounded, and $\{T_j(v)\}_{j=1}^\infty$ converges in W for a set of v's whose linear span is dense in V.

24 Shift operators, 2

Let R be the backward shift operator acting on sequences of real or complex numbers, as in Section 5. Thus if $a = \{a_j\}_{j=1}^\infty$ is such a sequence, then $R(a)$ is the sequence $\{a_{j+1}\}_{j=1}^\infty$. If n is a positive integer and R^n is the composition of n R's, then $R^n(a)$ is the sequence $\{a_{j+n}\}_{j=1}^\infty$. It is easy to see that R^n has operator norm equal to 1 on ℓ^p for every $n \ge 1$ and $1 \le p \le \infty$, and also on c_0 equipped with the ℓ^∞ norm. Moreover, $\{R^n\}_{n=1}^\infty$ converges to 0 in the strong operator topology as a sequence of bounded linear operators on ℓ^p when $1 \le p < \infty$, and on c_0 with the ℓ^∞ norm.

Now let T be the forward shift operator acting on sequences of real or complex numbers, so that the jth term of $T(a)$ is equal to a_{j-1} when $j \ge 2$ and to 0 when $j = 1$. Thus the jth term of $T^n(a)$ is equal to a_{j-n} when $j \ge n+1$ and to 0 when $1 \le j \le n$. In particular, $T^n(a) \to 0$ termwise as $n \to \infty$ for every sequence a of real or complex numbers. However, $\{T^n\}_{n=1}^\infty$ does not converge to 0 in the strong operator topology as a sequence of bounded linear operators on ℓ^p for any p, or on c_0 equipped with the ℓ^∞ norm. This is because

$$\|T^n(a)\|_p = \|a\|_p \tag{24.1}$$

for every $a \in \ell^p$, $1 \le p \le \infty$, and $n \ge 1$.

Similarly, forward and backward shift operators acting on doubly-infinite sequences preserve ℓ^p norms for every p, and so their iterates do not converge to 0 in the strong operator topology. If $a \in c_0(\mathbf{Z})$, then the iterates of the shift operators applied to a do converge to 0 termwise. This also works for $a \in \ell^p(\mathbf{Z})$ when $1 \le p < \infty$, since $\ell^p \subseteq c_0$ when $p < \infty$.

25 Multiplication operators

Let (X, μ) be a σ-finite measure space, and consider the corresponding Banach spaces $L^p(X)$, $1 \le p \le \infty$. For each measurable function b on X, let M_b be the corresponding multiplication operator defined by

$$M_b(f) = b\,f. \tag{25.1}$$

It is well known that this determines a bounded linear operator on $L^p(X)$ for some p if and only if b is essentially bounded. If b is essentially bounded, then M_b determines a bounded linear operator on L^p for each p, with operator norm equal to the L^∞ norm $\|b\|_\infty$ of b.

Suppose that $\{b_j\}_{j=1}^\infty$ is a sequence of essentially bounded measurable functions on X with uniformly bounded L^∞ norms. If $\{b_j\}_{j=1}^\infty$ converges pointwise almost everywhere to an essentially bounded measurable function b on X, then the dominated convergence theorem implies that the corresponding multiplication operators M_{b_j} converge to M_b in the strong operator topology as bounded linear operators on $L^p(X)$ when $1 \le p < \infty$. This also works when $\{b_j\}_{j=1}^\infty$ converges in measure to b on any measurable set in X with finite measure. Conversely, if $\{M_{b_j}\}_{j=1}^\infty$ converges to M_b in the strong operator topology as a sequence of bounded linear operators on $L^p(X)$ for some $p < \infty$, then $\{b_j\}_{j=1}^\infty$ converges to b in measure on every measurable set in X with finite measure.

For example, let X be the set $\mathbf{Z}_+$ of positive inetegers equipped with counting measure. In this case, all subsets of X are measurable, as are all functions on X, and only the empty set has measure 0. Also, $L^p(X)$ reduces to ℓ^p, essentially bounded functions are simply bounded in particular, and pointwise convergence almost everywhere or convergence in measure on sets of finite measure are the same as pointwise convergence on X. If $\{b_j\}_{j=1}^\infty$ is a uniformly bounded sequence of functions on $\mathbf{Z}_+$ that converges pointwise to a function b, then the corresponding sequence of multiplication operators $\{M_{b_j}\}_{j=1}^\infty$ converges to M_b in the strong operator topology on c_0 equipped with the ℓ^∞ norm as well as on ℓ^p, $1 \le p < \infty$.

26 The weak operator topology

Let V, W be vector spaces, both real or both complex, and equipped with norms. For each $v \in V$ and $\lambda \in W'$,

$$N_{v,\lambda}(T) = |\lambda(T(v))| \tag{26.1}$$

defines a seminorm on the vector space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W. The collection of all of these seminorms is a nice collection of seminorms on $\mathcal{BL}(V, W)$ in the sense of Section 1, because the Hahn–Banach theorem implies that for each $w \in W$ with $w \ne 0$ there is a $\lambda \in W'$ such that $\lambda(w) \ne 0$. The topology on $\mathcal{BL}(V, W)$ determined by this collection of seminorms is known as the *weak operator topology.*

If $V = \mathbf{R}$ or $\mathbf{C}$, as appropriate, then $\mathcal{BL}(V, W)$ can be identified with W as before, and the weak operator topology reduces to the weak topology on W. If $W = \mathbf{R}$ or $\mathbf{C}$, so that $\mathcal{BL}(V, W)$ is the same as V', then the weak operator topology reduces to the weak* topology. For any V, W, open subsets of $\mathcal{BL}(V, W)$ in the weak operator topology are also open sets in the strong operator topology, and hence in the topology determined by the operator norm. It is easy to see that closed balls in $\mathcal{BL}(V, W)$ with respect to the operator norm are closed sets in the weak operator topology.

If the linear span of $A \subseteq V$ is dense in V, and the linear span of $\Lambda \subseteq W'$ is dense in W' with respect to the associated dual norm, then $N_{v,\lambda}(T)$, $v \in A$, $\lambda \in \Lambda$, is also a nice collection of seminorms on $\mathcal{BL}(V, W)$. If $E \subseteq \mathcal{BL}(V, W)$ is bounded with respect to the operator norm, then the topology induced on E by the weak operator topology is the same as the topology determined by the seminorms $N_{v,\lambda}(T)$, $v \in A$, $\lambda \in \Lambda$. Hence the topology induced on E by the weak operator topology is metrizable when V and W' are separable, as in Section 7.

27 Convergence of sequences, 3

Let V, W be vector spaces, both real or both complex, and equipped with norms. A sequence $\{T_j\}_{j=1}^\infty$ of bounded linear mappings from V into W converges to another bounded linear mapping $T : V \to W$ in the weak operator topology when

$$\lim_{j \to \infty} \lambda(T_j(v)) = \lambda(T(v)) \tag{27.1}$$

for every $v \in V$ and $\lambda \in W'$. If $\{T_j\}_{j=1}^\infty$ is a uniformly bounded sequence of bounded linear mappings from V into W that converges to a linear mapping $T : V \to W$ in this way, then T is automatically bounded. This uses the Hahn–Banach theorem to control the norms of elements of w in terms of bounded linear functionals.

If $\{w_j\}_{j=1}^\infty$ is a sequence of elements of W such that $\{\lambda(w_j)\}_{j=1}^\infty$ is bounded in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for every $\lambda \in W'$, then $\{w_j\}_{j=1}^\infty$ is bounded with respect to the norm on W. This follows from the Banach–Steinhaus theorem in Section 10 applied to the linear functionals $\lambda \mapsto \lambda(w_j)$ on W', using the fact that W' is automatically complete, as in Section 11. This also uses the Hahn–Banach theorem to say that the norm of w_j is equal to the norm of $\lambda \mapsto \lambda(w_j)$ as a linear functional on W', as in Section 18. Suppose now that $\{T_j\}_{j=1}^\infty$ is a sequence of bounded linear mappings from V into W such that $\{\lambda(T_j(v))\}_{j=1}^\infty$ is a bounded sequence of real or complex numbers for every $v \in V$ and $\lambda \in W'$. In particular, this happens when $\{T_j\}_{j=1}^\infty$ converges in the weak operator topology. As in the previous remarks, this implies that $\{T_j(v)\}_{j=1}^\infty$ is a bounded sequence in W for every $v \in V$. If V is complete, then the Banach–Steinhaus theorem implies that the T_j's have uniformly bounded operator norm.

Let $\{T_j\}_{j=1}^\infty$ be a sequence of linear mappings from V into W, and let T be another linear mapping from V into W. For each $\lambda \in W'$, the set $A(\lambda)$ of $v \in V$

such that $\{\lambda(T_j(v))\}_{j=1}^{\infty}$ converges to $\lambda(T(v))$ is a linear subspace of V. If the T_j's are bounded linear mappings with uniformly bounded operator norms, and if T is bounded, then $A(\lambda)$ is also a closed set with respect to the norm on V. Under these conditions, the set of $\lambda \in W'$ such that $A(\lambda) = V$ is a closed linear subspace of W' with respect to the dual norm too. Of course, $\{T_j\}_{j=1}^{\infty}$ converges to T in the weak operator topology when $A(\lambda) = V$ for every $\lambda \in W'$.

If T is the forward shift operator acting on sequences of real or complex numbers, then $\{T^n\}_{n=1}^{\infty}$ converges to 0 in the weak operator topology as a sequence of bounded linear operators on ℓ^p when $1 < p < \infty$, and on c_0 equipped with the ℓ^∞ norm. The analogous statement also holds for forward and backward shift operators acting on doubly-infinite sequences. If (X, μ) is a σ-finite measure space and $1 \le p < \infty$, then the topology on $L^\infty(X)$ induced by the weak operator topology on the corresponding space of multiplication operators on $L^p(X)$ is the same as the weak* topology on $L^\infty(X)$ as the dual of $L^1(X)$. This uses the well-known identification of the dual of $L^p(X)$ with $L^q(X)$ when $1 \le p < \infty$ and $1/p + 1/q = 1$.

28 The weak* operator topology

Let V, Z be vector spaces, both real or both complex, and equipped with norms, and suppose that W is the dual of Z. For each $v \in V$ and $z \in Z$,

$$N_{v,z}(T) = |T(v)(z)| \tag{28.1}$$

defines a seminorm on the vector space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W. More precisely, if T is a linear mapping from V into W, $v \in V$, and $z \in Z$, then $T(v)$ is a bounded linear functional on Z, and $T(v)(z)$ is the value of this functional at z. As usual, this is a nice collection of seminorms in the sense of Section 1, and the topology that they describe may be described as the *weak* operator topology*.

For example, if $V = \mathbf{R}$ or $\mathbf{C}$, as appropriate, so that $\mathcal{BL}(V, W)$ can be identified with W, then the weak* operator topology reduces to the weak* topology on W as the dual of Z. If T is the forward shift operator acting on the space ℓ^∞ of bounded sequences of real or complex numbers, as in Section 5, then $\{T^n\}_{n=1}^{\infty}$ converges to 0 in the weak* operator topology, where the range is identified with the dual of ℓ^1. This uses the fact that T introduces a 0 in the first term of the sequence, and the analogous statement does not work for the shift operators on doubly-infinite sequences. Moreover, $\{T^n\}_{n=1}^{\infty}$ converges to 0 as a sequence of operators on ℓ^1 in the weak* operator topology, where the range is identified with the dual of c_0, and this works as well for shift operators on doubly-infinite sequences. If (X, μ) is a σ-finite measure space, then the topology induced on $L^\infty(X)$ by the weak* operator topology on the corresponding multiplication operators on $L^\infty(X)$ is the same as the weak* topology on $L^\infty(X)$, where $L^\infty(X)$ is identified with the dual of $L^1(X)$.

A sequence $\{T_j\}_{j=1}^{\infty}$ of bounded linear operators from V into W converges to a bounded linear operator $T : V \to W$ in the weak* operator topology when

$$\lim_{j \to \infty} T_j(v)(z) = T(v)(z) \tag{28.2}$$

for every $v \in V$ and $z \in Z$. If $\{T_j(v)(z)\}_{j=1}^{\infty}$ converges in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for every $v \in V$ and $z \in Z$, then the limit automatically defines a linear mapping $T : V \to W$. If the T_j's have uniformly bounded operator norms, then T is bounded too. If V, Z are complete, then one can use the Banach–Steinhaus theorem to show that a sequence $\{T_j\}_{j=1}^{\infty}$ of bounded linear mappings from V into W have uniformly bounded operator norms when $\{T_j(v)(z)\}_{j=1}^{\infty}$ is bounded in $\mathbf{R}$ or $\mathbf{C}$ for every $v \in V$ and $z \in Z$. More precisely, one can apply the Banach–Steinhaus theorem first to show that $\{T_j(v)\}_{j=1}^{\infty}$ is bounded in W for each $v \in V$, and then again to get the uniform boundedness of the T_j's.

It is easy to see that closed balls in $\mathcal{BL}(V, W)$ with respect to the operator norm are closed sets in the weak* operator topology. If $E \subseteq \mathcal{BL}(V, W)$ is bounded with respect to the operator norm and V, Z are separable, then the topology induced on E by the weak* operator topology is metrizable. A variant of the Banach–Alaoglu theorem implies that the closed unit ball in $\mathcal{BL}(V, W)$ with respect to the operator norm is compact in the weak* operator topology. This is equivalent to sequential compactness when V, Z are separable, because of metrizability. Of course, the weak* operator topology on $\mathcal{BL}(V, W)$ is the same as the weak operator topology when Z is reflexive.

29 Fourier series

Let $\mathbf{T}$ be the unit circle in the complex plane, which is to say the set of $z \in \mathbf{C}$ such that $|z| = 1$. The standard integral inner product for complex-valued functions on $\mathbf{T}$ is defined by

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{\mathbf{T}} f(z) \, \overline{g(z)} \, |dz|. \tag{29.1}$$

The functions z^j, $j \in \mathbf{Z}$, on $\mathbf{T}$ are orthonormal with respect to this inner product, because

$$\int_{\mathbf{T}} z^j \, |dz| = 0 \tag{29.2}$$

when $j \neq 0$, and the integral is equal to 2π when $j = 0$ since z^j reduces to 1. It is well known that these functions form an orthonormal basis for $L^2(\mathbf{T})$.

This leads to a natural one-to-one mapping from doubly-infinite sequences $a = \{a_j\}_{j=-\infty}^{\infty}$ of complex numbers in $\ell^2(\mathbf{Z})$ onto complex-valued functions in $L^2(\mathbf{T})$, namely,

$$a \mapsto \sum_{j=-\infty}^{\infty} a_j \, z^j. \tag{29.3}$$

More precisely, this doubly-infinite series can be treated as a sum of two ordinary infinite series, whose partial sums converge in the L^2 norm, as in Section 14.

Because of orthonormality, the standard inner product on $\ell^2(\mathbf{Z})$ corresponds exactly to the integral inner product on $L^2(\mathbf{T})$ under this mapping.

The (forward) shift operator that sends $a \in \ell^2(\mathbf{Z})$ to $\{a_{j-1}\}_{j=1}^\infty$ corresponds in this way to the multiplication operator on $L^2(\mathbf{T})$ associated to the function z. We have also considered shift operators on ℓ^p and multiplication operators on L^p, but they do not match up as well as when $p = 2$.

30 Adjoints

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let T be a bounded linear operator on V. For each $w \in V$,

$$v \mapsto \langle T(v), w\rangle \tag{30.1}$$

is a bounded linear functional on V. Hence there is a unique element $T^*(w)$ of V such that

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle \tag{30.2}$$

for every $v \in V$, as in Section 17. It is easy to see that $T^*(w)$ is linear in w, because of uniqueness. Let us check that T^* is also bounded, and that

$$\|T^*\|_{op} = \|T\|_{op}. \tag{30.3}$$

Because of (30.2),

$$|\langle v, T^*(w)\rangle| \le \|T\|_{op}\,\|v\|\,\|w\| \tag{30.4}$$

for every $v, w \in V$. This implies that T^* is bounded, with $\|T^*\|_{op} \le \|T\|_{op}$, and an analogous argument shows that $\|T\|_{op} \le \|T^*\|_{op}$. This operator T^* is known as the *adjoint* of T.

If T_1, T_2 are bounded linear operators on V, then

$$(T_1 + T_2)^* = T_1^* + T_2^* \tag{30.5}$$

and

$$(T_1 \circ T_2)^* = T_2^* \circ T_1^*. \tag{30.6}$$

If V is a real Hilbert space, then

$$(a\,T)^* = a\,T^* \tag{30.7}$$

for every $a \in \mathbf{R}$ and bounded linear operator T on V, while

$$(a\,T)^* = \overline{a}\,T^* \tag{30.8}$$

in the complex case. Thus $T \mapsto T^*$ is linear in the real case, and conjugate-linear in the complex case. Of course, $I^* = I$, where I denotes the identity operator on V. For any bounded linear operator T on V,

$$(T^*)^* = T. \tag{30.9}$$

Let (X, μ) be a measure space, and consider the Hilbert space $L^2(X)$ with the inner product defined by integration, as usual. If b is a bounded measurable function on X, then the operator M_b of multiplication by b is bounded on $L^2(X)$, as in Section 25. It is easy to see that the adjoint of M_b is itself in the real case, and is given by multiplication by $\overline{b}$ in the complex case.

Observe that $T \mapsto T^*$ is a homeomorphism from $\mathcal{BL}(V)$ onto itself with respect to the weak operator topology. However, this does not work for the strong operator topology. For example, if R is the backward shift operator on ℓ^2 as in Section 5, then one can check that its adjoint is the forward shift operator T. As in Section 24, $\{R^n\}_{n=1}^\infty$ converges to 0 in the strong operator topology, but $\{T^n\}_{n=1}^\infty$ does not. Note that $\{T^n\}_{n=1}^\infty$ does converge to 0 in the weak operator topology, as in Section 27.

31 Unitary transformations

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space. A bounded linear mapping $T : V \to V$ is *invertible* if it has a bounded inverse $T^{-1} : V \to V$, which is to say that T^{-1} is a bounded linear operator on V that satisfies

$$T \circ T^{-1} = T^{-1} \circ T = I, \tag{31.1}$$

and which is unique when it exists. If T_1, T_2 are invertible bounded linear operators on V, then $T_1 \circ T_2$ is also invertible, and

$$(T_1 \circ T_2)^{-1} = T_2^{-1} \circ T_1^{-1}. \tag{31.2}$$

If T is an invertible bounded linear operator on V, then its adjoint T^* is also invertible, and

$$(T^*)^{-1} = (T^{-1})^*. \tag{31.3}$$

A linear mapping T from V onto itself is said to be *unitary* if

$$\langle T(v), T(w) \rangle = \langle v, w \rangle \tag{31.4}$$

for every $v, w \in V$, which implies that

$$\|T(v)\| = \|v\| \tag{31.5}$$

for every $v \in V$. Thus unitary transformations are bounded in particular. Conversely, one can show that (31.5) implies (31.4), using polarization identities. Note that (31.5) implies that T has trivial kernel, and hence is one-to-one. It follows that unitary transformations are invertible, since they are surjective by definition. The inverse of a unitary transformation is clearly unitary as well. Compositions of unitary transformations are unitary too. As another characterization, a bounded linear mapping $T : V \to V$ is unitary if and only if it is invertible and

$$T^{-1} = T^*. \tag{31.6}$$

For example, forward and backward shift operators on $\ell^2(\mathbf{Z})$ are unitary, and in fact are inverses and adjoints of each other. Consider now the forward and backward shift operators T, R on $\ell^2(\mathbf{Z}_+)$, as in Section 5. Note that $R = T^*$, as in the previous section, and that

$$R \circ T = I. \tag{31.7}$$

Although T satisfies (31.4), it is not unitary, because it is not surjective, and indeed

$$T \circ R \neq I. \tag{31.8}$$

Similarly, R is not injective, but it is an isometry on the orthogonal complement of its kernel.

32 Self-adjoint operators

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space. A bounded linear mapping $A : V \to V$ is said to be *self-adjoint* if

$$A^* = A, \tag{32.1}$$

which is the same as

$$\langle A(v), w \rangle = \langle v, A(w) \rangle \tag{32.2}$$

for every $v, w \in V$. In the complex case, this implies that

$$\langle A(v), v \rangle = \langle v, A(v) \rangle = \overline{\langle A(v), v \rangle}, \tag{32.3}$$

so that $\langle A(v), v \rangle \in \mathbf{R}$ for every $v \in V$ when A is self-adjoint. For example, if (X, μ) is a measure space, then multiplication operators on $L^2(X)$ are self-adjoint in the real case, while multiplication operators associated to real-valued functions are self-adjoint in the complex case.

If A, B are bounded self-adjoint linear operators on V, then it is easy to see that $A + B$ is also self-adjoint. Furthermore, $r\,A$ is self-adjoint when A is self-adjoint and $r \in \mathbf{R}$. Hence the self-adjoint operators form a real-linear subspace of $\mathcal{BL}(V)$, and the restriction to real scalars is important in the complex case. Note that this subspace is closed in the weak operator topology.

Let W be a closed linear subspace of V, and consider the orthogonal projection P_W of V onto W, as in Section 15. Thus P_W is characterized by the conditions $P_W(v) \in W$ and $v - P_W(v) \in W^\perp$ for every $v \in V$. This implies that

$$\langle P_W(v), w \rangle = \langle P_W(v), P_W(w) \rangle = \langle v, P_W(w) \rangle \tag{32.4}$$

for every $v, w \in W$. In particular, P_W is self-adjoint.

33 Nonnegative self-adjoint operators

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space. A bounded self-adjoint linear mapping $A : V \to V$ is said to be *nonnegative* if

$$\langle A(v), v \rangle \geq 0 \tag{33.1}$$

for every $v \in V$, which may be expressed by $A \geq 0$. If A, B are bounded self-adjoint operators such that $A - B \geq 0$, then we put $A \geq B$.

If A, B are bounded nonnegative self-adjoint linear operators on V, then $A + B$ is nonnegative, and $r\,A$ is nonnegative for every real number $r \geq 0$. Thus the bounded nonnegative self-adjoint linear operators on V form a real convex cone in $\mathcal{BL}(V)$. This cone is also a closed set in the weak operator topology.

For example, if (X, μ) is a measure space, then multiplication operators associated to nonnegative real-valued functions are nonnegative on $L^2(X)$. It follows easily from (32.4) that orthogonal projections are nonnegative as well. If T is any bounded linear operator on V, then $T^* \circ T$ is self-adjoint, because

$$(T^* \circ T)^* = T^* \circ (T^*)^* = T^* \circ T. \tag{33.2}$$

Moreover, $T^* \circ T \geq 0$, because

$$\langle (T^* \circ T)(v), v \rangle = \langle T^*(T(v)), v \rangle = \langle T(v), T(v) \rangle \geq 0. \tag{33.3}$$

34 Cauchy–Schwarz, revisited

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let A be a bounded nonnegative self-adjoint linear operator on V. Thus $\langle A(v), w \rangle$ is like an inner product on V, except that $\langle A(v), v \rangle$ may be 0 even when $v \neq 0$. There is an analogue of the Cauchy–Schwarz inequality in this context, which states that

$$|\langle A(v), w \rangle| \leq \langle A(v), v \rangle^{1/2}\, \langle A(w), w \rangle^{1/2} \tag{34.1}$$

for every $v, w \in V$. This can be established in the same way as for ordinary inner products.

As a consequence,

$$\|A\|_{op} = \sup\{\langle A(v), v \rangle : v \in V,\ \|v\| \leq 1\}. \tag{34.2}$$

More precisely, the right side is less than or equal to $\|A\|_{op}$ by the ordinary Cauchy–Schwarz inequality, while the other direction uses (34.1). Similarly,

$$\|A(v)\| \leq \|A\|_{op}^{1/2}\, \langle A(v), v \rangle^{1/2} \tag{34.3}$$

for every $v \in V$, as one can see by by taking the supremum over $w \in V$ with $\|w\| \leq 1$ in (34.1).

Suppose that $A_1, A_2, \ldots$ is a sequence of bounded nonnegative self-adjoint linear operators on V that converges to 0 in the weak operator topology. The

Banach–Steinhaus theorem implies that the operator norms of the A_j's are uniformly bounded, although one could just as well include this as an additional hypothesis if V were not complete. Under these conditions, $\{A_j\}_{j=1}^\infty$ also converges to 0 in the strong operator topology, since $\|A_j(v)\|$ can be estimated in terms of $\langle A_j(v), v\rangle$, as in the previous paragraph.

Now suppose that $B_1, B_2, \ldots$ is a sequence of bounded self-adjoint linear operators on V which is increasing in the sense that

$$B_1 \le B_2 \le \cdots. \tag{34.4}$$

If the operator norms of the B_j's are uniformly bounded, then there is a bounded self-adjoint linear operator B on V such that $B_j \le B$ for each j and $\{B_j\}_{j=1}^\infty$ converges to B in the strong operator topology. Indeed, under these conditions $\{\langle B_j(v), v\rangle\}_{j=1}^\infty$ is a bounded monotone increasing sequence of real numbers for every $v \in V$, and therefore converges in $\mathbf{R}$. One can use this and polarization identities to show that $\{\langle B_j(v), w\rangle\}_{j=1}^\infty$ converges in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for every $v, w \in V$, and hence that $\{B_j\}_{j=1}^\infty$ converges in the weak operator topology. Convergence in the strong operator topology can be derived from this as in the preceding paragraph.

35 Continuity of compositions

Let V, W, and Z be a vector spaces, all real or all complex, and equipped with norms. If A is a bounded linear mapping from W into Z, then

$$T \mapsto A \circ T \tag{35.1}$$

is a bounded linear mapping from $\mathcal{BL}(V, W)$ into $\mathcal{BL}(V, Z)$ with respect to the corresponding operator norms. It is also continuous with respect to the strong and weak operator topologies. Similarly, if B is a bounded linear mapping from V into W, then

$$T \mapsto T \circ B \tag{35.2}$$

is a bounded linear mapping from $\mathcal{BL}(W, Z)$ into $\mathcal{BL}(V, Z)$ with respect to the corresponding operator norms, and is continuous with respect to the strong and weak operator topologies.

Now let $\{A_j\}_{j=1}^\infty$ be a sequence of bounded linear mappings from W into Z with uniformly bounded operator norms, and let $\{B_j\}_{j=1}^\infty$ be a sequence of bounded linear mappings from V into W. If $\{B_j\}_{j=1}^\infty$ converges to 0 in the strong operator topology, then $\{A_j \circ B_j\}_{j=1}^\infty$ also converges to 0 in the strong operator topology. If $\{A_j\}_{j=1}^\infty$ converges in the strong operator topology to a bounded linear mapping $A : W \to Z$, and $\{B_j\}_{j=1}^\infty$ converges in the strong operator topology to a bounded linear mapping $B : V \to W$, then $\{A_j \circ B_j\}_{j=1}^\infty$ converges to $A \circ B$ in the strong operator topology. This is because

$$A_j \circ B_j = A_j \circ (B_j - B) + A_j \circ B, \tag{35.3}$$

where the first term on the right converges to 0 and the second term converges to $A \circ B$ in the strong operator topology by the previous remarks. Similarly, if $\{A_j\}_{j=1}^\infty$ converges to a bounded linear mapping A in the weak operator topology, and $\{B_j\}_{j=1}^\infty$ converges to a bounded linear mapping B in the strong operator topology, then $\{A_j \circ B_j\}_{j=1}^\infty$ converges to $A \circ B$ in the weak operator topology.

However, if $\{A_j\}_{j=1}^\infty$ converges to A in the strong operator topology, and $\{B_j\}_{j=1}^\infty$ converges to B in the weak operator topology, then it may not be that $\{A_j \circ B_j\}_{j=1}^\infty$ converges to $A \circ B$ in the weak operator topology. For instance, let R, T be the backwards and forward shift operators acting on $\ell^p(\mathbf{Z}_+)$, $1 < p < \infty$, or on $c_0(\mathbf{Z}_+)$. If $A_j = R^j$ and $B_j = T^j$, then $\{A_j\}_{j=1}^\infty$ converges to 0 in the strong operator topology, $\{B_j\}_{j=1}^\infty$ converges to 0 in the weak operator topology, the A_j's and B_j's have operator norm equal to 1, and $A_j \circ B_j = I$ for each j.

There are analogous continuity statements for the weak* operator topology instead of the weak operator topology, under suitable conditions. If W, Z are dual spaces and one is interested in the continuity of $T \mapsto A \circ T$ with respect to the weak* operator topologies on $\mathcal{BL}(V, W)$ and $\mathcal{BL}(V, Z)$, then it is appropriate to ask that A be the dual of a bounded linear operator from the predual of Z into the predual of W. For the rest, it suffices that Z be a dual space, and the necessary changes are straightforward.

36 Multiplication operators, 2

Let (X, μ) be a σ-finite measure space, and let us check that the collection of multiplication operators M_b associated to bounded measurable functions b on X is closed in the weak operator topology on $\mathcal{BL}(L^p(X))$ for each p, $1 \le p \le \infty$. If $p = \infty$, then we can use the weak* operator topology, where $L^\infty(X)$ is identified with the dual of $L^1(X)$. To see this, suppose that T is a bounded linear operator on $L^p(X)$ which is in the closure of the multiplication operators with respect to the weak operator topology or weak* operator topology, as appropriate. Because multiplication operators commute with each other, it follows from the continuity of $T \mapsto A \circ T$ and $T \mapsto T \circ B$ as in the previous section that

$$T \circ M_b = M_b \circ T \tag{36.1}$$

for every $b \in L^\infty(X)$. If $p = \infty$, or $1 \le p < \infty$ and $\mu(X) < \infty$, then it follows that T is the same as multiplication by $T(1)$. Otherwise, suppose that E is a measurable set in X such that $0 < \mu(E) < \infty$, and let $\mathbf{1}_E(x)$ be the characteristic function of E on X, equal to 1 when $x \in E$ and to 0 when $x \in X \backslash E$. In this case, T is the same as multiplication by $T(\mathbf{1}_E)$ when acting on functions f equal to 0 on $X \backslash E$, and $T(\mathbf{1}_E)$ is equal to 0 on $X \backslash E$. To deal with arbitrary functions on X, one can apply this to a sequence of disjoint subsets of X whose union is X.

37 Complex Hilbert spaces

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space. A bounded linear operator A on V is said to be *anti-self-adjoint* if

$$A^* = -A. \tag{37.1}$$

In the complex case, this happens exactly when $A = i\,B$, where B is a bounded self-adjoint operator on V.

Suppose that T is any bounded linear operator on V. In both the real and complex cases, we can express T as

$$T = \frac{T + T^*}{2} + \frac{T - T^*}{2}, \tag{37.2}$$

where the first term on the right is self-adjoint, and the second term is anti-self-adjoint. If V is complex, then this implies that

$$T = T_1 + i\,T_2, \tag{37.3}$$

where T_1 and T_2 are both self-adjoint.

If V is a complex Hilbert space, and T is a bounded self-adjoint linear operator T on V, then

$$\langle T(v), v \rangle \in \mathbf{R} \tag{37.4}$$

for every $v \in V$, as in Section 32. Conversely, suppose that T is a bounded linear operator on V such that (37.4) holds for every $v \in V$, and let T_1, T_2 be bounded self-adjoint linear operators on V such that $T = T_1 + i\,T_2$, as in the previous paragraph. Thus T_1, T_2 have the same property, which implies that $\langle T_2(v), v \rangle = 0$ for every $v \in V$. Using polarization, one can show that $\langle T_2(v), w \rangle = 0$ for every $v, w \in V$, and hence that $T_2 = 0$. The same conclusion also follows from Section 34.

38 The C^* identity

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let T be a bounded linear operator on V. The C^* *identity* states that

$$\|T^* \circ T\|_{op} = \|T\|_{op}^2. \tag{38.1}$$

To see this, observe first that

$$\|T^* \circ T\|_{op} \le \|T^*\|_{op}\, \|T\|_{op} = \|T\|_{op}^2, \tag{38.2}$$

by basic properties of the operator norm. In the other direction,

$$\|T(v)\|^2 = \langle T(v), T(v) \rangle = \langle T^*(T(v)), v \rangle \le \|T^* \circ T\|_{op}\, \|v\|^2 \tag{38.3}$$

for every $v \in V$ implies that $\|T\|_{op}^2 \le \|T^* \circ T\|_{op}$, as desired.

Of course, the latter computation is basically the same as the proof that $T^* \circ T \geq 0$, as in Section 33. Note that the Cauchy–Schwarz inequality (34.1) follows easily from the usual version when $A = T^* \circ T$, since

$$\begin{aligned} |\langle (T^* \circ T)(v), w\rangle| &= |\langle T(v), T(w)\rangle| \le \|T(v)\| \, \|T(w)\| \\ &= \langle (T^* \circ T)(v), v\rangle^{1/2} \, \langle (T^* \circ T)(w), w\rangle^{1/2} \end{aligned} \tag{38.4}$$

for every $v, w \in V$. Similarly, the fact that the operator norm of $T^* \circ T$ is the same as the supremum of $\langle (T^* \circ T)(v), v\rangle$ over $v \in V$ with $\|v\| \le 1$ is implicit in the computations in the previous paragraph.

39 Finite-dimensional spaces

Let V be a real or complex vector space, and let N be a seminorm on V. The triangle inequality implies that

$$N(v) - N(w),\ N(w) - N(v) \le N(v - w) \tag{39.1}$$

and hence

$$|N(v) - N(w)| \le N(v - w) \tag{39.2}$$

for every $v, w \in V$. If $V = \mathbf{R}^n$ or $\mathbf{C}^n$ for some positive integer n, and $\|v\|$ denotes the standard Euclidean norm, then there is a nonnegative real number C such that

$$N(v) \le C \, \|v\| \tag{39.3}$$

for every $v \in V$. This can be shown by expressing v as a finite linear combination of the standard basis vectors, and applying the triangle inequality. It follows that N is continuous with respect to the standard topology defined by the Euclidean norm. If N is a norm on $\mathbf{R}^n$ or $\mathbf{C}^n$, then it follows that there is a $c > 0$ such that

$$N(v) \ge c \, \|v\| \tag{39.4}$$

for each v. Specifically, one can take c to be the minimum of N on the unit sphere associated to the Euclidean norm, which is attained because the sphere is compact.

If V is a finite-dimensional real or complex vector space of dimension n, then V is isomorphic to $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. The previous remarks show that the topology on V determined by any norm is equivalent to the topology induced by the standard topology on $\mathbf{R}^n$ or $\mathbf{C}^n$ by a linear isomorphism. It also follows that V is complete with respect to any norm, since $\mathbf{R}^n$ and $\mathbf{C}^n$ are complete with respect to the standard norms. If V is an infinite-dimensional vector space equipped with a norm, then the same statements can be applied to any finite-dimensional subspace W of V. In particular, W is automatically a closed subspace of V, because it is complete.

40 Continuous linear functionals

Let V be a real or complex vector space equipped with a nice collection of seminorms $\mathcal{N}$. Thus $\mathcal{N}$ determines a topology on V, and it makes sense to talk about continuous linear functionals on V. One can check that a linear functional λ on V is continuous if and only if there are finitely many seminorms $N_1, \ldots, N_l \in \mathcal{N}$ and a $C \geq 0$ such that

$$|\lambda(v)| \leq C \max(N_1(v), \ldots, N_l(v)) \tag{40.1}$$

for every $v \in V$. For if λ satisfies this condition, then it is easy to see that λ is continuous. Conversely, if λ is continuous at 0, then there is a neighborhood U of 0 in the topology on V determined by $\mathcal{N}$ such that

$$|\lambda(v)| < 1 \tag{40.2}$$

for every $v \in U$. This means that there are finitely many seminorms $N_1, \ldots, N_l$ in $\mathcal{N}$ and $r_1, \ldots, r_l > 0$ such that

$$\{v \in V : N_1(v) < r_1, \ldots, N_l(v) < r_l\} \subseteq \{v \in V : |\lambda(v)| < 1\}. \tag{40.3}$$

This implies the previous estimate, with $C = (\min(r_1, \ldots, r_l))^{-1}$.

As a special case, let Λ be a collection of linear functionals on V such that for each $v \in V$ with $v \neq 0$ there is a $\lambda \in \Lambda$ such that $\lambda(v) \neq 0$. In this case,

$$N_\lambda(v) = |\lambda(v)| \tag{40.4}$$

is a seminorm on V for each $\lambda \in \Lambda$, and these seminorms form a nice collection of seminorms on V. By construction, each $\lambda \in \Lambda$ is a continuous linear functional on V with respect to the topology determined by this collection of seminorms, as in any finite linear combination of the elements of Λ. Conversely, if μ is a continuous linear functional on V with respect to this collection of seminorms, then the characterization of continuity in the previous paragraph implies that the kernel of μ contains the intersection of the kernels of finitely many elements of Λ. This implies that μ is a linear combination of these elements of Λ.

Let $\mathcal{N}$ be any nice collection of seminorms on V again. Suppose that W is a closed linear subspace of W in the topology determined by $\mathcal{N}$, and that $v \in V$ is not an element of W. This implies that there are finitely many seminorms $N_1, \ldots, N_l \in \mathcal{N}$ and $r_1, \ldots, r_l > 0$ such that

$$\{u \in V : N_1(v - u) < r_1, \ldots N_l(v - u) < r_l\} \subseteq V \backslash W. \tag{40.5}$$

If $c = \min(r_1, \ldots, r_l)$, then it follows that

$$\max(N_1(v - w), \ldots, N_l(v - w)) \geq c \tag{40.6}$$

for every $w \in W$. Let λ be the linear functional on the linear span of W and v defined by

$$\lambda(w + t\, v) = t \tag{40.7}$$

for $w \in W$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. The previous estimate implies that

$$|\lambda(z)| \le c^{-1} \max(N_1(z), \ldots, N_l(z)) \tag{40.8}$$

for every z in the span of W and v, so that λ is a continuous linear functional on this subspace. The Hahn–Banach theorem works just as well for seminorms as for norms, and thus there is an extension of λ to a linear functional on V which satisfies the same condition, and hence is continuous.

41 Operator topologies, revisited

Let V, W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$, $\|w\|_W$. Suppose that ϕ is a continuous linear functional on $\mathcal{BL}(V, W)$ with respect to the strong operator topology. As in the previous section, there are finitely many vectors $v_1, \ldots, v_l \in V$ and a $C \ge 0$ such that

$$|\phi(T)| \le C \, \max(\|T(v_1)\|_W, \ldots, \|T(v_l)\|_W) \tag{41.1}$$

for every bounded linear mapping $T : V \to W$. We may as well ask that $v_1, \ldots, v_l$ be linearly independent, since otherwise we can use fewer vectors and still get an analogous estimate.

Let $\lambda_1, \ldots, \lambda_l$ be a dual basis of linear functionals on the span of $v_1, \ldots, v_l$, so that

$$\lambda_j(v_k) = 0 \text{ when } j \ne k, \text{ and } \lambda_j(v_j) = 1. \tag{41.2}$$

By the Hahn–Banach theorem, $\lambda_1, \ldots, \lambda_l$ may be extended to bounded linear functionals on V. If $w_1, \ldots, w_l$ are arbitrary elements of W, then

$$T(v) = \sum_{j=1}^{l} \lambda_j(v) \, w_j \tag{41.3}$$

is a bounded linear mapping from V into W such that

$$T(v_j) = w_j. \tag{41.4}$$

Let W^l be the vector space of l-tuples $(w_1, \ldots, w_l)$ of elements of W, with the norm

$$\|(w_1, \ldots, w_l)\|_{W^l} = \max(\|w_1\|_W, \ldots, \|w_l\|_W). \tag{41.5}$$

Note that $\phi(T)$ depends only on $T(v_1), \ldots, T(v_l)$, because $\phi(T) = 0$ when $T(v_1) = \cdots = T(v_l) = 0$, by (41.1). Every element of W^l can be expressed as $(T(v_1), \ldots, T(v_l))$ for some bounded linear mapping $T : V \to W$, as in the previous paragraph. Hence there is a unique linear functional ψ on W^l such that

$$\psi(T(v_1), \ldots, T(v_l)) = \phi(T) \tag{41.6}$$

for every bounded linear mapping $T : V \to W$, and which satisfies

$$|\psi(w_1, \ldots, w_l)| \le C \, \|(w_1, \ldots, w_l)\|_{W^l} \tag{41.7}$$

for every $(w_1, \ldots, w_l) \in W^l$, by (41.1). Thus ψ is a bounded linear functional on W^l, and it is easy to see that there are bounded linear functionals $\mu_1, \ldots, \mu_l$ on W such that

$$\psi(w_1, \ldots, w_l) = \mu_1(w_1) + \cdots + \mu_l(w_l) \tag{41.8}$$

for every $(w_1, \ldots, w_l) \in W^l$.

Equivalently,

$$\phi(T) = \mu_1(T(v_1)) + \cdots + \mu_l(T(v_l)) \tag{41.9}$$

for every bounded linear mapping $T : V \to W$. This shows that every continuous linear functional on $\mathcal{BL}(V, W)$ with respect to the strong operator topology is also continuous with respect to the weak operator topology. The converse is trivial, since the strong operator topology contains the weak operator topology. As a consequence, any closed linear subspace of $\mathcal{BL}(V, W)$ with respect to the strong operator topology is also closed with respect to the weal operator topology. This uses the Hahn–Banach theorem, as in the previous section.

42 Continuous functions

Let X be a compact Hausdorff topological space, and let $C(X)$ be the vector space of real or complex-valued continuous functions on X. If f is a continuous function on X, then its *supremum norm* is defined as usual by

$$\|f\|_{sup} = \sup\{|f(x)| : x \in X\}. \tag{42.1}$$

It is well known that $C(X)$ is complete with respect to the supremum norm.

Let μ be a positive Borel measure on X such that $\mu(X) < \infty$ and $\mu(U) > 0$ when U is a nonempty open set in X. In this case, the essential supremum or L^∞ norm of continuous function f on X with respect to μ is the same as the supremum norm of f. Hence the operator norm of multiplication by a continuous function b on $L^p(X)$ is equal to the supremum norm of b. In particular, the collection of multiplication operators associated to continuous functions on X is closed in the space of bounded linear operators on $L^p(X)$ with respect to the operator norm. However, multiplication operators associated to bounded Borel measurable functions on X are in the closure of the set of multiplication operators associated to continuous functions in the strong operator topology when $1 \leq p < \infty$, at least under suitable regularity conditions on the measure.

43 von Neumann algebras

Let V be a complex Hilbert space, and let $\mathcal{A}$ be an algebra of bounded linear operators on V. Thus $\mathcal{A}$ is a linear subspace of $\mathcal{BL}(V)$, and $A \circ B \in \mathcal{A}$ whenever $A, B \in \mathcal{A}$. We also ask that the identity operator I on V be contained in $\mathcal{A}$. If $A^* \in \mathcal{A}$ whenever $A \in \mathcal{A}$, then $\mathcal{A}$ is a $*$-algebra.

A *von Neumann algebra* of operators on V is a $*$-algebra $\mathcal{A} \subseteq \mathcal{BL}(V)$ which is closed in the weak operator topology. The algebra $\mathcal{BL}(V)$ of all bounded linear operators on V is automatically a von Neumann algebra. If (X, μ) is a σ-finite measure space, then the algebra of multiplication operators on $L^2(X)$ associated to bounded measurable functions is a von Neumann algebra.

Suppose that X is a compact Hausdorff space, and that μ is a positive finite Borel measure on X such that $\mu(U) > 0$ when U is a nonempty open set in X, as in the previous section. The collection of multiplication operators on $L^2(X)$ associated to continuous functions on X is a $*$-algebra that is closed in the topology determined by the operator norm, but it is not closed in the strong operator topology even when $X = [0, 1]$ and μ is Lebesgue measure. A $*$-algebra $\mathcal{A} \subseteq \mathcal{BL}(V)$ that is closed in the topology determined by the operator norm is a C^* *algebra*. More precisely, C^* algebras can be defined abstractly, without reference to bounded linear operators on a Hilbert space. In particular, $C(X)$ is already a C^* algebra with respect to the supremum norm.

44 Commutants

Let V be a real or complex vector space equipped with a norm, and let $\mathcal{E}$ be a collection of bounded linear operators on V. The *commutant* $\mathcal{C}(\mathcal{E})$ is the set of bounded linear operators T on V that commute with every element of $\mathcal{E}$, so that

$$T \circ A = A \circ T \tag{44.1}$$

for every $A \in \mathcal{E}$. For example, if $\mathcal{E}$ consists of scalar multiples of the identity operator I, then $\mathcal{C}(\mathcal{E}) = \mathcal{BL}(V)$. Similarly, $\mathcal{C}(\mathcal{BL}(V))$ is the set of scalar multiples of the identiy operator.

It is easy to check that $\mathcal{C}(\mathcal{E})$ is a subalgebra of $\mathcal{BL}(V)$ that contains the identity and is closed in the weak operator topology. If $\mathcal{A}$ is the algebra generated by $\mathcal{E}$ in $\mathcal{BL}(V)$, then $\mathcal{C}(\mathcal{A}) = \mathcal{C}(\mathcal{E})$. Similarly, the commutant of the closure of $\mathcal{E}$ in the weak operator topology is the same as $\mathcal{C}(\mathcal{E})$. If $\mathcal{E}_1 \subseteq \mathcal{E}_2$, then

$$\mathcal{C}(\mathcal{E}_2) \subseteq \mathcal{C}(\mathcal{E}_1). \tag{44.2}$$

Suppose now that V is a complex Hilbert space. If $\mathcal{E} \subseteq \mathcal{BL}(V)$, then put

$$\mathcal{E}^* = \{A^* : A \in \mathcal{E}\}. \tag{44.3}$$

It is easy to see that

$$\mathcal{C}(\mathcal{E}^*) = \mathcal{C}(\mathcal{E})^* \tag{44.4}$$

for any $\mathcal{E} \subseteq \mathcal{BL}(V)$, since $(A \circ B)^* = B^* \circ A^*$ for every $A, B \in \mathcal{BL}(V)$. If $\mathcal{E}^* = \mathcal{E}$, then it follows that $\mathcal{C}(E)$ is a von Neumann algebra. In particular, $\mathcal{C}(\mathcal{A})$ is a von Neumann algebra when $\mathcal{A} \subseteq \mathcal{BL}(V)$ is a $*$-algebra.

45 Second commutants

Let V be a vector space equipped with a norm. The second commutant $\mathcal{C}^2(\mathcal{E})$ of $\mathcal{E} \subseteq \mathcal{BL}(V)$ is the commutant $\mathcal{C}(\mathcal{C}(\mathcal{E}))$ of the commutant $\mathcal{C}(\mathcal{E})$ of $\mathcal{E}$. Thus

$$\mathcal{E} \subseteq \mathcal{C}^2(\mathcal{E}) \tag{45.1}$$

automatically, and one can check that

$$\mathcal{C}(\mathcal{C}^2(\mathcal{E})) = \mathcal{C}(\mathcal{E}). \tag{45.2}$$

If V is a complex Hilbert space, then von Neumann's *double commutant theorem* states that

$$\mathcal{C}^2(\mathcal{A}) = \mathcal{A} \tag{45.3}$$

when $\mathcal{A} \subseteq \mathcal{BL}(V)$ is a von Neumann algebra. If $\mathcal{A}$ is any $*$-algebra that contains the identity, then the theorem states that every element of $\mathcal{C}^2(\mathcal{A})$ is contained in the closure of $\mathcal{A}$ with respect to the strong operator topology. Hence $\mathcal{C}^2(\mathcal{A})$ is the same as the closure of $\mathcal{A}$ in $\mathcal{BL}(V)$ in the strong and weak operator topologies in this case.

Here are some of the ingredients in the proof. If T is a bounded linear operator on V and W is a closed linear subspace of V such that

$$T(W) \subseteq W, \tag{45.4}$$

then

$$T^*(W^\perp) \subseteq W^\perp, \tag{45.5}$$

where $W^\perp$ is the orthogonal complement of W in V. If $T^*(W) \subseteq W$, then it follows that $T(W^\perp) \subseteq W^\perp$. If P_W is the orthogonal projection of V onto W, then

$$T \circ P_W = P_W \circ T \tag{45.6}$$

if and only if $T(W) \subseteq W$ and $T(W^\perp) \subseteq W^\perp$.

Suppose that W is invariant under the $*$-algebra $\mathcal{A} \subseteq \mathcal{BL}(V)$, in the sense that $T(W) \subseteq W$ for every $T \in \mathcal{A}$. Because $\mathcal{A}$ is a $*$-algebra, $T^*(W) \subseteq W$ for every $T \in \mathcal{A}$, and hence $T(W^\perp) \subseteq W^\perp$ for every $T \in \mathcal{A}$. As in the previous paragraph, this implies that every element of $\mathcal{A}$ commutes with P_W, or

$$P_W \in \mathcal{C}(\mathcal{A}). \tag{45.7}$$

By definition of the second commutant, P_W commutes with every element of $\mathcal{C}^2(\mathcal{A})$, which implies that W and W^* are also invariant under $\mathcal{C}^2(\mathcal{A})$.

46 Invertibility

Let V be a real or complex vector space equipped with a norm $\|v\|$, and T be a bounded linear operator on V. By definition, T is invertible if there is a bounded linear operator T^{-1} on V such that

$$T^{-1} \circ T = T \circ T^{-1} = I, \tag{46.1}$$

where I is the identity operator on V. In particular, this implies that

$$\|v\| = \|T^{-1}(T(v))\| \leq \|T^{-1}\|_{op} \, \|T(v)\| \tag{46.2}$$

for every $v \in V$.

Conversely, suppose that there is a $\delta > 0$ such that

$$\|T(v)\| \geq \delta \, \|v\| \quad \text{for every } v \in V. \tag{46.3}$$

In this case, the kernel of T is trivial, and so T is one-to-one. Moreover, T is invertible as a bounded linear mapping from V onto $T(V)$, and the norm of the inverse of T as a linear mapping from $T(V)$ onto V is less than or equal to $1/\delta$. However, T may not be invertible as a linear operator on V, because $T(V)$ may be a proper linear subspace of V. For example, if T is the forward shift operator on $V = \ell^p(\mathbf{Z}_+)$ or $c_0(\mathbf{Z}_+)$, then T is an isometry of V onto $T(V) \neq V$.

If V is complete and T satisfies (46.3), then $T(V)$ is a closed linear subspace of V, basically because $T(V)$ is complete. For if $\{v_j\}_{j=1}^\infty$ is a sequence of vectors in V such that $\{T(v_j)\}_{j=1}^\infty$ converges in V, then (46.3) implies that $\{v_j\}_{j=1}^\infty$ is a Cauchy sequence in V, which converges to some $v \in V$ by completeness. Thus $\{T(v_j)\}_{j=1}^\infty$ converges to $T(v)$, which is in $T(V)$, as desired. If V is complete, T satisfies (46.3), and $T(V)$ is dense in V, then it follows that $T(V) = V$, T is invertible as a linear operator on V, and

$$\|T^{-1}\|_{op} \leq \frac{1}{\delta}. \tag{46.4}$$

If V is a Hilbert space, then the orthogonal complement of $T(V)$ in V is the same as the kernel of the adjoint T^* of T. Hence $T(V)$ is dense in V when the kernel of T^* is trivial. If T is self-adjoint and satisfies (46.3), then T is invertible, since the kernel of $T^* = T$ is trivial. Note that T^{-1} is self-adjoint when T is invertible and self-adjoint, because $(T^{-1})^* = (T^*)^{-1} = T^{-1}$.

47 Positivity

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let A be a bounded nonnegative self-adjoint linear operator on V. Let us say that A is *uniformly strictly positive* if there is a $\delta > 0$ such that

$$\langle A(v), v \rangle \geq \delta \, \|v\|^2 \tag{47.1}$$

for every $v \in V$. Using the Cauchy–Schwarz inequality, this implies that

$$\|A(v)\| \geq \delta \, \|v\| \tag{47.2}$$

for every $v \in V$, and it follows from the remarks in the previous section that A is invertible. Conversely, if A is invertible, then A is uniformly strictly positive, because of (34.3).

If A is invertible, then A^{-1} is also self-adjoint, as in the previous section. Let us check that A^{-1} is nonnegative too, so that

$$\langle A^{-1}(w), w \rangle \geq 0 \tag{47.3}$$

for every $w \in V$. Because A is invertible, it suffices to show that this holds when $w = A(v)$ for some $v \in V$. In this case,

$$\langle A^{-1}(w), w\rangle = \langle A^{-1}(A(v)), A(v)\rangle = \langle v, A(v)\rangle = \langle A(v), v\rangle \geq 0, \tag{47.4}$$

as desired.

If T is any bounded linear operator on V, then $T^* \circ T$ is self-adjoint and nonnegative, as in Section 33. If T is invertible, then T^* is invertible, and hence $T^* \circ T$ is invertible. More precisely, $T^* \circ T$ is invertible when T satisfies (46.3), because

$$\langle (T^* \circ T)(v), v\rangle = \langle T(v), T(v)\rangle = \|T(v)\|^2 \tag{47.5}$$

implies that $T^* \circ T$ is uniformly strictly positive. For example, $T^* \circ T = I$ when T is the forward shift operator on $\ell^2(\mathbf{Z}_+)$, even though T is not invertible on ℓ^2. In this example, T^* is the backward shift operator, and $T \circ T^*$ is nonnegative and self-adjoint but not invertible.

Similarly, if A, T are bounded self-adjoint linear operators on V and A is nonnegative, then $T^* \circ A \circ T$ is self-adjoint and nonnegative, because

$$(T^* \circ A \circ T)^* = T^* \circ A^* \circ (T^*)^* = T^* \circ A \circ T \tag{47.6}$$

and

$$\langle (T^* \circ A \circ T)(v), v\rangle = \langle A(T(v)), T(v)\rangle \tag{47.7}$$

for every $v \in V$. If A and T are also invertible, then $T^* \circ A \circ T$ is invertible too. If A is invertible and T satisfies (46.3), then $T^* \circ A \circ T$ is uniformly strictly positive, and hence invertible. If A is invertible and $T = A^{-1}$, then $T^* \circ A \circ T = A^{-1}$, which gives another way to look at the positivity of A^{-1}.

48 Invertibility, 2

Let $(V, \|v\|)$ be a real or complex Banach space, and suppose that T is a bounded linear operator on V with $\|T\|_{op} < 1$. For each $v \in V$,

$$\|v\| \leq \|v - T(v)\| + \|T(v)\| \leq \|v - T(v)\| + \|T\|_{op}\, \|v\|, \tag{48.1}$$

which implies that

$$\|v\| \leq \frac{1}{1 - \|T\|_{op}}\, \|v - T(v)\|. \tag{48.2}$$

In particular, $I - T$ is one-to-one on V, because it has trivial kernel. Let us show that $I - T$ maps V onto itself, and hence is invertible.

Consider the infinite series

$$\sum_{j=0}^{\infty} T^j(v), \tag{48.3}$$

where T^0 is interpreted as the identity operator I on V. This series converges absolutely, in the sense that $\sum_{j=0}^{\infty} \|T^j(v)\|$ converges as an infinite series of nonnegative real numbers, because

$$\sum_{j=0}^{\infty} \|T^j(v)\| \le \sum_{j=0}^{\infty} \|T\|_{op}^j \, \|v\| = \frac{1}{1 - \|T\|_{op}} \|v\|. \tag{48.4}$$

As in the context of series of real or complex numbers, one can use this to show that the partial sums $\sum_{j=0}^{n} T^j(v)$ form a Cauchy sequence in V, which converges by completeness. This means that the infinite series converges in V, and it is easy to see that the sum satisfies

$$(I - T)\Big(\sum_{j=0}^{\infty} T^j(v)\Big) = \sum_{j=0}^{\infty} T^j(v) - \sum_{j=0}^{\infty} T^{j+1}(v) = v. \tag{48.5}$$

Thus $I - T$ maps V onto V, so that $I - T$ is invertible on V, with

$$\|(I - T)^{-1}\|_{op} \le \frac{1}{1 - \|T\|_{op}}. \tag{48.6}$$

Alternatively, one can consider

$$\sum_{j=0}^{\infty} T^j \tag{48.7}$$

as an infinite series in $\mathcal{BL}(V)$, which converges absolutely because

$$\sum_{j=0}^{\infty} \|T^j\|_{op} \le \sum_{j=0}^{\infty} \|T\|_{op}^j = \frac{1}{1 - \|T\|_{op}^j}. \tag{48.8}$$

Again this implies that the partial sums $\sum_{j=0}^{\infty} T^j$ form a Cauchy sequence in $\mathcal{BL}(V)$ with respect to the operator norm, which converges by completeness. As usual,

$$(I - T)\left(\sum_{j=0}^{n} T^j\right) = \left(\sum_{j=0}^{n} T^j\right)(I - T) = I - T^{n+1} \tag{48.9}$$

for each nonnegative integer n, and

$$\|T^{n+1}\|_{op} \le \|T\|_{op}^{n+1} \to 0 \tag{48.10}$$

as $n \to \infty$, since $\|T\|_{op} < 1$. This shows that

$$(I - T)^{-1} = \sum_{j=0}^{\infty} T^j. \tag{48.11}$$

Suppose now that $(V, \langle v, w \rangle)$ is a Hilbert space, and that A is a bounded nonnegative self-adjoint linear operator on V which is also invertible. It will be convenient to make the normalization $\|A\|_{op} = 1$, which can always be arranged by multiplying A by a positive real number. As in the previous section, A is uniformly strictly positive, so that there is a $\delta > 0$ such that

$$\delta \, \|v\|^2 \le \langle A(v), v \rangle \le \|v\|^2 \tag{48.12}$$

for every $v \in V$. If $B = I - A$, then B is a bounded self-adjoint linear operator on V such that

$$0 \le \langle B(v), v \rangle \le (1 - \delta) \, \|v\|^2 \tag{48.13}$$

for every $v \in V$. Thus B is nonnegative and $\|B\|_{op} \le 1 - \delta$, by (34.2). Because $A = I - B$, the invertibility of A can also be seen as an instance of the discussion in the preceding paragraphs. In particular, $A^{-1} = \sum_{j=0}^{\infty} B^j$.

49 Invertibility, 3

Let V be a real or complex vector space equipped with a norm, and let $\mathcal{E}$ be a collection of bounded linear operators on V. If $A \in \mathcal{E}$ is invertible, then

$$A^{-1} \in \mathcal{C}^2(\mathcal{E}). \tag{49.1}$$

For if $T \in \mathcal{C}(\mathcal{E})$, then T commutes with A, and hence T commutes with A^{-1}.

Suppose that $\mathcal{A}$ is an algebra of bounded linear operators on V that contains the identity operator. If T is a bounded linear operator on V with $\|T\|_{op} < 1$ and V is complete, then $I - T$ is invertible on V, as in the previous section. If $T \in \mathcal{A}$ and $\mathcal{A}$ is closed in the topology determined by the operator norm, then

$$(I - T)^{-1} \in \mathcal{A}. \tag{49.2}$$

This is because $\sum_{j=0}^{n} T^j$ converges to $(I - T)^{-1}$ in the operator norm as $n \to \infty$.

Suppose now that V is a Hilbert space, and that $\mathcal{A}$ is a $*$-algebra of bounded linear operators on V that contains the identity operator and is closed with respect to the operator norm. If A is a bounded nonnegative self-adjoint linear operator on V which is invertible, and if $A \in \mathcal{A}$, then $A^{-1} \in \mathcal{A}$, by the remarks at the end of the previous section. If T is any element of $\mathcal{A}$, then $T^* \in \mathcal{A}$, and hence $T^* \circ T \in \mathcal{A}$. If T is invertible, then $T^* \circ T$ is invertible, and the preceding observation implies that $(T^* \circ T)^{-1} \in \mathcal{A}$. Hence

$$T^{-1} = (T^* \circ T)^{-1} \circ T^* \in \mathcal{A}. \tag{49.3}$$

If U is a unitary transformation on V, then $U^{-1} = U^*$. Thus $U^{-1} \in \mathcal{A}$ whenever $\mathcal{A}$ is a $*$-algebra that contains U. Also, $U^* \in \mathcal{C}^2(\mathcal{E})$ whenever $U \in \mathcal{E} \subseteq \mathcal{BL}(V)$, by (49.1).

50 The Hardy space

In this section, we are concerned with complex-valued functions on the unit circle $\mathbf{T}$, as in Section 29. The Hardy space $H^2 = H^2(\mathbf{T})$ may be defined as the set of $f \in L^2(\mathbf{T})$ such that

$$\int_{\mathbf{T}} f(z)\, z^j \, |dz| = 0 \tag{50.1}$$

for every positive integer j. Equivalently, $f \in H^2$ when $f \in L^2$ is orthogonal to the functions $z^{-j} = \overline{z}^j$, $j \in \mathbf{Z}_+$, with respect to the standard integral inner product on $\mathbf{T}$. Alternatively, H^2 is the closure of the linear span of z^j, $j \ge 0$, in L^2.

Every $f \in L^2$ corresponds to a Fourier series

$$\sum_{j=-\infty}^{\infty} a_j \, z^j, \tag{50.2}$$

where $\{a_j\}_{j=-\infty}^{\infty} \in \ell^2(\mathbf{Z})$. If $f \in H^2$, then this reduces to

$$\sum_{j=0}^{\infty} a_j \, z^j, \tag{50.3}$$

which also defines a convergent power series for $|z| < 1$. Thus elements of H^2 have natural extensions to holomorphic functions on the unit disk in $\mathbf{C}$, and this is another way to characterize the Hardy space.

Let $H^\infty = H^\infty(\mathbf{T})$ be the set of $f \in L^\infty(\mathbf{T})$ that satisfy (50.1) for every $j \in \mathbf{Z}_+$, which is the same as the intersection of L^∞ with H^2. This description in terms of integrals shows that H^∞ is a closed linear subspace of L^∞ with respect to the weak* topology, when L^∞ is identified with the dual of L^1. Let us check that

$$b\, f \in H^2 \tag{50.4}$$

whenever $f \in H^2$ and $b \in H^\infty$. Of course, $b\, f \in L^2$ when $f \in L^2$ and $b \in L^\infty$, and $f \mapsto b\, f$ is a bounded linear operator on L^2. If $b \in H^\infty$ and $f(z) = z^l$ for some $l \ge 0$, then it is easy to see that $b\, f \in H^2$. The same conclusion holds for any $f \in H^2$ by linearity and continuity, since H^2 is the closed linear span of the z^l's, $l \ge 0$. It follows too that $b\, f \in H^\infty$ when $b, f \in H^\infty$.

Let $\mathcal{A}$ be the algebra of bounded linear operators on L^2 corresponding to multiplication by elements of H^∞. This is a subalgebra of the algebra $\mathcal{B}$ of bounded linear operators on L^2 corresponding to multiplication by bounded measurable functions. We have already see that $\mathcal{B}$ is a closed subalgebra of the algebra of bounded linear operators on L^2 with respect to the weak operator topology, and it is easy to see that $\mathcal{A}$ is closed in the weak operator topology as well. However, $\mathcal{A}$ is not a $*$-algebra, because H^∞ is not invariant under complex conjugation. For example, if $b(z) = z$, then $b \in H^\infty$ and $1/b(z) = \overline{z}$ is an element of L^∞ but not H^∞. Multiplication by b defines an element of $\mathcal{A}$ that is also a unitary transformation on L^2, but whose inverse is not in $\mathcal{A}$. By the remarks in the previous paragraph, $\mathcal{A}$ can also be described as the subalgebra of $\mathcal{B}$ consisting of operators that map H^2 into itself.

51 Fourier series, 2

If f is an integrable complex-valued function on the unit circle $\mathbf{T}$ and j is an integer, then the jth Fourier coefficient of f is defined by

$$\widehat{f}(j) = \frac{1}{2\pi} \int_{\mathbf{T}} f(w)\, w^{-j}\, |dw|. \tag{51.1}$$

One can also think of this as the inner product of f with w^j, especially when $f \in L^2$. Note that

$$|\widehat{f}(j)| \le \frac{1}{2\pi} \int_{\mathbf{T}} |f(w)|\, |dw| \tag{51.2}$$

for each j, so that the Fourier coefficients of an integrable function are bounded.

Formally, the corresponding Fourier series is given by

$$\sum_{j=-\infty}^{\infty} \widehat{f}(j)\, z^j. \tag{51.3}$$

This is the same as

$$\sum_{j=0}^{\infty} \widehat{f}(j)\, z^j + \sum_{j=1}^{\infty} \widehat{f}(-j)\, \overline{z}^j \tag{51.4}$$

on the unit circle. However, an advantage of (51.4) is that these series automatically converge absolutely when $|z| < 1$, because of the boundedness of the Fourier coefficients.

The first series in (51.4) defines a holomorphic function on the unit disk, and the second series defines a conjugate-holomorphic function. In particular, their sum is a harmonic function on the unit disk. The second series is equal to 0 for every z with $|z| < 1$ if and only if $\widehat{f}(-j) = 0$ for each $j \ge 1$. Equivalently, (51.4) is holomorphic on the unit disk if and only if $\widehat{f}(-j) = 0$ when $j \ge 1$.

An advantage of (51.3) is that multiplication by z clearly corresponds to shifting the Fourier coefficients. If $\widehat{f}(-j) = 0$ for $j \ge 1$, then this extends to $|z| < 1$, but this does not work using (51.4) for arbitrary functions when $|z| < 1$. Similarly, if $\widehat{f_1}(-j) = \widehat{f_2}(-j) = 0$ for $j \ge 1$, then the product of f_1 and f_2 on the unit circle corresponds to the product of the corresponding power series on the unit disk under suitable conditions. This is related to the fact that the product of two holomorphic functions is holomorphic, while the product of harmonic functions is not normally harmonic.

52 Multiplication on H^2

If b is a bounded measurable function on the unit circle, then $f \mapsto b\, f$ defines a bounded linear operator on $L^2(\mathbf{T})$, with operator norm equal to the L^∞ norm of b. If $b \in H^\infty(\mathbf{T})$, then $H^2(\mathbf{T})$ is an invariant linear subspace of this operator, as in Section 50. In this case, we can also think of $f \mapsto b\, f$ as a bounded linear operator on H^2. Let us check that the operator norm of multiplication by $b \in H^\infty$ is the same on H^2 as on L^2. Of course, the operator norm on H^2 is automatically less than or equal to the operator norm on L^2.

By definition, the operator norm on H^2 is the supremum of

$$\Big(\frac{1}{2\pi}\int_{\mathbf{T}} |b(z)|^2\,|f(z)|^2\,|dz|\Big)^{1/2} \tag{52.1}$$

over $f \in H^2$ with

$$\Big(\frac{1}{2\pi}\int_{\mathbf{T}} |f(z)|^2\,|dz|\Big)^{1/2} \le 1. \tag{52.2}$$

Similarly, the operator norm on L^2 is equal to the supremum of (52.1) over $f \in L^2$ that satisfy (52.2). For the latter, we may as well restrict our attention to functions f given by finite sums of the form

$$f(z) = \sum_{j=-n}^{n} a_j\, z^j \tag{52.3}$$

since these functions are dense in L^2. If f is of this form, then

$$f_1(z) = z^n\, f(z) \in H^2, \tag{52.4}$$

and $|f_1(z)|^2 = |f(z)|^2$. Hence (52.1) and (52.2) are the same for f_1 as for f, and it follows that the operator norms on L^2 and H^2 are the same.

Suppose that A is a bounded linear operator on H^2 that commutes with multiplication by b for every $b \in H^\infty$. Actually, it suffices to ask that A commute with multiplication by z. If $a = A(1)$, then $a \in H^2$, and $A(f) = a\,f$ whenever f is a finite linear combination of the z^j's. Using the boundedness of A, one can show that $a \in H^\infty$, and that $A(f) = a\,f$ for every $f \in H^2$.

In particular, this implies that the algebra $\mathcal{A}_1$ of operators on H^2 defined by multiplication by functions in H^∞ is a closed subalgebra of the algebra of bounded linear operators on H^2 in the weak operator topology. This is a bit different from the situation discussed in Section 50, where the multiplication operators were considered as acting on L^2 instead of H^2. However, $\mathcal{A}_1$ is still not a von Neumann algebra, because it is not a $*$-algebra. Note that multiplication by $a \in H^\infty$ is invertible as an operator on H^2 if and only if $1/a \in H^\infty$, in contrast with the situation in Section 50.

53 The Poisson kernel

For the sake of completeness, let us briefly review some aspects of the Poisson kernel. This is defined for $z, w \in \mathbf{C}$ with $|z| < 1$ and $|w| = 1$ by

$$P(z,w) = \frac{1}{2\pi}\Big(\sum_{j=0}^{\infty} z^j\,\overline{w}^j + \sum_{j=1}^{\infty} \overline{z}^j\, w^j\Big). \tag{53.1}$$

Thus

$$\int_{\mathbf{T}} P(z,w)\, f(w)\, |dw| \tag{53.2}$$

is equal to (51.4) for every integrable function f on the unit circle. There is no problem with interchanging the order of summation and integration here, because the series in (53.1) converge uniformly in w for each $z \in \mathbf{C}$ with $|z| < 1$. In particular,

$$\int_{\mathbf{T}} P(z,w)\, |dw| = 1 \tag{53.3}$$

for every $z \in \mathbf{C}$ with $|z| < 1$.

Observe that

$$\sum_{j=0}^{\infty} z^j\, \overline{w}^j + \sum_{j=1}^{\infty} \overline{z}^j\, w^j = 2\, \mathrm{Re} \sum_{j=0}^{\infty} z^j\, \overline{w}^j - 1, \tag{53.4}$$

where $\mathrm{Re}\, a$ is the real part of a complex number a. Of course,

$$\sum_{j=0}^{\infty} z^j\, \overline{w}^j = \frac{1}{1 - z\,\overline{w}} = \frac{1 - \overline{z}\, w}{|1 - z\,\overline{w}|^2}, \tag{53.5}$$

and so

$$\begin{aligned} 2\sum_{j=0}^{\infty} z^j\, \overline{w}^j - 1 &= \frac{2 - 2\,\overline{z}\, w - |1 - z\,\overline{w}|^2}{|1 - z\,\overline{w}|^2} \\ &= \frac{1 - \overline{z}\, w + z\,\overline{w} - |z|^2}{|1 - z\,\overline{w}|^2}, \end{aligned} \tag{53.6}$$

using the fact that $|w| = 1$ in the last step. Hence

$$P(z,w) = \frac{1}{2\,\pi} \frac{1 - |z|^2}{|1 - z\,\overline{w}|^2}, \tag{53.7}$$

because $\mathrm{Re}(z\,\overline{w} - \overline{z}\, w) = 0$.

This formula implies that

$$P(z,w) \ge 0 \tag{53.8}$$

for every $z, w \in \mathbf{C}$ with $|z| < 1$ and $|w| = 1$, and that

$$P(z,w) \to 0 \tag{53.9}$$

as z tends to any point on the unit circle other than w, with uniform convergence outside of any neighborhood of w. If f is a continuous function on the unit circle, then one can use these properties of the Poisson kernel to show that (53.2) converges to $f(\zeta)$ as $z \to \zeta$ when $|\zeta| = 1$. Thus the Poisson integral (53.2) defines a continuous extension of f to the closed unit disk.

Let $P_r(f)(\zeta)$ be the Poisson integral (53.2) of f evaluated at $z = r\,\zeta$, where $0 \le r < 1$ and $|\zeta| = 1$. If f is a continuous function on the unit circle, then

$P_r(f)$ converges to f uniformly as $r \to 1$, because of uniform continuity. If $f \in L^p$, $1 \le p \le \infty$, then one can show that

$$\|P_r(f)\|_p \le \|f\|_p \tag{53.10}$$

for every $r \in [0, 1)$. This is straightforward when $p = 1$ or ∞, and otherwise can be seen as an integrated version of the triangle inequality. If $p < \infty$, then continuous functions are dense in L^p, and one can show that $P_r(f)$ converges to f in the L^p norm as $r \to 1$.

54 Fourier series, 3

The identification between the Poisson integral (53.2) and the series (51.4) when $|z| < 1$ can be re-expressed as

$$P_r(f)(\zeta) = \sum_{j=-\infty}^{\infty} r^{|j|}\, \widehat{f}(j)\, \zeta^j \tag{54.1}$$

for $|\zeta| = 1$ and $0 \le r < 1$. Of course, the Fourier series associated to f is the same as the right side of (54.1) with $r = 1$. If $r < 1$, then the right side of (54.1) is known as the Abel sum of the Fourier series of f associated to r. The convergence of $P_r(f)$ to f as $r \to 1$ can then be rephrased as the convergence of the Abel sums of the Fourier series of f to f.

If f is continuous, then $P_r(f)$ converges to f uniformly on the unit circle as $r \to 1$. For each $r < 1$, the partial sums of the series on the right side of (54.1) converge uniformly on the unit circle, because of the boundedness of the Fourier coefficients of f. In particular, finite linear combinations of the ζ^j's are dense in $C(\mathbf{T})$ with respect to the supremum norm. Similarly, finite linear combinations of the ζ^j's are dense in $L^p(\mathbf{T})$ when $1 \le p < \infty$. This can be used to show that the ζ^j's form an orthonormal basis for $L^2(\mathbf{T})$, as in Section 29.

If $f \in L^2$, then the orthonormality of the ζ^j's implies that

$$\|P_r(f)\|_2 = \Big(\sum_{j=-\infty}^{\infty} r^{2|j|}\, |\widehat{f}(j)|^2\Big)^{1/2}. \tag{54.2}$$

This is clearly less than or equal to

$$\|f\|_2 = \Big(\sum_{j=-\infty}^{\infty} |\widehat{f}(j)|^2\Big)^{1/2} \tag{54.3}$$

for each $r < 1$. Thus (53.10) is an extension of this to any p.

55 Hardy spaces, continued

The Hardy space $H^p = H^p(\mathbf{T})$ can be defined for $1 \le p \le \infty$ as the closed linear subspace of $L^p(\mathbf{T})$ consisting of functions f such that

$$\widehat{f}(-j) = 0 \text{ for each } j \in \mathbf{Z}_+. \tag{55.1}$$

In this case,

$$P_r(f)(\zeta) = \sum_{j=0}^{\infty} r^j \, \widehat{f}(j) \, \zeta^j. \tag{55.2}$$

If $p < \infty$, then the convergence of $P_r(f)$ to f as $r \to 1$ in L^p and the uniform convergence of the series on the right side of (55.2) imply that each $f \in H^p$ can be approximated in the L^p norm by finite linear combinations of the ζ^j's with $j \geq 0$. Of course, $\zeta^j \in H^p$ for each $j \geq 0$, and it follows that H^p is the closed linear span of the ζ^j's, $j \geq 0$, in L^p when $p < \infty$.

If $f \in L^p$ and $g \in L^q$, then Hölder's inequality implies that their product $f \, g$ is in L^r and

$$\|f \, g\|_r \leq \|f\|_p \, \|g\|_q, \tag{55.3}$$

where

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q}. \tag{55.4}$$

If $f \in H^p$, $1 \leq p \leq \infty$, then $f(\zeta) \, \zeta^l \in H^p$ for each $l \geq 0$. Hence $f \, g \in H^r$ when $g \in H^q$, $1 \leq q < \infty$, and $r \geq 1$, where r is as in (55.4), as one can show by approximating g by linear combinations of the ζ^l's. This also works when $q = \infty$ and $p < \infty$, by approximating f by linear combinations of the ζ^l's. If $p = q = \infty$, then the same argument implies that $f \, g \in H^r$ for every $r < \infty$, and it follows that $f \, g \in H^\infty$, since $f \, g \in L^\infty$.

Similarly, let $A = A(\mathbf{T})$ be the set of continuous functions f on the unit circle that satisfy (55.1). The uniform convergence of $P_r(f)$ to f as $r \to 1$ and of the partial sums of the right side of (55.2) imply that each $f \in A$ can be approximated uniformly by finite linear combinations of the ζ^j's with $j \geq 0$. As before, $\zeta^j \in A$ for each $j \geq 0$, and so A is the same as the closed linear span of the ζ^j's, $j \geq 0$, in $C(\mathbf{T})$. If $f, g \in A$, then one can use these approximations to show that $f \, g \in A$, as in the previous paragraph. In particular, A is a closed subalgebra of $C(\mathbf{T})$.

56 Multiplication, continued

The Fourier coefficients of a products of two functions f, g on the unit circle are given formally by

$$\widehat{f \, g}(n) = \sum_{j=-\infty}^{\infty} \widehat{f}(n-j) \, \widehat{g}(j), \tag{56.1}$$

as one can see by multiplying the corresponding Fourier series for f, g. It is easy to verify this formula when one of $f(z)$, $g(z)$ is an integrable function on the unit circle and the other is z^l for some integer l, or a linear combination of finitely many z^l's. One can also check that this formula holds when $f, g \in L^2(\mathbf{T})$. In this case, $f \, g \in L^1(\mathbf{T})$, and the sum in (56.1) converges absolutely for each $n \in \mathbf{Z}$, because the Fourier coefficients of f, g are square-summable. Thus both sides of (56.1) make sense, and one can show that they are equal using the convergence of Fourier series in L^2.

Suppose that f, g are of analytic type, in the sense that

$$\widehat{f}(-j) = \widehat{g}(-j) = 0 \tag{56.2}$$

for every positive integer j. Thus (56.1) reduces to

$$\widehat{f\,g}(n) = \sum_{j=0}^{n} \widehat{f}(n-j)\,\widehat{g}(j), \tag{56.3}$$

and there is no problem about the convergence of the sum on the right. If $f \in H^p$, $g \in H^q$, $1/r = 1/p + 1/q$, and $r \ge 1$, then $f\,g \in H^r$, and one can get (56.3) using approximation arguments.

Remember that the *Cauchy product* of a pair of infinite series $\sum_{j=0}^{\infty} a_j$, $\sum_{l=0}^{\infty} b_l$ is the infinite series $\sum_{n=0}^{\infty} c_n$, where

$$c_n = \sum_{j=0}^{n} a_{n-j}\,b_j. \tag{56.4}$$

Formally,

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{j=0}^{\infty} a_j\Big)\Big(\sum_{l=0}^{\infty} b_l\Big), \tag{56.5}$$

and more precisely

$$\sum_{n=0}^{\infty} c_n\,z^n = \Big(\sum_{j=0}^{\infty} a_j\,z^j\Big)\Big(\sum_{l=0}^{\infty} b_l\,z^l\Big) \tag{56.6}$$

as a product of formal power series. If $\sum_{j=0}^{\infty} a_j$, $\sum_{l=0}^{\infty} b_l$ converge absolutely, then it is well known that $\sum_{n=0}^{\infty} c_n$ also converges absolutely and satisfies (56.5). In particular, if $\sum_{j=0}^{\infty} a_j\,z^j$ and $\sum_{l=0}^{\infty} b_l\,z^l$ converge absolutely when $|z| < 1$, then $(\sum_{n=0}^{\infty} c_n\,z^n$ also converges absolutely when $|z| < 1$ and satisfies (56.6).

If f, g are of analytic type, then we have the associated power series

$$\sum_{j=0}^{\infty} \widehat{f}(j)\,z^j, \quad \sum_{l=0}^{\infty} \widehat{g}(l)\,z^l \tag{56.7}$$

on the unit disk, and (56.3) says exactly that

$$\sum_{n=0}^{\infty} \widehat{f\,g}(n)\,z^n \tag{56.8}$$

is their Cauchy product. If $f \in H^p$, $g \in H^q$, and $1/r = 1/p + 1/q \le 1$, then $f\,g \in H^r$, and these power series converge absolutely on the unit disk. The product of the series in (56.7) is equal to the one in (56.8) when $|z| < 1$, as in the previous paragraph. Note that H^p spaces can be defined directly in terms of holomorphic functions on the unit disk for all $p > 0$, and have analogous properties in terms of products.

57 Convolution on Z

If $a = \{a_j\}_{j=-\infty}^{\infty}$, $b = \{b_j\}_{j=-\infty}^{\infty}$ are doubly-infinite sequences of complex numbers, then their convolution $a * b$ is the doubly-infinite sequence defined formally by

$$(a * b)_n = \sum_{j=-\infty}^{\infty} a_{n-j}\, b_j. \tag{57.1}$$

Let us say that a has *finite support* if $a_j = 0$ for all but finitely many j. If at least one of a, b has finite support and the other is arbitrary, then $a * b$ is defined. If both a and b have finite support, then $a * b$ has finite support too. If $a_{-j} = b_{-j} = 0$ for each $j \geq 1$, then $(a * b)_n$ is the same as the Cauchy product (56.4), and makes sense without additional restriction on a, b.

Suppose that $a \in \ell^p(\mathbf{Z})$ and $b \in \ell^q(\mathbf{Z})$, where $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$. By Hölder's inequality, the series in the definition of $(a*b)_n$ converges absolutely for each $n \in \mathbf{Z}$, and is uniformly bounded in n, with

$$\|a * b\|_\infty \leq \|a\|_p \, \|b\|_q. \tag{57.2}$$

If $1 < p, q < \infty$, then one can also check that

$$a * b \in c_0(\mathbf{Z}). \tag{57.3}$$

This works as well when one of a, b is in $\ell^1(\mathbf{Z})$ and the other is in $c_0(\mathbf{Z})$.

It is easy to see that convolution of sequences is commutative, which is to say that

$$a * b = b * a \tag{57.4}$$

when the convolutions are defined. This especially clear when the convolution is expressed by

$$(a * b)_n = \sum_{\substack{j,l \in \mathbf{Z} \\ j+l=n}} a_j\, b_l. \tag{57.5}$$

Of course, $a * b$ is also linear in a and b.

Similarly, convolution is associative, in the sense that

$$(a * b) * c = a * (b * c) \tag{57.6}$$

under suitable conditions so that the convolutions are defined. More precisely,

$$((a * b) * c)_n = (a * (b * c))_n = \sum_{\substack{j,k,l \in \mathbf{Z} \\ j+k+l=n}} a_j\, b_k\, c_l. \tag{57.7}$$

For instance, this certainly works when a, b, c have finite support, or at least two of them have finite support. It also works when $a_{-j} = b_{-j} = c_{-j} = 0$ for each $j \geq 1$.

Suppose that $a, b \in \ell^1(\mathbf{Z})$. Thus a, b are bounded, the series in the definition of $(a * b)_n$ converges absolutely for each $n \in \mathbf{Z}$, and

$$|(a * b)_n| \le \sum_{j=-\infty}^{\infty} |a_{n-j}|\,|b_j|. \tag{57.8}$$

This implies that

$$\sum_{n=-\infty}^{\infty} |(a * b)_n| \le \Big(\sum_{j=-\infty}^{\infty} |a_j|\Big)\Big(\sum_{l=-\infty}^{\infty} |b_l|\Big), \tag{57.9}$$

by interchanging the order of summation, so that $a * b \in \ell^1(\mathbf{Z})$ too, and

$$\|a * b\|_1 \le \|a\|_1\,\|b\|_1. \tag{57.10}$$

In particular, $(a * b) * c$ and $a * (b * c)$ are defined when $a, b, c \in \ell^1(\mathbf{Z})$, and one can show that they are the same, as in the previous paragraph. Hence $\ell^1(\mathbf{Z})$ is an associative algebra with respect to convolution.

Let $\delta(l) = \{\delta_j(l)\}_{j=-\infty}^{\infty}$ be defined for each $l \in \mathbf{Z}$ by

$$\begin{aligned} \delta_j(l) &= 1 \quad \text{when } j = l \\ &= 0 \quad \text{when } j \ne l. \end{aligned} \tag{57.11}$$

If a is any doubly-infinite sequence of complex numbers, then

$$(a * \delta(l))_n = (\delta(l) * a)_n = a_{n-l} \tag{57.12}$$

for each $l, n \in \mathbf{Z}$. Thus $a * \delta(l)$ is the same as a shifted l steps. In particular,

$$a * \delta(0) = \delta(0) * a = a. \tag{57.13}$$

This shows that $\delta(0)$ is the multiplicative identity element in $\ell^1(\mathbf{Z})$ as an algebra with respect to convolution.

Suppose that $a \in \ell^p(\mathbf{Z})$, $1 \le p \le \infty$, and $b \in \ell^1(\mathbf{Z})$. Thus a is bounded, and $a * b$ is well defined. It is not too difficult to show that $a * b \in \ell^p(\mathbf{Z})$, with

$$\|a * b\|_p \le \|a\|_p\,\|b\|_1. \tag{57.14}$$

This basically follows from the triangle inequality for the ℓ^p norm, since $a * \delta(l)$ is a shift of a for each l, and hence has the same ℓ^p norm as a.

58 Convolution on Z, 2

If $a \in \ell^(\mathbf{Z})$, then the Fourier transform of a is the function on the unit circle defined by

$$\widehat{a}(z) = \sum_{j=-\infty}^{\infty} a_j\, z^j. \tag{58.1}$$

More precisely, this series converges uniformly to a continuous function on the unit circle, and

$$\sup_{z \in \mathbf{T}} |\widehat{a}(z)| \le \|a\|_1. \tag{58.2}$$

Using the orthonormality of the z^j's with respect to the standard integral inner product on the unit circle, one can check that the jth Fourier coefficient of $\widehat{a}$ is equal to a_j for each $j \in \mathbf{Z}$.

If f is a function on the unit circle whose Fourier coefficients $\widehat{f}(j)$ happen to be summable, then the Fourier transform of $\{\widehat{f}(j)\}_{j=-\infty}^{\infty}$ is given by the Fourier series associated to f, and is the same as f. This follows from the convergence of the Abel sums of the Fourier series of f to f, as in Section 54. More precisely, the Abel sums converge to the ordinary sum when the series converges.

If $a, b \in \ell^1(\mathbf{Z})$, then

$$(\widehat{a * b})(z) = \widehat{a}(z)\,\widehat{b}(z) \tag{58.3}$$

for each $z \in \mathbf{T}$. Indeed,

$$(\widehat{a * b})(z) = \sum_{n=-\infty}^{\infty} (a * b)_n \, z^n \tag{58.4}$$

can be expanded into

$$\sum_{n=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a_{n-j}\, b_j\, z^n = \sum_{n=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (a_{n-j}\, z^{n-j})\,(b_j\, z^j). \tag{58.5}$$

Because of absolute convergence, we can interchange the order of summation, to get that this is equal to

$$\sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} (a_{n-j}\, z^{n-j})\,(b_j\, z^j) = \sum_{j=-\infty}^{\infty} \widehat{a}(z)\, b_j\, z^j = \widehat{a}(z)\,\widehat{b}(z), \tag{58.6}$$

as desired. Of course, this is basically the same as multiplying the series associated to $\widehat{a}$, $\widehat{b}$ and collecting terms, as in Section 56.

If $a \in \ell^2(\mathbf{Z})$, then the series in the definition of the Fourier transform of a converges in $L^2(\mathbf{T})$, because of the orthonormality of the z^j's. The latter also implies that

$$\Big(\frac{1}{2\pi} \int_{\mathbf{T}} |\widehat{a}(z)|^2 \, |dz|\Big)^{1/2} = \|a\|_2. \tag{58.7}$$

As before, the jth Fourier coefficient of $\widehat{a}$ is equal to a_j for each $j \in \mathbf{Z}$.

59 Convolution on T

The convolution of two functions f, g on the unit circle is defined by

$$(f * g)(z) = \frac{1}{2\pi} \int_{\mathbf{T}} f(z\, w^{-1})\, g(w)\, |dw|. \tag{59.1}$$

This actually makes sense as an integrable function on $\mathbf{T}$ when f, g are integrable functions, because

$$\int_{\mathbf{T}}\int_{\mathbf{T}} |f(z\,w^{-1})|\,|g(w)|\,|dw|\,|dz| = \int_{\mathbf{T}}\int_{\mathbf{T}} |f(z\,w^{-1})|\,|g(w)|\,|dz|\,|dw| \tag{59.2}$$

by Fubini's theorem, which reduces to

$$\Big(\int_{\mathbf{T}} |f(z)|\,|dz|\Big)\Big(\int_{\mathbf{T}} |g(w)|\,|dw|\Big), \tag{59.3}$$

since Lebesgue measure on $\mathbf{T}$ is invariant under rotations. It is easy to see that convolution is commutative, using the change of variables $w \mapsto z\,w^{-1}$ in the integral defining $(f * g)(z)$. One can also check that convolution is associative, directly from the definition.

Let us check that the Fourier coefficients of $f*g$ are the same as the product of the Fourier coefficients of f, g, which is to say that

$$\widehat{(f * g)}(j) = \widehat{f}(j)\,\widehat{g}(j) \tag{59.4}$$

for each $j \in \mathbf{Z}$. By definition,

$$\begin{aligned} \widehat{(f * g)}(j) &= \frac{1}{2\pi}\int_{\mathbf{T}} (f * g)(z)\, z^{-j}\,|dz| \\ &= \frac{1}{(2\pi)^2}\int_{\mathbf{T}}\int_{\mathbf{T}} f(z\,w^{-1})\,g(w)\,z^{-j}\,|dw|\,|dz|. \end{aligned} \tag{59.5}$$

Using Fubini's theorem again, this can be re-expressed as

$$\frac{1}{(2\pi)^2}\int_{\mathbf{T}}\int_{\mathbf{T}} f(z\,w^{-1})\,(z\,w^{-1})^{-j}\,g(w)\,w^{-j}\,|dz|\,|dw|, \tag{59.6}$$

which reduces to

$$\Big(\frac{1}{2\pi}\int_{\mathbf{T}} f(z)\,z^{-j}\,|dz|\Big)\Big(\frac{1}{2\pi}\int_{\mathbf{T}} g(w)\,w^{-j}\,|dw|\Big) = \widehat{f}(j)\,\widehat{g}(j). \tag{59.7}$$

It is convenient to take the L^p norm of a function $f \in L^p(\mathbf{T})$, $1 \le p < \infty$, to be

$$\|f\|_p = \Big(\frac{1}{2\pi}\int_{\mathbf{T}} |f(z)|^p\,|dz|\Big)^{1/p}. \tag{59.8}$$

The L^∞ norm can be taken to be the same as usual. If $f \in L^p$, $g \in L^q$, and $1/p + 1/q = 1$, then the convolution $f * g$ is bounded, with

$$\|f * g\|_\infty \le \|f\|_p\,\|g\|_q, \tag{59.9}$$

by Hölder's inequality. One can also check that $f * g$ is a continuous function on the unit circle in this case. If $f \in L^p$, $1 \le p \le \infty$, and $g \in L^1$, then one can show that $f * g \in L^p$, and that

$$\|f * g\|_p \le \|f\|_p\,\|g\|_1. \tag{59.10}$$

This follows from Fubini's theorem as at the beginning of the section when $p = 1$, and is straightforward when $p = \infty$. Otherwise, this is basically an integrated version of the triangle inequality for the L^p norm, since $f(z\,w^{-1})$ is in L^p as a function of z and with the same norm as f for each $w \in \mathbf{T}$.

Put

$$p_r(z) = \frac{1 - r^2}{|1 - r\,z|^2} \tag{59.11}$$

for $z \in \mathbf{T}$ and $0 \le r < 1$. If f is an integrable function on the unit circle, then

$$P_r(f) = p_r * f, \tag{59.12}$$

where $P_r(f)$ is the Poisson integral of f, as in Section 53.

Part II
Topological groups

60 Topological groups

A *topological group* is a group G equipped with a topology such that the group operations are continuous. It is customary to ask that the set consisting of only the identity element e be a closed set, which implies that all one-element sets in the group are closed, by continuity of translations. One can use continuity of the group operations to show that G is also Hausdorff, and even regular.

If $a, b \in G$ and $E \subseteq G$, then we put

$$a\,E = \{a\,x : x \in E\}, \quad E\;b = \{x\,b : x \in E\}. \tag{60.1}$$

Similarly, if $A, B \subseteq G$, then we put

$$A\,B = \{a\,b : a \in A,\, b \in B\}. \tag{60.2}$$

Thus

$$A\,B = \bigcup_{a \in A} a\,B = \bigcup_{b \in B} A\,b, \tag{60.3}$$

and it follows that $A\,B$ is an open set whenever A or B is an open set.

Let U be an open set that contains e. For any set E, $E\,U$ and $U\,E$ are open sets that contain E, and

$$\overline{E} \subseteq E\,U,\; U\,E. \tag{60.4}$$

Moreover, $\overline{E}$ is equal to the intersection of $E\,U$ or $U\,E$ over all neighborhoods U of E.

Suppose that H is a subgroup of G that is also an open set. Hence the cosets $a\,H$ and $H\,b$ of H are open sets too. By elementary group theory, the complement of H is a union of cosets. It follows that H is automatically a closed set as well.

61 Locally compact groups

We shall be primarily concerned here with topological groups G that are locally compact. Because of continuity of translations, local compactness is equivalent to asking that there be an open set U such that $e \in U$ and $\overline{U}$ is compact. We may also ask that U be symmetric in the sense that $x^{-1} \in U$ when $x \in U$, since otherwise we can take the intersection of U with $U^{-1} = \{x^{-1} : x \in U\}$.

For each positive integer n, let U^n be the set of products $x_1 \cdots x_n$ of n elements $x_1, \ldots, x_n$ of U. This is the same as the product $U \cdots U$ of n U's, in the notation of the previous section. Thus U^n is an open set for each n. Also,

$$\overline{U^n} \subseteq U^{n+1} \tag{61.1}$$

for every n.

If A, B are compact sets, then $A\,B$ is compact as well. This follows from continuity of the group operation, and compactness of $A \times B$ in the product topology. Hence the product K_n of n $\overline{U}$'s is compact for each n. Of course, $U_n \subseteq K_n$, and so

$$\overline{U^n} \subseteq K_n \tag{61.2}$$

is compact too. Note that continuity of the group operation implies

$$K_n \subseteq \overline{U^n}. \tag{61.3}$$

It is easy to see that

$$H = \bigcup_{n=1}^{\infty} U^n. \tag{61.4}$$

is a subgroup of G. This is also an open set in G, since U is open. Moreover, H is σ-compact, meaning that it is the union of a sequence of compact sets, by the remarks in the previous paragraph.

62 Uniform continuity

Let G be a topological group, let f be a continuous real or complex-valued function on G, and let K be a compact set in G. Also let $\epsilon > 0$ be given. For each $x \in K$, there is an open set $U(x)$ such that $e \in U(x)$ and

$$|f(x) - f(y)| < \frac{\epsilon}{2} \tag{62.1}$$

when $y \in x\,U(x)$. By continuity of the group operation at e, there is another open set $U_1(e)$ such that $e \in U_1(x)$ and

$$U_1(x)\,U_1(x) \subseteq U(x). \tag{62.2}$$

Consider the covering of K consisting of the open sets $x\,U_1(x)$, $x \in K$. By compactness, there are finitely many elements $x_1, \ldots, x_n$ of K such that

$$K \subseteq \bigcup_{j=1}^{n} x_j\,U_1(x_j). \tag{62.3}$$

Put

$$U_1 = \bigcap_{j=1}^{n} U_1(x_j), \tag{62.4}$$

which is also a neighborhood of e.

Let us check that

$$|f(x) - f(y)| < \epsilon \tag{62.5}$$

for every $x \in K$ and $y \in x\,U_1$. Let $x \in K$ be given, and choose $j \in \{1, \ldots, n\}$ such that $x \in x_j\,U_1(x_j)$. If $y \in x\,U_1$, then

$$y \in x_j\,U_1(x_j)\,U_1 \subseteq x_j\,U_1(x_j)\,U_1(x_j) \subseteq x_j\,U(x_j). \tag{62.6}$$

This implies that

$$|f(x) - f(y)| \le |f(x) - f(x_j)| + |f(x_j) - f(y)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{62.7}$$

as desired.

Similarly, there is a neighborhood V_1 of e for which the same conclusion holds when $x \in K$ and $y \in V_1\,x$. This can also be derived from the previous assertion, applied to $f(x^{-1})$ and K^{-1}. If f has compact support, then these uniform continuity conditions hold on the whole group.

Note that G is locally compact as soon as there is a nonempty open set with compact closure. In particular, G is locally compact when there is a continuous function with compact support that is not identically zero.

63 Haar measure

If G is a locally compact topological group, then there are *Haar measures* H_L, H_R on G that are invariant under translations on the left and right, respectively. More precisely, H_L and H_R are Borel measures on G such that

$$H_L(K),\ H_R(K) < \infty \tag{63.1}$$

when $K \subseteq G$ is compact and

$$H_L(U),\ H_R(U) > 0 \tag{63.2}$$

when $U \subseteq G$ is nonempty and open. Translation-invariance means that

$$H_L(a\,E) = H_L(E), \quad H_R(E\,b) = H_R(E) \tag{63.3}$$

for every $a, b \in G$ and Borel set $E \subseteq G$. These measures also enjoy some standard regularity properties, which make them unique up to multiplication by positive real constants.

If G is a Lie group, so that G is a smooth manifold and the group operations are smooth mappings, then the Haar measures correspond to smooth left and right-invariant volume forms, which are uniquely determined by their values at the identity element. If G is equipped with the discrete topology, then one can use counting measure for both the left and right Haar measures. Of course, left and right invariance are the same for an abelian group. It is well known that Haar measure on a compact group is also invariant under both left and right translations, as we shall see in a moment.

Observe that $H_L(E^{-1})$ is invariant under translations of E on the right, and $H_R(E^{-1})$ is invariant under translations on the left. Thus $H_L(E^{-1})$, $H_R(E^{-1})$ are positive constant multiples of $H_R(E)$, $H_L(E)$, respectively. By multiplying the Haar measures by suitable constants, if necessary, we may also ask that

$$H_R(E) = H_L(E^{-1}). \tag{63.4}$$

Observe also that $H_R(a\,E)$ is invariant under translations of E on the right for each $a \in G$, and $H_L(E\,b)$ is invariant under translations of E on the left for each $b \in G$. Hence there are positive real-valued functions ϕ, ψ on G such that

$$H_R(a\,E) = \phi(a)\,H_R(E), \quad H_L(E\,b) = \psi(b)\,H_L(E). \tag{63.5}$$

If G is compact, then $\phi(a) = \psi(b) = 1$, as one can see by taking $E = G$.

These relations imply that

$$\int_G f(a^{-1}\,x)\,dH_R(x) = \phi(a)\int_G f(x)\,dH_R(x) \tag{63.6}$$

and

$$\int_G f(x\,b^{-1})\,dH_L(x) = \psi(b)\int_G f(x)\,dH_L(x) \tag{63.7}$$

for every continuous real-valued function f on G with compact support. By applying these identities to a nonnegative function f that is positive somewhere, and using uniform continuity as in the previous section, we get that ϕ, ψ are continuous functions. It is easy to see from their definitions that ϕ, ψ are also homomorphisms from G into the positive real numbers as a group with respect to multiplication. Moreover, $\psi = 1/\phi$, by (63.4). Note that ϕ and ψ do not depend on the choice of Haar measure, which is to say that they are unaffected by multiplication of Haar measure by a positive constant.

One can also check that

$$\phi(x)^{-1}\,dH_R(x), \quad \psi(x)^{-1}\,dH_L(x) \tag{63.8}$$

determine left and right-invariant measures on G, respectively, and hence are constant multiples of $dH_L(x)$, $dH_R(x)$. With the normalization (63.4), these constants are equal to 1, because these measures are approximately the same on sets E near e with positive finite measure and $E^{-1} = E$. This uses the fact that ϕ, ψ are continuous and equal to 1 at the identity. In particular, H_L, H_R are mutually absolutely continuous.

Suppose that α is an automorphism of G as a topological group, which is to say a group automorphism that is also a homeomorphism. Thus

$$H_R(\alpha(E)), \quad H_L(\alpha(E)) \tag{63.9}$$

have the same properties as Haar measure, and are therefore constant multiples of $H_R(E)$, $H_L(E)$. If G is compact, then these multiples are equal to 1, as one can see by taking $E = G$. Similarly, if G is discrete, then these multiples are equal to 1. If α is an inner automorphism, so that $\alpha(x) = a\,x\,a^{-1}$ for some $a \in G$, then this reduces to the earlier discussion.

Note that

$$I_R(f) = \int_G f(x)\,dH_R(x), \quad I_L(f) = \int_G f(x)\,dH_L(x) \tag{63.10}$$

are positive linear functionals on the space of continuous real-valued functions with compact support on G that are invariant under translations on the right and left, respectively. Conversely, one can start with positive linear functionals of this type, and use the Riesz representation theorem to get positive measures from them. Thus the existence and uniqueness of Haar measure can also be formulated in terms of these linear functionals, also known as Haar integrals.

64 Convolution, revisited

Let G be a locally compact topological group with left and right-invariant Haar measures H_L, H_R satisfying (63.4), as in the previous section. Under suitable integrability conditions, the left and right convolutions of functions f, g on G are defined by

$$(f *_L g)(x) = \int_G f(y^{-1}\,x)\,g(y)\,dH_L(y) \tag{64.1}$$

and

$$(f *_R g)(x) = \int_G f(x\,z^{-1})\,g(z)\,dH_R(z), \tag{64.2}$$

respectively. Of course, these are the same when G is commutative. Using the change of variables $y \mapsto y^{-1}$ and (63.4), we get that

$$(f *_L g)(x) = \int_G f(y\,x)\,g(y^{-1})\,dH_R(y), \tag{64.3}$$

and hence, using the change of variables $y \mapsto y\,x^{-1}$,

$$(f *_L g)(x) = \int_G f(y)\,g(x\,y^{-1})\,dH_R(y) = (g *_R f)(x). \tag{64.4}$$

If G is commutative, then it follows that convolution is commutative.

Moreover,

$$(f *_L g) *_R h = (f *_R h) *_L g. \tag{64.5}$$

This says that left and right convolution operators acting on f commute with each other. To see this, one can simply express the convolutions as integrals and apply Fubini's theorem, because left and right translations on the group automatically commute with each other. This implies the usual associativity conditions

$$(f *_L g) *_L h = f *_L (g *_L h) \tag{64.6}$$

and

$$(f *_R g) *_R h = f *_R (g *_R h), \tag{64.7}$$

using the commutativity conditions described in the previous paragraph.

As in the previous situations, convolutions with f are basically combinations of translations of f. Because of possible noncommutativity, one has to be careful to distinguish between translations on the left and on the right. However, left and right translations have the nice feature of commuting with each other.

One can also simply look at linear transformations acting on functions on the group that commute with left or right translations. This includes convolutions with functions, measures, or other types of objects, as appropriate.

65 Convolution of measures

Let G be a locally compact topological group. Under suitable conditions, the convolution $\mu * \nu$ of Borel measures μ, ν on G can be defined by

$$\int_G \phi \, d(\mu * \nu) = \int_G \int_G \phi(x\,y)\, d\mu(x)\, d\nu(y), \tag{65.1}$$

where $x\,y$ uses the group operation on the right side. For example, if μ, ν are real or complex measures with finite total mass, then this can be used to define $\mu*\nu$ as a measure with finite total mass. If μ, ν have compact support, then $\mu*\nu$ has compact support. If one of μ, ν has compact support and the other is locally finite, then $\mu * \nu$ can be defined as a locally finite measure, by considering only functions ϕ with compact support. Note that it is not necessary to distinguish between left and right convolution in this definition, but convolution is not commutative unless G is commutative. However, associativity follows easily from the definition.

As a very basic example, consider the point mass δ_a at a, so that $\delta_a(E) = 1$ when $a \in E$ and $\delta_a(E) = 0$ when $a \notin E$. It is easy to see that

$$\delta_a * \delta_b = \delta_{a\,b} \tag{65.2}$$

for each $a, b \in G$. If e is the identity element of G, then

$$\delta_e * \mu = \mu * \delta_e = \mu \tag{65.3}$$

for every measure μ. Otherwise,

$$(\mu * \delta_a)(E) = \mu(E\,a^{-1}), \quad (\delta_a * \nu)(E) = \nu(a^{-1}\,E). \tag{65.4}$$

The convolution of a function and a measure can also be defined directly as a function, by

$$(f * \nu)(x) = \int_G f(x\,y^{-1})\, d\nu(y) \tag{65.5}$$

and

$$(65.6) \qquad (\mu * f)(y) = \int_G f(x^{-1}\, y)\, d\mu(x).$$

Equivalently, if H_R is a right-invariant Haar measure, then

$$(65.7) \qquad (f * \nu)\, dH_R = (f\, dH_R) * \nu,$$

because

$$(65.8) \qquad \begin{aligned} \int_G \phi(x)\,(f * \nu)(x)\, dH_R(x) &= \int_G \int_G \phi(x)\, f(x\, y^{-1})\, d\nu(y)\, dH_R(x) \\ &= \int_G \int_G \phi(x\, y)\, f(x)\, dH_R(x)\, d\nu(y). \end{aligned}$$

Similarly, if H_L is a left-invariant Haar measure, then

$$(65.9) \qquad (\mu * f)\, dH_L = \mu * (f\, dH_L),$$

because

$$(65.10) \qquad \begin{aligned} \int_G \phi(y)\,(\mu * f)(y)\, dH_L(y) &= \int_G \int_G \phi(y)\, f(x^{-1}\, y)\, d\mu(x)\, dH_L(y) \\ &= \int_G \int_G \phi(x\, y)\, f(y)\, d\mu(x)\, dH_L(y). \end{aligned}$$

As usual,

$$(65.11) \qquad (\mu * f) * \nu = \mu * (f * \nu),$$

since left and right translations commute with each other. If δ_a is the point mass at a, then

$$(65.12) \qquad (\delta_a * f)(y) = f(a^{-1}\, y), \quad (f * \delta_a)(x) = f(x\, a^{-1}).$$

If f, g are functions on G, then their right and left convolutions can be described by

$$(65.13) \qquad f *_R g = f * (g\, dH_R), \quad f *_L g = (g\, dH_L) * f.$$

Hence

$$(65.14) \qquad (f *_R g)\, dH_R = (f\, dH_R) * (g\, dH_R)$$

and

$$(65.15) \qquad (f *_L g)\, dH_L = (g\, dH_L) * (f\, dH_L),$$

as in the previous paragraph.

66 Locally compact Hausdorff spaces

Let X be a locally compact Hausdorff topological space. A continuous real or complex-valued function f on X *vanishes at infinity* if for every $\epsilon > 0$ there is a compact set $K \subseteq X$ such that

$$|f(x)| < \epsilon \tag{66.1}$$

when $x \in X \backslash K$. In particular, this implies that f is bounded on X.

As usual, the space $C_0(X)$ of continuous functions on X vanishing at infinity is a vector space with respect to pointwise addition and scalar multiplication. It is also a Banach space with respect to the supremum norm

$$\|f\|_{sup} = \sup\{|f(x)| : x \in X\}. \tag{66.2}$$

Continuous functions on X with compact support are dense in $C_0(X)$.

It is well known that every bounded linear functional λ on $C_0(X)$ can be represented in a unique way as

$$\lambda(f) = \int_X f \, d\mu, \tag{66.3}$$

where μ is a real or complex regular Borel measure on X, as appropriate. The dual norm of λ is equal to

$$\|\mu\| = |\mu|(X), \tag{66.4}$$

where $|\mu|$ is the positive Borel measure associated to μ. This representation permits $\lambda(f)$ to be extended to integrable functions f on X with respect to $|\mu|$ in the usual way. In particular, $\lambda(f)$ is defined for bounded measurable functions f on X.

Note that continuous functions vanishing at infinity on a locally compact group satisfies the same uniform continuity conditions as mentioned earlier for functions with compact support.

67 Simple estimates

If μ, ν are real or complex measures on a locally compact group G, then

$$\|\mu * \nu\| \le \|\mu\| \, \|\nu\|. \tag{67.1}$$

For that matter,

$$\|\mu \times \nu\| \le \|\mu\| \, \|\nu\|, \tag{67.2}$$

where $\mu \times \nu$ is the corresponding product measure on $G \times G$. Similarly, if f is a bounded measurable function on G, then $\mu * f$, $f * \nu$ are bounded, and satisfy

$$\sup_{y \in G} |(\mu * f)(y)| \le \|\mu\| \left(\sup_{y \in G} |f(y)|\right) \tag{67.3}$$

and

$$\sup_{x \in G} |(f * \nu)(x)| \le \Big(\sup_{x \in G} |f(x)|\Big) \, \|\nu\|. \tag{67.4}$$

Suppose now that $1 \le p < \infty$, and that H_L, H_R are left and right-invariant Haar measures on G, respectively. If $f \in L^p(H_L)$, then $\mu * f \in L^p(H_L)$, and

$$\|\mu * f\|_{L^p(H_L)} \le \|\mu\| \, \|f\|_{L^p(H_L)}. \tag{67.5}$$

As usual, this is basically an integrated version of the triangle inequality. If $p = 1$, then this can be derived from Fubini's theorem, since

$$|(\mu * f)(y)| \le \int_G |f(x^{-1} \, y)| \, d|\mu|(x), \tag{67.6}$$

or seen as a special case of the previous estimate for convolution of measures.

A nice way to deal with $p > 1$ is to use Jensen's inequality when $\|\mu\| \le 1$ to get that

$$|(\mu * f)(y)|^p \le \int_G |f(x^{-1} \, y)|^p \, d|\mu|(x). \tag{67.7}$$

Integrating with respect to H_L and applying Fubini's theorem, we get that

$$\begin{aligned} \int_G |(\mu * f)(y)|^p \, dH_L(y) &\le \int_G \int_G |f(x^{-1} \, y)|^p \, d|\mu|(x) \, dH_L(y) \\ &\le \int_G \int_G |f(x^{-1} \, y)|^p \, dH_L(y) \, d|\mu|(x) \\ &= \int_G \int_G |f(y)|^p \, dH_L(y) \, d|\mu|(x) \\ &\le \int_G |f(y)|^p \, dH_L(y). \end{aligned} \tag{67.8}$$

This also uses the left-invariance of H_L and the hypothesis that $\|\mu\| \le 1$ at the end.

If, in addition, $g \in L^1(H_L)$, then $f *_L g \in L^p(H_L)$, and

$$\|f *_L g\|_{L^p(H_L)} \le \|f\|_{L^p(H_L)} \, \|g\|_{L^1(H_L)}. \tag{67.9}$$

This follows from the previous discussion applied to $\mu = g \, dH_L$. Similarly, if $f \in L^p(H_R)$, then $f * \nu \in L^p(H_R)$, and

$$\|f * \nu\|_{L^p(H_R)} \le \|f\|_{L^p(H_R)} \, \|\nu\|. \tag{67.10}$$

If $g \in L^1(H_R)$ as well, then $f *_R g \in L^p(H_R)$, and

$$\|f *_R g\|_{L^p(H_R)} \le \|f\|_{L^p(H_R)} \, \|g\|_{L^1(H_R)}, \tag{67.11}$$

by the preceding estimate applied to $\nu = g \, dH_R$.

68 Translation operators

Let G be a locally compact topological group. For each $a \in G$, the left and right translation operators L_a, R_a acting on functions on G are defined by

$$L_a(f)(x) = f(a^{-1}\,x), \quad R_a(f)(x) = f(x\,a). \tag{68.1}$$

Equivalently,

$$L_a(f) = \delta_a * f, \quad R_a(f) = f * \delta_{a^{-1}}. \tag{68.2}$$

As usual, left and right translations commute with each other, so that

$$L_a \circ R_b = R_b \circ L_a \tag{68.3}$$

for every $a, b \in G$.

Note that

$$\begin{aligned} (L_a(L_b(f)))(x) &= (L_b(f))(a^{-1}\,x) = f(b^{-1}\,a^{-1}\,x) \\ &= f((a\,b)^{-1}\,x) = L_{a\,b}(f)(x), \end{aligned} \tag{68.4}$$

which implies that

$$L_a \circ L_b = L_{a\,b}. \tag{68.5}$$

Similarly,

$$R_a \circ R_b = R_{a\,b}, \tag{68.6}$$

because

$$(R_a(R_b(f)))(x) = (R_b(f))(x\,a) = f(x\,a\,b) = R_{a\,b}(f)(x). \tag{68.7}$$

This can also be seen in terms of convolution with δ measures.

Of course, these translation operators determine isometries of $C_0(G)$ onto itself. Moreover,

$$a \mapsto L_a,\ R_a \tag{68.8}$$

are continuous mappings from G into the space of bounded linear operators on $C_0(G)$ with the strong operator topology, because of the uniform continuity properties of functions in $C_0(G)$. Similarly, L_a, R_a determine isometries of $L^p(H_L)$, $L^p(H_R)$ onto themselves, respectively, that are continuous in a relative to the strong operator topology on $\mathcal{BL}(L^p(H_L))$, $\mathcal{BL}(L^p(H_R))$ when $p < \infty$. As usual, the latter can be verified using the fact that continuous functions on G with compact support are dense in $L^p(H_L)$, $L^p(H_R)$ when $p < \infty$.

For each $a \in G$, L_a and R_a also determine bounded linear operators on $L^p(H_R)$ and $L^p(H_L)$, respectively, that are isometries multiplied by positive real numbers. This follows from the fact that translations on the left correspond to multiplying H_R by a positive real number, and similarly for translations on the right and H_L. As in the previous situation, L_a and R_a are continuous in a relative to the strong operator topology on $\mathcal{BL}(L^p(H_R))$ and $\mathcal{BL}(L^p(H_L))$ when $p < \infty$.

69 Uniform continuity, 2

Let G be a topological group, and let f be a real or complex-valued function on G. Let us say that f is *right-uniformly continuous* on G if for every $\epsilon > 0$ there is a neighborhood U of the identity element e such that

$$|f(x) - f(y)| < \epsilon \tag{69.1}$$

for every $x, y \in G$ with $y \in x\,U$. Similarly, f is *left-uniformly continuous* if for every $\epsilon > 0$ there is a neighborhood V of e such that this condition holds when $y \in V\,x$. We have seen before that continuous functions with compact support satisfy both of these conditions, as well as continuous functions that vanish at infinity on a locally compact group.

Equivalently, f is right-uniformly continuous if for every $\epsilon > 0$ there is a neighborhood U of e such that

$$|f(x) - f(x\,u)| < \epsilon \tag{69.2}$$

for every $x \in G$ and $u \in U$. Similarly, f is left-uniformly continuous if for every $\epsilon > 0$ there is a neighborhood V of e such that

$$|f(x) - f(v\,x)| < \epsilon \tag{69.3}$$

for every $x \in G$ and $v \in V$.

If f is right-uniformly continuous on G, then every left translate $L_a(f)$ is right-uniformly continuous, using the same neighborhood U of e for a given ϵ as for f. Every right translate $R_a(f)$ is also right-uniformly continuous, because

$$|f(x\,a) - f(x\,w\,a)| < \epsilon \tag{69.4}$$

when $w \in a\,U\,a^{-1}$ and f satisfies (69.2). Similarly, if f is left-uniformly continuous, then every right translate of f is left-uniformly continuous with the same neighborhood V of e for a given ϵ as for f. A left translate $L_a(f)$ of f is left-uniformly continuous, with

$$|f(a^{-1}\,x) - f(a^{-1}\,z\,x)| < \epsilon \tag{69.5}$$

when $z \in a\,V\,a^{-1}$ and f satisfies (69.3).

Let W be a neighborhood of e, and let K be a compact set in G. By the continuity of the group operations, there are neighborhoods W_1, W_2 of e such that

$$x\,y\,x^{-1} \in W \tag{69.6}$$

for every $x \in W_1$ and $y \in W_2$. The sets $W_1\,z$, $z \in K$, form an open covering of K, and so there are finitely many elements $z_1, \ldots, z_n$ of K such that

$$K \subseteq \bigcup_{l=1}^{n} W_1\,z_l. \tag{69.7}$$

If $W_3 = \bigcap_{l=1}^n z_l^{-1} W_2 z_l$, then W_3 is a neighborhood of e, and

$$a\, w\, a^{-1} \in W \tag{69.8}$$

for every $a \in K$ and $w \in W_3$. This implies a uniform version of right-uniform continuity for the right translates $R_a(f)$ of a right-uniformly continuous function f when $a \in K$, and a uniform version of left-uniform continuity for the left translates $L_a(f)$ of a left-uniformly continuous function f when $a \in K$.

70 Continuity and convolution

Let G be a locally compact topological group, let f be a real or complex-valued continuous function on G, and let μ be a Borel measure on G. We would like to consider continuity properties of $f * \mu$ and $\mu * f$, under appropriate conditions. For instance, if f has compact support and μ is locally finite, then it is easy to see that $f * \mu$ and $\mu * f$ are continuous, using uniform continuity of f. In particular, $f *_L g$ and $f *_R g$ are continuous when f is a continuous function with compact support and g is locally integrable.

Suppose now that μ is a finite measure with compact support, and that f is an arbitrary continuous function on G. Again one can check that $f * \mu$ and $\mu * f$ are continuous, using uniform continuity of f on compact sets. If f is left or right-uniformly continuous, then $f * \mu$ and $\mu * f$ have the same property. This uses the uniform versions of uniform continuity of translates of f discussed in the previous section. In some situations, one may only have uniform versions of uniform continuity of translations of f over compact sets, and this is sufficient.

If μ is a finite regular measure with arbitrary support, then $f * \mu$ and $\mu * f$ are defined for bounded continuous functions f. In this case, there are compact sets whose complements have arbitrarily small $|\mu|$-measure, and so μ can be approximated in the measure norm by measures with compact support. This permits one to get analogous continuity and uniform continuity conditions for $f * \mu$ and $\mu * f$ as when μ has compact support. More precisely, the boundedness of f implies that this approximation of μ corresponds to uniform approximation of the convolutions.

Of course, $f * \mu$ and $\mu * f$ have compact support when f and μ both have compact support. If f has compact support and μ is a finite regular measure with arbitrary support, then $f * \mu$ and $\mu * f$ vanish at infinity, because they can be approximated uniformly by functions with compact support as in the previous paragraph. If f vanishes at infinity and μ is a finite regular measure, then $f * \mu$ and $\mu * f$ vanish at infinity, as one can see by approximating f uniformly by functions with compact support.

71 Convolution operators on C_0

Let G be a locally compact topological group, and suppose that T is a bounded linear operator from $C_0(G)$ into itself that commutes with left translations.

More precisely, this means that

$$T \circ L_a = L_a \circ T \tag{71.1}$$

for every $a \in G$. For example, if there is a real or complex regular Borel measure ν on G such that

$$T(f) = f * \nu \tag{71.2}$$

for every $f \in C_0(G)$, then T has this property. Conversely, let us check that every such operator can be represented in this manner.

The main point is that

$$\lambda(f) = T(f)(e) \tag{71.3}$$

defines a bounded linear functional on $C_0(G)$. Hence there is a regular real or complex Borel measure μ on G, as appropriate, such that

$$\lambda(f) = \int_G f \, d\mu \tag{71.4}$$

for every $f \in C_0(G)$. This implies that (71.2) holds, with $\nu(E) = \mu(E^{-1})$, because T commutes with left translations.

Similarly, convolution with a finite regular measure on the left defines a bounded linear operator on $C_0(G)$ that commutes with translations on the right. Conversely, every bounded linear operator on $C_0(G)$ that commutes with right translations can be represented in this way.

Note that the same argument would work for bounded linear mappings from $C_0(G)$ into the space of bounded continuous functions on G, equipped with the supremum norm. Also, the representation (71.2) is unique for regular Borel measures ν, as in the case of bounded linear functionals on $C_0(G)$.

72 Compactness and σ-compactness

Let G be a locally compact topological group. If G is compact, then G has finite left and right invariant Haar measure. Conversely, let us check that finite Haar measure implies compactness. Let U be a neighborhood of e with compact closure. If $x_1, \ldots x_n$ are elements of G such that the corresponding translates $x_1 \, U, \ldots, x_n \, U$ are pairwise disjoint, then

$$H_L\Big(\bigcup_{i=1}^n x_i \, U\Big) = \sum_{i=1}^n H_L(x_i \, U) = n \, H_L(U). \tag{72.1}$$

This leads to an upper bound on n, since U has positive measure. Suppose now that n is as large as possible. This means that for each $y \in G$ there is an i, $1 \le i \le n$, such that

$$y \, U \cap x_i \, U \ne \emptyset, \tag{72.2}$$

and hence

$$y \in x_i \, U \, U^{-1}. \tag{72.3}$$

By hypothesis, $\overline{U}$ is compact, which implies that $\overline{U}^{-1}$ and thus $\overline{U}\,\overline{U}^{-1}$ are compact, because of the continuity of the group operations. It follows that G is compact, since G is the union of the compact sets $x_i \overline{U}\,\overline{U}^{-1}$, $1 \le i \le n$. Of course, the same argument would work with right-invariant Haar measure and right translations instead of left-invariant Haar measure and left translations.

As in Section 61, there is an open subgroup H of G which is σ-compact. If H has only finitely or countably many cosets in G, then G is also σ-compact. Conversely, if G is σ-compact, then H has only finitely or countably many cosets in G. This is because every compact set in G is contained in the union of finitely many cosets of H.

Suppose that E is a Borel set in G with finite left or right-invariant Haar measure. For each $\epsilon > 0$, there are only finitely many left or right cosets of H, as appropriate, whose intersection with E has Haar measure greater than ϵ. It follows that there are only finitely or countably many left or right cosets of H, as appropriate, whose intersection with E has positive Haar measure.

If G is σ-compact, then the left and right-invariant Haar measures on G are σ-finite. Conversely, if G has σ-finite Haar measure, then the remarks in the previous paragraph imply that H has only finitely or countably many cosets, and hence that G is σ-compact.

73 Equicontinuity

Let G be a topological group, and let $\mathcal{E}$ be a collection of real or complex-valued functions on G. We say that $\mathcal{E}$ is *equicontinuous* at a point x if for every $\epsilon > 0$ there is a neighborhood U of x such that

$$|f(x) - f(y)| < \epsilon \tag{73.1}$$

for every $f \in \mathcal{E}$ and $y \in U$. Let us say that $\mathcal{E}$ is *right-uniformly equicontinuous* on a set E if for every $\epsilon > 0$ there is a neighborhood U of e such that the same condition holds for every $f \in \mathcal{E}$, $x \in E$, and $y \in x\,U$. Similarly, $\mathcal{E}$ is *left-uniformly equicontinuous* on E if for every $\epsilon > 0$ there is a neighborhood V of e such that this condition holds for every $f \in \mathcal{E}$, $x \in E$, and $y \in V\,x$.

If E is compact and $\mathcal{E}$ is equicontinuous at every element of E, then one can check that $\mathcal{E}$ is left and right-uniformly equicontinuous on E. This follows from the same argument as for a single function, as in Section 62. Of course, $\mathcal{E}$ is equicontinuous at every element of E when $\mathcal{E}$ is left or right-uniformly equicontinuous on E.

A function f is right-uniformly continuous on G if and only if the collection of left translates $L_a(f)$, $a \in G$, is equicontinuous at e, in which case the collection of left translates of f is right-uniformly equicontinuous on G. In particular, the collection of left translates is equicontinuous at any element b of G, which implies that the left translates of $R_b(f)$ are equicontinuous at e, and that $R_b(f)$ is right-uniformly continuous, as in Section 69. Moreover, the left translates of f are left-uniformly equicontinuous on every compact set, by the remarks in the previous paragraph, which corresponds to another observation in Section 69.

Similarly, f is left-uniformly continuous on G if and only if the collection of right translates $R_a(f)$, $a \in G$, is equicontinuous at e, in which case it is left-uniformly equicontinuous on G. This implies that the right translates of f are equicontinuous at every point, and right-uniformly equicontinuous on compact sets.

74 Locally compact spaces, 2

Let X be a locally compact Hausdorff topological space, and let μ be a positive Borel measure on X. A standard regularity condition asks that for each open set U in X,

$$\mu(U) = \sup\{\mu(K) : K \subseteq U \text{ is compact}\}. \tag{74.1}$$

If μ is a real or complex Borel measure on X, then one would ask that the associated positive measure $|\mu|$ have this property. This is one of the regularity properties in the conclusion of the Riesz representation theorem for positive linear functionals on the space of continuous functions with compact support on X, as well as bounded linear functionals on $C_0(X)$.

If μ is a positive Borel measure on X with this property, and if $\{U_\alpha\}_{\alpha \in A}$ is a family of open subsets of X such that $\mu(U_\alpha) = 0$ for each $\alpha \in A$, then

$$\mu\Big(\bigcup_{\alpha \in A} U_\alpha\Big) = 0. \tag{74.2}$$

To see this, it suffices to show that

$$\mu(K) = 0 \tag{74.3}$$

for every compact set $K \subseteq \bigcup_{\alpha \in A} U_\alpha$. By compactness, there are finitely many indices $\alpha_1, \ldots, \alpha_n \in A$ such that $K \subseteq \bigcup_{j=1}^n U_{\alpha_j}$, and so

$$\mu(K) \le \sum_{j=1}^n \mu(U_{\alpha_j}) = 0, \tag{74.4}$$

as desired. In particular, if W is the union of all of the open sets $U \subseteq X$ such that $\mu(U) = 0$, then $\mu(W) = 0$.

The *support* $\operatorname{supp} \mu$ of μ is defined to be the complement of the open set W just mentioned. Thus $\operatorname{supp} \mu$ is a closed set in X, $\mu(X \setminus \operatorname{supp} \mu) = 0$, and $\mu(V) > 0$ whenever V is an open set in X such that $V \cap \operatorname{supp} \mu \ne \emptyset$. For any positive Borel measure μ on X, one can say that μ has compact support when there is a compact set in X whose complement has μ-measure 0. This is equivalent to the compactness of $\operatorname{supp} \mu$ when μ satisfies the regularity condition under consideration.

Let G be a locally compact topological group, and let H be an open subgroup of G. If μ is a positive Borel measure on G with compact support, then μ is identically 0 on the complement of finitely many cosets of H. If μ is any positive finite measure on G, then there are only finitely many cosets of H with μ-measure at least ϵ for any $\epsilon > 0$. Hence there are only finitely or countably many cosets of H with positive μ-measure. The same conclusion holds when μ is σ-finite. If μ also satisfies the regularity condition described before, then the support of μ is contained in the union of only finitely or countably many cosets of H. In particular, this can be applied to a σ-compact open subgroup, as in Section 61. It follows that the support of μ is σ-compact when μ is σ-finite and satisfies this regularity property.

75 Convolution operators on L^1

Let G be a locally compact topological group with right-invariant Haar measure H_R, and let T be a bounded linear operator on $L^1(H_R)$ that commutes with translations on the left, so that

$$T \circ L_a = L_a \circ T \tag{75.1}$$

for every $a \in G$. As in Section 71,

$$T(f) = f * \nu \tag{75.2}$$

has these properties when ν is a real or complex regular Borel measure on G. We would like to show that every such operator can be represented in this way. It is easy to see that ν has to be unique, by considering continuous functions f with compact support.

For each neighborhood U of e in G, let ϕ_U be a nonnegative real-valued continuous function on G with compact support contained in U and

$$\int_G \phi_U \, dH_R = 1. \tag{75.3}$$

It suffices to consider a local base for the topology of G at e, and thus sequences of U's and ϕ's when there is a countable local base. The boundedness of T on $L^1(H_R)$ implies that

$$\|T(\phi_U)\|_{L^1(H_R)} \le \|T\|_{op} \tag{75.4}$$

for each U, where $\|T\|_{op}$ is the operator norm of T on $L^1(H_R)$. Equivalently,

$$\lambda_U(f) = \int_G f(x)\, T(\phi_U)(x)\, dH_R(x) \tag{75.5}$$

is a bounded linear functional on $C_0(G)$, with dual norm bounded by $\|T\|_{op}$.

Now we would like to use compactness of closed balls in the dual of $C_0(G)$ with respect to the weak* topology. In nice situations, we could use sequential compactness, and otherwise one can work with nets, for instance. The bottom line is that there is a bounded linear functional λ on $C_0(G)$ with dual norm less than or equal to the operator norm of T which can be approximated in the weak* topology by λ_U's, where the U's can be taken to be arbitrarily small neighborhoods of e. More precisely, for each neighborhood V of e one can take $\Lambda(V)$ to be the set of λ_U's with $U \subseteq V$, and $\overline{\Lambda(V)}$ to be the closure of $\Lambda(V)$ in the weak* topology. If $V_1, \ldots, V_n$ are finitely many neighborhoods of e, then their intersection is too, and

$$\Lambda\Big(\bigcap_{j=1}^{n} V_j\Big) = \bigcap_{j=1}^{n} \Lambda(V_j), \tag{75.6}$$

which implies that

$$\overline{\Lambda\Big(\bigcap_{j=1}^{n} V_j\Big)} \subseteq \bigcap_{j=1}^{n} \overline{\Lambda(V_j)}. \tag{75.7}$$

In particular, the intersection of finitely many $\overline{\Lambda(V)}$'s is nonempty, and so the intersection of all of the $\overline{\Lambda(V)}$'s is nonempty, by compactness. The desired bounded linear functional λ can be taken to be any element of this intersection.

As usual, λ can be represented by a unique regular Borel measure ν on G, and we would like to show that T can be given by convolution on the right by ν, using the hypothesis that T commutes with translations on the left. If f is a continuous function with compact support on G, then the latter implies that

$$T(\phi_U *_L f) = T(\phi_U) *_L f \tag{75.8}$$

for each U, and hence

$$T(f *_R \phi_U) = f *_R T(\phi_U). \tag{75.9}$$

For sufficiently small U, $f *_R \phi_U$ is close to f in $L^1(H_R)$, and so $T(f *_R \phi_U)$ is close to $T(f)$ in $L^1(H_R)$. For some very small U, λ_U is close to λ in the weak* topology, which can be used to approximate $f * \nu$ by $f *_R T(\phi_U)$ uniformly on compact sets. Thus one can show that (75.2) holds when f is a continuous function with compact support, and hence for all $f \in L^1(H_R)$. Analogous arguments can be applied to bounded linear operators on $L^1(H_L)$ that commute with translations on the right. This can also be reduced to the previous situation using $f(x) \mapsto f(x^{-1})$.

76 Simple estimates, 2

Let G be a locally compact topological group, and let $1 < p, q < \infty$ be conjugate exponents, so that $1/p + 1/q = 1$. Using Hölder's inequality, we can estimate the convolution $f *_L g$ pointwise by

$$|(f *_L g)(x)| \le \Big(\int_G |f(y^{-1}\,x)|^p \, dH_L(y)\Big)^{1/p} \Big(\int_G |g(y)|^q \, dH_L(y)\Big)^{1/q}. \tag{76.1}$$

Because of (63.4), we can make the change of variables $y \mapsto y^{-1}$ to get that

$$\int_G |f(y^{-1}\,x)|^p \, dH_L(y) = \int_G |f(y\,x)|^p \, dH_R(y) = \int_G |f(y)|^p \, dH_R(y), \tag{76.2}$$

where translation-invariance of Haar measure is applied in the second step. Thus $f *_L g$ is bounded when $f \in L^p(H_R)$ and $g \in L^q(H_L)$, and

$$\sup_{x\in G} |(f *_L g)(x)| \le \|f\|_{L^p(H_R)} \, \|g\|_{L^q(H_L)}. \tag{76.3}$$

Under these conditions, $f *_L g$ is a continuous function that vanishes at infinity. The continuity of $f *_L g$ follows by approximating f in $L^p(H_R)$ by continuous functions with compact support. To get that $f *_L g$ vanishes at infinity, one can also approximate g in $L^q(H_L)$ by continuous functions with compact support.

Similarly, $f *_R g$ is bounded when $f \in L^p(H_L)$ and $g \in L^q(H_R)$, and

$$\sup_{x\in G} |(f *_R g)(x)| \le \|f\|_{L^p(H_L)} \, \|g\|_{L^q(H_R)}. \tag{76.4}$$

Moreover, $f *_R g$ is a continuous function that vanishes at infinity under these conditions.

As in Section 64, $f *_L g = g *_R f$. It is easy to see that the preceding statements are consistent with this fact.

77 Convolution operators into C_0

Let G be a locally compact topological group, and let $1 < p, q < \infty$ be conjugate exponents again. If $g \in L^q(H_L)$, then

$$T(f) = f *_L g \tag{77.1}$$

defines a bounded linear operator from $L^p(H_R)$ into $C_0(G)$ that commutes with translations on the right. Conversely, let us show that any bounded linear operator T from $L^p(H_R)$ into $C_0(G)$ that commutes with translations on the right is of this form. Note that translations on the right preserve both the $L^p(H_R)$ and supremum norms.

As in Section 71,

$$\lambda(f) = T(f)(e) \tag{77.2}$$

is a bounded linear functional on $L^p(H_R)$ under these conditions. Hence there is a $g_1 \in L^q(H_R)$ such that

$$\lambda(f) = \int_G f(y)\, g_1(y)\, dH_R(y) \tag{77.3}$$

for every $f \in L^p(H_R)$. Equivalently, $g(y) = g_1(y^{-1}) \in L^q(H_L)$, and

$$\lambda(f) = \int_G f(y^{-1})\, g(y)\, dH_L(y). \tag{77.4}$$

This implies that (77.1) holds at e, and therefore at every point, because T commutes with translations on the right. This argument works as well for bounded linear mappings from $L^p(H_R)$ into the space of bounded continuous functions with the supremum norm.

Similarly, $T(f) = f *_R g$ defines a bounded linear operator from $L^p(H_L)$ into $C_0(G)$ that commutes with translations on the left when $g \in L^q(H_R)$, as in the previous section. Conversely, if T is a bounded linear operator from $L^p(H_L)$ into $C_0(G)$ that commutes with translations on the left, then T can be represented in this way for some $g \in L^q(H_R)$. The same statement holds for bounded linear operators from $L^q(H_L)$ into the space of bounded continuous functions with the supremum norm.

Note that the identification of the dual of L^p with L^q works for arbitrary measures spaces when $1 < p < \infty$, even if they may not be σ-finite. In the present setting, it is somewhat easier to reduce to the σ-finite case, using a σ-compact open subgroup H, as in Section 61. The left or right cosets of H form a natural partition of the group into σ-finite pieces, and one can begin by representing a bounded linear functional on L^p by an element of L^q on each of the cosets. One can show that the restriction of the bounded linear functional on L^p to functions supported on cosets of H is nonzero for only finitely or countably many cosets of H, because the L^q norms of combinations of the representing functions on the cosets is bounded by the dual norm of the linear functional. The finitely or countably many nonzero representing functions on the corresponding cosets may be combined into an element of L^q that represents the whole linear functional on L^p, using the fact that elements of L^p are also combinations of their restrictions to finitely or countably many cosets.

78 L^1 and L^∞

Let G be a locally compact topological group, and suppose for the moment that G is σ-compact, so that Haar measure is σ-finite. If $f \in L^1(H_R)$ and $g \in L^\infty$, then $f *_L g$ is bounded, and

$$\sup_{x \in G} |(f *_L g)(x)| \le \|f\|_{L^1(H_R)} \, \|g\|_\infty. \tag{78.1}$$

Similarly, if $f \in L^1(H_L)$ and $g \in L^\infty$, then $f *_R g$ is bounded, and

$$\sup_{x \in G} |(f *_R g)(x)| \le \|f\|_{L^1(H_L)} \, \|g\|_\infty. \tag{78.2}$$

In both cases, one can check that the convolution of f and g is continuous, by approximating f in the L^1 norm by continuous functions with compact support. However, the convolution may not vanish at infinity.

Conversely, if T is a bounded linear mapping from $L^1(H_R)$ into the space of bounded continuous functions on G with the supremum norm that commutes with translations on the right, then there is a $g \in L^\infty$ such that $T(f) = f *_L g$. This follows from the same argument as in the previous section, since the dual of L^1 may be identified with L^∞. Similarly, if T is a bounded linear mapping from $L^1(H_L)$ into the space of bounded continuous functions with the supremum norm that commutes with translations on the left, then there is a $g \in L^\infty$ such that $T(f) = f *_R g$.

Analogous arguments can be employed without σ-compactness, but one should then be more careful about the dual of L^1. As in the previous section, one can use the cosets of an open σ-compact subgroup to localize to σ-finite pieces in a convenient way. Alternatively, if λ is a bounded linear functional on $L^1(H_L)$, then $f *_L \lambda$ can be defined for $f \in L^1(H_R)$ by

$$(f *_L \lambda)(x) = \lambda(f_{x,L}), \tag{78.3}$$

where $f_{x,L}(y) = f(y^{-1}\,x)$ is an element of $L^1(H_L)$ as a function of y. If λ is a bounded linear functional on $L^1(H_R)$, then $f *_R \lambda$ can be defined for $f \in L^1(H_L)$ by

$$(f *_R \lambda)(x) = \lambda(f_{x,R}), \tag{78.4}$$

where $f_{x,R}(z) = f(x\,z^{-1})$ is an element of $L^1(H_R)$ as a function of z. It is easy to see that these satisfy the same properties as before, and that bounded linear mappings from $L^1(H_R)$ or $L^1(H_L)$ into the space of bounded continuous mappings that commute with translations on the right or left, respectively, can be represented in this way.

In the same way, $f *_L \lambda$ can be defined directly when λ is a bounded linear functional on $L^p(H_L)$ and $f \in L^p(H_R)$, and $f *_R \lambda$ can be defined directly when λ is a bounded linear functional on $L^p(H_R)$ and $f \in L^p(H_L)$. These expressions reduce to $f *_L g$, $f *_R g$ when λ is given by integration of a function times $g \in L^q(H_L), L^q(H_R)$ with respect to H_L, H_R, respetively.

79 Conjugation by multiplication

Let G be a locally compact topological group, let T be a linear mapping from continuous functions with compact support on G to locally-integrable functions on G, and let ρ be a continuous complex-valued function on G such that $\rho(x) \neq 0$ for each $x \in G$. Consider the linear mapping

$$T_\rho(f) = \rho\, T(\rho^{-1}\, f) \tag{79.1}$$

also from continuous functions with compact support on G to locally-integrable functions on G.

Suppose that ρ is a homomorphism from G into the group of nonzero complex numbers with respect to multiplication, so that

$$\rho(x\,y) = \rho(x)\,\rho(y) \tag{79.2}$$

for every $x, y \in G$. If T commutes with translations on the left or on the right, then it is easy to see that T_ρ has the same property.

Let ϕ, ψ be the continuous homomorphisms from G into the positive real numbers described in Section 63, so that

$$dH_L = \phi^{-1}\, dH_R, \quad dH_R = \psi^{-1}\, dH_L. \tag{79.3}$$

If $1 \le p < \infty$, then

$$f \mapsto \phi^{-1/p} f \tag{79.4}$$

defines an isometric linear mapping from $L^p(H_L)$ onto $L^p(H_R)$. Similarly,

$$f \mapsto \psi^{-1/p} f \tag{79.5}$$

is an isometric linear mapping from $L^p(H_R)$ onto $L^p(H_L)$.

If T is a bounded linear operator on $L^p(H_R)$, then

$$T_\rho,\ \rho = \phi^{1/p}, \tag{79.6}$$

is a bounded linear operator on $L^p(H_L)$, with the same operator norm. If instead T is a bounded linear operator on $L^p(H_L)$, then

$$T_\rho,\ \rho = \psi^{1/p}, \tag{79.7}$$

is a bounded linear operator on $L^p(H_R)$, with the same operator norm. If the initial operator T commutes with translations on the left or on the right, then the new operator T_ρ has the same property, by the earlier remarks. Note that

$$\psi = 1/\phi, \tag{79.8}$$

which implies that the isometric linear mappings in the preceding paragraph are inverses of each other, as are the transformations between bounded linear operators on $L^p(H_L)$ and $L^p(H_R)$ given here.

80 L^1 into L^p

Let G be a locally compact topological group, and fix $1 \le p \le \infty$. If $f \in L^1(H_R)$ and $g \in L^p(H_R)$, then $f *_L g \in L^p(H_R)$, and

$$\|f *_L g\|_{L^p(H_R)} \le \|f\|_{L^1(H_R)} \, \|g\|_{L^p(H_R)}. \tag{80.1}$$

This is because $f *_L g = g *_R f$, as in Section 64, and the latter can be estimated as in Section 67. Thus

$$T(f) = f *_L g \tag{80.2}$$

defines a bounded linear operator from $L^1(H_R)$ into $L^p(H_R)$ that commutes with translations on the right. Note that translations on the right determine isometries on $L^1(H_R)$, $L^p(H_R)$.

Conversely, suppose that T is a bounded linear mapping from $L^1(H_R)$ into $L^p(H_R)$ that commutes with translations on the right, where $1 < p < \infty$. For each neighborhood U of e, let ϕ_U be a nonnegative real-valued continuous function with compact support contained in U and

$$\int_G \phi_U \, dH_R = 1, \tag{80.3}$$

as in Section 75. Thus

$$\|T(\phi_U)\|_{L^p(H_R)} \le \|T\|_{op}, \tag{80.4}$$

where $\|T\|_{op}$ is the operator norm of T from $L^1(H_R)$ into $L^p(H_R)$. If q is the exponent conjugate to p, then

$$\lambda_U(f) = \int_G f(x)\, T(\phi_U)(x)\, dH_R(x) \tag{80.5}$$

is a bounded linear functional on $L^q(H_R)$, with dual norm bounded by $\|T\|_{op}$. By compactness of closed balls in the dual of $L^q(H_R)$ with respect to the weak* topology, there is a bounded linear functional λ on $L^q(H_R)$ which can be approximated in the weak* topology by λ_U's for some arbitrarily small neighborhoods U of e, as in Section 75.

In the present situation, there is a $g \in L^p(H_R)$ such that

$$\lambda(f) = \int_G f(x)\, g(x)\, dH_R(x) \tag{80.6}$$

for every $f \in L^q(H_R)$. Because T commutes with translations on the right, one can check that

$$T(\phi_U *_R f) = T(\phi_U) *_R f \tag{80.7}$$

for every continuous function f with compact support, and hence

$$T(f *_L \phi_U) = f *_L T(\phi_U). \tag{80.8}$$

Observe that

$$\int_G \phi_U \, dH_L \tag{80.9}$$

is arbitrarily close to 1 for sufficiently small U, because H_L and H_R are approximately the same near e, as in Section 63. This implies that $f *_L \phi_U$ is close to f in $L^1(H_R)$ when U is sufficiently small, and so $T(f *_L \phi_U)$ is close to $T(f)$ in $L^p(H_R)$. One can also approximate $f *_L g$ by $f *_L T(\phi_U)$ uniformly on compact sets for some arbitrarily small U, since λ is approximated by λ_U for some arbitrarily small U in the weak* topology. This uses the fact that left and right-invariant Haar measures are mutually absolutely continuous, with continuous densities. Thus (80.2) holds when f is a continuous function with compact support, and hence for every $f \in L^1(H_R)$.

The same arguments can be applied when $p = \infty$, at least if G is σ-compact. Otherwise, one should be more careful about the duality between L^1 and L^∞. One can also look at bounded linear operators from $L^1(H_L)$ into $L^p(H_L)$ that commute with translations on the left. As before, $f \mapsto f *_R g$ satisfies these properties when $g \in L^p(H_L)$, and one can go in the other direction when $p > 1$.

81 Integral pairings

Let G be a locally compact topological group, and let $1 \le p, q \le \infty$ be conjugate exponents. If $f \in L^p(H_R)$ and $g \in L^q(H_L)$, then put

$$(f, g)_L = \int_G f(y^{-1})\, g(y)\, dH_L(y). \tag{81.1}$$

This makes sense because of Hölder's inequality and (63.4), and satisfies

$$|(f, g)_L| \le \|f\|_{L^p(H_R)}\, \|g\|_{L^q(H_L)}. \tag{81.2}$$

Similarly, if $f \in L^p(H_L)$ and $g \in L^q(H_R)$, then we can put

$$(f, g)_R = \int_G f(y^{-1})\, g(y)\, dH_R(y), \tag{81.3}$$

and this satisfies

$$|(f, g)_R| \le \|f\|_{L^p(H_L)}\, \|g\|_{L^q(H_R)}, \tag{81.4}$$

by Hölder's inequality. Using (63.4) and the change of variables $y \mapsto y^{-1}$, we also get that

$$(f, g)_R = (g, f)_L. \tag{81.5}$$

If $1 < p, q < \infty$, then one can use $(f, g)_L$ to identify $L^p(H_R)$ with the dual of $L^q(H_L)$. This is a bit more natural than the identification of $L^p(H_R)$ with the dual of $L^q(H_R)$ in the argument in the preceding section, and there was an analogous point in Section 77. One can also identify $L^\infty(H_R)$ with the dual of $L^1(H_L)$ when G is σ-compact. Otherwise, one can use the dual of $L^1(H_L)$ as a substitute for $L^\infty(H_R)$ as in Section 78, which is quite convenient for the $p = \infty$ case of the argument in the previous section. Of course, one can also use $(f, g)_R$ to identify $L^p(H_L)$ with the dual of $L^q(H_R)$ when $1 < p, q < \infty$, and when $q = 1$, $p = \infty$, and G is σ-compact.

Observe that

$$(R_a(f), g)_L = (f, L_{a^{-1}}(g))_L \tag{81.6}$$

and

$$(L_a(f), g)_R = (f, R_{a^{-1}}(g))_R, \tag{81.7}$$

where L_a, R_a are the operators corresponding to translations on the left and on the right as in Section 68, respectively. There are analogous formulae in which translations on the left and on the right are exchanged, with additional constant factors. This uses the fact that translations on the left send H_R to constant multiples of itself, and translations on the right send H_L to constant multiples of itself, as in Section 63.

82 Operators on L^2

Let G be a locally compact topological group, and consider the algebra of bounded linear operators on $L^2(H_R)$ that commute with translations on the right. It is easy to see that this is a von Neumann algebra. This uses the fact that translations on the right determine unitary transformations on $L^2(H_R)$ to get that this algebra is self-adjoint. One can also check that the algebra of bounded linear operators on $L^2(H_R)$ that commute with translations on the left is a von Neumann algebra, because translations on the left correspond to constant multiples of unitary transformations on $L^2(H_R)$. Similarly, the algebras of bounded linear operators on $L^2(H_L)$ that commute with translations on the left or on the right are von Neumann algebras.

The algebras of bounded linear operators on $L^2(H_R)$, $L^2(H_L)$ that commute with translations on the right are equivalent to each other in a simple way, as in Section 79. The algebras of bounded linear operators on $L^2(H_R)$, $L^2(H_L)$ that commute with translations on the left are equivalent in the same way. The mapping $x \mapsto x^{-1}$ determines an isometry between $L^2(H_R)$ and $L^2(H_L)$, and exchanges translations on the right and left. Thus these various algebras amount to basically the same thing.

83 Dual operators

Let V, W be vector spaces, both real or both complex, and equipped with norms $\|\cdot\|_V$, $\|\cdot\|_W$. Let V', W' be their dual spaces of bounded linear functionals with the dual norms $\|\cdot\|_{V'}$, $\|\cdot\|_{W'}$, as in Section 4. If T is a bounded linear operator from V into W, then the *dual operator* $T' : W' \to V'$ sends a bounded linear functional λ on W to the bounded linear functional $T'(\lambda)$ on V defined by

$$T'(\lambda)(v) = \lambda(T(v)). \tag{83.1}$$

It is easy to see from the definitions that

$$\|T'(\lambda)\|_{V'} \le \|T\|_{op} \, \|\lambda\|_{W'}, \tag{83.2}$$

so that T' is a bounded linear operator from W' to V' whose operator norm is less than or equal to the operator norm $\|T\|_{op}$ of T as a bounded linear operator from V into W. Using the Hahn–Banach theorem, one can check that the operator norms of T and T' are the same.

Suppose that V_1, V_2, V_3 are vector spaces, all real or all complex, and equipped with norms. If $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ are bounded linear operators, then

$$(T_2 \circ T_1)' = T_1' \circ T_2' \tag{83.3}$$

as bounded linear operators from V_3' into V_1'. The dual of the identity operator on a vector space V is equal to the identity operator on the dual space V'. Of course, the mapping that sends a bounded linear operator $T : V \to W$ to its dual $T' : W' \to V'$ is linear as a mapping from $\mathcal{BL}(V, W)$ into $\mathcal{BL}(W', V')$.

If V is a real Hilbert space, then the dual space V' can be identified with V using the inner product, and the dual T' of a bounded linear operator T on V is the same as its adjoint T^*, as in Section 30. In the complex case, the dual operator may not be quite the same as the adjoint, because of complex conjugation. Now let (X, μ) be a measure space, and let $1 < p, q < \infty$ be conjugate exponents, so that the dual of $L^p(X)$ can be identified with $L^q(X)$ in the standard way. If T is a bounded linear operator on $L^p(X)$, then the corresponding dual operator T' on $L^q(X)$ is characterized by

$$\int_X T(f)(x)\, g(x)\, d\mu(x) = \int_X f(x)\, T'(g)(x)\, d\mu(x) \tag{83.4}$$

for every $f \in L^p(X)$ and $g \in L^q(X)$.

Let G be a locally compact topological group. The dual of a right-translation operator R_a on $L^p(H_R)$ corresponds as in the previous paragraph to $R_{a^{-1}}$ on $L^q(H_R)$, and the dual of a left-translation operator L_a on $L^p(H_L)$ corresponds to $L_{a^{-1}}$ on $L^q(H_L)$. The dual of L_a as an operator on $L^p(H_R)$ is a constant multiple of $L_{a^{-1}}$ on $L^q(H_R)$, and the dual of R_a as an operator on $L^p(H_L)$ is a constant multiple of $R_{a^{-1}}$ on $L^q(H_L)$. Alternatively, the duals of $L^p(H_R)$, $L^p(H_L)$ can be identified with $L^q(H_L)$, $L^q(H_R)$, respectively, as in Section 81. Using these identifications, the duals of translation operators exchange translations on the right with translations on the left, and vice-versa, again with additional constant factors in some cases.

84 Operators on L^p

Let G be a locally compact topological group, and let $1 < p < \infty$ be given. The algebras of bounded linear operators on $L^p(H_R)$ or $L^p(H_L)$ that commute with translations on the right or on the left are closed with respect to the weak operator topology on $\mathcal{BL}(L^p(H_R))$ or $\mathcal{BL}(L^p(H_L))$, as appropriate. The algebras of bounded linear operators on $L^p(H_R)$ or $L^p(H_L)$ that commute with translations on the right are equivalent to each other through conjugation as in Section 79, as are the algebras of bounded linear operators on $L^p(H_R)$ or $L^p(H_L)$ that commute with translations on the left. The mapping $x \mapsto x^{-1}$ also determines an isometric equivalence between $L^p(H_R)$ and $L^p(H_L)$ that exchanges translations on the right and on the left, as in Section 82.

Let $1 < q < \infty$ be the exponent conjugate to p, so that $1/p + 1/q = 1$. If T is a bounded linear operator on $L^p(H_R)$, then there is a corresponding dual bounded linear operator T' on $L^q(H_R)$, as in the previous section. Similarly, if T is a bounded linear operator on $L^p(H_L)$, then there is a corresponding dual bounded linear operator T' on $L^q(H_L)$. In either case, if T commutes with translations on the left or on the right, then the dual operator T' has the same property. This follows from (83.3) and the description of the duals of translation operators mentioned in the previous section.

Alternatively, the duals of $L^p(H_R)$ and $L^p(H_L)$ may be identified with $L^q(H_L)$ and $L^q(H_R)$, respectively, as in Section 81. Using these identifications, the dual of a bounded linear operator on $L^p(H_R)$ corresponds to a bounded linear operator on $L^q(H_L)$, and the dual of a bounded linear operator on $L^p(H_L)$ corresponds to a bounded linear operator on $L^q(H_R)$. The dual of a bounded linear operator that commutes with translations on the right then commutes with translations on the left, and vice-versa. Of course, these identifications are related to the previous ones through the mapping $x \mapsto x^{-1}$.

85 Group representations

A *representation* of a group G on a vector space V is a homomorphism from G into the group of invertible linear mappings on V. If V is a real or complex vector space equipped with a norm, then one may restrict one's attention to bounded linear mappings on V with bounded inverses. If G is a topological group, then it is natural to ask that the representation satisfy additional continuity conditions.

The right and left translation operators described in Section 68 determine representations of a group G on spaces of functions on G, known as *regular* representations. If G is a topological group, then one may consider translation operators acting on continuous functions, for instance. Bounded continuous functions form a Banach space with respect to the supremum norm, on which translation operators act isometrically. If G is locally compact, then one can consider translation operators acting on continuous functions that vanish at infinity, or on L^p spaces with respect to Haar measure.

Let V be a real or complex vector space with a norm, and let $\mathcal{BL}(V)$ be the space of bounded linear operators on V, as usual. A representation of a topological group on V might be continuous with respect to the topology on $\mathcal{BL}(V)$ determined by the operator norm. This is a natural condition when V is finite-dimensional, but otherwise it is often too restrictive. One can consider the strong or weak operator topologies on $\mathcal{BL}(V)$ instead, which are of course the same as the topology associated to the operator norm when V is finite-dimensional.

The right and left regular representations of a locally compact group on the space of continuous functions that vanish at infinity equipped with the supremum norm is continuous with respect to the strong operator topology, because of the uniform continuity properties of these functions. Similarly, the right and left regular representations of any topological group on the space of right or left-uniformly continuous functions equipped with the supremum norm, respectively, is continuous relative to the strong operator topology. The regular representations of a locally compact group on L^p spaces associated to Haar measure are also continuous relative to the strong operator topology when $1 \le p < \infty$. These statements do not work for the topologies determined by the operator norm even for the compact abelian group $\mathbf{T}$ of complex numbers with modulus 1 under multiplication.

86 Local uniform boundedness

Let V, W be vector spaces, both real or both complex, and equipped with norms. Also let X be a topological space, and

$$x \mapsto T_x \tag{86.1}$$

be a mapping from X into the space $\mathcal{BL}(V, W)$ of bounded linear mappings from V into W. If $K \subseteq X$ is compact and (86.1) is continuous with respect to the strong operator topology on $\mathcal{BL}(V, W)$, then

$$K_v = \{T_x(v) : x \in K\} \tag{86.2}$$

is a compact set in W for each $v \in V$. In particular, K_v is a bounded set in W, and it follows from the Banach–Steinhaus theorem in Section 10 that

$$\{T_x : x \in K\} \tag{86.3}$$

is a bounded set in $\mathcal{BL}(V, W)$ with respect to the operator norm when V is complete. Similarly, if (86.1) is continuous with respect to the weak operator topology, then

$$K_{v,\lambda} = \{\lambda(T_x(v)) : x \in K\} \tag{86.4}$$

is a compact set of real or complex numbers, as appropriate, for each $v \in V$ and bounded linear functional λ on W. Using the Banach–Steinhaus and Hahn–Banach theorems, one can then show that K_v is a bounded set in W for each $v \in V$. If V is complete, then another application of the Banach–Steinhaus theorem implies that (86.3) is a bounded set in $\mathcal{BL}(V, W)$ with respect to the operator norm.

Another way to look at the continuity of (86.1) is in terms of the continuity of $T_x(v)$ as a function of both x and v. The continuity of (86.1) as a mapping from X into $\mathcal{BL}(V, W)$ with the strong operator topology is the same as the continuity of $T_x(v)$ in x and v separately. The continuity of $T_x(v)$ as a mapping from $X \times V$ with the product topology into W is a stronger condition a priori. This stronger condition implies that (86.1) is locally uniformly bounded, in the sense that each element of X has a neighborhood on which T_x has uniformly bounded operator norm. Conversely, if (86.1) is locally uniformly bounded and continuous as a mapping from X into $\mathcal{BL}(V, W)$ with the strong operator topology, then it is easy to see that $T_x(v)$ is continuous as a mapping from $X \times V$ into W. If X is locally compact, V is complete, and (86.1) is continuous as a mapping from X into $\mathcal{BL}(V, W)$ with the strong operator topology, then it follows from the earlier observations that (86.1) is locally uniformly bounded. Thus continuity of (86.1) as a mapping from X into $\mathcal{BL}(V, W)$ with the strong operator topology is equivalent to continuity of $T_x(v)$ as a mapping from $X \times V$ into W when X is locally compact and V is complete.

There is an analogous relationship between the continuity of (86.1) from X into $\mathcal{BL}(V, W)$ with the weak operator topology and the continuity of $\lambda(T_x(v))$ as a function of $x \in X$, $v \in V$, and $\lambda \in W'$. The boundedness of $T_x : V \to W$

implies that $\lambda(T_x(v))$ is continuous as a function of $(v, \lambda) \in V \times W'$ for each $x \in X$, and continuity with respect to the weak operator topology means that $\lambda(T_x(v))$ is continuous in x for each $v \in V$, $\lambda \in W'$. The continuity of $\lambda(T_x(v))$ as a function of $(x, v, \lambda) \in X \times V \times W'$ is a stronger condition a priori, which implies that T_x is locally uniformly bounded in particular. Conversely, if (86.1) is locally uniformly bounded and continuous in the weak operator topology, then $\lambda(T_x(v))$ is continuous as a function of (x, v, λ). If X is locally compact, V is complete, and (86.1) is continuous with respect to the weak operator topology, then the earlier remarks imply that T_x is locally uniformly bounded, and hence that $\lambda(T_x(v))$ is continuous as a function of (x, v, λ).

Suppose that there is a countable local base for the topology of X at a point $p \in X$. If (86.1) is not uniformly bounded on a neighborhood of p, then there is a sequence $\{p_l\}_{l=1}^{\infty}$ of elements of X that converges to p such that $\|T_{p_l}\|_{op} \to \infty$ as $l \to \infty$. However, the Banach–Steinhaus theorem implies that $\|T_{p_l}\|_{op}$ is uniformly bounded when V is complete, (86.1) is continuous with respect to the strong or weak operator topology, and $\{p_l\}_{l=1}^{\infty}$ converges to p. It follows that (86.1) is locally uniformly bounded when it is continuous with respect to the strong or weak operator topology, V is complete, and X has a countable local base for its topology at each point.

87 Uniform convexity

Let V be a real or complex vector space with a norm $\|v\|$. We say that V is *locally uniformly convex at* $u \in V$ *with* $\|u\| = 1$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|u - v\| < \epsilon \tag{87.1}$$

for every $v \in V$ that satisfies $\|v\| = 1$ and

$$\left\|\frac{u + v}{2}\right\| > 1 - \delta. \tag{87.2}$$

Of course,

$$\left\|\frac{u + v}{2}\right\| \le 1 \tag{87.3}$$

under these conditions by the triangle inequality. Similarly, v is *locally uniformly convex* if it is locally uniformly convex at every $u \in V$ with $\|u\| = 1$, and V is *uniformly convex* if for every $\epsilon > 0$ there is a $\delta > 0$ such that the previous property holds uniformly in u. Inner product spaces are uniformly convex because of the parallelogram law, and L^p spaces are uniformly convex when $1 < p < \infty$ by Clarkson's inequalities.

Suppose that V is locally uniformly convex at $u \in V$ with $\|u\| = 1$. The Hahn–Banach theorem implies that there is a bounded linear functional λ on V with dual norm 1 such that

$$\lambda(u) = 1. \tag{87.4}$$

Let ϵ, δ be as in the previous paragraph, and suppose that $v \in V$, $\|v\| = 1$, and

$$|\lambda(v) - \lambda(u)| < 2\delta. \tag{87.5}$$

Thus

$$\Big|\lambda\Big(\frac{u+v}{2}\Big) - 1\Big| < \delta, \tag{87.6}$$

and so

$$\Big\|\frac{u+v}{2}\Big\| \ge \Big|\lambda\Big(\frac{u+v}{2}\Big)\Big| > 1 - \delta. \tag{87.7}$$

This shows that the topology on the unit sphere $\{v \in V : \|v\| = 1\}$ induced by the weak topology on V is the same as the topology induced by the norm when V is locally uniformly convex. However, this does not mean that the unit sphere is a closed set in the weak topology. Of course, the closed unit ball in V is a closed set in the weak topology.

A linear mapping T of V onto itself is said to be an *isometry* if

$$\|T(v)\| = \|v\| \tag{87.8}$$

for every $v \in V$. If V is locally uniformly convex, then it follows from the remarks in the previous paragraph that the topology on the set of isometries on V induced by the weak operator topology is the same as the topology induced by the strong operator topology. The same argument works for the collection of linear isometric embeddings of a vector space into a locally uniformly convex space. As before, this does not mean that the isometries form a closed set in the weak operator topology. The set of isometric linear embeddings is closed in the strong operator topology, and the set of linear mappings with operator norm less than or equal to 1 is closed in the weak operator topology.

Let X be a topological space, and let f be a function on X with values in V. If f is continuous at a point $p \in X$ with respect to the weak topology, $\|f\|$ is continuous at p, and V is locally uniformly convex at $f(p)$, then one can show that f is continuous at p with respect to the topology on V associated to the norm. Instead of continuity of $\|f\|$ at p, it suffices to ask that $\|f\|$ be upper semicontinuous at p, in the sense that for every $\eta > 0$ there is a neighborhood U of p such that

$$\|f(x)\| < \|f(p)\| + \eta \tag{87.9}$$

for every $x \in U$. If T_x is a function on X with values in bounded linear mappings on V such that $T_p = I$ for some $p \in X$, T_x is continuous at p in the weak operator topology, $\|T_x\|_{op}$ is upper semicontinuous at p, and V is locally uniformly convex, then T_x is continuous at p with respect to the strong operator topology. This follows from the previous statement applied to $f(x) = T_x(v)$ for each $v \in V$, and this also works when the T_x's are bounded linear mappings into a locally uniformly convex space such that T_p is an isometric linear embedding.

88 Bounded representations

A representation ρ of a group G on a vector space V with a norm is said to be *bounded* if there is a $C \geq 1$ such that

$$\|\rho_g\|_{op} \leq C \tag{88.1}$$

for each $g \in G$. This implies that

$$\|\rho_g^{-1}\|_{op} \leq C \tag{88.2}$$

for each $g \in G$, since $\rho_g^{-1} = \rho_{g^{-1}}$. If G is a compact topological group, V is complete, and ρ is continuous with respect to the strong or weak operator topology on $\mathcal{BL}(V)$, then ρ is automatically bounded, as in Section 86. Note that ρ is bounded with $C = 1$ if and only if ρ_g is an isometry on V for each $g \in G$. In particular, a *unitary representation* is a representation that acts by unitary transformations on a Hilbert space.

Suppose that ρ is a representation of a compact topological group G on a Hilbert space $(V, \langle v, w\rangle)$ which is continuous with respect to the strong operator topology. Thus ρ is bounded, and we can put

$$\langle v, w\rangle_1 = \int_G \langle \rho_g(v), \rho_g(w)\rangle \, dH(g), \tag{88.3}$$

where H is Haar measure. Continuity of ρ with respect to the strong operator topology ensures that $\langle \rho_g(v), \rho_g(w)\rangle$ is continuous for every $v, w \in V$, so that the integral makes sense. This defines an inner product on V that satisfies

$$\langle \rho_x(v), \rho_x(w)\rangle_1 = \langle v, w\rangle_1 \tag{88.4}$$

for every $x \in G$ and $v, w \in V$, because of translation-invariance. Hence ρ is unitary with respect to $\langle v, w\rangle_1$, which is the famous "unitary trick" of Weyl. More precisely, if the operator norm of ρ is bounded by C with respect to $\|v\| = \langle v, v\rangle^{1/2}$, then

$$C^{-1} \, \|v\| \leq \|v\|_1 \leq \|v\| \tag{88.5}$$

for every $v \in V$, where $\|v\|_1 = \langle v, v\rangle_1^{1/2}$. This implies that $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent norms on V.

Suppose now that ρ is continuous with respect to the weak operator topology, and that V is infinite-dimensional and separable. If $e_1, e_2, \ldots$ is an orthonormal basis for V, then

$$\langle \rho_g(v), \rho_g(w)\rangle = \sum_{l=1}^{\infty} \langle \rho_g(v), e_l\rangle \, \langle e_l, \rho_g(w)\rangle. \tag{88.6}$$

Continuity with respect to the weak operator topology implies that

$$\langle \rho_g(v), e_l\rangle, \quad \langle e_l, \rho_g(w)\rangle \tag{88.7}$$

are continuous for each l, and hence that $\langle \rho_g(v), \rho_g(w) \rangle$ is the pointwise limit of a sequence of continuous functions which is therefore Borel measurable. Thus we can still make sense of the integral, and the rest of the argument works as before.

If A is a bounded linear operator on V, then

$$\langle A(v), w \rangle \tag{88.8}$$

is a bilinear form on V in the real case and Hermitian-bilinear in the complex case, meaning that it is linear in v and either linear or conjugate-linear in w, as appropriate. The boundedness of A implies that

$$|\langle A(v), w \rangle| \le \|A\|_{op} \, \|v\| \, \|w\| \tag{88.9}$$

for every $v, w \in V$. Conversely, it is well known that every bounded bilinear form on V corresponds to a bounded linear operator A on V in this way, as a consequence of the bounded linear functionals on V in terms of the inner product. Self-adjointness and positivity of A can also be described in terms of the associated bilinear form. Applying this to the inner product $\langle v, w \rangle_1$ defined before, one can show that there is a bounded invertible linear mapping T on V such that $T \circ \rho_g \circ T^{-1}$ is unitary with respect to $\langle v, w \rangle$.

89 Dual representations

Let ρ be a representation of a group G on a real or complex vector space V with a norm $\| \cdot \|_V$, and let V' be the dual space of bounded linear functionals on V with the dual norm $\| \cdot \|_{V'}$. The corresponding *dual representation* ρ' of G on V' is defined by

$$\rho'_g = (\rho_{g^{-1}})'. \tag{89.1}$$

As in Section 83, $T' : V' \to V'$ denotes the dual of a bounded linear operator $T : V \to V$. This is applied to $\rho_{g^{-1}} = (\rho_g)^{-1}$ instead of ρ_g in the definition of the dual representation because of (83.3). Note that $(T^{-1})' = (T')^{-1}$ when T is invertible.

Remember that the operator norm of T' is equal to the operator norm of T on V, because of the Hahn–Banach theorem. Hence the dual of a bounded representation is also bounded. Similarly, the dual of a representation consisting of isometries has the same property. If V is a real Hilbert space, then the dual of V can be identified with V using the inner product, and the dual of a bounded linear operator on V can be identified with its adjoint. This is a bit more complicated in the complex case, because of complex conjugation.

Observe that $T \mapsto T'$ is a continuous linear mapping from $\mathcal{BL}(V)$ with the weak operator topology into $\mathcal{BL}(V')$ with the weak* operator topology. If ρ is continuous with respect to the weak operator topology, then it follows that ρ' is continuous with respect to the weak* operator topology. If ρ also consists of isometries on V, so that ρ' consists of isometries on V', and V' is locally uniformly convex, then one can argue as in Section 87 to get that ρ' is continuous with respect to the strong operator topology.

If V is a Hilbert space, then the adjoint of ρ can be defined directly by

$$\rho_g^* = (\rho_{g^{-1}})^*. \tag{89.2}$$

As before, this is applied to $\rho_{g^{-1}}$ instead of ρ_g because of (30.6). Note that ρ is a unitary representation if and only if $\rho = \rho^*$. The continuity properties of ρ and ρ^* are obviously the same in this case.

90 Invertibility, 4

Let V be a real or complex vector space with a norm $\|v\|$, and let A be a bounded linear operator on V with a bounded inverse. Thus

$$\|v\| \le \|A^{-1}\|_{op}\, \|A(v)\| \tag{90.1}$$

for every $v \in V$. If B is a bounded linear operator on V such that

$$\|A^{-1}\|_{op}\, \|A - B\|_{op} < 1, \tag{90.2}$$

then

$$\begin{aligned} \|v\| &\le \|A^{-1}\|_{op}\, \|B(v)\| + \|A^{-1}\|_{op}\, \|(A - B)(v)\| \\ &\le \|A^{-1}\|_{op}\, \|B(v)\| + \|A^{-1}\|_{op}\, \|A - B\|_{op}\, \|v\| \end{aligned} \tag{90.3}$$

implies that

$$\|v\| \le \frac{\|A^{-1}\|_{op}}{1 - \|A^{-1}\|_{op}\, \|A - B\|_{op}}\, \|B(v)\|. \tag{90.4}$$

If B is invertible, then it follows that

$$\|B^{-1}\|_{op} \le \frac{\|A^{-1}\|_{op}}{1 - \|A^{-1}\|_{op}\, \|A - B\|_{op}}. \tag{90.5}$$

Of course, (90.4) implies that B is one-to-one, so that B is invertible when $B(V) = V$. One can show that B maps V onto a dense linear subspace of V under these conditions, which implies that B is invertible when V is complete. One can also express B as

$$B = A\,(I - A^{-1}\,(A - B)), \tag{90.6}$$

and use the fact that $I - A^{-1}\,(A - B)$ is invertible since

$$\|A^{-1}\,(A - B)\|_{op} \le \|A^{-1}\|_{op}\, \|A - B\|_{op} < 1, \tag{90.7}$$

as in Section 48.

If B is invertible, then

$$B^{-1} - A^{-1} = B^{-1}\,A\,A^{-1} - B^{-1}\,B\,A^{-1} = B^{-1}\,(A - B)\,A^{-1}, \tag{90.8}$$

and hence

$$\|B^{-1} - A^{-1}\|_{op} \le \frac{\|A^{-1}\|_{op}^2}{1 - \|A^{-1}\|_{op}\,\|A - B\|_{op}}\,\|A - B\|_{op}. \tag{90.9}$$

This shows that $A \mapsto A^{-1}$ is a continuous mapping on the group of invertible operators on V with respect to the topology associated to the operator norm, so that this group is a topological group with respect to this topology. If V is complete, then this group is an open subset of $\mathcal{BL}(V)$, as in the previous paragraph. If V is finite-dimensional, then this group is locally compact as well.

Using (90.8), one can also check that $A \mapsto A^{-1}$ is continuous with respect to the strong operator topology on any set of invertible operators whose inverses have uniformly bounded operator norms. Similarly, composition of operators is continuous with respect to the strong operator topology on any set of bounded linear operators with uniformly bounded operator norm. It follows that a group of invertible operators with uniformly bounded operator norms is a topological group with respect to the strong operator topology.

91 Uniform continuity, 3

Let G, H be topological groups, and let ϕ be a function on G with values in H. We say that ϕ is *right-uniformly continuous* if for every neighborhood W of the identity element in H there is an neighborhood U of the identity element in G such that

$$\phi(y) \in \phi(x)\,W \tag{91.1}$$

for every $x, y \in G$ with $y \in x\,U$. Similarly, ϕ is *left-uniformly continuous* if for every neighborhood W of the identity element in H there is a neighborhood U of the identity element in G such that

$$\phi(y) \in W\,\phi(x) \tag{91.2}$$

for every $x, y \in G$ with $y \in U\,x$. If $\phi : G \to H$ is a continuous homomorphism, then ϕ is both right and left-uniformly continuous. If $H = \mathbf{R}$ or $\mathbf{C}$ as a group with respect to addition and equipped with the standard topology, then this is the same as uniform continuity as in Section 69.

Let ρ be a representation of a topological group G on a vector space V with a norm $\|v\|$. Thus ρ is a homomorphism from G into the group of invertible operators on V. If ρ is continuous with respect to the operator norm on $\mathcal{BL}(V)$, then ρ is right and left-uniformly continuous in the multiplicative sense described in the previous paragraph. If ρ is also bounded, then it follows that ρ is right and left-uniformly continuous as a mapping from G into $\mathcal{BL}(V)$ as a group with respect to addition and with the topology determined by the operator norm.

Suppose now that ρ is a bounded representation which is continuous with respect to the strong operator topology on $\mathcal{BL}(V)$. Thus ρ takes values in a bounded subgroup of the group of invertible operators on V, which is a topological group with respect to the strong operator topology, as in the previous section. Hence ρ is right and left-uniformly continuous in the multiplicative sense described before, and we would like to see what this means in terms of uniform continuity with respect to addition.

Let $w \in V$ and $\epsilon > 0$ be given, and consider

$$W = \{T \in \mathcal{BL}(V) : \|T(w) - w\| < \epsilon\}. \tag{91.3}$$

The continuity of ρ with respect to the strong operator topology implies that there is a neighborhood U of the identity in G such that

$$\rho_g \in W \tag{91.4}$$

when $g \in U$, or equivalently

$$\|\rho_g(w) - w\| < \epsilon \tag{91.5}$$

for every $g \in U$. Because ρ is bounded, there is a $C \geq 1$ such that

$$\|\rho_x\|_{op} \leq C \tag{91.6}$$

for every $x \in G$. It follows that

$$\|\rho_{x\,g}(w) - \rho_x(w)\| < C\,\epsilon \tag{91.7}$$

for every $x \in G$ and $g \in U$, so that $\rho_x(w)$ is right-uniformly continuous on G for each $w \in V$ with respect to addition on V.

92 Invariant subspaces

Let ρ be a representation of a group G on a vector space V. A linear subspace W of V is said to be *invariant* under ρ if

$$\rho_g(W) \subseteq W \tag{92.1}$$

for every $g \in G$. This is equivalent to saying that

$$\rho_g(W) = W \tag{92.2}$$

for every $g \in G$, since (92.1) applied to g^{-1} implies that $\rho_g^{-1}(W) \subseteq W$. If V is equipped with a norm and the representation consists of bounded linear operators on V, then it is customary to restrict one's attention to invariant subspaces that are closed. Note that the closure of an invariant subspace is also invariant under these conditions. A closed subspace of a complete space is complete as well. At any rate, the restriction of ρ_g to an invariant linear subspace W of V defines a representation of G on W.

If ρ is a unitary representation on a Hilbert space V, and W is a closed linear subspace of V which is invariant under ρ, then the orthogonal complement $W^\perp$ of W in V is invariant under ρ too. However, even on a finite-dimensional vector space V, one can have an invertible linear transformation T and a linear subspace $W \subseteq V$ such that $T(W) = W$ and there is no linear subspace of V which is complementary to W and invariant under T. This corresponds to invertible matrices with nontrivial Jordan canonical form. Moreover,

$$n \mapsto T^n \tag{92.3}$$

then defines a representation of the additive group $\mathbf{Z}$ of integers on V under which W is invariant, and for which there is no invariant complement to W.

Let V be a real or complex vector space with a norm, and let W be a closed linear subspace of V. If $\widetilde{W}$ is the set of bounded linear functionals λ on V such that $\lambda(w) = 0$ for every $w \in W$, then it is easy to see that $\widetilde{W}$ is a closed linear subspace of the dual V' of V. For each bounded linear mapping T on V such that $T(W) \subseteq W$, one can check that the dual mapping T' on V' satisfies $T'(\widetilde{W}) \subseteq \widetilde{W}$. If W is invariant under a representation ρ on V, then it follows that $\widetilde{W}$ is invariant under the dual representation ρ' on V'.

One can also look at the quotient V/W of V modulo W, equipped with the usual quotient norm. If T is a bounded linear mapping on V such that $T(W) \subseteq W$, then T determines a bounded linear mapping on V/W in a natural way. If T is invertible and $T(W) = W$, then the induced mapping on V/W is invertible as well. If ρ is a representation of a group G on V and W is invariant under ρ, then ρ determines a representation of G on V/W too. Note that the dual of V/W can be identified with the linear subspace $\widetilde{W}$ of the dual V' of V described in the previous paragraph, since a bounded linear functional on V/W is basically the same as a bounded linear functional on V that vanishes on W. If T is a bounded linear mapping on V such that $T(W) \subseteq W$, then the dual of the associated mapping on V/W can be identified with the restriction of the dual T' of T to $\widetilde{W}$. If ρ is a representation of a group G on V for which W is an invariant subspace, then the dual of the associated representation on V/W corresponds to the restriction of the dual ρ' of ρ to $\widetilde{W}$.

93 Dual subspaces

Let V be a real or complex vector space with a norm, and let V' be the dual space of bounded linear functionals on V with the dual norm. If W is a linear subspace of V, then put

$$\widetilde{W} = \{\lambda \in V' : \lambda(w) = 0 \text{ for every } w \in W\}, \tag{93.1}$$

as in the previous section. This is a linear subspace of V' which is closed with respect to the weak* topology. If V is reflexive, then the weak and weak* topologies on V' are the same, and a closed linear subspace of V' with respect to the dual norm would also be closed in the weak* topology. The latter statement uses the fact that norm-closed linear subspaces are closed in the weak topology, by the Hahn–Banach theorem.

If L is a linear subspace of V', then put

$$\widehat{L} = \{v \in V : \lambda(v) = 0 \text{ for every } \lambda \in L\}. \tag{93.2}$$

This is a closed linear subspace of V, and

$$L \subseteq \widetilde{(\widehat{L})} \tag{93.3}$$

automatically. If L_1 is the weak* closure of L in V', then it is easy to see that $\widehat{L_1} = \widehat{L}$. One can check that $\widetilde{(\widehat{L})}$ is equal to the weak* closure of L, so that

$$L = \widetilde{(\widehat{L})} \tag{93.4}$$

if and only if L is closed in the weak* topology. Equivalently, if $\lambda \in V'$ is not an element of the weak* closure of L, then there is a $v \in V$ such that $v \in \widehat{L}$ and $\lambda(v) \neq 0$.

Let T be a bounded linear mapping on V, and let T' be the corresponding dual mapping on V'. If $T'(L) \subseteq L$ for some linear subspace L of V', then

$$T(\widehat{L}) \subseteq \widehat{L}. \tag{93.5}$$

Note that T' is continuous with respect to the weak* topology on V', and hence T' maps the weak* closure of L into itself when $T'(L) \subseteq L$. Thus one might as well ask that L be closed in the weak* topology on V'. If ρ is a representation of a group G on V and $L \subseteq V'$ is invariant under the dual representation ρ', then it follows that $\widehat{L}$ is invariant under ρ.

We also have that

$$W \subseteq \widehat{(\widetilde{W})} \tag{93.6}$$

for every linear subspace W of V. Of course, $\widetilde{\overline{W}} = \widetilde{W}$, where $\overline{W}$ is the closure of W in V with respect to the norm. If $v \in V \backslash \overline{W}$, then the Hahn–Banach theorem implies that there is a $\lambda \in \widetilde{W}$ such that $\lambda(v) \neq 0$. This means that v is not an element of $\widehat{(\widetilde{W})}$, and it follows that $\widehat{(\widetilde{W})} = \overline{W}$. In particular,

$$W = \widehat{(\widetilde{W})} \tag{93.7}$$

when W is a closed linear subspace of V.

94 Irreducibility

A representation ρ of a group G on a vector space V is said to be *irreducible* if there are no nontrivial invariant subspaces. Here we suppose that V is equipped with a norm and the representation consists of bounded linear operators on V, and we are concerned only with closed invariant subspaces of V. Equivalently, ρ is irreducible if for every $v \in V$ with $v \neq 0$,

$$\operatorname{span}\{\rho_g(v) : g \in G\} \tag{94.1}$$

is dense in V. If V has dimension equal to 1, then any representation on V is automatically irreducible.

Let λ be a bounded linear functional on V, and consider the linear mapping T_λ from V into functions on G defined by

$$T_\lambda(v)(g) = \lambda(\rho_g(v)). \tag{94.2}$$

Observe that

$$T_\lambda(v)(g\,h) = \lambda(\rho_{g\,h}(v)) = \lambda(\rho_g(\rho_h(v)) = T_\lambda(\rho_h(v))(g) \tag{94.3}$$

for every $g, h \in G$ and $v \in V$. Thus

$$R_h(T_\lambda(v)) = T_\lambda(\rho_h(v)) \tag{94.4}$$

for every $h \in G$ and $v \in V$, which can be expressed as well by

$$R_h \circ T_\lambda = T_\lambda \circ \rho_h \tag{94.5}$$

for every $h \in G$. This may be described by saying that T_λ *intertwines* ρ and the right regular representation of G.

Consider the kernel of T_λ,

$$\{v \in V : T_\lambda(v)(g) = 0 \text{ for every } g \in G\}. \tag{94.6}$$

This is a closed linear subspace of V contained in the kernel of λ,

$$\{v \in V : \lambda(v) = 0\}, \tag{94.7}$$

because $T_\lambda(v)(e) = \lambda(v)$. Hence the kernel of T_λ is a proper subspace of V when $\lambda \neq 0$. If ρ is irreducible, then it follows that the kernel of T_λ is trivial. The intertwining property in the previous paragraph implies that $T_\lambda(V)$ is invariant under translations on the right on G.

Note that $T_\lambda(v)$ is a continuous function on G for each $v \in V$ when G is a topological group and ρ is continuous with respect to the weak operator topology. If ρ is also a bounded representation, then T_λ is a bounded linear mapping from V into the space of bounded continuous functions on G with the supremum norm. If ρ is bounded and continuous with respect to the strong operator topology, then $T_\lambda(v)$ is right-uniformly continuous on G for every v in V, as in Section 91. Similarly, if ρ is bounded and the dual representation ρ' is continuous with respect to the strong operator topology, then $T_\lambda(v)$ is left-uniformly continuous on G for every $v \in V$. Of course, these continuity conditions on ρ are the same when V is finite-dimensional.

95 Vector-valued integration

Let V be a real or complex vector space with a norm $\|v\|_V$, and let μ be a real or complex measure on a measurable space X, according to whether V is real or complex. If A is a measurable subset of X, then the corresponding characteristic function $\mathbf{1}_A(x)$ on X is equal to 1 when $x \in A$ and to 0 when $x \in X \backslash A$. Let

$$s(x) = \sum_{j=1}^{n} v_j \, \mathbf{1}_{A_j} \tag{95.1}$$

be a measurable V-valued simple function on X, so that $A_1, \ldots, A_n \subseteq X$ are measurable, and $v_1, \ldots, v_n \in V$. The integral of $s(x)$ with respect to μ is given by

$$\int_X s(x) \, d\mu(x) = \sum_{j=1}^{n} v_j \, \mu(A_j). \tag{95.2}$$

It is easy to check that this does not depend on the particular representation (95.1) of s. If λ is a bounded linear functional on V, then

$$\lambda\Big(\int_X s(x) \, d\mu(x)\Big) = \int_X \lambda(s(x)) \, d\mu(x). \tag{95.3}$$

Similarly,

$$\Big\| \int_X s(x) \, d\mu(x) \Big\|_V \le \int_X \|s(x)\|_V \, d|\mu|(x), \tag{95.4}$$

as one can see by reducing to the case where $A_1, \ldots, A_n$ are pairwise disjoint.

Suppose now that X is a locally compact Hausdorff topological space, and that μ is a regular Borel measure on X. Let f be a continuous V-valued function on X, which implies that $\|f(x)\|_V$ is a continuous real-valued function on X. Hence the integrability condition

$$\int_X \|f(x)\|_V \, d|\mu|(x) < \infty \tag{95.5}$$

makes sense, and is satisfied when f is bounded or μ has compact support. If V is complete, then we would like to show that the integral of f with respect to μ can be defined in such a way that

$$\lambda\Big(\int_X f(x) \, d\mu(x)\Big) = \int_X \lambda(f(x)) \, d\mu(x) \tag{95.6}$$

for every bounded linear functional λ on V, and

$$\Big\| \int_X f(x) \, d\mu(x) \Big\|_V \le \int_X \|f(x)\|_V \, d|\mu|(x). \tag{95.7}$$

Let $\epsilon > 0$ be given, and let K be a compact set in X such that

$$\int_{X \backslash K} \|f(x)\|_V \, d|\mu|(x) < \frac{\epsilon}{2}. \tag{95.8}$$

For each $x \in K$, there is an open set $U(x)$ in X such that $x \in U(x)$ and

$$\|f(x) - f(y)\|_V < \frac{\epsilon}{2(1 + |\mu|(K))} \tag{95.9}$$

for every $y \in U(x)$. By compactness, there are finitely many elements $x_1, \ldots, x_n$ of K for which

$$K \subseteq \bigcup_{j=1}^{n} U(x_n). \tag{95.10}$$

Using this, it is easy to get a Borel measurable V-valued simple function s on X that satisfies

$$\|f(x) - s(x)\|_V < \frac{\epsilon}{2(1 + |\mu|(K))} \tag{95.11}$$

for every $x \in K$ and $s(x) = 0$ when $x \in X \backslash K$, and hence

$$\int_X \|f(x) - s(x)\|_V \, d|\mu|(x) < \epsilon. \tag{95.12}$$

Applying this to $\epsilon = 1/l$, $l \in \mathbf{Z}_+$, we get a sequence of Borel measurable V-valued simple functions $s_1, s_2, \ldots$ on X such that

$$\lim_{l \to \infty} \int_X \|f(x) - s_l(x)\|_V \, d|\mu|(x) = 0. \tag{95.13}$$

In particular, $\{s_l\}_{l=1}^{\infty}$ is a Cauchy sequence with respect to the L^1 norm for V-valued functions on X, in the sense that

$$\lim_{k,l \to \infty} \int_X \|s_k(x) - s_l(x)\|_V \, d|\mu|(x) = 0. \tag{95.14}$$

Because of (95.4), this implies in turn that

$$\lim_{k,l \to \infty} \Big\| \int_x s_k(x) \, d\mu(x) - \int_X s_l(x) \, d\mu(x) \Big\|_V = 0, \tag{95.15}$$

so that the integrals of the s_l's form a Cauchy sequence in V. If V is complete, then this sequence converges, and we put

$$\int_X f(x) \, d\mu(x) = \lim_{l \to \infty} \int_X s_l(x) \, d\mu(x). \tag{95.16}$$

It is easy to see that this limit does not depend on the sequence of simple functions approximating f. More precisely, suppose that $\{s_l'\}_{l=1}^{\infty}$ is another sequence of Borel measurable V-valued simple functions such that

$$\lim_{l \to \infty} \int_X \|f(x) - s_l'(x)\|_V \, d|\mu|(x) = 0. \tag{95.17}$$

This implies that

$$\lim_{l \to \infty} \int_X \|s_l(x) - s_l'(x)\|_V \, d|\mu|(x) = 0, \tag{95.18}$$

and hence that

$$\lim_{l \to \infty} \Big\| \int_X s_l(x) \, d\mu(x) - \int_X s_l'(x) \, d\mu(x) \Big\|_V = 0, \tag{95.19}$$

which ensures that

$$\lim_{l \to \infty} \int_X s_l(x)\, d\mu(x) = \lim_{l \to \infty} \int_X s_l'(x)\, d\mu(x), \tag{95.20}$$

as desired. This definition of the integral of f also satisfies (95.6) and (95.7), because of the corresponding properties of simple functions. Note that the integral of f is uniquely determined by (95.6), by the Hahn–Banach theorem.

96 Strongly continuous representations

Let $(V, \|v\|_V)$ be a real or complex Banach space, and let ρ be a representation of a locally compact topological group G by bounded linear operators on V which is continuous with respect to the strong operator topology on $\mathcal{BL}(V)$. Thus ρ is uniformly bounded on compact subsets of G by the Banach–Steinhaus theorem, as in Section 86. Let μ be a finite regular Borel measure on G which is real or complex according to whether V is real or complex, as in the previous section. We would like to define the convolution of μ with $v \in V$ with respect to ρ by

$$\mu *_\rho v = \int_G \rho_g(v)\, d\mu(g), \tag{96.1}$$

under suitable integrability conditions.

More precisely, we can do this when

$$\int_G \|\rho_g(v)\|_V\, d|\mu|(g) < \infty. \tag{96.2}$$

This condition holds in particular when μ has compact support, and when ρ is bounded. It is also implied by

$$\int_G \|\rho_g\|_{op}\, d|\mu|(g) < \infty. \tag{96.3}$$

Note that the operator norm of ρ is lower semicontinuous, in the sense that $\|\rho_g\|_{op} > t$ determines an open set in G for every $t \in \mathbf{R}$. This follows from continuity of ρ even with respect to the weak operator topology, and it implies that $\|\rho_g\|_{op}$ is Borel measurable.

In addition,

$$\|\mu *_\rho v\|_V \le \int_G \|\rho_g(v)\|_V\, d|\mu|(g) \tag{96.4}$$

when (96.2) holds, by (95.7). Hence $\mu *_\rho v$ defines a bounded linear operator on V when (96.3) holds, with operator norm bounded by the integral in (96.3).

If ρ is the left regular representation on an appropriate space of functions, then $\mu *_\rho v$ corresponds to convolution with μ on the left, as discussed previously. If ρ is the right regular representation, then $\mu *_\rho v$ corresponds to convolution with $\widetilde{\mu}(E) = \mu(E^{-1})$ on the right.

97 Dual representations, 2

Let ρ be a representation of a locally compact group G by bounded linear operators on a real or complex vector space V with a norm $\|v\|_V$, and let ρ' be the corresponding dual representation on the dual space V' of bounded linear functionals on V, as in Section 89. Thus ρ' is continuous with respect to the weak* operator topology when ρ is continuous with respect to the weak operator topology. Also let μ be a finite regular Borel measure on G which is real or complex, according to whether V is real or complex. We can try to define the convolution $\mu *_{\rho'} \lambda$ of μ with $\lambda \in V'$ with respect to ρ' by

$$(\mu *_{\rho'} \lambda)(v) = \int_G \rho'_g(\lambda)(v)\, d\mu(g) = \int_G \lambda(\rho_{g^{-1}}(v))\, d\mu(g). \tag{97.1}$$

More precisely, we want to define $\mu *_{\rho'} \lambda$ as a bounded linear functional on V, and (97.1) is supposed to be the value of this linear functional at $v \in V$.

The integrand $\rho'_g(\lambda)(v) = \lambda(\rho_{g^{-1}}(v))$ is continuous by hypothesis, but we should still be careful about its integrability with respect to $|\mu|$. This certainly holds when

$$\int_G \|\rho'_g\|_{op}\, d|\mu|(g) = \int_G \|\rho_{g^{-1}}\|_{op}\, d|\mu|(g) < \infty. \tag{97.2}$$

Here $\|\rho'_g\|_{op}$ is the operator norm of ρ'_g on V', which is equal to the operator norm $\|\rho_{g^{-1}}\|_{op}$ of $\rho_{g^{-1}}$ on V. As in the previous section, this condition holds when μ has compact support, and when ρ is bounded. This condition implies that

$$|(\mu *_{\rho'} \lambda)(v)| \le \Big(\int_G \|\rho_{g^{-1}}\|_{op}\, d|\mu|(g)\Big)\, \|\lambda\|_{V'}\, \|v\|_V, \tag{97.3}$$

so that $\mu *_{\rho'} \lambda$ is a bounded linear functional on V.

If ρ is continuous with respect to the strong operator topology on V, then

$$(\mu *_{\rho'} \lambda)(v) = \lambda(\widetilde{\mu} *_\rho v), \tag{97.4}$$

where $\widetilde{\mu}(E) = \mu(E^{-1})$ and $\widetilde{\mu} *_\rho v$ is as in the previous section. This assumes the integrability of $\|\rho_{g^{-1}}\|_{op}$ with respect to $|\mu|$ as in the preceding paragraph, which is the same as the integrability of $\|\rho_g\|_{op}$ with respect to $|\widetilde{\mu}|$. Thus $\widetilde{\mu} *_\rho v$ is defined, and $\lambda(\widetilde{\mu} *_\rho v)$ can be evaluated as in (95.6). Equivalently, the operator $\mu *_{\rho'} \lambda$ on V' is the dual of the operator $\widetilde{\mu} *_\rho v$ on V under these conditions.

For example, this applies to regular representations on L^∞ as the dual of L^1, or to finite regular Borel measures as the dual of $C_0(G)$. These representations are not continuous with respect to the strong operator topology even on the unit circle. If V is reflexive, so that V can be identified with the dual of its own dual space V', then this can also be a nice way to deal with convolutions on V.

98 Compositions

Let ρ be a representation of a locally compact topological group G by bounded linear operators on a real or complex Banach space V. Also let μ, ν be finite regular Borel measures on G that are real or complex, according to whether V is real or complex. We would like to show that

$$\mu *_\rho (\nu *_\rho v) = (\mu * \nu) *_\rho v \tag{98.1}$$

for each $v \in V$, under suitable conditions. Let us begin with the case where ρ is continuous in the strong operator topology as in Section 96, and then consider the analogous statement for dual representations as in Section 97.

If μ and ν have compact support, then their convolution $\mu * \nu$ does too, and additional condition is required. Similarly, if ρ is bounded, then no additional conditions on μ, ν are necessary. Otherwise, suppose that

$$\int_G \|\rho_g\|_{op} \, d|\mu|(g), \ \int_G \|\rho_g\|_{op} \, d|\nu|(g) < \infty. \tag{98.2}$$

This implies that

$$\begin{aligned} & \int_G \int_G \|\rho_{gh}\|_{op} \, d|\mu|(g) \, d|\nu|(h) \\ \le \ & \int_G \int_G \|\rho_g\|_{op} \, \|\rho_h\|_{op} \, d|\mu|(g) \, d|\nu|(h) \\ = \ & \Big(\int_G \|\rho_g\|_{op} \, d|\mu|(g) \Big) \Big(\int_G \|\rho_h\|_{op} \, d|\nu|(h) \Big) < \infty, \end{aligned} \tag{98.3}$$

and hence $\mu*\nu$ has the same property. Thus $\nu *_\rho v$, $\mu *_\rho (\nu *_\rho v)$, and $(\mu * \nu) *_\rho v$ are defined under these integrability conditions when ρ is continuous in the strong operator topology.

It suffices to show that

$$\lambda(\mu *_\rho (\nu *_\rho v)) = \lambda((\mu * \nu) *_\rho v) \tag{98.4}$$

for every bounded linear functional λ on V and $v \in V$. Using (95.6), we get that

$$\begin{aligned} \lambda(\mu *_\rho (\nu *_\rho v)) & = \int_G \lambda(\rho_g(\nu *_\rho v)) \, d\mu(g) \\ & = \int_G \int_G \lambda(\rho_g(\rho_h(v))) \, d\mu(g) \, d\nu(h) \\ & = \int_G \int_G \lambda(\rho_{gh}(v)) \, d\mu(g) \, d\nu(h). \end{aligned} \tag{98.5}$$

Similarly,

$$\begin{aligned} \lambda((\mu * \nu) *_\rho v) & = \int_G \lambda(\rho_g(v)) \, d(\mu * \nu)(g) \\ & = \int_G \int_G \lambda(\rho_{gh}(v)) \, d\mu(g) \, d\nu(h), \end{aligned} \tag{98.6}$$

as desired.

Suppose now that ρ is continuous with respect to the weak operator topology on V, and let us show that

$$\mu *_{\rho'} (\nu *_{\rho'} \lambda) = (\mu * \nu) *_{\rho'} \lambda \tag{98.7}$$

for every $\lambda \in V'$ under suitable conditions. As before, it suffices to ask that $\|\rho_{g^{-1}}\|_{op}$ be integrable with respect to $|\mu|$, $|\nu|$, and hence $|\mu * \nu|$. Of course, (98.7) is the same as

$$(\mu *_{\rho'} (\nu *_{\rho'} \lambda))(v) = ((\mu * \nu) *_{\rho'} \lambda)(v) \tag{98.8}$$

for every $\lambda \in V'$ and $v \in V$. This follows from (97.1) in practically the same way as in the previous situation.

99 Unbounded continuous functions

Let Y be a locally compact Hausdorff topological space, and let $C(Y)$ be the space of real or complex-valued continuous functions on Y. If K is a nonempty compact subset of Y, then

$$N_K(f) = \sup\{|f(x)| : x \in K\} \tag{99.1}$$

defines a seminorm on $C(Y)$, the supremum seminorm associated to K. As in Section 1, the collection of these seminorms determines a topology on $C(Y)$. If Y is compact, then the seminorm associated to $K = Y$ is the same as the supremum norm, and $C(Y)$ is a Banach space.

In this section, we would like to focus on the case where Y is not compact. Thus continuous functions on Y need not be bounded, and the topology on Y is not determined by a single norm. If Y is σ-compact, so that Y is the union of a sequence of compact sets, then the topology on Y may be described by a sequence of seminorms. More precisely, one can use σ-compactness and local compactness to show that there is a sequence of compact subsets $K_1, K_2, \ldots$ of Y such that $Y = \bigcup_{l=1}^\infty K_l$ and K_l is contained in the interior of K_{l+1} for each l. Hence each compact set $K \subseteq Y$ is contained in the interior of K_l for some l, because K is contained in the union of the interiors of finitely many K_l's by compactness.

Suppose that λ is a continuous linear functional on $C(Y)$ with respect to the topology determined by the supremum seminorms on compact sets. As in Section 40, there are finitely many compact sets $K_1, \ldots, K_n \subseteq Y$ and a $C \geq 0$ such that

$$|\lambda(f)| \leq C \max(N_{K_1}(f), \ldots, N_{K_n}(f)) \tag{99.2}$$

for every $f \in C(Y)$. Equivalently,

$$|\lambda(f)| \leq C\, N_K(f), \tag{99.3}$$

where $K = K_1 \cup \cdots \cup K_n$. In particular,

$$\lambda(f) = 0 \tag{99.4}$$

when $f = 0$ on K. As usual, λ can be represented by a real or complex regular Borel measure on Y with support contained in K, as appropriate.

Now let X be another topological space, and suppose that Φ is a continuous mapping from X into $C(Y)$. Let $\phi(x, y)$ be the function on $X \times Y$ whose value at (x, y) is the same as the value of $\Phi(x) \in C(Y)$ at y. Thus $\phi(x, y)$ is continuous in y for each $x \in X$, because $\Phi(x) \in C(Y)$. The continuity of Φ as a mapping from X into $C(Y)$ means that for each compact set $K \subseteq Y$, the family of functions on X of the form $\phi(x, y)$ with $y \in K$ is equicontinuous at each point in X. It is easy to see that $\phi(x, y)$ is a continuous function on $X \times Y$ with respect to the product topology under these conditions, and conversely that a continuous function $\phi(x, y)$ on $X \times Y$ determines a continuous mapping $\Phi : X \to C(Y)$ in this way.

Suppose that X is locally compact and that μ is a real or complex regular Borel measure with compact support on X, as appropriate. Of course, one could just as well take X to be the support of μ, so that X is compact. In this case, the integral of a continuous mapping $\Phi : X \to C(Y)$ with respect to μ can be defined directly by

$$\Big(\int_X \Phi \, d\mu\Big)(y) = \int_X \phi(x, y) \, d\mu(x). \tag{99.5}$$

It is easy to see that this is a continuous function of y when $\phi(x, y)$ is a continuous function on $X \times Y$ and μ has compact support, because the family of functions on Y of the form $\phi(x, y)$ with x in a compact subset of X is equicontinuous at each point in Y.

If λ is a continuous linear functional on $C(Y)$, then

$$\lambda\Big(\int_X \Phi \, d\mu\Big) = \int_X \lambda(\Phi) \, d\mu. \tag{99.6}$$

Note that $\lambda(\Phi)$ is a continuous real or complex-valued function on X when $\Phi : X \to C(Y)$ is continuous, so that the right side of the equation is well-defined. If λ is represented by a regular Borel measure ν with compact support on Y, then (99.6) reduces to

$$\int_X \int_Y \phi(x, y) \, d\mu(x) \, d\nu(y) = \int_Y \int_X \phi(x, y) \, d\nu(y) \, d\mu(x). \tag{99.7}$$

One can also check that

$$N_K\Big(\int_X \Phi \, d\mu\Big) \le \int_X N_K(\Phi) \, d|\mu| \tag{99.8}$$

for each compact set $K \subseteq Y$.

If Y is compact, then $C(Y)$ is a Banach space, and this discussion of the integral of a continuous function with values in $C(Y)$ is consistent with the one in Section 95. Otherwise, the restriction of continuous functions on Y to a nonempty compact set $K \subseteq Y$ defines a continuous linear mapping from

$C(Y)$ into $C(K)$, where $C(K)$ is equipped with the supremum norm. Thus continuous mappings into $C(Y)$ lead to continuous mappings into $C(K)$ by composition. This is a convenient way to compare functions with values in $C(Y)$ with functions with values in a Banach space, as in Section 95.

Let G be a locally compact topological group. It is easy to see that the usual right and left translation operators on $C(G)$ are continuous with respect to the topology determined by the supremum seminorms on compact sets. The right and left regular representations of G on $C(G)$ are also continuous with respect to the analogue of the strong operator topology on $C(G)$ in this context, in the sense that $a \mapsto L_a(f), R_a(f)$ are continuous mappings from G into $C(G)$ for each continuous function f on G, as in Section 62. Convolutions of regular Borel measures with compact support on G and continuous functions on G may be considered as convolutions associated to the regular representations on $C(G)$, as in the previous sections.

100 Bounded continuous functions

Let Y be a locally compact Hausdorff topological space, and let $C_b(Y)$ be the space of bounded continuous real or complex-valued functions on Y. Of course, this is the same as the space $C(Y)$ of all continuous functions on Y when Y is compact. It is well known that $C_b(Y)$ is a Banach space with respect to the supremum norm

$$\|f\|_{sup} = \sup\{|f(y)| : y \in Y\}. \tag{100.1}$$

The space $C_0(Y)$ of continuous functions on Y that vanish at infinity is a closed linear subspace of $C_b(Y)$.

Suppose that μ is a positive Borel measure on Y such that $\mu(K) < \infty$ when $K \subseteq Y$ is compact and $\mu(U) > 0$ when $U \subseteq Y$ is nonempty and open. This implies that the essential supremum norm of a bounded continuous function b on Y with respect to μ is the same as the supremum norm. This is also the same as the operator norm of the multiplication operator associated to b on $L^p(Y)$, $1 \le p \le \infty$. In particular, the multiplication operators on $L^p(Y)$ associated to bounded continuous functions form a closed linear subspace of the space of all bounded linear operators on $L^p(Y)$ with respect to the operator norm.

One can also consider multiplication by $b \in C_b(Y)$ as a bounded linear operator on $C_0(Y)$ or $C_b(Y)$, whose operator norm is again equal to $\|b\|_{sup}$. If T is a bounded linear operator on $C_b(Y)$ that commutes with multiplication operators, then

$$T(f) = T(1)\, f \tag{100.2}$$

for every $f \in C_b(Y)$, so that T is a multiplication operator too. Similarly, one can show that a bounded linear operator T on $C_0(Y)$ that commutes with multiplication operators is a multiplication operator as well.

If $\{b_j\}_{j=1}^{\infty}$ is a uniformly bounded sequence of continuous functions on Y that converges uniformly on compact subsets of Y to a function b, then b is also a bounded continuous function on Y. It is easy to check that the corresponding sequence of multiplication operators on $C_0(Y)$ converges to the multiplication operator associated to b in the strong operator topology. More precisely, the topology on a bounded set $E \subseteq C_b(Y)$ with respect to the supremum norm induced by the strong operator topology on the corresponding multiplication operators on $C_0(Y)$ is the same as the topology induced on E by the topology on $C(Y)$ described in the previous section. However, the strong operator topology on multiplication operators on $C_b(Y)$ is the same as the topology determined by the operator norm, because of the way that a multiplication operator T on $C_b(Y)$ is determined by $T(1)$.

Let X be another topological space, and let $\phi(x, y)$ be a continuous real or complex-valued function on $X \times Y$ with respect to the product topology. Thus $\phi(x, y)$ determines a continuous mapping $\Phi : X \to C(Y)$ as in the previous section, where the value of $\Phi(x)$ at $y \in Y$ is $\phi(x, y)$. Suppose that $\phi(x, y)$ is uniformly bounded in y for each $x \in X$, which is the same as saying that $\Phi(x) \in C_b(Y)$. This does not necessarily mean that Φ is a continuous mapping from X into $C_b(Y)$ with respect to the supremum norm, unless Y is compact.

Observe that

$$\|\Phi(x)\|_{sup} = \sup\{|\phi(x, y)| : y \in Y\} \tag{100.3}$$

is lower semicontinuous as a function of x under these conditions. This is because

$$\|\Phi(x)\|_{sup} > t \tag{100.4}$$

if and only if $|\phi(x, y)| > t$ for some $y \in Y$, in which case $|\phi(w, y)| > t$ for w in a neighborhood of x. Thus (100.4) implies that $\|\Phi(w)\|_{sup} > t$ on a neighborhood of x, as desired.

Suppose that X is a locally compact Hausdorff space, and that μ is a real or complex regular Borel measure on X, as appropriate. The lower semicontinuity of $\|\Phi(x)\|_{sup}$ implies that it is Borel measurable, and we would like to ask now that it also be integrable with respect to $|\mu|$. If $\phi(x, y)$ is uniformly bounded in x and y, then this follows from the finiteness of $|\mu|(X)$. Under this integrability condition, we can define the integral of Φ with respect to μ pointwise on Y, as in (99.5). The resulting function on Y is uniformly bounded, with supremum norm less than or equal to the integral of $\|\Phi(x)\|_{sup}$ with respect to $|\mu|$, as usual.

If $\Phi : X \to C_b(Y)$ is continuous with respect to the supremum norm, then the integral of Φ can be defined as in Section 95, and it is easy to see that this is consistent with the pointwise definition. Otherwise, one can restrict functions on Y to a compact set $K \subseteq Y$, to get a continuous mapping from X into $C(K)$, to which the discussion in Section 95 can be applied. Alternatively, if the support of μ is compact, then the remarks in the previous section are applicable. The integral of Φ over X can be approximated by integrals of Φ over compact subsets of X under the integrability condition in the preceding paragraph, where the remainder has small supremum norm on Y. In particular, the integral of Φ with respect to μ determines a continuous function on Y under these conditions.

Let G be a locally compact topological group. The right and left regular representations of G on $C_b(G)$ are uniformly bounded, and they are continuous with respect to the strong operator topology associated to the topology on $C(G)$

described in the previous section. Continuity in the strong operator topology associated to the supremum norm on $C_b(G)$ requires uniform continuity of the functions on which the regular representations act. Convolution of a finite regular Borel measure on G with a bounded continuous function can be seen as an integral of translates of the function in the sense of the preceding paragraphs.

101 Discrete sets and groups

Let X be a discrete set, which is to say a set with the discrete topology. In practice, one might be especially interested in the case where X has only finitely or countably many elements. We also take X to be equipped with counting measure, so that integrals on X are reduced to sums. This can be handled directly, as follows.

If $f(x)$ is a nonnegative real-valued function on X, then

$$\sum_{x \in X} f(x) \tag{101.1}$$

can be defined as the supremum of the subsums

$$\sum_{x \in E} f(x) \tag{101.2}$$

over all finite subsets E of X. This is interpreted as being $+\infty$ when the finite sums are unbounded. We say that f is *summable* on X when the subsums (101.2) are bounded, so that (101.1) is finite. If f, g are nonnegative real-valued summable functions on X and a, b are nonnegative real numbers, then it is easy to see that $a\,f + b\,g$ is summable too, and that

$$\sum_{x \in X} \big(a\,f(x) + b\,g(x)\big) = a \sum_{x \in X} f(x) + b \sum_{x \in X} g(x). \tag{101.3}$$

If f is a real or complex-valued function on X, then f is said to be summable on X when $|f|$ is summable on X. The sum (101.1) of a summable real or complex-valued function f on X can be defined by first expressing f as a linear combination of nonnegative real-valued summable functions and then applying the previous definition to those functions. Note that linear combinations of summable real or complex-valued functions on X are summable, and that the sum (101.1) is linear in f over the real or complex numbers.

Because X is discrete, the support of a function f on X is simply the set of $x \in X$ such that $f(x) \neq 0$. If f has finite support, then f is obviously summable, and (101.1) reduces to a finite sum. The support of any summable function on X has only finitely or countably many elements, and the sum (101.1) can be identified with an infinite series when f has infinite support. Summability of the function corresponds to absolute convergence of the series, which implies that the sum is invariant under rearrangements.

Suppose now that $(V, \|v\|_V)$ is a real or complex Banach space, and that f is a V-valued function on X. Under these conditions, we say that f is summable on X if $\|f(x)\|_V$ is summable as a nonnegative real-valued function on X. Let $\ell^1(X, V)$ be the space of V-valued summable functions on X. It is easy to see that this is a vector space over the real or complex numbers, as appropriate, and that

$$\|f\|_1 = \|f\|_{\ell^1(X,V)} = \sum_{x \in X} \|f(x)\|_V \tag{101.4}$$

defines a norm on $\ell^1(X, V)$. One can also use completeness of V to show that $\ell^1(X, V)$ is a Banach space with respect to this norm.

If f is a summable V-valued function on X, then one would like to define the sum (101.1) as an element of V, and with the property that

$$\Big\| \sum_{x \in X} f(x) \Big\|_V \le \sum_{x \in X} \|f(x)\|_V. \tag{101.5}$$

If f has finite support, then (101.1) reduces to a finite sum, and (101.5) follows from the triangle inequality. The collection of V-valued functions with finite support on X is a dense linear subspace of $\ell^1(X, V)$, and this version of the sum determines a bounded linear mapping from this dense linear subspace into V. As in Section 9, the completeness of V implies that there is a unique extension of this mapping to a bounded linear mapping from $\ell^1(X, V)$ into V with norm 1, which can be used as the definition of (101.1) for any $f \in \ell^1(X, V)$. In particular, this approach to the sum can also be used when $V = \mathbf{R}$ or $\mathbf{C}$. Alternatively, the sum can be treated in terms of infinite series, as in the previous paragraph.

If λ is a bounded linear functional on V and f is a V-valued summable function on X, then $\lambda \circ f$ is a summable real or complex-valued function on X, as appropriate, because

$$\sum_{x \in X} |\lambda(f(x))| \le \sum_{x \in X} \|\lambda\|_{V'} \, \|f(x)\|_V \le \|\lambda\|_{V'} \, \|f\|_{\ell^1(X,V)}, \tag{101.6}$$

where $\|\lambda\|_{V'}$ denotes the dual norm of λ. Moreover,

$$\lambda\Big(\sum_{x \in X} f(x) \Big) = \sum_{x \in X} \lambda(f(x)). \tag{101.7}$$

This is clear when f has finite support, and otherwise one can approximate f in $\ell^1(X, V)$ by functions with finite support. Note that this discussion of the sum of a summable V-valued function on X may be considered as a special case of the integrals in Section 95, since discrete spaces are locally compact and arbitrary functions on them are continuous.

Now let Γ be a discrete group, which is to say a group equipped with the discrete topology. In practice, one may be especially interested in groups with only finitely or countably many elements, including finitely-generated groups. As in Section 63, counting measure on Γ is obviously invariant under translations on the right and on the left, and may be used as Haar measure.

Let $(V, \|v\|_V)$ be a real or complex Banach space, and suppose that ρ is a representation of Γ by bounded linear operators on V. Because Γ is discrete, one need not be concerned with continuity properties of ρ. Also let f be a summable real or complex-valued function on Γ, as appropriate. The restriction of f to any set $E \subseteq \Gamma$ is summable as well, and so we get a measure μ_f on Γ defined by

$$\mu_f(E) = \sum_{x \in E} f(x). \tag{101.8}$$

As in Section 96, we can define the convolution of μ_f with $v \in V$ with respect to ρ by

$$\mu_f *_\rho v = \sum_{x \in \Gamma} f(x)\, \rho_x(v) \tag{101.9}$$

under the summability condition

$$\sum_{x \in \Gamma} |f(x)|\, \|\rho_x(v)\|_V < \infty. \tag{101.10}$$

This summability condition holds for every $v \in V$ when

$$\sum_{x \in \Gamma} |f(x)|\, \|\rho_x\|_{op} < \infty, \tag{101.11}$$

where $\|\rho_x\|_{op}$ is the operator norm of ρ_x on V, as usual. In particular, this condition is satisfied when ρ is bounded. More precisely, (101.10) says exactly that $f(x)\, \rho_x(v)$ is a summable V-valued function on Γ, and so its sum can be defined as before. Similarly, (101.11) says that $f(x)\, \rho_x$ is a summable $\mathcal{BL}(V)$-valued function on Γ, which implies that its sum can be defined as an element of $\mathcal{BL}(V)$ as before.

102 Group algebras

Let G be a group. The corresponding *group algebra* over the real or complex numbers consists of formal sums

$$\sum_{j=1}^{l} r_j\, g_j, \tag{102.1}$$

where $g_1, \ldots, g_l \in G$ and $r_1, \ldots, r_l \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. Addition and scalar multiplication are defined termwise, so that elements of the group form a basis of the group algebra as a vector space. Multiplication is defined as a bilinear extension of group multiplication, which can be given explicitly by the equation

$$\Big(\sum_{j=1}^{l} r_j\, g_j\Big) \Big(\sum_{k=1}^{n} t_k\, h_k\Big) = \sum_{j=1}^{l} \sum_{k=1}^{n} r_j\, t_k\, g_j\, h_k. \tag{102.2}$$

It follows from associativity of the group operation that multiplication in the group algebra is associative too. If the group is commutative, then the group algebra is commutative as well. The identity element of the group is also the multiplicative identity element of the group algebra.

Suppose that ρ is a representation of G on a real or complex vector space V. If V is equipped with a norm, then we may ask that ρ act by bounded linear operators on V. If $\sum_{j=1}^{l} r_j\, g_j$ is an element of the corresponding real or complex group algebra, as appropriate, then we get a linear operator on V defined by

$$v \mapsto \sum_{j=1}^{l} r_j\, \rho_{g_j}(v). \tag{102.3}$$

Equivalently, this determines a mapping

$$\sum_{j=1}^{l} r_j\, g_j \mapsto \sum_{j=1}^{l} r_j\, \rho_{g_j} \tag{102.4}$$

from the group algebra into the space of linear operators on V. It is easy to see that this mapping is a homomorphism from the group algebra into the algebra of linear mappings on V with respect to composition.

Real or complex regular Borel measures on a locally compact group also form an associative algebra with respect to convolution. Elements of the group may be identified with Dirac measures on the group, so that linear combinations of elements of the group are identified with linear combinations of Dirac measures. With respect to this identification, convolution of these measures is the same as the previous definition of multiplication. The convolution of a measure with a representation of the group is also compatible with the mapping described in the preceding paragraph.

Alternatively, one can consider convolution algebras of integrable functions with respect to right or left-invariant Haar measure on a locally compact group. This is basically the same a looking at measures on the group that are absolutely continuous with respect to right or left-invariant Haar measure. On discrete groups, this reduces to convolution algebras of summable functions. Summable functions include functions with finite support as a dense linear subspace, which can be identified with linear combinations of group elements as before.

Part III
Self-adjoint linear operators

103 Self-adjoint operators, 2

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. Put

$$\alpha(A) = \inf\{\langle A(v), v\rangle : v \in V, \|v\| = 1\}. \tag{103.1}$$

Thus $A - \alpha(A)\, I$ is a nonnegative self-adjoint operator which is not uniformly strictly positive in the sense of Section 47. Conversely, one can check that $\alpha(A)$ is uniquely determined by these two conditions. Equivalently, $A - \lambda\, I$ is uniformly strictly positive when $\lambda \in \mathbf{R}$ satisfies $\lambda < \alpha(A)$, while $A - \lambda\, I$ is not nonnegative when $\lambda > \alpha(A)$. As in Section 47, a bounded nonnegative self-adjoint linear operator on V is invertible if and only if it is uniformly strictly positive. Hence $A - \alpha(A)\, I$ is not invertible, and $A - \lambda\, I$ is invertible when $\lambda < \alpha(A)$. If

$$\sigma(A) = \{\lambda \in \mathbf{R} : A - \lambda\, I \text{ is not invertible}\}, \tag{103.2}$$

then it follows that

$$\alpha(A) = \min \sigma(A). \tag{103.3}$$

In particular, $A \geq 0$ if and only if $A + t\, I$ is invertible for every $t > 0$, which is the same as saying that $\min \sigma(A) \geq 0$.

Similarly, put

$$\beta(A) = \sup\{\langle A(v), v\rangle : v \in V, \|v\| = 1\}. \tag{103.4}$$

Note that

$$\alpha(A) \leq \beta(A) \tag{103.5}$$

and

$$-\beta(A) = \alpha(-A). \tag{103.6}$$

Also,

$$\alpha(r\, A) = r\, \alpha(A), \quad \beta(r\, A) = r\, \beta(A) \tag{103.7}$$

for each nonnegative real number r, and

$$\alpha(A + t\, I) = \alpha(A) + t, \quad \beta(A + t\, I) = \beta(A) + t \tag{103.8}$$

for every $t \in \mathbf{R}$. Of course,

$$\alpha(I) = \beta(I) = 1, \tag{103.9}$$

and if B is another bounded self-adjoint linear operator on V, then

$$\alpha(A) + \alpha(B) \leq \alpha(A + B), \quad \beta(A + B) \leq \beta(A) + \beta(B). \tag{103.10}$$

If $\alpha(A) = \beta(A)$, then one can check that

$$A = \alpha(A)\, I = \beta(A)\, I, \tag{103.11}$$

by reducing to the case where $\alpha(A) = \beta(A) = 0$ and applying the Cauchy–Schwarz inequality as in Section 34.

Observe that

$$\sigma(-A) = -\sigma(A) = \{-\lambda : \lambda \in \sigma(A)\}. \tag{103.12}$$

More precisely, this is the same as saying that

$$(-A) - (-\lambda)\, I = -(A - \lambda\, I) \tag{103.13}$$

is invertible if and only if $A - \lambda I$ is invertible. Using (103.6), we get that

$$\beta(A) = \max \sigma(A). \tag{103.14}$$

It is easy to see that $\sigma(r\, I) = \{r\}$ for each $r \in \mathbf{R}$. Conversely, if $\sigma(A) = \{r\}$ for some $r \in \mathbf{R}$, then $\alpha(A) = \beta(A) = r$, and $A = r\, I$.

If V is a complex Hilbert space, A is a bounded self-adjoint linear operator on V, and λ is a complex number with nonzero imaginary part, then $A - \lambda I$ is automatically invertible on V. To see this, remember that

$$\langle A(v), v \rangle \in \mathbf{R} \tag{103.15}$$

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

$$\operatorname{Im}\langle (A - \lambda\, I)(v), v \rangle = (\operatorname{Im} \lambda)\, \|v\|^2, \tag{103.16}$$

and therefore

$$\|(A - \lambda\, I)(v)\| \ge |\operatorname{Im} \lambda|\, \|v\|, \tag{103.17}$$

for every $v \in V$. Of course, $A - \overline{\lambda}\, I$ has the same property, which shows that the kernel of $(A - \lambda\, I)^* = A - \overline{\lambda}\, I$ is trivial. The invertibility of $A - \lambda\, I$ when $\operatorname{Im} \lambda \ne 0$ now follows as in Section 46.

104 Invertibility and compositions

Let V be a real or complex vector space, and let A, B be linear transformations on V. If A, B are invertible, then so is their composition $A \circ B$, and

$$(A \circ B)^{-1} = B^{-1} \circ A^{-1}. \tag{104.1}$$

Conversely, if $A \circ B$ is invertible, then A is invertible on the right and B is invertible on the left, because

$$A \circ (B \circ (A \circ B)^{-1}) = (A \circ B) \circ (A \circ B)^{-1} = I \tag{104.2}$$

and similarly

$$((A \circ B)^{-1} \circ A) \circ B = (A \circ B)^{-1} \circ (A \circ B) = I. \tag{104.3}$$

This is sufficient to imply the invertibility of A, B when V has finite dimension, but not when V is infinite-dimensional, as one can see from examples of shift operators on spaces of sequences of real or complex numbers.

However, if A and B commute, then the invertibility of $A \circ B$ does imply the invertibility of A, B. In this case, A and B commute with $A \circ B$, and hence with $(A \circ B)^{-1}$. This implies that the one-sided inverses in the previous paragraph are actually two-sided inverses.

If V is equipped with a norm, then we may be especially interested in bounded linear operators on V, and operators with bounded inverses. If A, B

are commuting bounded linear operators on V such that $A \circ B$ has a bounded inverse, then the same argument shows that A and B have bounded inverses.

Suppose now that V is a complex Hilbert space, and that A is a bounded self-adjoint linear operator on V. Thus

$$(A - \lambda\, I) \circ (A - \overline{\lambda}\, I) = (A - (\operatorname{Re} \lambda)\, I)^2 + (\operatorname{Im} \lambda)^2\, I \tag{104.4}$$

for each complex number λ. Observe that $A - (\operatorname{Re} \lambda)\, I$ is self-adjoint, which implies that $(A - (\operatorname{Re} \lambda)\, I)^2$ is self-adjoint and nonnegative. Hence (104.4) is uniformly strictly positive when $\operatorname{Im} \lambda \neq 0$, and is therefore invertible. Because $A - \lambda\, I$ and $A - \overline{\lambda}\, I$ commute, it follows that they are also invertible when $\operatorname{Im} \lambda \neq 0$, as in the previous section.

105 A few simple applications

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. Note that A^2 is self-adjoint and nonnegative, which implies that

$$\sigma(A^2) \subseteq [0, \infty). \tag{105.1}$$

We would like to check that

$$\sigma(A^2) = \{\lambda^2 : \lambda \in \sigma(A)\}. \tag{105.2}$$

The main point is that

$$A^2 - \lambda^2\, I = (A + \lambda\, I) \circ (A - \lambda\, I) \tag{105.3}$$

for every real number λ. This implies that $A^2 - \lambda^2\, I$ is invertible if and only if $A + \lambda\, I$ and $A - \lambda\, I$ are invertible, by the remarks in the previous section. Of course, $A + \lambda\, I$ and $A - \lambda\, I$ obviously commute with each other. It follows that $A^2 - \lambda^2\, I$ is not invertible if and only if at least one of $A + \lambda\, I$ and $A - \lambda\, I$ is not invertible. This is exactly the same as (105.2), as desired.

Let us now use this to show that

$$\|A\|_{op} = \max\{|\lambda| : \lambda \in \sigma(A)\}. \tag{105.4}$$

Observe that

$$\max\{|\lambda| : \lambda \in \sigma(A)\} = \max(-\alpha(A), \beta(A)), \tag{105.5}$$

since $\alpha(A) = \min \sigma(A)$ and $\beta(A) = \max \sigma(A)$, as in Section 103. The Cauchy–Schwarz inequality and the definition of the operator norm imply that

$$|\alpha(A)|, |\beta(A)| \leq \|A\|_{op}. \tag{105.6}$$

If A is nonnegative, then $\beta(A) = \|A\|_{op}$, as in Section 34. It is easy to see that

$$\beta(A^2) = \|A\|_{op}^2, \tag{105.7}$$

because

$$\langle A^2(v), v\rangle = \langle A(v), A(v)\rangle = \|A(v)\|^2 \tag{105.8}$$

for each $v \in V$. Hence

$$\max \sigma(A^2) = \|A\|_{op}^2, \tag{105.9}$$

because $\max \sigma(A^2) = \beta(A^2)$. Thus (105.2) implies that

$$\|A\|_{op}^2 = \max\{\lambda^2 : \lambda \in \sigma(A)\}, \tag{105.10}$$

which is the same as (105.4).

Suppose now that A is invertible, and let us determine $\sigma(A^{-1})$. In this case, $0 \notin \sigma(A), \sigma(A^{-1})$, and we would like to show that

$$\sigma(A^{-1}) = \{\lambda^{-1} : \lambda \in \sigma(A)\}. \tag{105.11}$$

This time, we use the identity

$$A^{-1} - \lambda^{-1} = -(\lambda\, A)^{-1}(A - \lambda\, I) \tag{105.12}$$

for each real number $\lambda \neq 0$. Because $(\lambda\, A)^{-1}$ is invertible and commutes with $A - \lambda\, I$, it follows from the remarks in the previous section that $A^{-1} - \lambda^{-1}\, I$ is invertible if and only if $A - \lambda\, I$ is invertible. Equivalently, $A^{-1} - \lambda^{-1}\, I$ is not invertible if and only if $A - \lambda\, I$ is not invertible, which is what we wanted.

In particular, this applies to the situation where A is uniformly strictly positive. This is the same as saying that $\alpha(A) > 0$, and we also have that $A^{-1} \geq 0$, as in Section 47. It is easy to see that

$$\|A^{-1}\|_{op} \leq \alpha(A)^{-1}, \tag{105.13}$$

as in Sections 46 and 47. Using the identification of $\sigma(A^{-1})$ in the preceding paragraph, we get that

$$\|A^{-1}\|_{op} = \alpha(A)^{-1}. \tag{105.14}$$

More precisely, $\|A^{-1}\|_{op} = \beta(A^{-1}) = \max \sigma(A^{-1})$ is equal to the reciprocal of $\min \sigma(A) = \alpha(A)$.

106 Polynomials

Let

$$p(z) = a_n\, z^n + a_{n-1}\, z^{n-1} + \cdots + a_1\, z + a_0 \tag{106.1}$$

be a polynomial of degree n with complex coefficients, which is to say that $a_0, \ldots, a_n \in \mathbf{C}$ and $a_n \neq 0$. The fundamental theorem of algebra implies that there are complex numbers $a, z_1, \ldots, z_n$ such that $a \neq 0$ and

$$p(z) = a \prod_{j=1}^{n} (z - z_j). \tag{106.2}$$

More precisely, $z_1, \ldots, z_n$ are exactly the zeros of $p(z)$ as a function on the complex plane, with their appropriate multiplicities, and a is the same as the leading coefficient a_n of $p(z)$.

Now let V be a complex vector space, and let T be a linear transformation on V. We can define $p(T)$ as a linear transformation on V by

$$p(T) = a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I, \tag{106.3}$$

where T^j is the composition of j T's, which is interpreted as being the identity operator I when $j = 0$. Using (106.2), this is the same as

$$p(T) = a \prod_{j=1}^{n} (T - z_j I), \tag{106.4}$$

where the product represents the composition of the factors $T - z_j I$. The order of the factors in the product does not matter, because they commute with each other. As in Section 104, $p(T)$ is invertible if and only if $T - z_j I$ is invertible for $j = 1, \ldots, n$.

Suppose now that p is a real polynomial, in the sense that the coefficients $a_0, \ldots, a_n$ are real numbers. This is equivalent to the conditions that

$$\overline{p(z)} = p(\overline{z}) \tag{106.5}$$

for each $z \in \mathbf{C}$, or that

$$p(x) \in \mathbf{R} \tag{106.6}$$

for every $x \in \mathbf{R}$. In this case, the zeros of $p(z)$ as a function on $\mathbf{C}$ consist of the real zeros $x_1, \ldots, x_l$, the complex zeros $\zeta_1, \ldots, \zeta_r$ in the upper half plane $\operatorname{Im} \zeta > 0$, and the complex conjugates of the latter $\overline{\zeta_1}, \ldots, \overline{\zeta_r}$, so that $n = l + 2r$. This leads to the expression

$$p(x) = a \prod_{j=1}^{l} (x - x_j) \prod_{k=1}^{r} ((x - \operatorname{Re} \zeta_k)^2 + (\operatorname{Im} \zeta_k)^2) \tag{106.7}$$

for $x \in \mathbf{R}$.

If V is a real vector space, T is a linear transformation on V, and p is a real polynomial, then we can define $p(T)$ as a linear transformation on V in the same way as before. We can also use (106.7) to express $p(T)$ as a product of commuting factors,

$$p(T) = a \prod_{j=1}^{l} (T - x_j I) \prod_{k=1}^{r} ((T - \operatorname{Re} \zeta_k I)^2 + (\operatorname{Im} \zeta_k)^2 I). \tag{106.8}$$

If V is a real or complex Hilbert space, T is a bounded linear operator on V, and p is a real polynomial, then

$$p(T^*) = p(T)^*. \tag{106.9}$$

In particular, $p(T)$ is self-adjoint when T is self-adjoint and p is real.

107 Nonnegative real polynomials

Let p be a polynomial with real coefficients such that

$$p(x) \geq 0 \tag{107.1}$$

for each $x \in \mathbf{R}$. In this case, the real roots of p have even multiplicity, and the fundamental theorem of algebra implies that

$$p(x) = a \prod_{j=1}^{l} ((x - x_j)^2 + y_j^2) \tag{107.2}$$

for some real numbers $a, x_1, y_1, \ldots, x_l, y_l$, with $a \geq 0$. By expanding the product of sums into a sum of products, we can express p as

$$p(x) = \sum_{k=1}^{r} q_k(x)^2, \tag{107.3}$$

where $q_1, \ldots, q_k$ are also polynomials with real coefficients. If $p(x) > 0$ for every $x \in \mathbf{R}$, then $y_1, \ldots, y_l \neq 0$, and we can take one of the q_k's to be a nonzero constant. Conversely, a polynomial of the form (107.3) is automatically nonnegative, and strictly positive when one of the q_k's is a nonzero constant.

Here is a more elementary derivation of (107.3), using induction. The base case concerns constant polynomials, and is clear. If p is a nonnegative real polynomial of positive degree, then $p(x) \to \infty$ as $|x| \to \infty$, and the minimum of p on $\mathbf{R}$ is attained at some $x_0 \in \mathbf{R}$. This implies that

$$p(x) = (x - x_0)^2 \, p_1(x) + p(x_0) \tag{107.4}$$

for some nonnegative real polynomial $p_1(x)$. More precisely, this uses the fact that $p(x) - p(x_0)$ is a nonnegative real polynomial, so that its root x_0 has even multiplicity. If the multiplicity is larger than 2, or if $p(x) - p(x_0)$ has other roots, then additional factors can be extracted from $p(x) - p(x_0)$, but the main point is that the degree of p_1 is strictly less than the degree of p. Hence p_1 is a sum of squares by induction, and it follows that p is a sum of squares too. If p is strictly positive, then $p(x_0) > 0$, and one of the terms can be taken to be the square of a nonzero constant.

Let V be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. If p is a nonnegative real polynomial, then it follows from (107.3) that $p(A)$ is a sum of squares of self-adjoint operators on V, and hence

$$p(A) \geq 0. \tag{107.5}$$

If p is strictly positive, then

$$p(A) \geq \delta \, I \tag{107.6}$$

for some $\delta > 0$, and $p(A)$ is invertible.

Suppose now that p is a real polynomial which is not identically zero, and let $x_1, \ldots, x_l$ be the real roots of p, with their appropriate multiplicity. Thus

$$p(x) = \pm p_+(x) \prod_{j=1}^{l} (x - x_j), \tag{107.7}$$

where p_+ is a real polynomial with no real roots, and the sign can be chosen so that p_+ is positive. This uses the fact that a continuous non-vanishing real-valued function on the real line has constant sign, by the intermediate value theorem. If A is a bounded self-adjoint linear operator on a real or complex Hilbert space V, then $p_+(A)$ is invertible on V, as in the previous paragraph. It follows from the remarks in Section 104 that $p(A)$ is invertible if and only if $A - x_j\, I$ is invertible for $j = 1, \ldots, l$.

108 Polynomials and operators

Let V be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. Also let p be a polynomial with real coefficients, so that $p(A)$ is self-adjoint on V as well. We would like to check that

$$\sigma(p(A)) = \{p(\lambda) : \lambda \in \sigma(A)\}. \tag{108.1}$$

As in the previous section, $p(A)$ is invertible if and only if $A - \lambda\, I$ is invertible for each $\lambda \in \mathbf{R}$ such that $p(\lambda) = 0$. Thus $p(A)$ is not invertible if and only if $A - \lambda\, I$ is not invertible for some $\lambda \in \mathbf{R}$ such that $p(\lambda) = 0$. Equivalently, $0 \in \sigma(p(A))$ if and only if $p(\lambda) = 0$ for some $\lambda \in \sigma(A)$. If $\mu \in \mathbf{R}$, then we can apply the preceding statement to $\widetilde{p}(x) = p(x) - \mu$ to get that $\mu \in \sigma(p(A))$ if and only if $p(\lambda) = \mu$ for some $\lambda \in \sigma(A)$, which is the same as (108.1).

Note that $\sigma(A)$ is automatically a closed set in the real line, because the set of $\lambda \in \mathbf{R}$ such that $A - \lambda\, I$ is invertible is an open set, as in Section 90. Since $\alpha(p(A)) = \min \sigma(p(A))$, as in Section 103, we get that

$$\alpha(p(A)) = \min\{p(\lambda) : \lambda \in \sigma(A)\}. \tag{108.2}$$

In particular, $p(A) \geq 0$ if and only if $p(\lambda) \geq 0$ for every $\lambda \in \sigma(A)$. Similarly,

$$\|p(A)\|_{op} = \max\{|p(\lambda)| : \lambda \in \sigma(A)\} \tag{108.3}$$

by (105.4) applied to $p(A)$.

We can extend these results to real rational functions without poles in $\sigma(A)$, as follows. Let q be a real polynomial such that $q(\lambda) \neq 0$ for every $\lambda \in \sigma(A)$. Thus $q(A)$ is an invertible self-adjoint linear operator on V, which implies that $q(A)^{-1}$ is self-adjoint as well. If p is any real polynomial, then $p(A)$ and $q(A)$ commute, and so $p(A)$ commutes with $q(A)^{-1}$ too. This implies that $p(A)\, q(A)^{-1}$ is self-adjoint, because the composition of commuting self-adjoint operators is also self-adjoint.

Observe that $p(A)\,q(A)^{-1}$ is invertible if and only if $p(A)$ is invertible, since $q(A)^{-1}$ is invertible. Similarly, for each $\mu \in \mathbf{R}$,

$$p(A)\,q(A)^{-1} - \mu\, I = (p(A) - \mu\, q(A))\, q(A)^{-1} \tag{108.4}$$

is invertible if and only if $p(A) - \mu\, q(A)$ is invertible. Equivalently, (108.4) is not invertible exactly when $p(A) - \mu\, q(A)$ is not invertible. As before, this happens if and only if $p(\lambda) - \mu\, q(\lambda) = 0$ for some $\lambda \in \sigma(A)$, which is the same as $\mu = p(\lambda)/q(\lambda)$. Hence $\sigma(p(A)\,q(A)^{-1})$ consists of the real numbers of the form $p(\lambda)/q(\lambda)$ with $\lambda \in \sigma(A)$, which can be used to analyze the operator norm and nonnegativity of $p(A)\,q(A)^{-1}$ as in the previous situation.

109 Functional calculus

Let V be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. Thus $\sigma(A)$ is a nonempty compact set in the real line, and we let $C(\sigma(A))$ be the space of continuous real-valued functions on $\sigma(A)$. Also let $\mathcal{P}(\sigma(A))$ be the linear subspace of $C(\sigma(A))$ consisting of the restrictions of real polynomials to $\sigma(A)$. The Lebesgue–Stone–Weierstrass theorem implies that $\mathcal{P}(\sigma(A))$ is dense in $C(\sigma(A))$ with respect to the supremum norm. Note that the restriction of a real polynomial $p(x)$ to $\sigma(A)$ uniquely determines $p(x)$ when $\sigma(A)$ has infinitely many elements. Even if $\sigma(A)$ has only finitely many elements, $p(A)$ is uniquely determined by the restriction of p to $\sigma(A)$. This is because $p(A) = 0$ when $p(\lambda) = 0$ for every $\lambda \in \sigma(A)$, by (108.3).

It follows that

$$f \mapsto f(A) \tag{109.1}$$

defines a linear mapping from $\mathcal{P}(\sigma(A))$ into the vector space of bounded self-adjoint linear operators on V. Moreover, (109.1) is an isometry with respect to the supremum norm on $\mathcal{P}(\sigma(A))$ and the operator norm on $\mathcal{BL}(V)$, because of (108.3). This implies that (109.1) has a unique extension to an isometric linear embedding of $C(\sigma(A))$ with the supremum norm into the space of bounded self-adjoint linear operators on V with the operator norm, as in Section 9. If $f_1, f_2 \in C(\sigma(A))$, then one can check that this embedding sends their product to the composition of the corresponding operators $f_1(A)$, $f_2(A)$ on V, or

$$(f_1\, f_2)(A) = f_1(A)\, f_2(A), \tag{109.2}$$

because of the analogous statement for polynomials. In particular, if $f(x) = p(x)/q(x)$, where p, q are real polynomials and $q(\lambda) \neq 0$ for each $\lambda \in \sigma(A)$, then $f(A) = p(A)\,q(A)^{-1}$.

If f is a continuous real-valued function on $\sigma(A)$ such that $f(\lambda) \neq 0$ for each $\lambda \in \sigma(A)$, then $1/f$ is also a continuous function on $\sigma(A)$. It follows that $f(A)$ is invertible on V, with $f(A)^{-1} = (1/f)(A)$. Conversely, if $f(A)$ is invertible as a linear operator on V, then $f(\lambda) \neq 0$ for every $\lambda \in \sigma(A)$. For if $f(\lambda_0) = 0$ for some $\lambda_0 \in \sigma(A)$, then f can be approximated uniformly on $\sigma(A)$ by polynomials p such that $p(\lambda_0) = 0$. As before, $p(A)$ is not invertible when p has this property.

However, if $f(A)$ is invertible, then any bounded linear operator on V which is sufficiently close to $f(A)$ in the operator norm is invertible, as in Section 90. This is a contradiction, since the operator norm of $f(A) - p(A)$ is equal to the supremum norm of $f - p$ on $\sigma(A)$.

This implies that

$$\sigma(f(A)) = \{f(\lambda) : \lambda \in \sigma(A)\} \tag{109.3}$$

for every $f \in C(\sigma(A))$. More precisely, $0 \in \sigma(f(A))$ if and only if $f(\lambda) = 0$ for some $\lambda \in \sigma(A)$, by the remarks in the preceding paragraph. Similarly, $\mu \in \sigma(f(A))$ if and only if $\mu = f(\lambda)$ for some $\lambda \in \sigma(A)$, by the previous argument applied to $\widetilde{f}(x) = f(x) - \mu$. As a consequence, note that

$$\alpha(f(A)) = \min\{f(\lambda) : \lambda \in \sigma(A)\}, \tag{109.4}$$

since $\alpha(f(A)) = \min \sigma(f(A))$, as in Section 103. In particular, $f(A) \geq 0$ if and only if $f(\lambda) \geq 0$ for each $\lambda \in \sigma(A)$.

110 Step functions

For each positive integer j, let f_j be the continuous piecewise-linear function on the real line defined by

$$\begin{aligned} f_j(x) &= 0 && \text{when } x \leq -2^{-j}, \\ &= 1 + 2^j\, x && \text{when } -2^{-j} \leq x \leq 0 \\ &= 1 && \text{when } x \geq 0. \end{aligned} \tag{110.1}$$

Thus

$$0 \leq f_{j+1}(x) \leq f_j(x) \leq 1 \tag{110.2}$$

for every $x \in \mathbf{R}$ and $j \geq 1$, and $\{f_j\}_{j=1}^\infty$ converges pointwise on $\mathbf{R}$ to the step function f defined by

$$\begin{aligned} f(x) &= 0 && \text{when } x < 0 \\ &= 1 && \text{when } x \geq 0. \end{aligned} \tag{110.3}$$

Moreover,

$$f_{j+1}(x) \leq f_j(x)^2 \leq f_j(x), \tag{110.4}$$

because

$$(1 + 2^j\, x)^2 = 1 + 2^{j+1}\, x + 2^{2j}\, x^2 \geq 1 + 2^{j+1}\, x \tag{110.5}$$

for each x and j.

Let V be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. If $f_j(A)$ is the self-adjoint operator on V associated to f_j as in the previous section, then

$$0 \leq f_{j+1}(A) \leq f_j(A) \leq I \tag{110.6}$$

for each $j \geq 1$, because of (110.2). As in Section 34, $\{f_j(A)\}_{j=1}^\infty$ converges in the strong operator topology to a bounded self-adjoint linear operator P on V that satisfies

$$0 \leq P \leq I. \tag{110.7}$$

Basically, this is a way to define $f(A)$, even though f is not continuous on $\mathbf{R}$ and hence perhaps not on $\sigma(A)$.

Similarly, (110.4) implies that

$$f_{j+1}(A) \leq f_j(A)^2 \leq f_j(A), \tag{110.8}$$

which is the same as

$$\langle f_{j+1}(A)(v), v\rangle \leq \langle f_j(A)(v), f_j(A)(v)\rangle \leq \langle f_j(A)(v), v\rangle \tag{110.9}$$

for each $v \in V$ and $j \geq 1$, because $f_j(A)$ is self-adjoint. Hence

$$\langle P(v), v\rangle = \langle P(v), P(v)\rangle, \tag{110.10}$$

by the strong convergence of $\{f_j(A)\}_{j=1}^\infty$ to P. Using the self-adjointness of P and polarization, we get that

$$\langle P(v), w\rangle = \langle P(v), P(w)\rangle = \langle P^2(v), w\rangle \tag{110.11}$$

for each $v, w \in V$, which is to say that $P^2 = P$. This corresponds to the fact that $f^2 = f$ on $\mathbf{R}$.

Thus P is a bounded self-adjoint projection operator on V, which means that P is the orthogonal projection of V onto $W = P(V)$. More precisely, W can also be characterized as the kernel of $I - P$, which makes it clear that W is closed. Because P is self-adjoint, it is easy to see that the kernel of P is the same as the orthogonal complement of W. For each $v \in V$, $P(v) \in W$, and $v - P(v)$ is in the kernel of P. This shows that $P(v)$ is the same as the orthogonal projection of v onto W.

111 Composition of functions

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space V, let ψ be a continuous real-valued function on $\sigma(A)$, and let ϕ be a continuous real-valued function on

$$\sigma(\psi(A)) = \psi(\sigma(A)). \tag{111.1}$$

Thus the composition

$$(\phi \circ \psi)(x) = \phi(\psi(x)) \tag{111.2}$$

of ϕ and ψ is a continuous real-valued function on $\sigma(A)$. Let us check that

$$(\phi \circ \psi)(A) = \phi(\psi(A)), \tag{111.3}$$

where $\psi(A)$ and $(\phi \circ \psi)(A)$ are defined using the functional calculus applied to A, and $\phi(\psi(A))$ is defined using the functional calculus applied to $\psi(A)$. If ϕ is a polynomial, then this follows from the fact that the functional calculus is an algebra homomorphism from continuous functions on $\sigma(A)$ into bounded linear operators on V. If ϕ is an arbitrary continuous function on $\psi(\sigma(A))$, then ϕ can be approximated by polynomials uniformly on $\psi(\sigma(A))$, and one can use the isometric property of the functional calculus with respect to the supremum norm to pass to the limit.

As another way to look at this,

$$\phi \mapsto \phi \circ \psi \tag{111.4}$$

defines an algebra homomorphism from $C(\psi(\sigma(A)))$ into $C(\sigma(A))$, which is also an isometry with respect to the supremum norm. The functional calculus determines algebra homomorphisms from $C(\psi(\sigma(A)))$ and $C(\sigma(A))$ into $\mathcal{BL}(V)$. The statement in the previous paragraph is equivalent to the identification of the homomorphism from $C(\psi(\sigma(A)))$ into $\mathcal{BL}(V)$ with the composition of the homomorphisms from $C(\psi(\sigma(A)))$ into $C(\sigma(A))$ and from $C(\sigma(A))$ into $\mathcal{BL}(V)$.

For example, let t be a real number, and consider

$$\psi(x) = x - t. \tag{111.5}$$

Thus $\psi(A) = A - t\,I$, and

$$\sigma(A - t\,I) = \{\lambda - t : \lambda \in \sigma(A)\}. \tag{111.6}$$

If $\phi(x)$ is a continuous real-valued function on $\sigma(A - t\,I)$, then

$$\phi_t(x) = \phi(x - t) \tag{111.7}$$

is a continuous real-valued function on $\sigma(A)$, and

$$\phi_t(A) = \phi(A - t\,I). \tag{111.8}$$

In particular, we can apply this to the continuous piecewise-linear functions $f_j(x)$ on $\mathbf{R}$ described in the previous section. If $f_{j,t}(x) = f_j(x - t)$, then we get that $f_{j,t}(A) = f_j(A - t\,I)$. As in Section 110, $\{f_j(A - t\,I)\}_{j=1}^\infty$ converges in the strong operator topology to an orthogonal projection P_t on V. This is the same as saying that $\{f_{j,t}(A)\}_{j=1}^\infty$ converges strongly to P_t.

112 Step functions, continued

Let a, b be real numbers with $a < b$, and let n be a positive integer. Put

$$t_k = a + \frac{k}{n}(b - a) \tag{112.1}$$

for $k = 0, 1, \ldots, n$, so that

$$a = t_0 < t_1 < \cdots < t_n = b \tag{112.2}$$

and $t_k - t_{k-1} = (b-a)/n$ when $k \ge 1$. Consider

$$\phi_j(x) = a + \frac{b-a}{n} \sum_{k=1}^{n} f_j(x - t_k), \tag{112.3}$$

where $f_j(x)$ is as in Section 110.

Let us restrict our attention to sufficiently large j, in the sense that

$$2^{-j} \le \frac{b-a}{n}. \tag{112.4}$$

If $k \ge 1$, then

$$\begin{aligned} f_j(x - t_k) &= 0 \quad \text{when } x \le t_{k-1} \\ &= 1 \quad \text{when } x \ge t_k. \end{aligned} \tag{112.5}$$

Thus

$$\begin{aligned} \phi_j(x) &= a \quad \text{when } x \le a, \\ &= b \quad \text{when } x \ge b, \\ &= t_k \quad \text{when } x = t_k, \end{aligned} \tag{112.6}$$

and

$$t_{k-1} \le \phi_j(x) \le t_k \tag{112.7}$$

when $t_{k-1} \le x \le t_k$, $k \ge 1$. More precisely,

$$\phi_j(x) \le x \tag{112.8}$$

when $a \le x \le b$. Hence

$$0 \le x - \phi_j(x) \le \frac{b-a}{n} \tag{112.9}$$

on $[a, b]$, and

$$\max_{a \le x \le b} |\phi_j(x) - x| \le \frac{b-a}{n}. \tag{112.10}$$

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space V with $\sigma(A) \subseteq [a, b]$. Since $\phi_j(x)$ is a continuous real-valued function on the real line, we get a bounded self-adjoint linear operator $\phi_j(A)$ on V as in Section 109. The estimates in the previous paragraph imply that

$$\phi_j(A) \le A \tag{112.11}$$

and

$$\|\phi_j(A) - A\|_{op} \le \frac{b-a}{n}. \tag{112.12}$$

If $f_{j,k}(x) = f_j(x - t_k)$, then $f_{j,k}(A)\}_{j=1}^{\infty}$ converges in the strong operator topology to an orthogonal projection P_k on V, as in the preceding section. Hence $\{\phi_j(A)\}_{j=1}^{\infty}$ converges strongly to

$$A_n = a\, I + \frac{b-a}{n} \sum_{k=1}^{n} P_k. \tag{112.13}$$

Of course, t_k and ϕ_j depend on n as well, but we have suppresed this from the notation for simplicity. Using (112.12), we also get that

$$\|A_n - A\|_{op} \le \frac{b-a}{n}. \tag{112.14}$$

113 Algebras of operators

Let V be a real or complex vector space, and let $\mathcal{A}$ be an algebra of linear operators on V that includes the identity operator I. If $A \in \mathcal{A}$ and p is a polynomial with real or complex coefficients, as appropriate, then $p(A) \in \mathcal{A}$.

Suppose that V is a real or complex Hilbert space, and that $\mathcal{A}$ is an algebra of bounded linear operators on V that contains I and is closed with respect to the operator norm. If $A \in \mathcal{A}$ is self-adjoint and f is a continuous real-valued function on $\sigma(A)$, then $f(A) \in \mathcal{A}$. This is because $f(A)$ can be approximated in the operator norm by polynomials of A, as in Section 109.

If $\mathcal{A}$ is closed with respect to the strong operator topology, then $\mathcal{A}$ also contains the orthogonal projections associated to a self-adjoint operator $A \in \mathcal{A}$ as in Sections 110 and 111. By construction, these projections were obtained as strong limits of sequences of continuous functions of A, which are contained in $\mathcal{A}$ as in the preceding paragraph.

If V is complex and $\mathcal{A}$ is a $*$-algebra, then every $T \in \mathcal{A}$ can be expressed as $A + i\,B$, where $A, B \in \mathcal{A}$ are self-adjoint, as in Section 37. The discussion in Section 112 implies that A, B can be approximated in the operator norm by linear combinations of orthogonal projections that are also elements of $\mathcal{A}$. It follows that every element of $\mathcal{A}$ can be approximated in this way.

114 Irreducibility, 2

Let V be a real or complex vector space, and let $\mathcal{E}$ be a collection of linear transformations on V. A linear subspace W of V is said to be *invariant* under $\mathcal{E}$ if

$$T(W) \subseteq W \tag{114.1}$$

for every $T \in \mathcal{E}$. Thus V, $\{0\}$ are automatically invariant under $\mathcal{E}$, and $\mathcal{E}$ is said to be *irreducible* if these are the only invariant subspaces. Note that this includes the case of group representations discussed in Section 94, by taking $\mathcal{E}$ to be the collection of operators in the representation. As before, we normally suppose here that V is equipped with a norm, and that $\mathcal{E}$ consists of bounded linear operators on V, and we are only concerned with closed linear subspaces of V.

If $\mathcal{A}$ is the algebra of operators on V generated by $\mathcal{E}$, then $\mathcal{A}$ and $\mathcal{E}$ have the same invariant subspaces. In particular, $\mathcal{A}$ is irreducible if and only if $\mathcal{E}$ is irreducible. If V is a Hilbert space, $\mathcal{E}$ consists of bounded linear operators on V, and $\mathcal{E}^* = \mathcal{E}$, then $\mathcal{A}^* = \mathcal{A}$, which is to say that $\mathcal{A}$ is a $*$-algebra. Observe that $\mathcal{E}^* = \mathcal{E}$ when $\mathcal{E}$ consists of the operators in a unitary group representation.

Suppose that V is equipped with a norm, and that $\mathcal{E}$ consists of bounded linear transformations on V. It is easy to see that the closure of $\mathcal{E}$ with respect to the operator norm has the same closed invariant subspaces as $\mathcal{E}$. This also works for the closure of $\mathcal{E}$ with respect to the strong operator topology. Using the Hahn–Banach theorem, one can check that this holds for the weak operator topology too. Of course, one might as well take $\mathcal{E}$ to be a linear subspace of $\mathcal{BL}(V)$, for which the closure in the strong and weak operator topologies is the same, as in Section 41.

Suppose now that V is a Hilbert space, $\mathcal{E}$ consists of bounded linear operators on V, and $\mathcal{E}^* = \mathcal{E}$. If W is a closed linear subspace of V which is invariant under $\mathcal{E}$, then the orthogonal complement $W^\perp$ of W is also invariant under $\mathcal{E}$, as in Section 45. This implies that the orthogonal projection P_W of V onto W is in the commutant $\mathcal{C}(\mathcal{E})$ of $\mathcal{E}$. Conversely, if $P_W \in \mathcal{C}(\mathcal{E})$, then W and $W^\perp$ are invariant under $\mathcal{E}$. It follows that $\mathcal{E}$ is irreducible if and only if $\mathcal{C}(\mathcal{E})$ does not contain any orthogonal projections other than the trivial ones corresponding to $W = V, \{0\}$.

As in Section 44, $\mathcal{C}(\mathcal{E})$ is a $*$-subalgebra of $\mathcal{BL}(V)$ when $\mathcal{E} \subseteq \mathcal{BL}(V)$ satisfies $\mathcal{E}^* = \mathcal{E}$. Moreover, $\mathcal{C}(\mathcal{E})$ is closed with respect to the weak operator topology. If $\mathcal{E}$ is irreducible, then it follows from the discussion in the previous section that every self-adjoint operator $A \in \mathcal{C}(\mathcal{E})$ can be expressed as $A = r\, I$ for some $r \in \mathbf{R}$. Otherwise, there would be nontrivial orthogonal projections in $\mathcal{C}(\mathcal{E})$, and hence nontrivial invariant subspaces of $\mathcal{E}$. Conversely, if the only self-adjoint operators in $\mathcal{C}(\mathcal{E})$ are multiples of the identity operator, then the only orthogonal projections in $\mathcal{C}(\mathcal{E})$ are I, 0, and $\mathcal{E}$ is irreducible.

If V is complex, then every element of $\mathcal{C}(\mathcal{E})$ is a linear combination of self-adjoint operators in $\mathcal{C}(\mathcal{E})$. Under these conditions, $\mathcal{E}$ is irreducible if and only if $\mathcal{C}(\mathcal{E})$ consists of only complex multiples of the identity operator.

The analogous argument does not work in the real case. There are even counterexamples, as in the next sections.

115 Restriction of scalars

Let V be a complex vector space. We can also consider V to be a real vector space, by restricting scalar multiplication to real numbers. Let this real version of V be denoted V_r.

Let $\mathcal{A}$ be an algebra of linear transformations on V that contains the identity operator I. More precisely, $\mathcal{A}$ is supposed to be an algebra over the complex numbers, and so $\mathcal{A}$ contains $i\, I$ too. Complex-linear mappings on V can also be considered as real-linear mappings on V_r, so that $\mathcal{A}$ can be seen as an algebra of linear transformations on V_r as well. If W is a real-linear subspace of V which is invariant under $\mathcal{A}$, then W is invariant under $i\, I$ in particular, and hence is a complex-linear subspace of V. If $\mathcal{A}$ is irreducible as an algebra of complex-linear transformations on V, then it follows that $\mathcal{A}$ is also irreducible as an algebra of real-linear transformations on V_r.

If $\|v\|$ is a norm on V with respect to the complex numbers, then $\|v\|$ is also a norm on V_r with respect to the real numbers. The corresponding metric

$$d(v,w) = \|v - w\| \tag{115.1}$$

is the same in either case, and determines the same topology on V. A bounded complex-linear transformation on V is also a bounded real-linear transformation on V_r with respect to this norm, and the remarks in the previous paragraph can be applied to algebras of bounded linear mappings and closed linear subspaces. In particular, the algebra $\mathcal{BL}(V)$ of bounded complex-linear mappings on V may be considered as a subalgebra of the algebra $\mathcal{BL}(V_r)$ of bounded real-linear mappings on V_r. Of course, $\mathcal{BL}(V)$ is irreducible as an algebra of operators on V, and it follows that $\mathcal{BL}(V)$ is also irreducible as an algebra of operators on V_r.

If $\langle v, w\rangle$ is an inner product on V, then it is easy to see that

$$\langle v, w\rangle_r = \operatorname{Re}\langle v, w\rangle \tag{115.2}$$

is an inner product on V_r that determines the same norm

$$\|v\| = \langle v, v\rangle^{1/2} = \langle v, v\rangle_r^{1/2}. \tag{115.3}$$

If T is a bounded complex-linear mapping on V, then the adjoint T^* of T with respect to $\langle v, w\rangle$ is the same as the adjoint of T as a bounded real-linear mapping on V_r with respect to $\langle v, w\rangle_r$. Hence the algebra $\mathcal{BL}(V)$ of bounded complex-linear operators on V is also a $*$-algebra as an algebra of bounded real-linear operators on V_r. Note that complex-linear mappings on V are characterized among real-linear mappings by the property of commuting with $i\,I$, so that the commutant of $\mathcal{BL}(V)$ consists of complex multiples of the identity operator, even as a subalgebra of $\mathcal{BL}(V_r)$. By contrast, the commutant of $\mathcal{BL}(V_r)$ consists of real multiples of the identity.

116 Quaternions

A *quaternion* x may be expressed as

$$x = x_1 + x_2\, i + x_3\, j + x_4\, k, \tag{116.1}$$

where x_1, x_2, x_3, x_4 are real numbers,

$$i^2 = j^2 = k^2 = -1, \tag{116.2}$$

and

$$i\, j = -j\, i = k, \tag{116.3}$$

which implies that

$$i\, k = -k\, i = -j, \quad j\, k = -k\, j = i. \tag{116.4}$$

More precisely, the quaternions $\mathbf{H}$ form an associative algebra over the real numbers which contains the real numbers as a subalgebra. The real number 1 is also the multiplicative identity element for $\mathbf{H}$, and real numbers commute with arbitrary quaternions. As a vector space over the real numbers, $\mathbf{H}$ has dimension 4, and $1, i, j, k$ is a basis for $\mathbf{H}$.

The norm of a quaternion (116.1) is defined as usual by

$$|x| = (x_1^2 + x_2^2 + x_3^2 + x_4^2)^{1/2}. \tag{116.5}$$

If $x^* = x_1 - x_2\, i - x_3\, j - x_4\, k$, then it is easy to see that

$$x\, x^* = x^*\, x = |x|^2. \tag{116.6}$$

Moreover,

$$(x\, y)^* = y^*\, x^* \tag{116.7}$$

for every $x, y \in \mathbf{H}$. Therefore

$$|x\, y|^2 = x\, y\, (x\, y)^* = x\, y\, y^*\, x^* = x\, x^*\, |y|^2 = |x|^2\, |y|^2, \tag{116.8}$$

and hence $|x\, y| = |x|\, |y|$.

The standard inner product on $\mathbf{H}$ as a real vector space is defined by

$$\langle x, y \rangle = x_1\, y_1 + x_2\, y_2 + x_3\, y_3 + x_4\, y_4. \tag{116.9}$$

This is the same as the real parts of $x^*\, y$, $x\, y^*$, $y\, x^*$, and $y^*\, x$. For each $w \in \mathbf{H}$,

$$T_w(x) = w\, x \tag{116.10}$$

is a linear transformation on $\mathbf{H}$ as a real vector space whose adjoint with respect to the standard inner product is equal to T_{w^*}. We also have that

$$T_v(T_w(x)) = v\, T_w(x) = v\, w\, x = T_{v\, w}, \tag{116.11}$$

for every $v, w, x \in \mathbf{H}$.

Thus the collection of linear transformations T_w, $w \in \mathbf{H}$, is a $*$-algebra of operators on $\mathbf{H}$, since $T_v + T_w = T_{v+w}$. Note that each nonzero quaternion x is invertible in $\mathbf{H}$, with $x^{-1} = x^*/|x|^2$. This implies that the algebra of T_w's is irreducible on $\mathbf{H}$. The commutant of this algebra consists of the linear operators of the form $x \mapsto x\, z$, $z \in \mathbf{H}$.

117 Real $*$-algebras

Let V be a real Hilbert space, and let $\mathcal{B}$ be a $*$-algebra of bounded linear operators on V that contains the identity I. Suppose that every self-adjoint operator in $\mathcal{B}$ is of the form $r\, I$ for some $r \in \mathbf{R}$. In particular, this happens when $\mathcal{B}$ is the commutant of an irreducible $*$-algebra of bounded linear operators on V. Every bounded linear operator T on V can be expressed as $T_1 + T_2$, where

$T_1 = (T+T^*)/2$ is self-adjoint and $T_2 = (T-T^*)/2$ is anti-self-adjoint. If $T \in \mathcal{B}$, then $T^* \in \mathcal{B}$, and so $T_1, T_2 \in \mathcal{B}$ too. By hypothesis, $T_1 = r\,I$ for some $r \in \mathbf{R}$, and we would like to know more about T_2. If A is any bounded anti-self-adjoint linear operator on V, then A^2 is self-adjoint, and moreover $A^2 \le 0$, because

$$\langle A^2(v), v\rangle = -\langle A(v), A(v)\rangle = -\|A(v)\|^2 \le 0 \tag{117.1}$$

for every $v \in V$. If $A \in \mathcal{B}$, then $A^2 \in \mathcal{B}$, and it follows that $A^2 = c\,I$ for some $c \in \mathbf{R}$ with $c \le 0$. Note that $A^2 = 0$ implies that $A = 0$, by the previous computation. Thus $c < 0$ when $A \ne 0$.

If R, T are bounded linear operators on V, then $R\,T^* + T\,R^*$ is self-adjoint, and is therefore a multiple of the identity when $R, T \in \mathcal{B}$. Let (R, T) be the real number defined by

$$(R, T)\,I = \frac{R\,T^* + T\,R^*}{2} \tag{117.2}$$

when $R, T \in \mathcal{B}$. One can check that this is an inner product on $\mathcal{B}$, for which the anti-self-adjoint elements of $\mathcal{B}$ are orthogonal to I, and $(I, I) = 1$. More precisely, $(T, T) \ge 0$ because $T\,T^*, T^*\,T \ge 0$ as self-adjoint linear operators, and $(T, T) = 0$ implies that $T\,T^* = T^*\,T = 0$, which implies that $T = 0$.

If A, B are anti-self-adjoint elements of $\mathcal{B}$, then $(A, B) = 0$ if and only if

$$A\,B = -B\,A. \tag{117.3}$$

This implies that $C = A\,B \in \mathcal{B}$ is anti-self-adjoint as well, and that

$$A\,C = -C\,A, \quad B\,C = -C\,B. \tag{117.4}$$

Thus $(A, C) = (B, C) = 0$. Also, $C^2 = -A^2\,B^2$, so that $C \ne 0$ when $A, B \ne 0$. If $T \in \mathcal{B}$ is anti-self-adjoint and orthogonal to A and B, then

$$A\,T = -T\,A \quad \text{and} \quad B\,T = -T\,B, \tag{117.5}$$

which implies that

$$C\,T = T\,C. \tag{117.6}$$

If T is also orthogonal to C, then

$$C\,T = -T\,C, \tag{117.7}$$

and hence

$$C\,T = 0. \tag{117.8}$$

If $C \ne 0$, then C^2 is a nonzero multiple of the identity, so that C is invertible, and

$$T = 0. \tag{117.9}$$

Of course, it may be that every element of $\mathcal{B}$ is a multiple of the identity. If the linear subspace of $\mathcal{B}$ consisting of anti-self-adjoint operators has dimension 1, then $\mathcal{B}$ is isomorphic to the complex numbers as an algebra over $\mathbf{R}$. Otherwise, suppose that $\mathcal{B}$ contains two linearly independent anti-self-adjoint operators A, B. Without loss of generality, we may suppose that $(A, B) = 0$. As in the previous paragraph, $C = A\,B \in \mathcal{B}$ is nonzero, anti-self-adjoint, and orthogonal to A and B. Moreover, every anti-self-adjoint element of $\mathcal{B}$ orthogonal to A, B, and C is equal to 0. Thus the anti-self-adjoint elements of $\mathcal{B}$ are spanned by A, B, and C, which means that $\mathcal{B}$ is spanned by A, B, C, and I. One can also normalize A, B so that $A^2 = B^2 = -I$, which implies that $C^2 = -I$. It follows that $\mathcal{B}$ is isomorphic to the quaternions as an algebra over $\mathbf{R}$ in this case.

118 Another approach

Let V be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. If $\sigma(A)$ contains only one element, then $\alpha(A) = \beta(A)$, and A is a multiple of the identity operator on V. Otherwise, $\sigma(A)$ has at least two elements, and there are continuous real-valued functions f_1, f_2 that are not identically 0 on $\sigma(A)$ and satisfy

$$f_1(\lambda)\, f_2(\lambda) = 0 \tag{118.1}$$

for every $\lambda \in \sigma(A)$. This implies that $f_1(A), f_2(A) \neq 0$ and

$$f_1(A)\, f_2(A) = 0, \tag{118.2}$$

so that the kernel of $f_1(A)$ is a nontrivial proper closed linear subspace of V.

Suppose that T is a bounded linear operator on V that commutes with A. Thus T commutes with $p(A)$ for any polynomial p with real coefficients. If f is a continuous real-valued function on $\sigma(A)$, then T commutes with $f(A)$ as well, since $f(A)$ can be approximated by polynomials of A in the operator norm. In particular, T commutes with $f_1(A)$, which implies that the kernel of $f_1(A)$ is invariant under T. If $\mathcal{E}$ is a collection of bounded linear operators on V and $A \in \mathcal{C}(\mathcal{E})$, then it follows that the kernel of $f_1(A)$ is invariant under $\mathcal{E}$.

If $\mathcal{E}$ is an irreducible collection of bounded linear operators on V and A is a bounded self-adjoint linear operator on V in the commutant of $\mathcal{E}$, then we may conclude that $A = r\,I$ for some $r \in \mathbf{R}$, as before. Of course, this is especially interesting when $\mathcal{E}^* = \mathcal{E}$, so that $\mathcal{C}(\mathcal{E})$ is a $*$-algebra, and hence contains the self-adjoint parts of its elements.

If V has finite dimension, and A is a self-adjoint linear transformation on V, then A is diagonalizable. If A is not a multiple of the identity, then the kernel of $A - \lambda\,I$ is a proper nontrivial linear subspace of V for some $\lambda \in \mathbf{R}$. This does not necessarily work in infinite dimensions, since there are self-adjoint multiplication operators A such that the kernel of $A - \lambda\,I$ is always trivial.

119 Finite dimensions

Let V be a vector space over the real or complex numbers, let T be a linear transformation on V, and let W be the kernel of T. If R is another linear transformation on V that commutes with T, then it is easy to check that

$$R(W) \subseteq W. \tag{119.1}$$

Similarly, if $Z = T(V)$, then $R(Z) \subseteq Z$. If V is equipped with a norm and T is a bounded linear transformation on V, then W is automatically a closed linear subspace of V, but $T(V)$ may not be. In this case, it is better to take Z to be the closure of $T(V)$, and we still have $R(Z) \subseteq Z$ when $R \in \mathcal{BL}(V)$ commutes with T.

If $\mathcal{E}$ is a collection of linear transformations on V, and $T \in \mathcal{C}(\mathcal{E})$, then W and Z are invariant under $\mathcal{E}$. If $\mathcal{E}$ is irreducible, then it follows that $W = V$ or $W = \{0\}$ and $Z = V$ or $\{0\}$. More precisely, $W = \{0\}$ if and only if $Z = \{0\}$, and so either $W = Z = \{0\}$ or $W = Z = V$. If V has finite dimension, then we can simply say that $T = 0$ or T is invertible on V. Otherwise, irreducibility is normally defined in terms of closed linear subspaces of V, and we get that $T = 0$ or T is one-to-one and $T(V)$ is dense in V.

Suppose that V is a finite-dimensional complex vector space, and let T be a linear mapping on V. It is well known that there is a $\lambda_0 \in \mathbf{C}$ such that $T - \lambda_0 I$ has a nontrivial kernel, because the determinant of $T - \lambda I$ is a polynomial in λ and thus has a root. If $\mathcal{E}$ is an irreducible collection of linear transformations on V and $T \in \mathcal{C}(\mathcal{E})$, then $T - \lambda I \in \mathcal{C}(\mathcal{E})$ for each $\lambda \in \mathbf{C}$, and hence $T = \lambda_0 I$.

Now suppose that V is a finite-dimensional real vector space. If $\mathcal{E}$ is an irreducible collection of linear transformations on V, then every nonzero element of $\mathcal{C}(\mathcal{E})$ is invertible as a linear mapping on V. It is easy to check that the inverse of an invertible element of $\mathcal{C}(\mathcal{E})$ is also an element of $\mathcal{C}(\mathcal{E})$. Of course, the dimension of $\mathcal{C}(\mathcal{E})$ as a real vector space is less than or equal to the dimension of the algebra of all linear transformations on V, which is equal to the square of the dimension of V. Thus $\mathcal{C}(\mathcal{E})$ is a finite-dimensional division algebra over the real numbers, and is therefore isomorphic as an abstract algebra over $\mathbf{R}$ to the real numbers, the complex numbers, or the quaternions, by a well-known theorem in algebra.

120 Multiplication operators, 3

Let (X, μ) be a σ-finite measure space, and let b be a bounded measurable function on X. As in Section 25, the corresponding multiplication operator M_b is bounded on $L^p(X)$ for $1 \le p \le \infty$, with operator norm equal to the L^∞ norm of b. Let us restrict our attention to the case where $p = 2$ and b is real-valued, so that M_b is a self-adjoint operator on a Hilbert space.

It is easy to see that $\alpha(M_b)$, $\beta(M_b)$ are the essential infimum and supremum of b on X. Equivalently,

$$\alpha(M_b) \le b(x) \le \beta(M_b) \tag{120.1}$$

for almost every $x \in X$, and $\alpha(M_b)$ is the largest real number for which the first condition holds, and $\beta(M_b)$ is the smallest real number for which the second condition holds. In particular,

$$\max(-\alpha(M_b), \beta(M_b)) = \|b\|_\infty, \tag{120.2}$$

and $M_b \ge 0$ as a self-adjoint operator if and only if $b(x) \ge 0$ for almost every $x \in X$. If there is a $\delta > 0$ such that

$$|b(x)| \ge \delta \tag{120.3}$$

for almost every $x \in X$, then $1/b$ also determines and L^∞ function on X, and the corresponding multiplication operator is the inverse of M_b. Conversely, if M_b has a bounded inverse, then one can check that b satisfies this property for some $\delta > 0$. Similarly, $M_b - \lambda\, I = M_{b-\lambda}$ has a bounded inverse for some $\lambda \in \mathbf{R}$ if and only if there is a $\delta > 0$ such that

$$|b(x) - \lambda| \ge \delta \tag{120.4}$$

for almost every $x \in X$. Thus $\sigma(M_b)$ consists of the $\lambda \in \mathbf{R}$ for which there is no such $\delta > 0$, also known as the essential range of b.

Suppose that X is a locally compact Hausdorff topological space, and that μ is a positive Borel measure on X such that $\mu(K) < \infty$ when $K \subseteq X$ is compact and $\mu(U) > 0$ when $U \subseteq X$ is open and nonempty. If b is a bounded continuous function on X, then the L^∞ norm of b is equal to the supremum norm of b, and $\alpha(M_b)$, $\beta(M_b)$ are equal to the infimum and supremum of b on X. The essential range of b reduces to the closure of the set $b(X)$ of values of b in $\mathbf{R}$ in this case. If X is compact, then the maximum and minimum of b are attained on X, and M_b is invertible when $b(x) \ne 0$ for every $x \in X$. Hence $M_b - \lambda\, I = M_{b-\lambda}$ is not invertible exactly when $b(x) = \lambda$ for some $x \in X$, so that $\sigma(M_b) = b(X)$.

If f is a polynomial with real coefficients, then

$$f(M_b) = M_{f(b)}. \tag{120.5}$$

This also works for continuous real-valued functions f on the essential range of b, by suitable approximation arguments. It is easy to check directly that

$$\|f \circ b\|_\infty = \max\{|f(\lambda)| : \lambda \in \sigma(M_b)\} \tag{120.6}$$

and

$$\sigma(f(M_b)) = f(\sigma(M_b)). \tag{120.7}$$

If b is a bounded continuous function on a locally compact Hausdorff space X, as in the preceding paragraph, then $f(b)$ is also a bounded continuous function on X.

121 Measures

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space $(V, \langle v, w\rangle)$. If f is a continuous real-valued function on $\sigma(A)$ and $v \in V$, then we put

$$L_v(f) = \langle f(A)(v), v\rangle. \tag{121.1}$$

Thus L_v is a linear functional on $C(\sigma(A))$, and

$$|L_v(f)| \le \|f(A)\|_{op} \, \|v\|^2 = \max\{|f(\lambda)| : \lambda \in \sigma(A)\} \, \|v\|^2. \tag{121.2}$$

If $f \ge 0$ on $\sigma(A)$, then $f(A) \ge 0$, and hence

$$L_v(f) \ge 0. \tag{121.3}$$

The Riesz representation theorem implies that there is a unique positive Borel measure μ_v on $\sigma(A)$ such that

$$L_v(f) = \int_{\sigma(A)} f(\lambda) \, d\mu_v(\lambda) \tag{121.4}$$

for every $f \in C(\sigma(A))$. Remember that $f(A) = I$ when $f \equiv 1$, and so

$$\mu_v(\sigma(A)) = L_v(1) = \|v\|^2. \tag{121.5}$$

In particular, the dual norm of L_v as a bounded linear functional on $C(\sigma(A))$ with respect to the supremum norm is equal to $\|v\|^2$. If $f \in C(\sigma(A))$, then $f^2 \ge 0$, and in fact

$$L_v(f^2) = \langle f(A)(v), f(A)(v)\rangle = \|f(A)(v)\|^2, \tag{121.6}$$

because $f(A)$ is self-adjoint. Equivalently,

$$\|f_1(A)(v) - f_2(A)(v)\|^2 = \int_{\sigma(A)} (f_1(\lambda) - f_2(\lambda))^2 \, d\mu_v(\lambda) \tag{121.7}$$

for every $f_1, f_2 \in C(\sigma(A))$.

Suppose that μ is a positive finite Borel measure with compact support on the real line, and consider the Hilbert space $L^2(\mu)$, with the standard inner product defined by integration. By hypothesis, $b(x) = x$ is bounded on the support of μ, and so the corresponding multiplication operator M_b is bounded on $L^2(\mu)$. This operator is also self-adjoint, since b is real-valued, with $\sigma(M_b)$ equal to the support of μ. If f is a continuous real-valued function on the support of μ, then $f(M_b)$ reduces to multiplication by $f(b(x)) = f(x)$, as in the previous section. In this case, $d\mu_v = |v|^2 \, d\mu$ for each $v \in L^2(\mu)$.

If b is a bounded measurable real-valued function on a σ-finite measure space (X, μ), then the situation is a bit more complicated. The main point is to transform integrals on X to integrals on the essential range of b. Each continuous real-valued function f on the essential range of b leads to a bounded measurable function $f(b(x))$ on X. If $v \in L^2(X)$, then $|v|^2 \, d\mu$ is a finite measure on X, and $f(b(x))$ is integrable with respect to this measure. This integral can also be expressed as the integral of f times a positive Borel measure on the essential range of b, and this measure is the same as μ_v.

122 Measures, 2

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space $(V, \langle v, w \rangle)$ again, as in the previous section. If f is a continuous real-valued function on $\sigma(A)$ and $v, w \in V$, then we put

$$L_{v,w}(f) = \langle f(A)(v), w \rangle. \tag{122.1}$$

Thus $L_{v,w}$ is a real-linear mapping from $C(\sigma(A))$ into $\mathbf{R}$ or $\mathbf{C}$, depending on whether V is a real or complex Hilbert space, and

$$|L_{v,w}(f)| \le \|f(A)\|_{op} \, \|v\| \, \|w\| = \max\{|f(\lambda)| : \lambda \in \sigma(A)\} \, \|v\| \, \|w\|. \tag{122.2}$$

The Riesz representation theorem implies that there is a unique real or complex Borel measure $\mu_{v,w}$ on $\sigma(A)$, as appropriate, such that

$$L_{v,w}(f) = \int_{\sigma(A)} f(\lambda) \, d\mu_{v,w}(\lambda) \tag{122.3}$$

for every $f \in C(\sigma(A))$. Clearly $L_{v,v} = L_v$, and so $\mu_{v,v} = \mu_v$. Using polarization identities, $L_{v,w}$ can be expressed as a linear combination of L_z's for suitable $z \in V$. Similarly, $\mu_{v,w}$ can be expressed as a linear combination of μ_z's.

123 Complex-valued functions

Let $(V, \langle v, w \rangle)$ be a complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. If f is a continuous complex-valued function on $\sigma(A)$, then f can be expressed as $f = f_1 + i\, f_2$, where f_1, f_2 are continuous real-valued functions on $\sigma(A)$. In this case, we can define $f(A)$ by

$$f(A) = f_1(A) + i\, f_2(A). \tag{123.1}$$

This is a bounded linear operator on V, whose adjoint is given by

$$f(A)^* = f_1(A) - i\, f_2(A) = \overline{f}(A). \tag{123.2}$$

Thus

$$f(A)^* \, f(A) = f_1(A)^2 + f_2(A)^2 = |f|^2(A). \tag{123.3}$$

Because $|f|^2$ is a continuous real-valued function on $\sigma(A)$, we have that

$$\| |f|^2(A) \|_{op} = \max\{|f(\lambda)|^2 : \lambda \in \sigma(A)\}. \tag{123.4}$$

This implies that

$$\|f(A)\|_{op} = \max\{|f(\lambda)| : \lambda \in \sigma(A)\}, \tag{123.5}$$

since $\|f(A)\|_{op}^2 = \|f(A)^* \, f(A)\|_{op}$, by the C^* identity. As in the previous section, $L_{v,w}(f) = \langle f(A)(v), w \rangle$ defines a linear functional on the space of continuous complex-valued functions on $\sigma(A)$, and satisfies

$$|L_{v,w}(f)| \le \max\{|f(\lambda)| : \lambda \in \sigma(A)\} \, \|v\| \, \|w\|. \tag{123.6}$$

Similarly,

$$\begin{aligned}(123.7)\qquad \|f(A)(v)\|^2 &= \langle f(A)(v), f(A)(v)\rangle = \langle f(A)^*(f(A)(v)), v\rangle \\ &= L_v(|f|^2) = \int_{\sigma(A)} |f(\lambda)|^2 \, d\mu_v(\lambda),\end{aligned}$$

as in Section 121.

124 Pointwise convergence

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space $(V, \langle v, w\rangle)$, and suppose that $\{f_j\}_{j=1}^\infty$ is a uniformly bounded sequence of continuous real-valued functions on $\sigma(A)$ that converges pointwise to a real-valued function f on $\sigma(A)$. Thus f is bounded and Borel measurable, and the dominated convergence theorem implies that

$$(124.1)\qquad \lim_{j\to\infty} \langle f_j(A)(v), w\rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_{v,w}(\lambda)$$

for every $v, w \in V$. Hence $\{f_j(A)\}_{j=1}^\infty$ converges in the weak operator topology to a bounded self-adjoint linear operator on V that might be denoted $f(A)$. Equivalently, $f(A)$ may be defined by

$$(124.2)\qquad \langle f(A)(v), w\rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_{v,w}.$$

Using (121.7), one can show that $\{f_j(A)\}_{j=1}^\infty$ converges to $f(A)$ in the strong operator topology. More precisely, the dominated convergence theorem implies that

$$(124.3)\qquad \lim_{j\to\infty} \int_{\sigma(A)} (f_j(\lambda) - f(\lambda))^2 \, d\mu_v(\lambda) = 0$$

for each $v \in V$. Hence

$$(124.4)\qquad \|f_j(A)(v) - f_k(A)(v)\|^2 = \int_{\sigma(A)} (f_j(\lambda) - f_k(\lambda))^2 \, d\mu_v(\lambda) \to 0$$

as $j, k \to \infty$, so that $\{f_j(A)(v)\}_{j=1}^\infty$ is a Cauchy sequence in V which therefore converges. Of course, the limit is the same as the weak limit $f(A)(v)$.

125 Borel functions

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space $(V, \langle v, w\rangle)$, and let f be a bounded real-valued Borel measurable function on $\sigma(A)$. There is a unique bounded self-adjoint linear operator $f(A)$ on V such that

$$(125.1)\qquad \langle f(A)(v), w\rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_{v,w}(\lambda)$$

for every $v, w \in V$. Observe that $L_{v,w}$ is linear in v and either linear or conjugate-linear in w, depending on whether V is real or complex. Hence $\mu_{v,w}$ is also linear in v and linear or conjugate-linear in w, as appropriate. Similarly, $\mu_{w,v}$ is equal to $\mu_{v,w}$ in the real case and to $\overline{\mu_{v,w}}$ in the complex case, by the corresponding properties of $L_{v,w}$. Moreover,

$$\left| \int_{\sigma(A)} f(\lambda)\, d\mu_{v,w}(\lambda) \right| \le \sup\{|f(\lambda)| : \lambda \in \sigma(A)\} \, \|v\| \, \|w\| \tag{125.2}$$

for every $v, w \in V$, so that

$$\|f(A)\|_{op} \le \sup\{|f(\lambda)| : \lambda \in \sigma(A)\}. \tag{125.3}$$

If $f(\lambda) \ge 0$ for every $\lambda \in \sigma(A)$, then $f(A) \ge 0$.

126 Approximations

Let A be a bounded self-adjoint linear operator on a real or complex Hilbert space $(V, \langle v, w \rangle)$, and let f be a bounded real-valued Borel measurable function on $\sigma(A)$. Also let $v_1, w_1, \ldots, v_n, w_n \in V$ be given, so that

$$\mu = \sum_{l=1}^{n} |\mu_{v_l, w_l}| \tag{126.1}$$

is a positive Borel measure on $\sigma(A)$. For each $\epsilon > 0$, there is a continuous real-valued function g on $\sigma(A)$ such that

$$\mu(\{\lambda \in \sigma(A) : f(\lambda) \ne g(\lambda)\}) < \epsilon \tag{126.2}$$

and

$$\max\{|g(\lambda)| : \lambda \in \sigma(A)\} \le \sup\{|f(\lambda)| : \lambda \in \sigma(A)\}. \tag{126.3}$$

This is a well known theorem in measure theory, which implies that

$$\begin{aligned} \sum_{l=1}^{n} \left| \int_{\sigma(A)} f(\lambda)\, d\mu_{v_l,w_l}(\lambda) - \int_{\sigma(A)} g(\lambda)\, d\mu_{v_l,w_l}(\lambda) \right| & \\ \le \ \sum_{l=1}^{n} \int_{\sigma(A)} |f(\lambda) - g(\lambda)| \, d|\mu_{v_l,w_l}|(\lambda) & \\ \le \ 2\epsilon \sup\{|f(\lambda)| : \lambda \in \sigma(A)\}. & \end{aligned} \tag{126.4}$$

In particular, $f(A)$ can be approximated by continuous functions of A in the weak operator topology, since

$$\sum_{l=1}^{n} |\langle f(A)(v_l), w_l \rangle - \langle g(A)(v_l), w_l \rangle| \le 2\epsilon \sup\{|f(\lambda)| : \lambda \in \sigma(A)\}. \tag{126.5}$$

This can also be used to estimate L^2 norms of $f-g$. More precisely,

$$(126.6)\quad \sum_{l=1}^{n} \int_{\sigma(A)} (f(\lambda)-g(\lambda))^2\, d|\mu_{v_l,w_l}|(\lambda) \le 4\epsilon \sup\{|f(\lambda)|^2 : \lambda \in \sigma(A)\}.$$

This is especially helpful when $v_l = w_l$, in which case $\mu_{v_l,w_l} = \mu_{v_l}$ is already positive.

As an application of this, let us check that

$$(126.7)\qquad \Big|\int_{\sigma(A)} f(\lambda)\, d\mu_{v,w}(\lambda)\Big| \le \Big(\int_{\sigma(A)} f(\lambda)^2\, d\mu_v\Big)^{1/2} \|w\|$$

for every $v, w \in V$. If g is a continuous real-valued function on $\sigma(A)$, then

$$(126.8)\qquad \begin{aligned} \Big|\int_{\sigma(A)} g(\lambda)\, d\mu_{v,w}(\lambda)\Big| &= |\langle g(A)(v), w\rangle| \le \|g(A)(v)\|\, \|w\| \\ &= \Big(\int_{\sigma(A)} g(\lambda)^2\, d\mu_v\Big)^{1/2} \|w\|. \end{aligned}$$

To get (126.7) from this, one can approximate f by continuous functions g as before, with $n = 2$, $v_1 = w_1 = v_2 = v$, and $w_2 = w$. Hence

$$(126.9)\qquad |\langle f(A)(v), w\rangle| \le \Big(\int_{\sigma(A)} f(\lambda)^2\, d\mu_v\Big)^{1/2} \|w\|,$$

and therefore

$$(126.10)\qquad \|f(A)(v)\| \le \Big(\int_{\sigma(A)} f(\lambda)^2\, d\mu_v\Big)^{1/2}.$$

Let us apply this to $f-g$, where $g \in C(\sigma(A))$ approximates f as before, with $v_l = w_l$ for $l = 1, \ldots, n$. Using (126.6), we get that

$$(126.11)\quad \begin{aligned} \sum_{l=1}^{n} \|f(A)(v_l) - g(A)(v_l)\|^2 &\le \sum_{l=1}^{n} \int_{\sigma(A)} (f(\lambda)-g(\lambda))^2\, d\mu_{v_l}(\lambda) \\ &\le 4\epsilon \sup\{|f(\lambda)|^2 : \lambda \in \sigma(A)\}. \end{aligned}$$

This shows that $f(A)$ can be approximated in the strong operator topology by continuous functions of A. Using this approximation with $n = 1$ and $v_1 = v$, we get that

$$(126.12)\qquad \|f(A)(v)\| = \Big(\int_{\sigma(A)} f(\lambda)^2\, d\mu_v(\lambda)\Big)^{1/2}$$

for each $v \in V$, by reducing to the case of continuous functions, for which the identity was already established. Similarly, if f_1 and f_2 are bounded real-valued Borel measurable functions on $\sigma(A)$, then one can check that

$$(126.13)\qquad (f_1\, f_2)(A) = f_1(A)\, f_2(A),$$

by reducing to the case of continuous functions.

127 Densities

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. Also let f be a continuous real or complex-valued function on $\sigma(A)$, depending on whether V is real or complex. Thus

$$|\langle f(A)(v), w \rangle| \le \|f(A)(v)\| \, \|w\|, \tag{127.1}$$

by the Cauchy–Schwarz inequality, and so

$$|L_{v,w}(f)| \le \Big(\int_{\sigma(A)} |f(\lambda)|^2 \, d\mu_v \Big)^{1/2} \|w\|. \tag{127.2}$$

It follows that $L_{v,w}$ determines a bounded linear functional on $L^2(\mu_v)$. The Riesz representation theorem implies that there is a unique $h_w \in L^2(\mu_v)$ such that

$$L_{v,w}(f) = \int_{\sigma(A)} f(\lambda) \, h_w(\lambda) \, d\mu_v(\lambda) \tag{127.3}$$

in the real case, and

$$L_{v,w}(f) = \int_{\sigma(A)} f(\lambda) \, \overline{h_w(\lambda)} \, d\mu_v(\lambda) \tag{127.4}$$

in the complex case. Hence $d\mu_{v,w} = h_w \, d\mu_v$ in the real case, and $d\mu_{v,w} = \overline{h_w} \, d\mu_v$ in the complex case. Of course, h_w also depends on v, even if that has been suppressed from the notation.

128 Eigenvalues

Let V be a real or complex vector space, and let A be a linear transformation on V. A real or complex number λ, as appropriate, is an *eigenvalue* of A if there is a $v \in V$ such that $v \ne 0$ and

$$A(v) = \lambda \, v, \tag{128.1}$$

in which case v is an *eigenvector* of A. Thus λ is an eigenvalue of A if and only if $A - \lambda \, I$ has nontrivial kernel. In particular, $A - \lambda \, I$ is not invertible when λ is an eigenvalue of A, and the converse holds when V has finite dimension. If $v \in V$ is an eigenvector of A with eigenvalue λ and p is a polynomial with real or complex coefficients, as appropriate, then

$$p(A)(v) = p(\lambda) \, v, \tag{128.2}$$

so that v is an eigenvector of $p(A)$ with eigenvalue $p(\lambda)$.

Suppose now that V is a Hilbert space, and that A is a bounded self-adjoint linear operator on V. Suppose also that λ is an eigenvalue of A, and that $v \in V$

is an eigenvector of A associated to λ. Note that $\lambda \in \mathbf{R}$ even when V is complex, because

$$\lambda \langle v, v\rangle = \langle A(v), v\rangle = \langle v, A(v)\rangle = \overline{\lambda}\, \langle v, v\rangle \tag{128.3}$$

implies that $\overline{\lambda} = \lambda$ when $v \neq 0$. Of course, $\lambda \in \sigma(A)$, since $A - \lambda I$ is not invertible, as in the previous paragraph.

Let f be a continuous real-valued function on $\sigma(A)$. It is easy to see that

$$f(A)(v) = f(\lambda)\, v, \tag{128.4}$$

by approximating f by polynomials and applying the earlier statement in that case. Hence v is an eigenvector of $f(A)$ with eigenvalue $f(\lambda)$. In particular, $f(\lambda) \in \sigma(f(A))$, as in Section 109.

It follows that

$$L_v(f) = f(\lambda)\, \|v\|^2, \tag{128.5}$$

which means that $\mu_v = \|v\|^2\, \delta_\lambda$, where δ_λ is the Dirac measure at λ. Similarly,

$$L_{v,w}(f) = f(\lambda)\, \langle v, w\rangle \tag{128.6}$$

for each $w \in V$, and so $\mu_{v,w} = \langle v, w\rangle\, \delta_\lambda$. If f is a bounded Borel measurable function on $\sigma(A)$, then we get that

$$\langle f(A)(v), w\rangle = \int_{\sigma(A)} f\, d\mu_{v,w} = f(\lambda)\, \langle v, w\rangle \tag{128.7}$$

for every $w \in V$. Thus (128.4) holds in this case as well.

129 Another limit

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let A be a bounded self-adjoint linear operator on V. Suppose that

$$0 \leq A \leq I, \tag{129.1}$$

which implies that $\|A\|_{op} \leq 1$, as in Section 34. Of course, A^n is also a bounded self-adjoint linear operator on V for each positive integer n, with

$$\|A^n\|_{op} \leq \|A\|_{op}^n \leq 1. \tag{129.2}$$

If $\|A\|_{op} < 1$, then $\{A^n\}_{n=1}^\infty$ converges to 0 in the operator norm.

Let us check that

$$0 \leq A^{n+1} \leq A^n \leq I \tag{129.3}$$

for each n. If $n = 2k$ is an even integer, then $A^n \geq 0$, because

$$\langle A^{2k}(v), v\rangle = \langle A^k(v), A^k(v)\rangle = \|A^k(v)\|^2 \geq 0. \tag{129.4}$$

If $n = 2l + 1$ is odd, then

$$\langle A^{2l+1}(v), v\rangle = \langle A(A^l(v)), A^l(v)\rangle \geq 0, \tag{129.5}$$

because $A \geq 0$. Moreover,

$$\langle A^{2l+1}(v), v\rangle = \langle A^{l+1}(v), A^l(v)\rangle \leq \|A^{l+1}(v)\|\,\|A^l(v)\|, \tag{129.6}$$

by the Cauchy–Schwarz inequality, and so

$$\langle A^{2l+1}(v), v\rangle \leq \|A\|_{op}\,\|A^l(v)\|^2 \leq \|A^l(v)\|^2 = \langle A^{2l}(v), v\rangle, \tag{129.7}$$

which shows that $A^{2l+1} \leq A^{2\,l}$. Similarly,

$$\langle A^{2l+2}(v), v\rangle = \|A(A^l(v))\|^2 \leq \|A\|_{op}\,\langle A(A^l(v)), A^l(v)\rangle, \tag{129.8}$$

as in Section 34, and thus

$$\langle A^{2l+2}(v), v\rangle \leq \langle A^{l+1}(v), A^l(v)\rangle = \langle A^{2l+1}(v), v\rangle, \tag{129.9}$$

which is the same as $A^{2l+2} \leq A^{2l+1}$. This shows that $A^{n+1} \leq A^n$ for each n, and $A^n \leq I$ because $\|A^n\|_{op} \leq 1$. Alternatively, (129.1) implies that $\sigma(A) \subseteq [0,1]$, and these properties of A^n can be obtained from the fact that

$$0 \leq \lambda^{n+1} \leq \lambda^n \leq 1 \tag{129.10}$$

when $0 \leq \lambda \leq 1$.

As in Section 34, it follows that $\{A^n\}_{n=1}^\infty$ converges in the strong operator topology to a bounded self-adjoint linear operator P on V such that

$$0 \leq P \leq I. \tag{129.11}$$

We also have that

$$P^2 = P, \tag{129.12}$$

because $\{A^{2n}\}_{n=1}^\infty$ is a subsequence of $\{A^n\}_{n=1}^\infty$ and hence converges in the strong operator topology to the same limit. Thus P is the orthogonal projection of V onto a closed linear subspace W. Similarly, $\{A^{n+1}\}_{n=1}^\infty$ converges to P in the strong operator topology, and so

$$A\,P = P. \tag{129.13}$$

This implies that W is contained in the kernel of $A-I$. If $v \in V$ is an eigenvector of A with eigenvalue 1, then

$$A^n(v) = v \tag{129.14}$$

for each n, and hence

$$P(v) = v. \tag{129.15}$$

This shows that W is equal to the kernel of $A - I$.

Note that $f_n(\lambda) = \lambda^n$ converges pointwise on $[0,1]$ to the function defined by $f(\lambda) = 0$ when $0 \leq \lambda < 1$ and $f(1) = 1$. As in Section 124, the dominated convergence theorem implies that $\{A^n\}_{n=1}^\infty$ converges in the strong operator topology to $f(A)$, where the latter is the same as in Section 125. Thus $f(A) = P$.

130 Kernels and projections

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let $\mathcal{A}$ be a $*$-algebra of bounded linear operators on V that contains the identity operator I. If $T \in \mathcal{A}$, then $T^* \in \mathcal{A}$, and hence $T^* T \in \mathcal{A}$. The kernel of T is obviously contained in the kernel of $T^* T$, and in fact the two are the same, because

$$\langle (T^* T)(v), v \rangle = \langle T(v), T(v) \rangle = \|T(v)\|^2. \tag{130.1}$$

Suppose also that $\|T\|_{op} \le 1$, which can always be arranged by multiplying T by a positive real number, without affecting the kernel of T.

Consider $A = I - T^* T \in \mathcal{A}$. This is a bounded self-adjoint linear operator on V which satisfies $0 \le A \le I$. As in the previous section, $\{A^n\}_{n=1}^\infty$ converges in the strong operator topology to the orthogonal projection P of V onto the kernel of $A - I$, which is the same as the kernel of T. If $\mathcal{A}$ is closed with respect to the strong operator topology, then $P \in \mathcal{A}$.

Similarly, if $T_1, \ldots, T_n \in \mathcal{A}$, then the intersection of their kernels is the same as the kernel of $T_1^* T_1 + \cdots + T_n^* T_n$. If P is the orthogonal projection of V onto the intersection of the kernels of $T_1, \ldots, T_n$ and $\mathcal{A}$ is closed with respect to the strong operator topology, then $P \in \mathcal{A}$. Conversely, if $P \in \mathcal{A}$ is a projection, then $P(V)$ is equal to the kernel of $I - P \in \mathcal{A}$.

Let T be any element of $\mathcal{A}$, and let W be the closure of $T(V)$ in V. This is a closed linear subspace of V, whose orthogonal complement is the same as the kernel of T^*. If Q is the orthogonal projection of V onto $W^\perp$ and $\mathcal{A}$ is closed with respect to the strong operator topology, then $Q \in \mathcal{A}$. In this case, $I - Q$ is the orthogonal projection of V onto W, and $I - Q \in \mathcal{A}$.

131 Projections and subspaces

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let P be an orthogonal projection on V. Thus

$$\langle P(v), v \rangle = \langle P(v), P(v) \rangle = \|P(v)\|^2 \tag{131.1}$$

for each $v \in V$. If Q is another orthogonal projection on V, then it follows that $P \le Q$ as self-adjoint linear operators on V if and only if

$$\|P(v)\| \le \|Q(v)\| \tag{131.2}$$

for every $v \in V$.

Now let W, Z be closed linear subspaces of V, and let P_W, P_Z be the orthogonal projections of V onto W, Z, respectively. If $W \subseteq Z$, then it is easy to see that $P_W \le P_Z$ as self-adjoint linear operators on V. More precisely, if $Y = W^\perp \cap Z$, and P_Y is the orthogonal projection of V onto Y, then

$$P_W + P_Y = P_Z. \tag{131.3}$$

In particular,

$$\|P_Z - P_W\|_{op} = \|P_Y\|_{op} \tag{131.4}$$

is either 0 or 1. As another approach, one can check that $P_W = P_W\, P_Z$ when $W \subseteq Z$, because $Z^\perp \subseteq W^\perp$, and hence

$$\|P_W(v)\| = \|P_W(P_Z(v))\| \le \|P_Z(v)\| \tag{131.5}$$

for each $v \in V$.

Conversely, if $P_W \le P_Z$, then $W \subseteq Z$. To see this, it suffices to check that $Z^\perp \subseteq W^\perp$. Of course, $W^\perp$, $Z^\perp$ are the kernels of P_W, P_Z, respectively. As before, the hypothesis that $P_W \le P_Z$ implies that $\|P_W(v)\| \le \|P_Z(v)\|$ for each $v \in V$, and hence the kernel of P_Z is contained in the kernel of P_W, as desired. Alternatively, if $w \in W$, then $P_W(w) = w$, and so $\|P_W(w)\| = \|w\|$. This implies that $\|P_Z(w)\| = \|w\|$, since $\|P_W(w)\| \le \|P_Z(w)\|$ by hypothesis and $\|P_Z(w)\| \le \|w\|$ for every $w \in V$. Hence $w \in Z$, because $\|P_Z(w)\| < \|w\|$ otherwise.

Suppose that $W \subseteq Z$, and

$$\langle P_W(v), v\rangle = \langle P_Z(v), v\rangle \tag{131.6}$$

for some $v \in V$, which is the same as

$$\|P_W(v)\| = \|P_Z(v)\|. \tag{131.7}$$

If $Y = Z \cap W^\perp$, as before, then $P_W(v)$ and $P_Y(v)$ are orthogonal to each other, and hence

$$\|P_Z(v)\|^2 = \|P_W(v)\|^2 + \|P_Y(v)\|^2. \tag{131.8}$$

It follows that $P_Y(v) = 0$, and so $P_W(v) = P_Z(v)$.

132 Sequences of subspaces

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let $W_1, W_2, W_3, \ldots$ be a sequence of closed linear subspaces of V such that

$$W_j \subseteq W_{j+1} \tag{132.1}$$

for each j. Thus $\bigcup_{j=1}^\infty W_j$ is a linear subspace of V, and

$$W = \overline{\bigcup_{j=1}^\infty W_j}, \tag{132.2}$$

is a closed linear subspace of V. Also let P_j be the orthogonal projection of V onto W_j, and let P be the orthogonal projection of V onto W. As in the previous section,

$$P_j \le P_{j+1} \le P \tag{132.3}$$

for each j.

Let us check that

$$\lim_{j \to \infty} P_j(v) = P(v) \tag{132.4}$$

for each $v \in V$, so that $\{P_j\}_{j=1}^\infty$ converges to P in the strong operator topology. If $v \in W_l$ for some l, then $P_j(v) = P(v) = v$ when $j \geq l$. This implies that (132.4) holds for every $v \in \bigcup_{l=1}^\infty W_l$, and hence for every $v \in W$, since the P_j's have uniformly bounded operator norm. Similarly, if $v \in W^\perp$, then $P_j(v) = P(v) = 0$ for each j. Thus (132.4) holds for every $v \in V$, because of linearity.

Suppose now that $Y_1, Y_2, \ldots$ is a sequence of closed linear subspaces of V such that

$$Y_{j+1} \subseteq Y_j \tag{132.5}$$

for each j, and consider

$$Y = \bigcap_{j=1}^\infty Y_j. \tag{132.6}$$

If Q_j, Q are the orthogonal projections of V onto Y_j, Y, respectively, then

$$Q \leq Q_{j+1} \leq Q_j \tag{132.7}$$

for each j. Again we have that $\{Q_j\}_{j=1}^\infty$ converges to Q in the strong operator topology. This can be derived from the previous case applied to $W_j = Y_j^\perp$, so that $W = Y^\perp$, $P_j = I - Q_j$, and $P = I - Q$. Alternatively, one can argue as in Section 34 that $\{Q_j\}_{j=1}^\infty$ converges in the strong operator topology to a bounded self-adjoint linear operator R on V. Because $Q_j^2 = Q_j$ for each j, $R^2 = R$, and so R is the orthogonal projection of V onto a closed linear subspace Z of V. We also have that

$$Q \leq R \leq Q_j \tag{132.8}$$

for each j, by (132.7), which implies that $Y \subseteq Z \subseteq Y_j$ for each j, as in the previous section. Hence $Z \subseteq \bigcap_{j=1}^\infty Y_j = Y$, and therefore $Y = Z$, as desired.

By contrast, if $\{P_j\}_{j=1}^\infty$ converges to P in the operator norm, then $P_j = P$ for all but finitely many j, because $\|P_j - P\|_{op} = 0$ or 1 for each j, as in the previous section. Similarly, if $\{Q_j\}_{j=1}^\infty$ converges to Q in the operator norm, then $Q_j = Q$ for all but finitely many j. Of course, it is important that these are monotone sequences of projections for this argument.

133 Families of subspaces

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let P_W be the orthogonal projection of V onto a closed linear subspace W of V, as usual. Suppose that $\mathcal{E}$ is a nonempty collection of closed linear subspaces of V with the property that if $W_1, \ldots, W_n$ are finitely many elements of $\mathcal{E}$, then there is a $W \in \mathcal{E}$ such that

$$\bigcup_{l=1}^n W_l \subseteq W. \tag{133.1}$$

For instance, it may simply be that the closure of the linear span of $W_1, \ldots, W_n$ is an element of $\mathcal{E}$. This condition implies that $\bigcup_{W \in \mathcal{E}} W$ is a linear subspace of V, whose closure will be denoted by Z.

Let us check that P_Z can be approximated in the strong operator topology by P_W, $W \in \mathcal{E}$. Thus if $v_1, \ldots, v_n$ are finitely many elements of V, then we would like to show that there is a $W \in \mathcal{E}$ such that $P_Z(v_l)$ is approximated by $P_W(v_l)$, $1 \le l \le n$. We may as well suppose that $v_l \in Z$ for each l, since $P_W(v) = P_Z(v) = 0$ for every $W \in \mathcal{E}$ when $v \in Z^\perp$. Hence $P_Z(v_l) = v_l$ for each l, and we would like to choose $W \in \mathcal{E}$ such that v_l is approximated by $P_W(v_l)$ for each l. Equivalently, we would like to choose $W \in \mathcal{E}$ such that each v_l is approximated by an element of W. This is easy to do, by the definition of Z, and the hypothesis on $\mathcal{E}$.

Suppose now that $\mathcal{E}_1$ is a nonempty collection of closed linear subspaces of V such that if $Y_1, \ldots, Y_n$ are finitely many elements of $\mathcal{E}_1$, then

$$\bigcap_{l=1}^{n} Y_l \in \mathcal{E}_1. \tag{133.2}$$

It would also be sufficient to ask that $\bigcap_{l=1}^n Y_l$ contain a subspace of V that is an element of $\mathcal{E}_1$. If

$$X = \bigcap_{Y \in \mathcal{E}_1} Y, \tag{133.3}$$

then P_X can be approximated in the strong operator topology by P_Y, $Y \in \mathcal{E}_1$. This follows from the previous case applied to $\mathcal{E} = \{Y^\perp : Y \in \mathcal{E}_1\}$, because $P_{Y^\perp} = I - P_Y$ for each closed linear subspace Y of V.

Let $\mathcal{A}$ be a $*$-algebra of bounded linear operators on V that includes the identity operator, and let $\mathcal{E}_1$ be a nonempty collection of kernels of elements of $\mathcal{A}$. As in Section 130, the intersection of the kernels of finitely many elements of $\mathcal{A}$ is also the kernel of an element of $\mathcal{A}$. Thus we may as well ask that the intersection of finitely many elements of $\mathcal{E}_1$ also be an element of $\mathcal{E}_1$. If $\mathcal{A}$ is closed with respect to the strong operator topology, then $P_Y \in \mathcal{A}$ for each $Y \in \mathcal{E}_1$, as in Section 130 again. It follows from the remarks in the previous paragraph that the orthogonal projection of V onto the intersection (133.3) is also contained in $\mathcal{A}$ under these conditions.

134 The positive square root

Let $(V, \langle v, w \rangle)$ be a real or complex Hilbert space, and let A be a bounded nonnegative self-adjoint linear operator on V. Thus $\sigma(A) \subseteq [0, \infty)$, and

$$f(x) = \sqrt{x} \tag{134.1}$$

is a continuous real-valued function on $\sigma(A)$. If $B = f(A)$ is as in Section 109, then B is a bounded nonnegative self-adjoint linear operator on V, and

$$B^2 = A. \tag{134.2}$$

Let us check that B is uniquely determined by these properties.

Suppose that C is a bounded linear operator on V that commutes with A, so that

$$A\,C = C\,A. \tag{134.3}$$

It is easy to see that C also commutes with B,

$$B\,C = C\,B, \tag{134.4}$$

since B can be approximated by polynomials of A. In particular, C commutes with A when $C^2 = A$, and hence with B too.

Suppose now that C is nonnegative, self-adjoint, and satisfies $C^2 = A$. For each $v \in V$,

$$\langle A(v), v\rangle = \langle B^2(v), v\rangle = \langle B(v), B(v)\rangle = \|B(v)\|^2, \tag{134.5}$$

and similarly

$$\langle A(v), v\rangle = \|C(v)\|^2. \tag{134.6}$$

Hence $A(v) = 0$ implies that $B(v) = C(v) = 0$. Conversely, $A(v) = 0$ when $B(v) = 0$ or $C(v) = 0$, because $A = B^2 = C^2$.

This shows that A, B, and C have the same kernel, which will be denoted W. Observe that

$$A(W^\perp),\ B(W^\perp),\ C(W^\perp) \subseteq W^\perp, \tag{134.7}$$

by self-adjointness. If $v \in W^\perp$ and $v \neq 0$, then

$$A(v),\ B(v),\ C(v) \neq 0, \tag{134.8}$$

since $v \notin W$. Because of positivity,

$$\langle B(v), v\rangle,\ \langle C(v), v\rangle > 0, \tag{134.9}$$

as in Section 34, and hence

$$\langle B(v), v\rangle + \langle C(v), v\rangle > 0. \tag{134.10}$$

It follows that $B(v) + C(v) \neq 0$, which means that the kernel of $B + C$ is equal to W as well.

Using the fact that B and C commute, we get that

$$(B + C)(B - C) = B^2 - C^2 = 0. \tag{134.11}$$

This implies that $B - C = 0$ on $W^\perp$ and hence on V, because $B - C$ maps $W^\perp$ into itself and $B + C$ is injective on $W^\perp$, as in the previous paragraph. Thus $B = C$, as desired.

As an application, suppose that A_1 and A_2 are bounded self-adjoint linear operators on V that commute with each other. Thus their product $A_1 A_2$ is also self-adjoint. If $A_1 \geq 0$, then there is a a bounded nonnegative self-adjoint linear operator B_1 on V such that $B_1^2 = A_1$, as before. Moreover, B_1 commutes with A_2, and so

$$\langle (A_1 A_2)(v), v\rangle = \langle (B_1\, A_2\, B_1)(v), v\rangle = \langle A_2(B_1(v)), B_1(v)\rangle \tag{134.12}$$

for every $v \in V$. If $A_2 \geq 0$, then it follows that $A_1 A_2 \geq 0$, so that the product of commuting bounded nonnegative self-adjoint linear operators is also nonnegative.

Alternatively, if $A_2 \geq 0$, then there is a bounded nonnegative self-adjoint linear operator B_2 on V such that $B_2^2 = A_2$. Because B_1 commutes with A_2, B_1 commutes with B_2 too. Hence $B_1 B_2$ is self-adjoint and $(B_1 B_2)^2 = A_1 A_2$, which implies that $A_1 A_2 \geq 0$.

135 Positive self-adjoint part

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let T be a bounded linear operator on V. As in Section 37, T can be expressed as $A + B$, where

$$A = \frac{T + T^*}{2} \tag{135.1}$$

is self-adjoint, and

$$B = \frac{T - T^*}{2} \tag{135.2}$$

is anti-self-adjoint. In the complex case, $B = i\, C$, where C is self-adjoint. It follows that

$$\langle T(v), v\rangle = \langle A(v), v\rangle \tag{135.3}$$

for every $v \in V$ in the real case, and

$$\mathrm{Re}\langle T(v), v\rangle = \langle A(v), v\rangle \tag{135.4}$$

in the complex case. If A is uniformly strictly positive, as in Section 47, then there is a $\delta > 0$ such that

$$\delta\, \|v\|^2 \leq |\langle T(v), v\rangle| \leq \|T(v)\|\, \|v\|, \tag{135.5}$$

and hence $\|T(v)\| \geq \delta\, \|v\|$ for every $v \in V$. Of course, $T^* = A - B$ has the same property, which implies that the kernel of T^* is trivial, and that $T(V)$ is dense in V. Thus T is invertible on V, as in Section 46.

Suppose that T commutes with T^*, which may be described by saying that T is *normal*, at least in the complex case. Using

$$\|T(v)\|^2 = \langle T(v), T(v)\rangle = \langle (T^*\, T)(v), v\rangle \tag{135.6}$$

and

$$\|T^*(v)\|^2 = \langle (T\, T^*)(v), v\rangle, \tag{135.7}$$

we get that

$$\|T(v)\| = \|T^*(v)\| \tag{135.8}$$

for every $v \in V$. Conversely, (135.8) implies that

$$\langle (T^*\,T)(v), v\rangle = \langle (T\,T^*)(v), v\rangle \tag{135.9}$$

for every $v \in V$, and hence that $T^*\,T = T\,T^*$, by polarization. Note that $T^*\,T$, $T\,T^*$ are self-adjoint.

Equivalently, T commutes with T^* when A, B commute. This implies that

$$T^*\,T = (A - B)(A + B) = A^2 - B^2, \tag{135.10}$$

and therefore

$$\begin{aligned} \|T(v)\|^2 &= \langle (A^2 - B^2)(v), v\rangle \\ &= \langle A(v), A(v)\rangle + \langle B(v), B(v)\rangle \\ &= \|A(v)\|^2 + \|B(v)\|^2, \end{aligned} \tag{135.11}$$

since B is anti-self-adjoint. If A or B is invertible, then

$$\|T(v)\| \ge \|T_1(v)\| \ge \delta\,\|v\| \tag{135.12}$$

for some $\delta > 0$ and each $v \in V$, and T^* satisfies the same condition by (135.8). It follows that T is invertible, as in Section 46.

By contrast, if T does not commute with T^*, then T may not be invertible even though A and B are invertible. There are counterexamples already for $V = \mathbf{R}^2$ or $\mathbf{C}^2$ with the standard inner product. One can take A to correspond to a 2×2 matrix $\left(\begin{smallmatrix} a_1 & 0 \\ 0 & a_2 \end{smallmatrix}\right)$ with $a_1, a_2 \in \mathbf{R}$ and $a_1, a_2 \ne 0$, and B to correspond to a matrix $\left(\begin{smallmatrix} 0 & b \\ -b & 0 \end{smallmatrix}\right)$ with $b \in \mathbf{R}$ and $b \ne 0$. Thus T corresponds to $\left(\begin{smallmatrix} a_1 & b \\ -b & a_2 \end{smallmatrix}\right)$, whose determinant is equal to

$$a_1\,a_2 + b^2. \tag{135.13}$$

This is automatically positive when a_1, a_2 have the same sign, but may vanish when a_1, a_2 have opposite sign.

136 Polar decompositions

Let $(V, \langle v, w\rangle)$ be a real or complex Hilbert space, and let T be a bounded linear operator on V. Thus $A = T^*\,T$ is self-adjoint and nonnegative, and has a nonnegative self-adjoint square root B. Observe that

$$\|B(v)\|^2 = \langle B(v), B(v)\rangle = \langle B^2(v), v\rangle = \langle (T^*\,T)(v), v\rangle, \tag{136.1}$$

and so

$$\|B(v)\| = \|T(v)\| \tag{136.2}$$

for each $v \in V$. In particular, the kernels of B and T are the same.

It follows that

$$R(B(v)) = T(v) \tag{136.3}$$

is a well-defined linear mapping from $B(V)$ into V, and satisfies

$$\|R(w)\| = \|w\| \tag{136.4}$$

for each $w \in B(V)$. As in Section 9, R has a unique extension to an isometric linear mapping of the closure of $B(V)$ into V. Put $R(w) = 0$ when $w \in B(V)^\perp$, so that R is a bounded linear operator on V. Note that $B(V)^\perp$ is equal to the kernel of B, since B is self-adjoint, which is the same as the kernel of T, as in the previous paragraph. Hence (136.3) holds for every $v \in V$.

If $t > 0$, then $A_t = T^* T + t I$ is self-adjoint, nonnegative, and invertible, and its nonnegative self-adjoint square root B_t is invertible as well. Consider

$$R_t = T B_t^{-1}. \tag{136.5}$$

Observe that

$$R_t^* R_t = B_t^{-1} T^* T B_t^{-1} = T^* T A_t^{-1}, \tag{136.6}$$

because B_t and hence B_t^{-1} commutes with $T^* T$. Thus

$$I - R_t^* R_t = t A_t^{-1} \ge 0, \tag{136.7}$$

which implies that $R_t^* R_t \le I$, and therefore $\|R_t\|_{op} \le 1$.

Let us check that R_t converges to R as $t \to 0$ in the strong operator topology, which is to say that

$$\lim_{t \to 0} R_t(w) = R(w) \tag{136.8}$$

for every $w \in V$. If w is in the kernel of T, then w is in the kernel of $T^* T$ too, $A_t(w) = t w$, and $B_t(w) = t^{1/2} w$, as in Section 128. This implies that $B_t^{-1}(w) = t^{-1/2} w$, and hence $R_t(w) = 0$, while $R(w) = 0$, because the kernel of T is the same as $B(V)^\perp$. Thus (136.8) holds trivially when $w \in B(V)^\perp$, and it remains to verify (136.8) when $w \in \overline{B(V)}$.

Because $\|R_t\|_{op} \le 1$ for each t, it suffices to show that (136.8) holds when $w \in B(V)$. If $w = B(v)$ for some $v \in V$, then (136.8) reduces to

$$\lim_{t \to 0} R_t(B(v)) = R(B(v)) = T(v), \tag{136.9}$$

or equivalently

$$\lim_{t \to 0} T(B_t^{-1}(B(v)) - v) = 0. \tag{136.10}$$

This is the same as

$$\lim_{t \to 0} B(B_t^{-1}(B(v)) - v) = 0, \tag{136.11}$$

because of (136.2).

Of course, B commutes with B_t^{-1}, since they are both functions of $T^* T$, and so

$$\begin{aligned} B B_t^{-1} B &= B_t^{-1} B^2 = B_t^{-1} T^* T \\ &= B_t^{-1}(T^* T + t I) - t B_t^{-1} = B_t - t B_t^{-1}. \end{aligned} \tag{136.12}$$

It is easy to see that $B_t \to B$ as $t \to 0$ in the operator norm, by properties of the functional calculus. The C^* identity implies that $\|B_t^{-1}\|_{op}^2 = \|B_t^{-2}\|_{op}$, which is the same as $\|A_t^{-1}\|_{op}$, and hence less than or equal to t^{-1} by the definition of A_t. Thus $\|B_t^{-1}\|_{op} \le t^{-1/2}$, which implies that $t\,\|B_t^{-1}\|_{op} \to 0$ as $t \to 0$, and (136.9) follows.

More precisely, this argument also shows that $R_t\,B \to R\,B = T$ as $t \to 0$ in the operator norm. If T is invertible, then $A = T^*\,T$ is invertible, $B = A^{1/2}$ is invertible, and $R = T\,B^{-1}$. In this case, one can show that $B_t^{-1} \to B_t^{-1}$ as $t \to 0$ in the operator norm, using basic properties of the functional calculus again. This implies that $R_t \to R$ as $t \to 0$ in the operator norm as well.

References

[1] W. Arveson, *An Invitation to C^*-Algebras*, Springer-Verlag, 1976.

[2] W. Arveson, *A Short Course on Spectral Theory*, Springer-Verlag, 2002.

[3] R. Beals, *Topics in Operator Theory*, University of Chicago Press, 1971.

[4] S. Berberian, *Measure and Integration*, Chelsea, 1970.

[5] S. Berberian, *Baer $*$-Rings*, Springer-Verlag, 1972.

[6] S. Berberian, *Lectures in Functional Analysis and Operator Theory*, Springer-Verlag, 1974.

[7] S. Berberian, *Introduction to Hilbert Space*, Chelsea, 1976.

[8] B. Blackadar, *Operator Algebras: Theory of C^*-Algebras and von Neumann Algebras*, Springer-Verlag, 2006.

[9] D. Blecher and C. Le Merdy, *Operator Algebras and their Modules — An Operator Space Approach*, Oxford University Press, 2004.

[10] K. Chandrasekharan, *A Course on Topological Groups*, Hindustan Book Agency, 1996.

[11] D. Cohn, *Measure Theory*, Birkhäuser, 1993.

[12] R. Coifman and G. Weiss, *Analyse Harmonique Non-Commutative sur Certains Espaces Homogènes*, Lecture Notes in Mathematics **242**, Springer-Verlag, 1971.

[13] R. Coifman and G. Weiss, *Transference Methods in Analysis*, American Mathematical Society, 1976.

[14] R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis*, Bulletin of the American Mathematical Society **83** (1977), 569–645.

[15] A. Connes, *Noncommutative Geometry*, Academic Press, 1994.

[16] J. Conway, *A Course in Functional Analysis*, 2nd edition, Springer-Verlag, 1990.

[17] J. Conway, *The Theory of Subnormal Operators*, American Mathematical Society, 1991.

[18] J. Conway, *A Course in Operator Theory*, American Mathematical Society, 2000.

[19] J. Dixmier, *C^*-Algebras*, translated from the second French edition by F. Jellet, North-Holland, 1977.

[20] J. Dixmier, *von Neumann Algebras*, with a preface by E. Lance, translated from the second French edition by F. Jellet, North-Holland, 1981.

[21] R. Douglas, *Banach Algebra Techniques in Operator Theory*, 2nd edition, Springer-Verlag, 1998.

[22] P. Duren, *Theory of H^p Spaces*, Academic Press, 1970.

[23] P. Duren and A. Schuster, *Bergman Spaces*, American Mathematical Society, 2004.

[24] E. Effros and Z.-J. Ruan, *Operator Spaces*, Oxford University Press, 2000.

[25] P. Eymard, *Moyennes Invariantes et Représentations Unitaires*, Lecture Notes in Mathematics **300**, Springer-Verlag, 1972.

[26] P. Fillmore, *A User's Guide to Operator Algebras*, Wiley, 1996.

[27] G. Folland, *A Course in Abstract Harmonic Analysis*, CRC Press, 1995.

[28] G. Folland, *Real Analysis*, 2nd edition, Wiley, 1999.

[29] J. Garnett, *Bounded Analytic Functions*, revised edition, Springer-Verlag, 2007.

[30] E. Glasner, *Ergodic Theory via Joinings*, American Mathematical Society, 2003.

[31] F. Greenleaf, *Invariant Means on Topological Groups and their Applications*, van Nostrand, 1969.

[32] P. Halmos, *Measure Theory*, Van Nostrand, 1950.

[33] P. Halmos, *A Hilbert Space Problem Book*, 2nd edition, Springer-Verlag, 1982.

[34] P. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, reprint of the 2nd edition, AMS Chelsea, 1998.

[35] E. Hewitt and K. Ross, *Abstract Harmonic Analysis*, Volume I, *Structure of Topological Groups, Integration Theory, Group Representations*, Volume II, *Structure and Analysis for Compact Groups, Analysis on Locally Compact Groups*, Springer-Verlag, 1970, 1979.

[36] E. Hewitt and K. Stromberg, *Real and Abstract Harmonic Analysis*, Springer-Verlag, 1975.

[37] P. Higgins, *Introduction to Topological Groups*, Cambridge University Press, 1974.

[38] K. Hoffman, *Banach Spaces of Analytic Functions*, Dover, 1988.

[39] N. Howes, *Modern Analysis and Topology*, Springer-Verlag, 1995.

[40] K. Jacobs, *Measure and Integral*, with an appendix by J. Kurzweil, Academic Press, 1978.

[41] R. Kadison and J. Ringrose, *Fundamentals of the Theory of Operator Algebras, I, Elementary Theory, II, Advanced Theory*, American Mathematical Society, 1997.

[42] R. Kadison and J. Ringrose, *Fundamentals of the Theory of Operator Algebras, III, Special Topics, Elementary Theory — An Exercise Approach, IV, Advanced Theory — An Exercise Approach*, Birkhäuser, 1991, 1992.

[43] S. Kantorovitz, *Introduction to Modern Analysis*, Oxford University Press, 2003.

[44] I. Kaplansky, *Rings of Operators*, Benjamin, 1968.

[45] Y. Katznelson, *An Introduction to Harmonic Analysis*, 3rd edition, Cambridge University Press, 2004.

[46] J. Kelley and T. Srinivasan, *Measure and Integral*, Springer-Verlag, 1988.

[47] A. Kharazishvili, *Transformation Groups and Invariant Measures*, World Scientific, 1998.

[48] J. Kingman and S. Taylor, *Introduction to Measure and Probability*, Cambridge University Press, 1966.

[49] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.

[50] A. Knapp, *Advanced Real Analysis*, Birkhäuser, 2005.

[51] P. Koosis, *Introduction to H_p Spaces*, 2nd edition, with two appendices by V. Havin, Cambridge University Press, 1998.

[52] E. Lorch, *Spectral Theory*, Oxford University Press, 1962.

[53] G. McCarty, *Topology: An Introduction with Application to Topological Groups*, Dover, 1988.

[54] M. Munroe, *Measure and Integration*, 2nd edition, Addison-Wesley, 1971.

[55] G. Murphy, *C^*-Algebras and Operator Theory*, Academic Press, 1990.

[56] L. Nachbin, *The Haar Integral*, Van Nostrand, 1965.

[57] J. von Neumann, *Invariant Measures*, American Mathematical Society, 1999.

[58] N. Nikolski, *Operators, Functions, and Systems: An Easy Reading*, Volume 1, *Hardy, Hankel, and Toeplitz*, Volume 2, *Model Operators and Systems*, translated from the French by A. Hartmann and revised by the author, American Mathematical Society, 2002.

[59] K. Parthasarathy, *Introduction to Probability and Measure*, Hindustan Book Agency, 2005.

[60] A. Paterson, *Amenability*, American Mathematical Society, 1988.

[61] A. Paterson, *Groupoids, Inverse Semigroups, and their Operator Algebras*, Birkhäuser, 1999.

[62] V. Paulsen, *Completely Bounded Maps and Dilations*, Wiley, 1986.

[63] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, 2002.

[64] G. Pedersen, *C^*-Algebras and their Automorphism Groups*, Academic Press, 1979.

[65] G. Pedersen, *Analysis Now*, Springer-Verlag, 1989.

[66] V. Peller, *Hankel Operators and their Applications*, Springer-Verlag, 2003.

[67] V. Pestov, *Dynamics of Infinite-Dimensional Groups*, American Mathematical Society, 2006.

[68] J.-P. Pier, *Amenable Locally Compact Groups*, Wiley, 1984.

[69] J.-P. Pier, *Amenable Banach Algebras*, Wiley, 1988.

[70] G. Pisier, *Similarity Problems and Completely Bounded Maps*, 2nd edition, Lecture Notes in Mathematics **1618**, Springer-Verlag, 2001.

[71] G. Pisier, *Introduction to Operator Space Theory*, Cambridge University Press, 2003.

[72] I. Raeburn, *Graph Algebras*, American Mathematical Society, 2005.

[73] M. Rao, *Measure Theory and Integration*, 2nd edition, Dekker, 2004.

[74] H. Reiter and J. Stegeman, *Classical Harmonic Analysis and Locally Compact Groups*, 2nd edition, Oxford University Press, 2000.

[75] C. Rickart, *General Theory of Banach Algebras*, van Nostrand, 1960.

[76] A. Robert, *Introduction to the Representation Theory of Compact and Locally Compact Groups*, Cambridge University Press, 1983.

[77] H. Royden, *Real Analysis*, 3rd edition, Macmillan, 1988.

[78] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, 1987.

[79] W. Rudin, *Fourier Analysis on Groups*, Wiley, 1990.

[80] W. Rudin, *Functional Analysis*, 2nd edition, McGraw-Hill, 1991.

[81] V. Runde, *Lectures on Amenability*, Lecture Notes in Mathematics **1774**, Springer-Verlag, 2004.

[82] R. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer-Verlag, 2002.

[83] S. Sakai, *C^*-Algebras and W^*-Algebras*, Springer-Verlag, 1998.

[84] D. Sarason, *Function Theory on the Unit Circle*, Department of Mathematics, Virginia Polytechnic Institute and State University, 1978.

[85] M. Simonnet, *Measures and Probabilities*, with a forward by C.-M. Marle, Springer-Verlag, 1996.

[86] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970.

[87] E. Stein, *Topics in Harmonic Analysis Related to the Littlewood–Paley Theory*, Princeton University Press, 1970.

[88] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, 1993.

[89] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[90] M. Stroppel, *Locally Compact Groups*, European Mathematical Society, 2006.

[91] M. Takesaki, *Theory of Operator Algebras I, II, III*, Springer-Verlag, 2002, 2003.

[92] N. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and Geometry on Groups*, Cambridge University Press, 1992.

[93] M. Väth, *Integration Theory: A Second Course*, World Scientific, 2002.

[94] K. Zhu, *An Introduction to Operator Algebras*, CRC Press, 1993.