An introduction to some aspects of functional analysis, 4: Dual spaces and linear mappings

Stephen Semmes
Rice University

Abstract

Some basic aspects of duals of Banach spaces and bounded linear mappings between them are discussed.

Contents

Part I
Dual spaces

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 $N(v)$ on V is said to be a *seminorm* if

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

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

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

for every $v, w \in V$. Here $|t|$ denotes the absolute value of $t \in \mathbf{R}$ in the real case, and the modulus of $t \in \mathbf{C}$ in the complex case. Note that $N(0) = 0$, by applying (1.1) with $t = 0$. If $N(v) > 0$ for each $v \in V$ with $v \neq 0$, then N is said to be a *norm* on V.

If N is a norm on V, then it is easy to see that

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

defines a metric on V, which leads to a topology on V in the usual way. It is well known and not difficult to check that addition and scalar multiplication on V are continuous with respect to this topology. Observe that

$$N(v) - N(w) \le N(v - w) \tag{1.4}$$

for every $v, w \in V$, by the triangle inequality (1.2), and similarly

$$N(w) - N(v) \le N(v - w). \tag{1.5}$$

This implies that

$$|N(v) - N(w)| \le N(v - w) \tag{1.6}$$

for every $v, w \in V$, and hence that N is a continuous function on V with respect to the metric (1.3).

Suppose for the moment that $V = \mathbf{R}^n$ or $\mathbf{C}^n$ for some positive integer n, and that N is a norm on V. In this case,

$$N(v) \le C\,|v| \tag{1.7}$$

for some nonnegative real number C and every $v \in V$, where $|v|$ denotes the standard Euclidean norm on $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. This can be verified by expressing v as a finite linear combination of the standard basis vectors in $\mathbf{R}^n$ or $\mathbf{C}^n$, and then using (1.1) and (1.2) to estimate $N(v)$ in terms of the coordinates of v. Combining this with (1.6), we get that N is a continuous function with respect to the standard topology on $\mathbf{R}^n$ or $\mathbf{C}^n$.

It is well known that closed and bounded subsets of $\mathbf{R}^n$ and $\mathbf{C}^n$ are compact with respect to the standard topology, and in particular that the unit sphere is compact. It is also well known that a continuous real-valued function on nonempty compact set attains its maximum and minimum on that set, so that N attains its minimum on the unit sphere. If c is the minimum of N on the unit sphere, then $c > 0$, because $N(v) > 0$ when $v \neq 0$. Thus $N(v) \ge c$ when $|v| = 1$, which implies that

$$N(v) \ge c\,|v| \tag{1.8}$$

for every $v \in \mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, because of the homogeneity property of norms. Using this and (1.7), it follows that the topology on $\mathbf{R}^n$ or $\mathbf{C}^n$ determined by N is the same as the standard topology.

2 Completeness

Let V be a real or complex vector space with a norm $\|v\|$, and let

$$d(v, w) = \|v - w\| \tag{2.1}$$

be the corresponding metric on V. Using this metric, one can define convergence of sequences in V, Cauchy sequences in V, and so on, in the usual way. In particular, V is complete if every Cauchy sequence of elements of V converges to an element of V, in which case V is said to be a *Banach space*. It is well known that $\mathbf{R}^n$ and $\mathbf{C}^n$ are complete with respect to their standard Euclidean metrics, for instance. It follows that $\mathbf{R}^n$ and $\mathbf{C}^n$ are complete with respect to the metrics associated to arbitrary norms on them, because of the equivalence of an arbitrary norm with the standard Euclidean norm, as in the previous section. Similarly, if V is a finite-dimensional real or complex vector space of dimension n, then there is a one-to-one linear mapping T from $\mathbf{R}^n$ or $\mathbf{C}^n$ onto V, as appropriate. If $\|v\|$ is a norm on V, then

$$N(z) = \|T(z)\| \tag{2.2}$$

defines a norm on $\mathbf{R}^n$ or $\mathbf{C}^n$. Because $\mathbf{R}^n$, $\mathbf{C}^n$ are complete with respect to any norm, it follows that V is complete with respect to $\|v\|$.

Let V be a real or complex vector space with a norm $\|v\|$ again, and let W be a linear subspace of V. Thus the restriction of $\|v\|$ to $v \in W$ defines a norm on W. Suppose that W is complete with respect to the restriction of $\|v\|$ to $v \in W$, and let us check that W is a closed set in V with respect to the metric (2.1) on V associated to this norm. Equivalently, this means that if $\{w_j\}_{j=1}^\infty$ is a sequence of elements of W that converges to some $v \in V$ with respect to this norm, then $v \in W$. Under these conditions, $\{w_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to this norm, and hence $\{w_j\}_{j=1}^\infty$ converges to some $w \in W$ because W is complete. Hence $v = w \in W$, as desired. In particular, it follows that finite-dimensional linear subspaces of V are closed subsets of V, since they are always complete.

Let V be a real or complex vector space with a norm $\|v\|$, and let $\sum_{j=1}^\infty a_j$ be an infinite series with $a_j \in V$ for each j. As usual, $\sum_{j=1}^\infty a_j$ is said to converge in V if the corresponding sequence of partial sums

$$A_n = \sum_{j=1}^n a_j \tag{2.3}$$

converges in V. If

$$\sum_{j=1}^\infty \|a_j\| \tag{2.4}$$

converges as an infinite series of nonnegative real numbers, then $\sum_{j=1}^\infty a_j$ is said to converge absolutely in V. This implies that the sequence $\{A_n\}_{j=1}^\infty$ of partial sums is a Cauchy sequence in V, because

$$\|A_n - A_l\| = \left\| \sum_{j=l+1}^{n} a_j \right\| \le \sum_{j=l+1}^{n} \|a_j\| \to 0 \tag{2.5}$$

when $n \ge l$ and $l \to \infty$. If V is complete, then it follows that $\{A_n\}_{n=1}^\infty$ converges in V, so that $\sum_{j=1}^\infty a_j$ converges in V.

Conversely, suppose that every absolutely convergent series with terms in V also converges in V, and let us show that V is complete. If $\{v_j\}_{j=1}^\infty$ is any Cauchy sequence in V, then there is a subsequence $\{v_{j_l}\}_{l=1}^\infty$ of $\{v_j\}_{j=1}^\infty$ such that

$$\|v_{j_{l+1}} - v_{j_l}\| < 2^{-l} \tag{2.6}$$

for each l, by standard arguments. This implies that

$$\sum_{l=1}^{\infty} (v_{j_{l+1}} - v_{j_l}) \tag{2.7}$$

converges absolutely in V, and hence that $\sum_{l=1}^\infty (v_{j_{l+1}} - v_{j_l})$ converges in V, by hypothesis. Equivalently,

$$\sum_{l=1}^{n} (v_{j_{l+1}} - v_{j_l}) = v_{j_{n+1}} - v_{j_1} \tag{2.8}$$

converges in V as $n \to \infty$, so that $\{v_{j_{n+1}}\}_{n=1}^\infty$ converges as a sequence in V. It follows that the whole sequence $\{v_j\}_{j=1}^\infty$ converges in V to the same limit, as desired, because $\{v_j\}_{j=1}^\infty$ is a Cauchy sequence in V.

3 Bounded linear functionals

Let V be a real or complex vector space with a norm $\|v\|$. Remember that a linear functional on V is a linear mapping from V into the real or complex numbers as a 1-dimensional real or complex vector space, as appropriate. A linear functional λ on V is said to be *bounded* if there is a nonnegative real number A such that

$$|\lambda(v)| \le A\,\|v\| \tag{3.1}$$

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

$$|\lambda(v) - \lambda(w)| = |\lambda(v - w)| \le A\,\|v - w\| \tag{3.2}$$

for every $v, w \in V$, and hence that λ is uniformly continuous on V with respect to the metric associated to the norm $\|v\|$. Conversely, if λ is continuous at 0 on V, then there is a $\delta > 0$ such that $|\lambda(v)| < 1$ for every $v \in V$ with $\|v\| < \delta$, and one can use this to show that (3.1) holds with $A = 1/\delta$.

If λ is a bounded linear functional on V, then put

$$\|\lambda\|_* = \sup\{|\lambda(v)| : v \in V,\ \|v\| \le 1\}. \tag{3.3}$$

Thus $\|\lambda\|_* \le A$ when λ satisfies (3.1), and conversely (3.1) holds with $A = \|\lambda\|_*$. Let V^* be the space of bounded linear functionals on V. If λ, μ are bounded linear functionals on V, then it is easy to see that $\lambda + \mu$ is also a bounded linear functional on V, and that

$$\|\lambda + \mu\|_* \le \|\lambda\|_* + \|\mu\|_*. \tag{3.4}$$

Similarly, if λ is a bounded linear functional on V and a is a real or complex number, as appropriate, then $a\,\lambda$ is also a bounded linear functional on V, and

$$\|a\,\lambda\|_* = |a|\,\|\lambda\|_*. \tag{3.5}$$

Note that $\lambda = 0$ when $\|\lambda\|_* = 0$. It follows that V^* is a vector space with respect to pointwise addition and scalar multiplication, and that $\|\lambda\|_*$ defines a norm on V^*. More precisely, V^* is known as the dual of V, and $\|\lambda\|_*$ is known as the dual norm on V^* associated to the norm $\|v\|$ on V.

If $V = \mathbf{R}^n$ or $\mathbf{C}^n$ for some positive integer n, equipped with the standard Euclidean norm, then it is easy to see that every linear functional on V is bounded. This also works when $V = \mathbf{R}^n$ or $\mathbf{C}^n$ is equipped with any norm, because of the equivalence of any norm with the standard norm in this case, as in Section 1. If V is any finite-dimensional real or complex vector space with a norm, then every linear linear functional on V is again continuous, because one can reduce to the case where $V = \mathbf{R}^n$ or $\mathbf{C}^n$ using a one-to-one linear mapping from $\mathbf{R}^n$ or $\mathbf{C}^n$ onto V, as appropriate. It is well known that the space of linear functionals on a finite-dimensional vector space V is also a finite-dimensional vector space, with dimension equal to the dimension of V.

Let V be any real or complex vector space with a norm $\|v\|$ again, and let us check that the corresponding dual space V^* is complete with respect to the dual norm $\|\lambda\|_*$. Let $\{\lambda_j\}_{j=1}^\infty$ be a Cauchy sequence of bounded linear functionals on V with respect to the dual norm, so that for each $\epsilon > 0$ there is an $L(\epsilon) \ge 1$ such that

$$\|\lambda_j - \lambda_l\|_* < \epsilon \tag{3.6}$$

for every $j, l \ge L(\epsilon)$. By definition of the dual norm, this implies that

$$|\lambda_j(v) - \lambda_l(v)| \le \epsilon\,\|v\| \tag{3.7}$$

for every $v \in V$ and $j, l \ge L(\epsilon)$. Hence $\{\lambda_j(v)\}_{j=1}^\infty$ is a Cauchy sequence in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for every $v \in V$. Because of the completeness of the real and complex numbers, it follows that $\{\lambda_j(v)\}_{j=1}^\infty$ converges to a real or complex number $\lambda(v)$, as appropriate, for every $v \in V$. It is easy to see that $\lambda(v)$ defines a linear functional on V, since $\lambda_j(v)$ is a linear functional on V for each j. Taking the limit as $j \to \infty$ in (3.7), we get that

$$|\lambda(v) - \lambda_l(v)| \le \epsilon\,\|v\| \tag{3.8}$$

for every $v \in V$ and $l \ge L(\epsilon)$. In particular, we can apply this with $\epsilon = 1$ and $l = L(1)$, to get that

$$|\lambda(v)| \le |\lambda(v) - \lambda_{L(1)}(v)| + |\lambda_{L(1)}(v)| \le \|v\| + \|\lambda_{L(1)}\|_*\,\|v\| \tag{3.9}$$

for every $v \in V$, which implies that λ is a bounded linear functional on V. We can also reformulate (3.8) as saying that

$$\|\lambda - \lambda_l\|_* \le \epsilon \tag{3.10}$$

for every $l \ge L(\epsilon)$, so that $\{\lambda_l\}_{l=1}^\infty$ converges to λ with respect to the dual norm on V^*, as desired.

4 The Hahn–Banach Theorem

Let V be a real or complex vector space, and let N be a seminorm on V. Also let W be a linear subspace of V, and let λ be a linear functional on W such that

$$|\lambda(w)| \le C\, N(w) \tag{4.1}$$

for some nonnegative real number C and every $w \in W$. Under these conditions, the Hahn–Banach theorem implies that there is an extension of λ to a linear functional on V that satisfies (4.1) for every $w \in V$, with the same constant C. Of course, this is trivial when $C = 0$, and otherwise it is easy to reduce to the case where $C = 1$. The proof is well known, and we shall not include it here.

As an application, suppose that $\|\cdot\|$ is a norm on V, and let v be a nonzero element of V. Let W be the linear span of v in V, which is the 1-dimensional subspace of V consisting of vectors of the form $t\,v$, where $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. Let λ be the linear functional on W defined by

$$\lambda(t\,v) = t\,\|v\| \tag{4.2}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, so that

$$|\lambda(t\,v)| = |t|\,\|v\| = \|t\,v\| \tag{4.3}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$. The Hahn–Banach theorem implies that there is an extension of λ to a linear functional on V such that

$$|\lambda(w)| \le \|w\| \tag{4.4}$$

for every $w \in V$. This extension is a bounded linear functional on V with dual norm equal to 1, since $\lambda(v) = \|v\|$ by construction.

Now let Z be a closed linear subspace of V with respect to $\|\cdot\|$, and put

$$N_Z(v) = \operatorname{dist}(v, Z) = \inf\{\|v - z\| : z \in Z\} \tag{4.5}$$

for each $v \in V$. This is the distance from v to Z in V with respect to $\|\cdot\|$, and it is easy to see that $N_Z(v) = 0$ if and only if $v \in Z$, because Z is closed. If $w, v \in V$ and $y, z \in Z$, then

$$N_Z(v + w) \le \|v + w - z - y\| \le \|v - z\| + \|w - y\|, \tag{4.6}$$

because $z + y \in Z$, and hence

$$N_z(v + w) \le N_z(v) + N_z(w), \tag{4.7}$$

by taking the infimum over $y, z \in Z$ in (4.6). Similarly, one can check that

$$N_Z(t\,v) = |t|\,N_Z(v) \tag{4.8}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, so that N_Z is a seminorm on V. Of course,

$$N_Z(v) \le \|v\| \tag{4.9}$$

for every $v \in V$, since we can always take $z = 0$ in (4.5).

Let v be an element of V not in Z, and let W be the linear span of Z and v in V. Thus W consists of the elements of V of the form $t\,v + z$ for some $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, and $z \in Z$, and every element of W has a unique representation of this type, since $v \notin Z$. Let λ be the linear functional on W given by

$$\lambda(t\,v + z) = t\,N_Z(v) \tag{4.10}$$

for each $t \in \mathbf{R}$ or $\mathbf{C}$ and $z \in Z$. Observe that

$$|\lambda(t\,v + z)| = |t|\,N_Z(v) = N_Z(t\,v) = N_Z(t\,v + z) \tag{4.11}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$ and $z \in Z$, using the definition of N_Z and the fact that Z is a linear subspace of V in the last step. The Hahn–Banach theorem implies that there is an extension of λ to a linear functional on V that satisfies

$$|\lambda(w)| \le N_Z(w) \le \|w\| \tag{4.12}$$

for every $w \in V$. In particular, λ is a bounded linear functional on V with dual norm less than or equal to 1. One can check that the dual norm of λ is actually equal to 1, by choosing $z \in Z$ such that $\|v + z\|$ approximates $N_Z(v)$, while $\lambda(v + z) = N_Z(v)$ for every $z \in Z$. Note that $\lambda(v) = N_Z(v) > 0$ and $\lambda(z) = 0$ for every $z \in Z$, by construction.

5 Quotient spaces

Let V be a real or complex vector space, and let W be a linear subspace of V. If $u, v \in V$ satisfy $u - v \in W$, then put $u \sim v$. This defines a relation on V which is reflexive in the sense that $v \sim v$ for every $v \in V$, because $v - v = 0$ is automatically an element of W. Similarly, this relation is symmetric on V, which means that $u \sim v$ is equivalent to $v \sim u$, because $u - v = -(v - u) \in W$ if and only if $v - u \in W$. Let us check that this relation is also transitive on V, so that $u \sim v$ and $v \sim z$ imply that $u \sim z$. In this case, $u - v \in W$ and $v - z \in W$, and hence

$$u - z = (u - v) + (v - z) \in W, \tag{5.1}$$

as desired. Thus $u \sim v$ is an equivalence relation on V, since it is reflexive, symmetric, and transitive. If $u, u' \in V$ satisfy $u \sim v$ and $u' \sim v'$, so that $u - v \in W$ and $u' - v' \in W$, then we get that

$$(u + u') - (v + v') = (u - v) + (u' - v') \in W, \tag{5.2}$$

which implies that

$$u + u' \sim v + v'. \tag{5.3}$$

Similarly, if $u, v \in V$ satisfy $u \sim v$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then

$$t\, u \sim t\, v, \tag{5.4}$$

because $u - v \in W$ and hence $t\, u - t\, v = t\,(u - v) \in W$.

Let v be any element of V, and put

$$v + W = \{v + w : w \in W\}. \tag{5.5}$$

This is the same as the set of $u \in V$ such that $u \sim v$, which is known as the equivalence class in V that contains v. Note that $v + W = v' + W$ for some $v, v' \in V$ if and only if $v \sim v'$. If $u \in V$ is an element of both $v + W$ and $v' + W$ for some $v, v' \in V$, then $u \sim v$ and $u \sim v'$, so that $v \sim v'$ and $v + W = v' + W$. It follows that any two equivalence classes in V are either the same or disjoint as subsets of V.

Let V/W be the set of these equivalence classes in V, and let q be the quotient mapping from V onto V/W that sends each $v \in V$ to the corresponding equivalence class $q(v) = v + W$. If $v, v' \in V$, then put

$$(v + W) + (v' + W) = (v + v') + W. \tag{5.6}$$

If $u, u' \in V$ satisfy $u + W = v + W$ and $u' + W = v' + W$, which is to say that $u \sim v$ and $u' \sim v'$, then we have seen that $u + u' \sim v + v'$, and hence

$$(u + u') + W = (v + v') + W. \tag{5.7}$$

This shows that addition on V/W is well-defined, in the sense that the sum of two equivalence classes does not depend on how the equivalence classes are represented. Similarly, if $v \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then we put

$$t\,(v + W) = t\, v + W, \tag{5.8}$$

and we have that

$$t\, u + W = t\, v + W \tag{5.9}$$

when $u \in V$ and $u + W = v + W$. Thus scalar multiplication on V/W is also well-defined, and it is well known and easy to check that V/W becomes a vector space over $\mathbf{R}$ or $\mathbf{C}$, as appropriate, with these definitions of addition and scalar multiplication, and with $W = 0 + W$ as the zero element in V/W. By construction, the natural quotient mapping q is a linear mapping from V onto V/W with kernel equal to W, since $q(v) = 0$ in V/W if and only if $v \in W$.

Let Z be another vector space over the real or complex numbers, depending on whether V is real or complex. Also let T be a linear mapping from V into Z whose kernel contains W, so that $T(w) = 0$ in Z for every $w \in W$. Thus

$$T(u) = T(v) \tag{5.10}$$

in Z when $u, v \in V$ satisfy $u \sim v$, which is to say that $u - v \in W$. This permits us to define a mapping $\widehat{T}$ from V/W into Z by

$$\widehat{T}(v + W) = T(v), \tag{5.11}$$

since (5.10) holds when $u + W = v + W$. More precisely, $\widehat{T}$ is a linear mapping from V/W into Z under these conditions, and the composition of $\widehat{T}$ with q is equal to T.

6 Quotient norms

Let V be a real or complex vector space with a norm $\|v\|_V$, and let W be a closed linear subspace of V. Put

$$N_W(v) = \mathrm{dist}(v, W) = \inf\{\|v - w\|_V : w \in W\} \tag{6.1}$$

as in Section 4, so that N_W is a seminorm on V and $N_W(v) = 0$ if and only if $v \in W$. Let V/W be the quotient of V by W, as in the previous section, and let q be the natural quotient mapping from V onto V/W. If $u, v \in V$ and $q(u) = q(v)$, then $u - v \in W$, and hence $N_W(u) = N_W(v)$. This permits us to define a nonnegative real-valued function $\|\cdot\|_{V/W}$ on V/W by putting

$$\|q(v)\|_{V/W} = N_W(v). \tag{6.2}$$

One can check that this defines a norm on V/W, because of the corresponding properties of N_W on V. This is known as the *quotient norm* on V/W associated to the norm $\|v\|_V$ on V.

By construction,

$$\|q(v)\|_{V/W} \le \|v\|_V \tag{6.3}$$

for every $v \in V$, and hence

$$\|q(u) - q(v)\|_{V/W} = \|q(u - v)\|_{V/W} \le \|u - v\|_V \tag{6.4}$$

for every $u, v \in V$. As usual, this implies that q is uniformly continuous as a mapping from V with the metric associated to the norm $\|\cdot\|_V$ onto V/W with the metric associated to the quotient norm $\|\cdot\|_{V/W}$. If $v \in V$ and r is a positive real number, then let $B(v, r)$ be the open ball in V with center v and radius r, so that

$$B(v, r) = \{u \in V : \|u - v\| < r\}. \tag{6.5}$$

Because of (6.4), q maps $B(v, r)$ into the open ball in V/W with center $q(v)$ and radius r with respect to the quotient norm $\|\cdot\|_{V/W}$. Let us check that q

actually maps $B(v, r)$ onto the open ball in V/W with center $q(v)$ and radius r. If $u \in V$ satisfies $\|q(u) - q(v)\|_{V/W} < r$, then there is a $w \in W$ such that $\|u + w - v\|_V < r$, because of the way that the quotient norm is defined. This implies that $u + w \in B(v, r)$ and $q(u + w) = q(u)$, so that $q(u)$ is in the image of $B(v, r)$ under q, as desired. It follows that q is an open mapping from V onto V/W, which means that q maps open subsets of V to open subsets of V/W.

Suppose now that V is complete with respect to $\|v\|_V$, and let us show that V/W is complete with respect to the corresponding quotient norm. As in Section 2, it suffices to show that every absolutely convergent infinite series with terms in V/W converges in V/W under these conditions. In this case, this means that if $\{v_j\}_{j=1}^\infty$ is a sequence of elements of V such that

$$\sum_{j=1}^\infty \|q(v_j)\|_{V/W} \tag{6.6}$$

converges as an infinite series of nonnegative real numbers, then $\sum_{j=1}^\infty q(v_j)$ converges in V/W. Because of the way that the quotient norm is defined, for each positive integer j there is a $w_j \in W$ such that

$$\|v_j - w_j\|_V < \|q(v_j)\|_{V/W} + 1/j^2. \tag{6.7}$$

This implies that $\sum_{j=1}^\infty \|v_j - w_j\|_V$ converges as an infinite series of nonnegative real numbers, since $\sum_{j=1}^\infty 1/j^2$ converges. It follows that $\sum_{j=1}^\infty (v_j - w_j)$ converges in V, because V is complete. Of course,

$$q\Big(\sum_{j=1}^n (v_j - w_j)\Big) = \sum_{j=1}^n q(v_j - w_j) = \sum_{j=1}^n q(v_j) \tag{6.8}$$

for each $n \ge 1$. Using this and the continuity of q, we get that

$$\lim_{n \to \infty} \sum_{j=1}^n q(v_j) = q\Big(\sum_{j=1}^\infty (v_j - w_j)\Big), \tag{6.9}$$

as desired.

Let Z be another linear subspace of V, and consider the linear subspace $W + Z$ of V spanned by W and Z, which is given by

$$W + Z = \{w + z : w \in W,\, z \in Z\}. \tag{6.10}$$

If Z is a finite-dimensional linear subspace of V, then its image $q(Z)$ in V/W under q also has finite dimension, with the dimension of $q(Z)$ being less than or equal to the dimension of Z. In particular, $q(Z)$ is a closed linear subspace of V/W with respect to the quotient norm when Z is finite-dimensional, as in Section 2. This implies that the inverse image $q^{-1}(q(Z))$ of $q(Z)$ under q is a closed linear subspace of V when Z has finite dimension, because q is continuous. Hence $W + Z$ is a closed linear subspace of V when W is a closed linear subspace of V and Z is a finite-dimensional linear subspace of V, since $W + Z = q^{-1}(q(Z))$.

7 Duals of quotient spaces

Let V be a real or complex vector space with a norm $\|v\|_V$, and let W be a closed linear subspace of V. Also let V/W be the corresponding quotient space, with the quotient mapping q from V onto V/W and the quotient norm $\|\cdot\|_{V/W}$ on V/W. If $\widehat{\lambda}$ is a bounded linear functional on V/W, then it is easy to see that

$$\lambda(v) = \widehat{\lambda}(q(v)) \tag{7.1}$$

is a bounded linear functional on V. More precisely, the dual norm $\|\lambda\|_{V^*}$ of λ on V with respect to $\|\cdot\|_V$ is less than or equal to the dual norm $\|\widehat{\lambda}\|_{(V/W)^*}$ of $\widehat{\lambda}$ on V/W with respect to the quotient norm $\|\cdot\|_{V/W}$, because of (6.3). Of course, $\lambda(v) = 0$ for every $v \in W$, because $q(v) = 0$ when $v \in W$, by construction.

Conversely, suppose that λ is a bounded linear functional on V such that $\lambda(v) = 0$ for every $v \in W$. As in Section 5, there is a linear functional $\widehat{\lambda}$ on V/W that satisfies (7.1) for every $v \in V$. Note that

$$|\lambda(v)| \le \|\lambda\|_{V^*}\, \|v\|_V \tag{7.2}$$

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

$$|\lambda(v)| = |\lambda(v - w)| \le \|\lambda\|_{V^*}\, \|v - w\|_V \tag{7.3}$$

for every $v \in V$ and $w \in W$, and hence that

$$|\lambda(v)| \le \|\lambda\|_{V^*}\, N_W(v) \tag{7.4}$$

for every $v \in V$, where $N_W(v)$ is as in the previous section. Equivalently,

$$|\widehat{\lambda}(q(v))| \le \|\lambda\|_{V^*}\, \|q(v)\|_{V/W} \tag{7.5}$$

for every $v \in V$, so that $\widehat{\lambda}$ is a bounded linear functional on V/W, with $\|\widehat{\lambda}\|_{(V/W)^*} \le \|\lambda\|_{V^*}$. It follows that

$$\|\widehat{\lambda}\|_{(V/W)^*} = \|\lambda\|_{V^*} \tag{7.6}$$

when λ and $\widehat{\lambda}$ are related as in (7.1), since the opposite inequality was obtained earlier.

Consider the linear subspace $W^\perp$ of V^* defined by

$$W^\perp = \{\lambda \in V^* : \lambda(v) = 0 \text{ for every } v \in W\}. \tag{7.7}$$

Note that this is a closed subset of V^* with respect to the dual norm $\|\lambda\|_{V^*}$. The preceding discussion shows that there is a natural isometric linear isomorphism from $(V/W)^*$ onto $W^\perp$, which sends $\widehat{\lambda} \in (V/W)^*$ to $\lambda = \widehat{\lambda} \circ q \in W^\perp$.

8 Duals of linear subspaces

Let M and N be metric spaces, let E be a dense subset of M, and let f be a uniformly continuous mapping from E into N. If N is complete, then it is well known that there is a unique extension of f to a uniformly continuous mapping from M into N. More precisely, only continuity of the extension is needed to get uniqueness. If E is not dense in M, then one can extend f to the closure $\overline{E}$ of E in M. Let us briefly sketch the proof of this fact. If $x \in \overline{E}$, then there is a sequence $\{x_j\}_{j=1}^\infty$ of elements of E that converges to x in M. In particular, $\{x_j\}_{j=1}^\infty$ is a Cauchy sequence of elements of E, and one can check that $\{f(x_j)\}_{j=1}^\infty$ is a Cauchy sequence in N, because f is uniformly continuous. If N is complete, then it follows that $\{f(x_j)\}_{j=1}^\infty$ converges in N. If $\{x_j'\}_{j=1}^\infty$ is another sequence of elements of E that converges to x, then one can use the uniform continuity of f again to show that $\{f(x_j')\}_{j=1}^\infty$ converges to the same point in N as $\{f(x_j)\}_{j=1}^\infty$. The extension of f to $\overline{E}$ is defined by taking $f(x)$ to be the common value of the limit in N, and one can show that this is uniformly continuous on $\overline{E}$ using the uniform continuity of f on E.

Let V be a real or complex vector space with a norm $\|v\|_V$, and let W be a linear subspace of V. If λ is a bounded linear functional on W, then λ is uniformly continuous on W, as in Section 3. Thus λ extends to a unique uniformly continuous mapping from the closure $\overline{W}$ of W in V into the real or complex numbers, as appropriate, since $\mathbf{R}$ and $\mathbf{C}$ are complete. It is easy to check that this extension is also a bounded linear functional on $\overline{W}$ under these conditions, with the same dual norm with respect to the restriction of $\|\cdot\|_V$ to $\overline{W}$ as the dual norm of λ on W. Let us restrict our attention to closed linear subspaces W of V from now on in this section.

If λ is a bounded linear functional on V, then the restriction $R_W(\lambda)$ of λ to W defines a bounded linear functional on W. More precisely, if $\|\lambda\|_{V^*}$ denotes the dual norm of a bounded linear functional λ on V with respect to $\|v\|_V$, and if $\|\mu\|_{W^*}$ denotes the dual norm of a bounded linear functional μ on W with respect to the restriction of $\|v\|_V$ to $v \in W$, then

$$\|R_W(\lambda)\|_{W^*} \le \|\lambda\|_{V^*} \tag{8.1}$$

for every $\lambda \in V^*$. Of course, $R_W(\lambda) = 0$ if and only if $\lambda \in W^\perp$, where $W^\perp$ is as in (7.7). If μ is any bounded linear functional on W, then the Hahn–Banach theorem implies that there is a $\lambda \in V^*$ such that $\mu = R_W(\lambda)$ and

$$\|\lambda\|_{V^*} = \|\mu\|_{W^*}. \tag{8.2}$$

In particular, R_W defines a linear mapping from V^* onto W^*.

Since $W^\perp$ is a closed linear subspace of V^*, we can define the quotient norm on $V^*/W^\perp$ associated to the dual norm $\|\lambda\|_{V^*}$ on V^* as before. Because the kernel of R_W is equal to $W^\perp$, there is a one-to-one linear mapping $\widehat{R}_W$ from $V^*/W^\perp$ onto W^* whose composition with the natural quotient mapping from V^* onto $V^*/W^\perp$ is equal to R_W. Using (8.1) and (8.2), one can check that $\widehat{R}_W$ is an isometry with respect to the quotient norm on $V^*/W^\perp$ just mentioned and the dual norm $\|\mu\|_{W^*}$ on W^*.

9 Second duals

Let V be a real or complex vector space with a norm $\|v\|$, and let V^* be the dual space of bounded linear functionals on V, with the corresponding dual norm $\|\lambda\|_*$. Repeating the process, we get the second dual space V^{**} of bounded linear functionals on V^*, with the corresponding dual norm $\|\cdot\|_{**}$. If $v \in V$, then

$$L_v(\lambda) = \lambda(v) \tag{9.1}$$

defines a linear functional on V^*, which is bounded because

$$|L_v(\lambda)| = |\lambda(v)| \le \|\lambda\|_* \, \|v\| \tag{9.2}$$

for each $\lambda \in V^*$. More precisely, this shows that $\|L_v\|_{**} \le \|v\|$, and we actually have that

$$\|L_v\|_{**} = \|v\| \tag{9.3}$$

for every $v \in V$. Indeed, if $v \neq 0$, then the Hahn–Banach theorem implies that there is a $\lambda \in V^*$ such that $\|\lambda\|_* = 1$ and $\lambda(v) = \|v\|$, as in Section 4. This shows that equality can hold in (9.2) with $\lambda \neq 0$, so that (9.3) holds, as desired. Thus the mapping $v \mapsto L_v$ is an isometric linear embedding of V into V^{**}.

If every element of V^{**} is of the form L_v for some $v \in V$, then V is said to be *reflexive*. It is easy to see that this happens when V is finite-dimensional, in which case V^* and V^{**} are also finite-dimensional, with the same dimension as V. Remember that dual spaces are automatically complete, as in Section 3. Thus V has to be complete in order to be reflexive, because V^{**} is complete.

Let W be a closed linear subspace of V, and let $W^\perp$ be the corresponding closed linear subspace of V^* as in (7.7). Repeating the process, we get a linear subspace $(W^\perp)^\perp$ of V^{**} which is closed with respect to $\|\cdot\|_{**}$. If $v \in W$, then $L_v(\lambda) = \lambda(v) = 0$ for every $\lambda \in W^\perp$, and hence $L_v \in (W^\perp)^\perp$. However, if $v \in V \backslash W$, then there is a $\lambda \in W^\perp$ such that $\lambda(v) \neq 0$, as in Section 4. This implies that $L_v \notin (W^\perp)^\perp$ when $v \in V \backslash W$, so that $L_v \in (W^\perp)^\perp$ if and only if $v \in W$.

If V is reflexive, then it follows that $(W^\perp)^\perp$ corresponds exactly to W under the natural isomorphism between V and V^{**}. In this case, one can check that W and V/W are both reflexive as well. Remember that the dual of W can be identified with $V^*/W^\perp$ in a natural way, and then the dual of $V^*/W^\perp$ can be identified with $(W^\perp)^\perp$. Similarly, the dual of V/W can be identified with $W^\perp$, and the dual of $W^\perp$ can be identified with $V^{**}/(W^\perp)^\perp$. If V is reflexive, then one can check that these identifications match up in the right way, to get that W and V/W are reflexive too.

10 Minkowski functionals

Let V be a real or complex vector space. As usual, a subset A of V is said to be *convex* if

$$t\,v + (1-t)\,w \in A \tag{10.1}$$

for every $v, w \in A$ and $t \in \mathbf{R}$ with $0 \le t \le 1$. Similarly, we say that A is *balanced* if

$$t\,A \subseteq A \tag{10.2}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|t| \le 1$, where $t\,A$ is the set of vectors of the form $t\,v$ with $v \in A$. If A is balanced and nonempty, then $0 \in A$, and the condition that (10.2) hold for $t \in \mathbf{R}$ with $0 \le t \le 1$ means that A is "star-like about 0", which is also implied by convexity when $0 \in A$. If A is star-like about 0, then A is balanced if and only if (10.2) holds when $t \in \mathbf{R}$ or $\mathbf{C}$ satisfies $|t| = 1$. In this case, we have that

$$t\,A = A \tag{10.3}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, with $|t| = 1$, since we can also apply (10.2) to t^{-1} when $|t| = 1$. If V is a real vector space, then (10.3) reduces to the condition that A be symmetric about the origin, in the sense that $-A = A$.

Suppose now that V is equipped with a norm $\|v\|$, and that A is a nonempty balanced convex open subset of V. In particular, A contains a ball of some positive radius around 0, since $0 \in A$ and A is an open set in V. The Minkowski functional on V corresponding to A is defined by

$$N_A(v) = \inf\{r > 0 : r^{-1}\,v \in A\} = \inf\{r > 0 : v \in r\,A\} \tag{10.4}$$

for each $v \in V$. More precisely, the infima are taken over all positive real numbers r with the indicated properties, and these properties hold when r is sufficiently large because A contains a ball around 0 in V. By construction,

$$N_A(t\,v) = t\,N_A(v) \tag{10.5}$$

for every $v \in V$ and positive real number t, and in fact

$$N_A(t\,v) = |t|\,N_A(v) \tag{10.6}$$

for every $v \in V$ and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, because A is balanced.

If $v \in V$ satisfies $N_A(v) < 1$, then $v \in r\,A$ for some $r < 1$, and hence $v \in A$, because $r\,A \subseteq A$. Conversely, if $v \in A$, then $r^{-1}\,v \in A$ when r is sufficiently close to 1, because A is an open set in V. This implies that $N_A(v) < 1$, and it follows that

$$A = \{v \in V : N_A(v) < 1\}. \tag{10.7}$$

Now let v and w be arbitrary elements of V, and let us check that

$$N_A(v + w) \le N_A(v) + N_A(w). \tag{10.8}$$

If $r_v > N_A(v)$ and $r_w > N_A(w)$, then $r_v^{-1}\,v, r_w^{-1}\,w \in A$, because A is star-like about 0. This implies that

$$(r_v + r_w)^{-1}\,(v + w) = \frac{r_v}{r_v + r_w}\,(r_v^{-1}\,v) + \frac{r_w}{r_v + r_w}\,(r_w^{-1}\,w) \in A, \tag{10.9}$$

because A is convex, so that

$$N_A(v + w) \le r_v + r_w. \tag{10.10}$$

Thus we get (10.8), since (10.10) holds for every $r_v > N_A(v)$ and $r_w > N_A(w)$.

This shows that $N_A(v)$ is a seminorm on V under these conditions. One can also check that

$$N_A(v) \le C\,\|v\| \tag{10.11}$$

for some $C \ge 0$ and every $v \in V$, because A contains a ball around 0 with respect to $\|v\|$. This implies that N_A is a continuous function on V with respect to $\|v\|$, because of (1.6). Conversely, if N is any seminorm on V, then it is easy to see that the open unit ball

$$\{v \in V : N(v) < 1\} \tag{10.12}$$

in V with respect to N is a balanced convex subset of V. If N is continuous with respect to $\|v\|$ on V, then (10.12) is an open set in V with respect to $\|v\|$.

11 Some separation theorems

Let V be a real or complex vector space with a norm $\|v\|$, and let A be a nonempty balanced convex open subset of V. Thus the Minkowski functional $N_A(v)$ corresponding to A is a seminorm on V that is bounded by a constant multiple of $\|v\|$, as in the previous section. Suppose that $v \in V\backslash A$, and let λ be the linear functional defined initially on the 1-dimensional linear subspace of V passing through v by

$$\lambda(t\,v) = t\,N_A(v) \tag{11.1}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate. Thus

$$|\lambda(t\,v)| = |t|\,N_A(v) = N_A(t\,v) \tag{11.2}$$

for every $t \in \mathbf{R}$ or $\mathbf{C}$, and the Hahn–Banach theorem implies that there is an extension of λ to a linear functional on V that satisfies

$$|\lambda(w)| \le N_A(w) \tag{11.3}$$

for every $w \in V$. In particular,

$$|\lambda(w)| < 1 \tag{11.4}$$

for each $w \in A$, $\lambda(v) = N_A(v) \ge 1$, and this extension is a bounded linear functional on V, because $N_A(w)$ is bounded by a constant multiple of $\|w\|$.

Suppose now that $E \subseteq V$ is nonempty, balanced, closed, and convex, and that $v \in V\backslash E$. Because E is closed and $v \notin E$,

$$\|v - y\| \ge r \tag{11.5}$$

for some $r > 0$ and every $y \in E$. Put

$$A = E + B(0,r) = \{y + z : y \in E,\ z \in V,\ \text{and } \|z\| < r\}. \tag{11.6}$$

Thus $v \in V \backslash A$, by (11.5), and one can check that A is a balanced convex open set in V, because of the corresponding porperties of E and $B(0, r)$.

As before, there is a bounded linear functional λ on V such that $|\lambda(v)| \geq 1$ and (11.4) holds for every $w \in A$. This implies that

$$|\lambda(y) + \lambda(z)| \leq 1 \tag{11.7}$$

for every $y \in E$ and $z \in V$ with $\|z\| < r$, by applying (11.4) to $w = y + z \in A$. Using this, one can check that

$$|\lambda(y)| \leq 1 - r \tag{11.8}$$

for every $y \in E$, by considering z's which are suitable scalar multiples of v.

Part II
Dual linear mappings

12 Bounded linear mappings

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, respectively. A linear mapping T from V into W is said to be *bounded* if there is a nonnegative real number A such that

$$\|T(v)\|_W \leq A\,\|v\|_V \tag{12.1}$$

for every $v \in V$. In this case,

$$\|T(u) - T(v)\|_W = \|T(u - v)\|_W \leq A\,\|u - v\|_V \tag{12.2}$$

for every $u, v \in V$, which implies that T is uniformly continuous with respect to the metrics corresponding to the norms on V and W. Conversely, if T is continuous at 0, then there is a $\delta > 0$ such that $\|T(v)\|_W < 1$ for every $v \in V$ with $\|v\|_V < \delta$, and one can check that (12.1) holds with $A = 1/\delta$. Note that bounded linear functionals on V are the same as bounded linear mappings from V into $\mathbf{R}$ or $\mathbf{C}$, respectively. If $V = \mathbf{R}^n$ or $\mathbf{C}^n$ with the standard norm for some positive integer n, then it is easy to see that any linear mapping T from V into any W is bounded. This also works when $V = \mathbf{R}^n$ or $\mathbf{C}^n$ is equipped with any norm, because of the equivalence of any norm with the standard norm, as in Section 1. Similarly, this works as well for any finite-dimensional vector space V, since V is isomorphic to $\mathbf{R}^n$ or $\mathbf{C}^n$ for some n.

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

$$\|T\|_{op} = \sup\{\|T(v)\|_W : v \in V,\ \|v\|_V \leq 1\}. \tag{12.3}$$

This is known as the operator norm of T, and it reduces to the dual norm of a bounded linear functional when $W = \mathbf{R}$ or $\mathbf{C}$, as approriate, with the standard norm. Equivalently, (12.1) holds with $A = \|T\|_{op}$, and $\|T\|_{op}$ is the smallest nonnegative real number with this property. If T' is another bounded linear mapping from V into W, then it is easy to see that $T + T'$ is bounded as well, and that

$$\|T + T'\|_{op} \le \|T\|_{op} + \|T'\|_{op}. \tag{12.4}$$

Similarly, if a is a real or complex number, as appropriate, then $a\,T$ is also a bounded linear mapping from V into W, and

$$\|a\,T\|_{op} = |a|\,\|T\|_{op}. \tag{12.5}$$

Thus 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 the operator norm defines a norm on $\mathcal{BL}(V, W)$. If W is complete, then one can show that $\mathcal{BL}(V, W)$ is complete with respect to the operator norm, in essentially the same way as for the dual space in Section 3.

Suppose that V_1, V_2, and V_3 are vector spaces, all real or all complex, and equipped with norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_3$, respectively. Let T_1 be a bounded linear mapping from V_1 into V_2, and let T_2 be a bounded linear mapping from V_2 into V_3. The composition $T_2 \circ T_1$ of T_1 and T_2 is the linear mapping from V_1 into V_3 defined by

$$(T_2 \circ T_1)(v) = T_2(T_1(v)) \tag{12.6}$$

for every $v \in V_1$. It is easy to see that $T_2 \circ T_1$ is also a bounded linear mapping from V_1 into V_3 under these conditions, and that

$$\|T_2 \circ T_1\|_{op,13} \le \|T_1\|_{op,12}\,\|T_2\|_{op,23}, \tag{12.7}$$

where the subscripts indicate the spaces and norms used in the corresponding operator norm.

13 Dual mappings

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, respectively. If T is a bounded linear mapping from V into W and λ is a bounded linear functional on W, then

$$T^*(\lambda) = \lambda \circ T \tag{13.1}$$

is a bounded linear functional on V, with

$$\|T^*(\lambda)\|_{V^*} \le \|T\|_{op}\,\|\lambda\|_{W^*}. \tag{13.2}$$

Here $\|\cdot\|_{V^*}$, $\|\cdot\|_{W^*}$ are the dual norms on the dual spaces V^*, W^* associated to the given norms on V, W, respectively, and $\|T\|_{op}$ is the corresponding operator norm of T, as in the previous section. It is easy to see that T^* defines a linear mapping from W^* into V^* which is bounded, with

$$\|T^*\|_{op,*} \le \|T\|_{op} \tag{13.3}$$

by (13.2), where $\|T^*\|_{op,*}$ is the operator norm of T^* with respect to the dual norms $\|\cdot\|_{W^*}$, $\|\cdot\|_{V^*}$.

Let us check that

$$\|T^*\|_{op,*} = \|T\|_{op}. \tag{13.4}$$

If v is any vector in V and λ is any bounded linear functional on W, then

$$|\lambda(T(v))| = |(T^*(\lambda))(v)| \le \|T^*(\lambda)\|_{V^*} \|v\|_V \le \|T^*\|_{op,*} \|\lambda\|_{W^*} \|v\|_V. \tag{13.5}$$

Using the Hahn–Banach theorem, there exist $\lambda \in W^*$ which give

$$\|T(v)\|_W \le \|T^*\|_{op,*} \|v\|_V. \tag{13.6}$$

This implies that $\|T\|_{op} \le \|T^*\|_{op,*}$, as desired.

Observe that

$$T \mapsto T^* \tag{13.7}$$

defines a linear mapping from $\mathcal{BL}(V, W)$ into $\mathcal{BL}(W, V)$. Suppose now that V_1, V_2, and V_3 are vector spaces, all real or all complex, and equipped with norms. If T_1 is a bounded linear mapping from V_1 into V_2 and T_2 is a bounded linear mapping from V_2 into V_3, then it is easy to see that

$$(T_2 \circ T_1)^* = T_1^* \circ T_2^* \tag{13.8}$$

as linear mappings from V_3^* into V_1^*. Indeed,

$$(T_2 \circ T_1)^*(\lambda) = \lambda \circ T_2 \circ T_1 = (T_2^*(\lambda)) \circ T_1 = T_1^*(T_2^*(\lambda)) \tag{13.9}$$

for every $\lambda \in V_3^*$.

Let V be a real or complex vector space equipped with a norm, and let W be a linear subspace of V. Let T be the natural inclusion mapping of W into V, which sends $w \in W$ to itself as an element of V. In this case, the corresponding dual mapping T^* sends a bounded linear functional λ on V to its restriction to W, as in Section 8. Similarly, if W is a closed linear subspace of V, then the natural quotient mapping q from V onto V/W is a bounded linear mapping with respect to the quotient norm on V/W. The dual q^* of q is an isometric embedding of $(V/W)^*$ into V^*, as in Section 7.

14 Invertibility

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, respectively. A bounded linear mapping T from V into W is said to be *invertible* if T is a one-to-one mapping from V onto W for which the corresponding inverse mapping $T^{-1} : W \to V$ is bounded. In this case, it is easy to see that the corresponding dual operator $T^* : W^* \to V^*$ is invertible, and that

$$(T^*)^{-1} = (T^{-1})^*. \tag{14.1}$$

More precisely, if I_V and I_W are the identity mappings on V and W, respectively, then their duals mappings are the same as the identity mappings I_{V^*} and I_{W^*} on V^* and W^*, respectively. Thus

$$T^{-1} \circ T = I_V, \quad T \circ T^{-1} = I_W \tag{14.2}$$

imply that

$$T^* \circ (T^{-1})^* = (T^{-1} \circ T)^* = I_{V^*}, \quad (T^{-1})^* \circ T^* = (T \circ T^{-1})^* = I_{W^*}, \tag{14.3}$$

as desired.

Conversely, suppose that T^* is invertible, and let us show that this implies that T is invertible when V is complete. By the argument in the previous paragraph, the dual T^{**} of T^* is invertible as a bounded linear mapping from the second dual V^{**} of V onto the second dual W^{**} of W. If V and W are reflexive, then it is easy to see that T^{**} corresponds exactly to T under the natural isomorphisms between V and V^{**} and W and W^{**}, and hence that T is invertible. Otherwise, T corresponds to the restriction of T^{**} to the image of the natural embedding of V into V^{**}, which takes values in the image of the natural embedding of W in W^{**}. This implies that

$$\|T(v)\|_W \ge c\,\|v\|_V \tag{14.4}$$

for some $c > 0$ and every $v \in V$, because of the analogous condition for T^{**} that follows from invertibility.

Let $\mathcal{N}(T^*)$ be the nullspace or kernel of T^*, which is the linear subspace of W^* consisting of the bounded linear functionals λ on W such that $T^*(\lambda) = 0$. Also let $T(V)$ be the image of T in W, which is the linear subspace of W consisting of vectors of the form $T(v)$ with $v \in V$. It is easy to see that

$$\mathcal{N}(T^*) = T(V)^\perp \tag{14.5}$$

for any bounded linear mapping $T : V \to W$, where $T(V)^\perp$ is as in (7.7). More precisely, $T^*(\lambda) = 0$ if and only if

$$(T^*(\lambda))(v) = \lambda(T(v)) = 0 \tag{14.6}$$

for every $v \in V$, which is the same as saying that $\lambda \in T(V)^\perp$. In particular, if T^* is one-to-one, then $\mathcal{N}(T^*) = \{0\}$, which implies that $T(V)$ is dense in W. Otherwise, if $T(V)$ is not dense in W, then there is a nonzero bounded linear functional on W equal to 0 on $T(V)$ as in Section 4, contradicting the fact that $T(V)^\perp = \mathcal{N}(T^*) = \{0\}$. Conversely, if $T(V)$ is dense in W, then $T(V)^\perp = \{0\}$, which implies that $\mathcal{N}(T^*) = \{0\}$ and hence that T^* is one-to-one.

A bounded linear mapping $T : V \to W$ satisfies (14.4) for some $c > 0$ if and only if T is invertible as a mapping from V onto $T(V)$. If V is complete, then it follows that $T(V)$ is complete as well. This implies that $T(V)$ is a closed linear subspace of W, as in Section 2. If T^* is invertible, then $\mathcal{N}(T^*) = \{0\}$, so that $T(V)$ is dense in W, as in the previous paragraph. Thus we get that $T(V) = W$ under these conditions, because $T(V)$ is both dense and closed in W. This shows that $T : V \to W$ is invertible when $T^* : W^* \to V^*$ is invertible and V is complete, as desired.

15 Open mappings

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, respectively. A bounded linear mapping T from V into W is said to be an open mapping if $T(U)$ is an open set in W for every open set U in V. In particular, if U is the open unit ball in V, then it follows that

$$B_W(0,c) \subseteq T(B_V(0,1)) \tag{15.1}$$

for some $c > 0$, where $B_V(v,r)$ and $B_W(w,r)$ are the open balls in V and W centered at v and w and with radius $r > 0$, respectively. Conversely, (15.1) implies that

$$B_W(T(v), c\,r) \subseteq T(B_V(v,r)) \tag{15.2}$$

for every $v \in V$ and $r > 0$, and hence that T is an open mapping. Note that open linear mappings are surjective, and that a bounded linear mapping $T : V \to W$ is invertible if and only if it is one-to-one and open.

Let T be a bounded linear mapping from V into W again, and let Z be a closed linear subspace of V that is contained in the kernel of T. Thus

$$\|T(v)\|_W = \|T(v-z)\|_W \le \|T\|_{op}\, \|v-z\|_V \tag{15.3}$$

for every $z \in Z$, which implies that

$$\|T(v)\|_W \le \|T\|_{op}\, N_Z(v) \tag{15.4}$$

for every $v \in V$, where $N_Z(v)$ is as in (4.5). If q is the natural quotient mapping from V onto V/Z, then there is a linear mapping $\widehat{T}$ from V/Z into W such that $\widehat{T} \circ q = T$, as in Section 5. The estimate (15.4) implies that $\widehat{T}$ is a bounded linear mapping from V/Z into W with respect to the quotient norm on V/Z, with operator norm less than or equal to the operator norm of T. Hence the operator norms of T and $\widehat{T}$ are equal, because the operator norm of T is automatically less than or equal to the operator norm of $\widehat{T}$.

Under these conditions, it is easy to see that T is an open mapping from V onto W if and only if $\widehat{T}$ is an open mapping from V/Z onto W. Observe also that the kernel of T is always a closed linear subspace of V when T is bounded, and that $\widehat{T}$ is one-to-one when Z is the kernel of T. In particular, it follows that $\widehat{T}$ is invertible when T is an open mapping and Z is the kernel of T.

16 A criterion for openness

Let V and W be vector spaces, both real or both complex, and with norms $\|v\|_V$ and $\|w\|_W$, respectively. Let T be a bounded linear mapping from V into W, and suppose that

$$B_W(0,c) \subseteq \overline{T(B_V(0,1))} \tag{16.1}$$

for some $c > 0$. Here $\overline{T(B_V(0,1))}$ is the closure of $T(B_V(0,1))$ in W, so that (16.1) automatically implies that

$$\overline{B}_W(0,c) \subseteq \overline{T(B_V(0,1))}, \tag{16.2}$$

where $\overline{B}_W(w,r)$ denotes the closed ball in W with center w and radius r. If V is complete, then this implies that

$$B_W(0,c) \subseteq T(B_V(0,1)), \tag{16.3}$$

and hence that T is an open mapping of V onto W. To show this, it suffices to consider the case where $c = 1$, since otherwise we can replace T with $c^{-1}\,T$.

If $c = 1$, then (16.2) implies that for each $w \in W$ with $\|w\|_W \le 1$ and every $\epsilon > 0$ there is a $v \in V$ such that $\|v\|_V \le 1$ and $\|w - T(v)\|_W < \epsilon$, by (16.2). Equivalently, for each $w \in W$ and $\epsilon > 0$ there is a $v \in V$ such that $\|v\|_V \le \|w\|_W$ and $\|w - T(v)\|_W < \epsilon$. Let $w \in W$ with $\|w\|_W < 1$ be given, and let $\{\epsilon_j\}_{j=1}^\infty$ be an infinite sequence of positive real numbers such that

$$\sum_{j=1}^\infty \epsilon_j < 1 - \|w\|_W. \tag{16.4}$$

Let v_1 be an element of V such that $\|v_1\|_V \le \|w\|_W$ and $\|w - T(v_1)\|_W < \epsilon_1$. Similarly, if $v_1, \ldots, v_n \in V$ have already been chosen for some positive integer n, then let v_{n+1} be an element of V such that

$$\|v_{n+1}\|_V \le \Big\| w - \sum_{j=1}^n T(v_j) \Big\|_W \tag{16.5}$$

and

$$\Big\| w - \sum_{j=1}^n T(v_j) - T(v_{n+1}) \Big\|_W < \epsilon_{n+1}. \tag{16.6}$$

Thus $\|v_{n+1}\|_V < \epsilon_n$ for each $n \ge 1$, and

$$\sum_{j=1}^\infty \|v_j\|_V < \|v_1\|_V + \sum_{j=1}^\infty \epsilon_j < \|w\|_W + (1 - \|w\|_W) = 1. \tag{16.7}$$

This implies that $\sum_{j=1}^\infty v_j$ converges absolutely, and hence converges in V, since V is complete. Put $v = \sum_{j=1}^\infty v_j$, so that $\|v\|_V < 1$, by (16.7). Of course,

$$T(v) = \sum_{j=1}^\infty T(v_j), \tag{16.8}$$

because T is bounded and hence continuous. It follows that

$$w = \lim_{n \to \infty} \sum_{j=1}^n T(v_j) = T(v), \tag{16.9}$$

as desired, using (16.6) and the fact that $\epsilon_n \to 0$ as $n \to \infty$ in the first step.

17 The open mapping theorem

Let V and W be vector spaces again, both real or both complex, and with norms $\|v\|_V$ and $\|w\|_W$, respectively. Also let T be a bounded linear mapping from V *onto* W. If V and W are complete, then Banach's open mapping theorem implies that T is an open mapping.

To see this, observe that

$$\bigcup_{n=1}^{\infty} T(B_V(0,n)) = T(V) = W, \tag{17.1}$$

and hence

$$\bigcup_{n=1}^{\infty} \overline{T(B_V(0,n))} = W. \tag{17.2}$$

Because W is complete, the Baire category theorem implies that $\overline{T(B_V(0,n))}$ has nonempty interior in W for some positive integer n. Of course, this implies that $\overline{T(B_V(0,r))}$ has nonempty interior for every $r > 0$, by linearity.

It is easy to see that

$$\overline{T(B_V(0,1/2))} - \overline{T(B_V(0,1/2))} \subseteq \overline{T(B_V(0,1))}, \tag{17.3}$$

where the left side is the set of differences of any two elements of $\overline{T(B_V(0,1/2))}$. If $\overline{T(B_V(0,1/2))}$ has an interior point in W, as in the previous paragraph, then it follows that 0 is in the interior of $\overline{T(B_V(0,1))}$. Equivalently, this means that (16.1) holds for some $c > 0$. If V is complete, then it follows that T is an open mapping, as in the previous section.

18 Openness and duality

Let V and W be vector spaces which are both real or both complex, as usual, and let $\|v\|_V$ and $\|w\|_W$ be norms on V and W, respectively. Also let $\|\cdot\|_{V^*}$ and $\|\cdot\|_{W^*}$ be the corresponding dual norms on V^* and W^*, respectively, and let T be a bounded linear mapping from V into W.

Suppose that T satisfies (16.1) and hence (16.2) for some $c > 0$. If λ is any bounded linear functional on W, then we get that

$$\sup\{|\lambda(w)| : w \in W,\ \|w\|_W \le c\} \le \sup\{|\lambda(T(v))| : v \in V,\ \|v\| \le 1\}. \tag{18.1}$$

This implies that

$$c\,\|\lambda\|_{W^*} \le \|T^*(\lambda)\|_{V^*}. \tag{18.2}$$

In particular, if T is an open mapping from V onto W, then this holds for some $c > 0$ and every $\lambda \in W^*$.

Conversely, suppose that (18.2) holds for some $c > 0$ and every $\lambda \in W^*$, and let us show that T satisfies (16.2). Put $E = \overline{T(B_V(0,1))}$, and let $w \in W \backslash E$ be given. Thus E is a closed set in W by construction, and it is easy to see that E is also balanced and convex, because $B_V(0,1)$ and hence $T(B_V(0,1))$ are balanced and convex. Under these conditions, there is a bounded linear functional λ on W such that

$$\sup_{x \in E} |\lambda(x)| < |\lambda(w)|, \tag{18.3}$$

as in Section 11. In this case,

$$\sup_{x \in E} |\lambda(x)| = \|T^*(\lambda)\|_{V^*} \ge c\,\|\lambda\|_{W^*}, \tag{18.4}$$

so that $c\,\|\lambda\|_{W^*} < |\lambda(w)| \le \|\lambda\|_{W^*}\,\|w\|_W$. Thus we get that $\|w\|_{W^*} > c$ for every $w \in W \backslash E$, which implies that $\overline{B}_W(0,c) \subseteq E$, as desired. If V is complete, then it follows that (16.3) holds, as in Section 16, and hence that T is an open mapping from V onto W.

There are analogous statements with the roles of T and T^* reversed. More precisely, suppose first that

$$c\,\|v\|_V \le \|T(v)\|_W \tag{18.5}$$

for some $c > 0$ and every $v \in V$, so that T is invertible as a mapping of V onto $T(V)$. This case is basically the same as an inclusion mapping, and one can check that T^* is an open mapping from W^* onto V^*, using the Hahn–Banach theorem. Conversely, suppose that T^* is an open mapping from W^* onto V^*, so that

$$B_{V^*}(0,c) \subseteq T^*(B_{W^*}(0,1)) \tag{18.6}$$

for some $c > 0$. Of course, V^* and W^* are automatically complete, and hence T^* is automatically an open mapping when $T^*(W^*) = V^*$, by the open mapping theorem. One could also start with a condition analogous to (16.1), as before. At any rate, one can use the same type of argument as in (18.1) and (18.2), to get that

$$c\,\|L\|_{V^{**}} \le \|T^{**}(L)\|_{W^{**}} \tag{18.7}$$

for every $L \in V^{**}$. This implies that (18.5) holds for every $v \in V$, by taking $L = L_v$ in the image of the natural embedding of V into V^{**} in (18.7).

19 Closed range

Let V be a real or complex vector space with a norm $\|v\|_V$. If Y is a linear subspace of V, then we let $Y^\perp$ be the linear subspace of V^* given by

$$Y^\perp = \{\lambda \in V^* : \lambda(y) = 0 \text{ for every } y \in Y\}, \tag{19.1}$$

as in (7.7). Similarly, if Z is a linear subspace of V^*, then we put

$${}^\perp Z = \{v \in V : \lambda(v) = 0 \text{ for every } \lambda \in Z\}. \tag{19.2}$$

This is a closed linear subspace of V which can be identified with $Z^\perp$ in V^{**} when V is reflexive, and otherwise ${}^\perp Z$ consists of the elements of V that correspond to elements of $Z^\perp$ under the standard embedding of V in V^{**}. Observe that

$$Y \subseteq {}^\perp(Y^\perp) \tag{19.3}$$

and

$$Z \subseteq ({}^\perp Z)^\perp, \tag{19.4}$$

by construction. Using the Hahn–Banach theorem, one can show that ${}^\perp(Y^\perp)$ is the same as the closure of Y in V with respect to the norm $\|v\|_V$, and in particular that $Y = {}^\perp(Y^\perp)$ when Y is a closed linear subspace of V with respect to the norm. Although $Y^\perp$ is a closed linear subspace of V^* with respect to the dual norm, a more precise statement is that $Y^\perp$ is a closed set with respect to the weak* topology on V^*. One can show that $({}^\perp Z)^\perp$ is the same as the closure of Z with respect to the weak* topology on V^*, so that $Z = ({}^\perp Z)^\perp$ when Z is already closed with respect to the weak* topology on V^*.

Let W be another vector space which is real or complex depending on whether V is real or complex, and equipped with a norm $\|w\|_W$. Let T be a bounded linear mapping from V into W, and remember that $\mathcal{N}(T^*) = T(V)^\perp$, as in (14.5). We also have that

$$\mathcal{N}(T) = {}^\perp T^*(W^*). \tag{19.5}$$

Indeed, $T(v) = 0$ if and only if $\lambda(T(v)) = 0$ for every $\lambda \in W^*$, because of the Hahn–Banach theorem. This is the same as saying that $(T^*(\lambda))(v) = 0$ for every $\lambda \in W^*$, which is equivalent to $v \in {}^\perp T^*(W^*)$, as desired.

If we take $Z = T^*(W^*)$ in (19.4), then we get that

$$T^*(W^*) \subseteq \mathcal{N}(T)^\perp \tag{19.6}$$

when T is any bounded linear mapping from V into W. Let us show that equality holds when V and W are complete and $T(V)$ is a closed linear subspace of W, which implies that $T(V)$ is also complete with respect to the restriction of the norm $\|w\|_W$ to $T(V)$. If $\mathcal{N}(T) = \{0\}$, then the open mapping theorem implies that (18.5) holds for some $c > 0$ and every $v \in V$, and the Hahn–Banach theorem implies that $T^*(W^*) = V^*$, as in the previous section. Otherwise, we can express T as the composition of the natural quotient mapping q from V onto $V/\mathcal{N}(T)$ with a bounded linear mapping $\widehat{T}$ from $V/\mathcal{N}(T)$ into W. Note that $\mathcal{N}(T)$ is automatically a closed linear subspace of V, because T is bounded and hence continuous, and that $V/\mathcal{N}(T)$ is complete. By construction, $\widehat{T} : V/\mathcal{N}(T) \to W$ is one-to-one and

$$\widehat{T}(V/\mathcal{N}(T)) = T(V) \tag{19.7}$$

is a closed linear subspace of W, so that

$$\widehat{T}^*(W^*) = (V/\mathcal{N}(T))^*, \tag{19.8}$$

as in the previous case. We also know that q^* maps $(V/\mathcal{N}(T))^*$ onto $\mathcal{N}(T)^\perp$, as in Section 7. It follows that

$$T^*(W^*) = \mathcal{N}(T)^\perp \tag{19.9}$$

under these conditions, because $T = \widehat{T} \circ q$ and hence $T^* = q^* \circ \widehat{T}^*$.

Let T be a bounded linear mapping from V into W again, and let Y be the closure of $T(V)$ in W. Thus

$$Y = {}^\perp\mathcal{N}(T^*), \tag{19.10}$$

because $Y^\perp = \mathcal{N}(T^*)$ by (14.5), and $Y = {}^\perp(Y^\perp)$ since Y is closed. If $T^*(W^*)$ is a closed linear subspace of V^* and V is complete, then $T(V)$ is a closed linear subspace of W, which is to say that $T(V) = Y$. To see this, suppose first that $\mathcal{N}(T^*) = \{0\}$, so that $Y = W$. In this case, the open mapping theorem implies that T^* satisfies (18.2) for some $c > 0$ and every $\lambda \in W^*$, because V^* and W^* are automatically complete, and hence $T^*(W^*)$ is complete since it is a closed linear subspace of V^*. It follows that T is an open mapping from V onto W when V is complete, as in the previous section. Otherwise, let T_1 be the same as T, but considered as a linear mapping from V into Y, and let T_2 be the obvious inclusion mapping from Y into W. Thus $T = T_2 \circ T_1$, which implies that

$$T^* = T_1^* \circ T_2^*. \tag{19.11}$$

As usual, $T_2^*(W^*) = Y^*$, because of the Hahn–Banach theorem, and hence

$$T_1^*(Y^*) = T_1^*(T_2^*(W^*)) = T^*(W^*). \tag{19.12}$$

In particular, $T_1^*(Y^*)$ is a closed linear subspace of V^*, because $T^*(W^*)$ is closed by hypothesis. By construction, $T_1(V) = T(V)$ is dense in Y, so that

$$\mathcal{N}(T_1^*) = T_1(V)^\perp = \{0\} \tag{19.13}$$

as subspaces of Y^*. If Y is complete, then the previous argument implies that $T(V) = T_1(V) = Y$, as desired.

20 Compact linear mappings

Let $(M, d(x, y))$ be a metric space. Remember that a subset E of M is said to be *totally bounded* if for each $\epsilon > 0$, E is contained in the union of finitely many balls of radius ϵ in M. It is easy to see that compact sets are totally bounded, and it is well known that a subset of a complete metric space if compact if and only if it is closed and totally bounded. Bounded subsets of $\mathbf{R}^n$ and $\mathbf{C}^n$ are totally bounded, and this also works for finite-dimensional real or complex vector spaces with any norm, by reducing to $\mathbf{R}^n$ or $\mathbf{C}^n$ with the standard norm in the usual way.

Let V and W be vector spaces, both real or both complex, and with norms $\|v\|_V$ and $\|w\|_W$, respectively. A linear mapping T from V into W is said to be *compact* if $T(B_V(0,1))$ is totally bounded in W, where $B_V(0,1)$ is the open unit ball in W. Thus compact linear mappings are bounded in particular, since totally bounded sets are bounded. If W is complete, then T is compact if and only if the closure $\overline{T(B_V(0,1))}$ of $T(B_V(0,1))$ in W is compact. If a bounded linear mapping $T : V \to W$ has finite rank, in the sense that $T(V)$ is a finite-dimensional linear subspace of W, then T is automatically compact, because bounded subsets of finite-dimensional spaces are totally bounded.

Let $\mathcal{CL}(V,W)$ be the space of compact linear mappings from V into W. It is easy to see that this is a linear subspace of the space $\mathcal{B}(V,W)$ of bounded linear mappings from V into W. More precisely, this uses the fact that if E_1 and E_2 are totally bounded subsets of W and $t \in \mathbf{R}$ or $\mathbf{C}$, as appropriate, then

$$t\,E_1 = \{t\,w : w \in E_1\} \tag{20.1}$$

and

$$E_1 + E_2 = \{w + z : w \in E_1, z \in E_2\} \tag{20.2}$$

are also totally bounded subsets of W. One can also check that $\mathcal{CL}(V,W)$ is a closed set in $\mathcal{BL}(V,W)$ with respect to the operator norm. Equivalently, if $\{T_j\}_{j=1}^\infty$ is a sequence of compact linear mappings from V into W that converges to a bounded linear mapping $T : V \to W$ with respect to the operator norm, then T is also compact.

In particular, if $\{T_j\}_{j=1}^\infty$ is a sequence of bounded linear mappings from V into W that converges to a bounded linear mapping $T : V \to W$ with respect to the operator norm, and if T_j has finite rank for each j, then T is compact. In many cases, every compact linear mapping from V into W can be approximated by bounded linear mappings with finite rank with respect to the operator norm in this way. This is especially simple in the context of Hilbert spaces, using orthogonal projections.

Let V_1, V_2, and V_3 be vector spaces, all real or all complex, and equipped with norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_3$, respectively. If $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$ are bounded linear mappings, and if either T_1 or T_2 is compact, then it is easy to see that the composition $T_2 \circ T_1$ is also compact. Similarly, if T_1 or T_2 can be approximated by bounded linear mappings with finite rank with respect to the operator norm, then $T_2 \circ T_1$ has the same property. Of course, if T_1 or T_2 has finite rank, then $T_2 \circ T_1$ does too.

Let V be a real or complex vector space with a norm $\|v\|$. If the open unit ball in V is totally bounded, then it is well known that V is finite-dimensional. Indeed, under these conditions, there are finitely many elements $y_1, \ldots, y_n$ of V such that $\|y_j\| < 1$ for $j = 1, \ldots, n$, and for each $v \in V$ with $\|v\| < 1$ there is a j such that $\|v - y_j\| < 1/2$. Let Y be the linear subspace of V spanned by the y_j's, so that Y is a finite-dimensional linear subspace of V with dimension less than or equal to n. If $v \in V$ and $\|v\| < r$ for some $r > 0$, then there is a $y \in Y$ such that $\|v - y\| < r/2$, by applying the previous approximation condition to v/r. This implies that for each positive integer l there is a $y \in Y$ such that $\|v - y\| < 2^{-l}\,r$, by repeating the process. Thus Y is dense in V, and we have seen that Y is also closed in V, because Y is a finite-dimensional linear subspace of V. Hence $V = Y$, so that V is finite-dimensional, as desired.

In particular, if $T : V \to W$ is compact and invertible, then W is finite-dimensional, which implies that V is finite-dimensional as well. Similarly, if $T : V \to W$ is compact and the restriction of T to a linear subspace Z of V is an invertible linear mapping onto $T(Z)$, then Z is finite-dimensional.

21 Compactness and duality

Let $(M, d(x, y))$ be a metric space, and let $\mathcal{E}$ be a collection of real or complex-valued functions on M which is equicontinuous in the sense that for each $\epsilon > 0$ there is a $\delta(\epsilon) > 0$ such that

$$|f(x) - f(y)| < \epsilon \tag{21.1}$$

for every $x, y \in M$ with $d(x, y) < \delta(\epsilon)$ and every $f \in \mathcal{E}$. If M is totally bounded, and if the elements of $\mathcal{E}$ are uniformly bounded, then it is well known that $\mathcal{E}$ is also totally bounded as a subset of the space of bounded continuous real or complex-valued functions on M with respect to the supremum norm. To see this, let $\epsilon > 0$ be given, and let A be a finite subset of M such that for every $x \in M$ there is a $y \in A$ that satisfies $d(x, y) < \delta(\epsilon/3)$. If $f, g \in \mathcal{E}$ satisfy

$$|f(y) - g(y)| < \epsilon/3 \tag{21.2}$$

for every $y \in A$, then one can check that

$$|f(x) - g(x)| < \epsilon \tag{21.3}$$

for every $x \in M$. This permits the total boundedness problem on M to be reduced to one on finite subsets of M, which can then be handled using the hypothesis that the functions in $\mathcal{E}$ be uniformly bounded.

Now let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, respectively. If T is a compact linear mapping from V into W, then the dual mapping T^* is compact as a linear mapping from W^* into V^*. To see this, let M be the image of the open unit ball in V under T, equipped with the restriction of the metric associated to $\|w\|_W$ on W. Also let $\mathcal{E}$ be the collection of functions on M which are the restrictions of bounded linear functionals on W with dual norm less than 1 to M. Thus M is totally bounded by hypothesis, and the equicontinuity of the elements of $\mathcal{E}$ follows from linearity and the boundedness of the dual norm. The elements of $\mathcal{E}$ are also uniformly bounded on M, because M is bounded in W. As in the previous paragraph, $\mathcal{E}$ is totally bounded with respect to the supremum norm on M. Using this, one can check that the image of the open unit ball in W^* under T^* is totally bounded in V^*, as desired.

Similarly, if T^* is compact, then T^{**} is compact as a linear mapping from V^{**} into W^{**}. This implies that T is compact, since T can be identified with the restriction of T^{**} to the image of the natural embedding of V in V^{**}.

Note that the dual of a bounded linear mapping with finite rank also has finite rank, basically because the dual of a finite-dimensional vector space has the same finite dimension. If $T : V \to W$ can be approximated by finite-rank operations with respect to the operator norm, then it follows that T^* has the same property. The converse can be obtained by considering T^{**} as before.

22 Alternate characterizations

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, as usual. Suppose that T is a compact linear mapping from V into W, and let $\epsilon > 0$ be given. By hypothesis, there are finitely many elements $z_1, \ldots, z_n$ of W such that

$$T(B_V(0,1)) \subseteq \bigcup_{j=1}^{n} B_W(z_j, \epsilon), \tag{22.1}$$

where $B_V(v,r)$ and $B_W(w,t)$ are the usual open balls in V and W, respectively. Let Z be the linear subspace of W spanned by $z_1, \ldots, z_n$, so that Z has finite dimension less than or equal to n. Under these conditions, the composition of T with the natural quotient mapping from W onto W/Z has operator norm less than or equal to ϵ with respect to the quotient norm on W/Z associated to the given norm on W.

Conversely, suppose that T is a bounded linear mapping from V into W, and that for each $\epsilon > 0$ there is a finite-dimensional linear subspace Z of W such that the composition of T with the natural quotient mapping from W onto W/Z has operator norm less than or equal to ϵ. This implies that $T(B_V(0,1))$ lies in an ϵ-neighborhood of Z in W. Because T is bounded, $T(B_V(0,1))$ is a bounded subset of W, and hence $T(B_V(0,1))$ lies within an ϵ-neighborhhood of a bounded subset of Z. Because Z is finite-dimensional, bounded subsets of Z are totally bounded, and it follows that $T(B_V(0,1))$ is contained in the union of finitely many balls of radius 2ϵ in W. Thus T is compact under these conditions.

Suppose now that T is a bounded linear mapping from V into W, and that for each $\epsilon > 0$ there is a closed linear subspace Y of V with finite codimension such that the restriction of T to Y has operator norm less than or equal to ϵ. More precisely, to say that Y has finite codimension in V means that the quotient space V/Y has finite dimension. Also, the condition that Y be closed in V is not a real restriction, since otherwise one can replace Y with its closure in V. Note that $(V/Y)^*$ has finite dimension equal to the dimension of V/Y in this situation, and remember that there is a natural isometric isomorphism between $(V/Y)^*$ and $Y^\perp \subseteq V^*$, as in Section 7. Let A be the natural inclusion mapping of Y into V, which sends every element of Y to itself as an element of V. The hypothesis that the operator norm of the restriction of T to Y be less than or equal to ϵ is equivalent to saying that $T \circ A$ has operator norm less than or equal to ϵ as a bounded linear mapping from Y into W. This implies that $(T \circ A)^*$ has operator norm less than or equal to ϵ as a bounded linear mapping from W^* into Y^*, as in Section 13. Of course, $(T \circ A)^* = A^* \circ T^*$, and we have seen that A^* corresponds exactly to the quotient mapping from V^* onto $Y^* \cong V^*/Y^\perp$. This shows that T^* satisfies the criterion for compactness discussed in the preceding paragraph, since $Y^\perp \cong (V/Y)^*$ has finite dimension. It follows that T is compact, as in the previous section.

Conversely, suppose that T is compact, and let $\epsilon > 0$ be given. Because T^* is compact as a linear mapping from W^* into V^*, there is a finite-dimensional linear subspace Z of V^* such that the composition of T^* with the natural quotient mapping from V^* onto V^*/Z has operator norm less than or equal to ϵ. Put $Y = {}^\perp Z$, so that Y is a closed linear subspace of V. By linear algebra, Y has finite codimension in V equal to the dimension of Z, and there is a natural isomorphism between $(V/Y)^*$ and Z. In particular, $Y^\perp = Z$ in this case, since $Z \subseteq Y^\perp$ automatically, and Z and $Y^\perp \cong (V/W)^*$ have the same finite dimension. Let A be the natural inclusion mapping of Y in V, as before. Thus $(T \circ A)^* = A^* \circ T^*$, where A^* corresponds exactly to the quotient mapping from V^* onto $Y^* \cong V^*/Y^\perp = V^*/Z$. By construction, $A^* \circ T^*$ has operator norm less than or equal to ϵ, which implies that $T \circ A$ has operator norm less than or equal to ϵ. This is the same as saying that the restriction of T to Y has operator norm less than or equal to ϵ, as desired.

23 Compactness and closed range

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$, $\|w\|_W$, respectively. Suppose that A is a bounded linear mapping from V into W with the property that

$$\|A(v)\|_W \ge c\,\|v\|_V \tag{23.1}$$

for some $c > 0$ and every $v \in V$. Also let T be a compact linear mapping from V into W. As in the previous section, there is a closed linear subspace Y with finite codimension in V such that

$$\|T(v)\|_W \le (c/2)\,\|v\|_V \tag{23.2}$$

for every $v \in Y$. Thus

$$\begin{aligned} c\,\|v\|_V \le \|A(v)\|_W &\le \|(A+T)(v)\|_W + \|T(v)\|_W \\ &\le \|(A+T)(v)\|_W + (c/2)\,\|v\|_V \end{aligned} \tag{23.3}$$

for every $v \in Y$, which implies that

$$(c/2)\,\|v\|_V \le \|(A+T)(v)\|_W \tag{23.4}$$

for every $v \in Y$.

If V is complete, then we have seen before that (23.1) implies that $A(V)$ is complete, and hence that $A(V)$ is a closed linear subspace of W. Similarly, (23.4) implies that $A+T$ maps Y onto a closed linear subspace of W when V is complete. Because Y has finite codimension in V, there is a finite-dimensional linear subspace X of V such that the span of X and Y is equal to V. This implies that $(A+T)(X)$ is a finite-dimensional linear subspace of W, and that $(A+T)(V)$ is the span of $(A+T)(X)$ and $(A+T)(Y)$. It follows that $(A+T)(V)$ is a closed linear subspace of W when V is complete, because it is the span of a closed subspace and a finite-dimensional subspace.

A similar argument works when A satisfies (23.1) on a closed linear subspace of V of finite codimension. Of course, if A satisfies (23.1) on some linear subspace of V, then it satisfies the same condition on the closure of that linear subspace. Note that the intersection of two linear subspaces of V with finite codimension also has finite codimension in V.

24 Schauder's fixed-point theorem

Let V be a real or complex vector space with a norm $\|v\|$, and let E be a nonempty totally bounded subset of V. Also let $\epsilon > 0$ be given, and let $x_1, \ldots, x_n$ be finitely many elements of E such that

$$E \subseteq \bigcup_{j=1}^{n} B(x_j, \epsilon). \tag{24.1}$$

Let $\phi_j(v)$ be a continuous real-valued function on V such that $\phi_j(v) > 0$ when $\|v - x_j\| < \epsilon$ and $\phi_j(v) = 0$ otherwise, for each $j = 1, \ldots, n$. One can take

$$\phi_j(v) = (\epsilon - \|v - x_j\|)_+, \tag{24.2}$$

for instance, where r_+ is the nonnegative part of a real number r, equal to r when $r \geq 0$ and to 0 when $r \leq 0$. Put

$$\psi_j(v) = \phi_j(v) \Big(\sum_{l=1}^{n} \phi_l(v) \Big)^{-1} \tag{24.3}$$

for each $v \in E$, which makes sense because $\sum_{l=1}^{n} \phi_l(v) > 0$ for every $v \in E$, by construction. Thus ψ_j is a continuous nonnegative real-valued function on E for each j, $\psi_j(v) > 0$ when $\|v - x_j\| < \epsilon$, and $\psi_j(v) = 0$ otherwise, because of the corresponding properties of ϕ_j. Moreover,

$$\sum_{j=1}^{n} \psi_j(v) = \Big(\sum_{j=1}^{n} \phi_j(v) \Big) \Big(\sum_{l=1}^{n} \phi_l(v) \Big)^{-1} = 1 \tag{24.4}$$

for each $v \in V$, so that the ψ_j's form a partition of unity on E.

Put

$$A_\epsilon(v) = \sum_{j=1}^{n} \psi_j(v)\, x_j \tag{24.5}$$

for each $v \in E$, which defines a continuous mapping from E into V. More precisely, $A_\epsilon(v)$ is an element of the convex hull of E for each $v \in E$, because of the properties of the ψ_j's discussed in the previous paragraph. Note that

$$A_\epsilon(v) - v = \sum_{j=1}^{n} \psi_j(v)\,(x_j - v) \tag{24.6}$$

for each $v \in E$, by (24.4), and hence

$$\|A_\epsilon(v) - v\| \le \sum_{j=1}^{n} \psi_j(v)\,\|x_j - v\|. \tag{24.7}$$

This implies that

$$\|A_\epsilon(v) - v\| < \epsilon \tag{24.8}$$

for every $v \in E$, since $\|x_j - v\| < \epsilon$ when $\psi_j(v) > 0$. Let W_ϵ be the linear span of $x_1, \ldots, x_n$ in V, so that W_ϵ is a finite-dimensional linear subspace of V, and $A_\epsilon(v) \in W_\epsilon$ for each $v \in E$.

Suppose now that E is compact and convex, and let f be a continuous mapping from E into itself. Thus A_ϵ maps E into itself, because E is convex, and in fact A_ϵ maps E into $E \cap W_\epsilon$. It follows that

$$f_\epsilon = A_\epsilon \circ f \tag{24.9}$$

is a continuous mapping from E into $E \cap W_\epsilon$, and in particular the restriction of f_ϵ to $E \cap W_\epsilon$ is a continuous mapping from $E \cap W_\epsilon$ into itself. Because W_ϵ has finite dimension, one can use Brouwer's fixed point theorem to show that there is a point v_ϵ in $E \cap W_\epsilon$ such that $f_\epsilon(v_\epsilon) = v_\epsilon$. Remember that W_ϵ is a closed linear subspace of V, since it has finite dimension, so that $E \cap W_\epsilon$ is a compact convex set in W_ϵ.

Observe that

$$\|f(v_\epsilon) - v_\epsilon\| = \|f(v_\epsilon) - f_\epsilon(v_\epsilon)\| = \|f(v_\epsilon) - A_\epsilon(f(v_\epsilon))\| < \epsilon, \tag{24.10}$$

by (24.8). Because E is compact, and hence sequentially compact, there is a sequence $\{\epsilon_j\}_{j=1}^\infty$ of positive real numbers converging to 0 such that $\{v_{\epsilon_j}\}_{j=1}^\infty$ converges to an element v of E with respect to the norm on V. Using (24.10) and the continuity of f, it follows that $f(v) = v$, so that f has a fixed point in E under these conditions.

25 Tychonoff's fixed-point theorem

Let V be a locally convex topological vector space over the real or complex numbers, let E be a nonempty compact convex subset of V, and let f be a continuous mapping from E into itself. Under these conditions, f has a fixed point in E, which is to say that there is a $v \in E$ such that $f(v) = v$. This can be shown in roughly the same way as in the previous section, with suitable modifications. In particular, one can get approximate fixed points for f in E with respect to compatible seminorms on V in essentially the same way as before. If there is a countable local base for the topology of V at 0, so that the topology on V is determined by a translation-invariant metric, then one can get a fixed point for f on E using sequential compactness as before. This also works when the topology on E induced by the one on V is metrizable, and otherwise one can use slightly more complicated compactness arguments. An advantage of this version is that one can have larger classes of compact sets, although one should also be careful about the corresponding continuity conditions.

To get a basic class of examples, let X be a nonempty compact Hausdorff topological space, and let $C(X)$ be the space of continuous real-valued functions on X, equipped with the supremum norm

$$\|f\| = \sup_{x \in X} |f(x)|. \tag{25.1}$$

Remember that a linear functional λ on $C(X)$ is said to be *nonnegative* if

$$\lambda(f) \geq 0 \tag{25.2}$$

for each $f \in C(X)$ such that $f(x) \geq 0$ for every $x \in X$. Let $\mathbf{1}_X(x)$ be the constant function on X equal to 1 for each $x \in X$. If λ is a nonnegative linear functional on $C(X)$ and $f \in C(X)$ satisfies $\|f\| \leq 1$, then $\mathbf{1}_X - f \geq 0$ and $\mathbf{1}_X + f \geq 0$ on X, and hence $\lambda(\mathbf{1}_X - f) \geq 0$ and $\lambda(\mathbf{1}_X + f) \geq 0$. This implies that $\lambda(f) \leq \lambda(\mathbf{1}_X)$ and $-\lambda(f) \leq \lambda(\mathbf{1}_X)$, so that

$$|\lambda(f)| \leq \lambda(\mathbf{1}_X). \tag{25.3}$$

It follows that a nonnegative linear functional λ on $C(X)$ is automatically bounded with respect to the supremum norm on $C(X)$, with dual norm less than or equal to $\lambda(\mathbf{1}_X)$. The dual norm of λ on $C(X)$ is in fact equal to $\lambda(\mathbf{1}_X)$, since $\|\mathbf{1}_X\| = 1$.

Let $\mathcal{P}(X)$ be the set of nonnegative linear functionals λ on $C(X)$ that satisfy $\lambda(\mathbf{1}_X) = 1$. Note that the elements of $\mathcal{P}(X)$ correspond exactly to regular Borel probability measures on X, by the Riesz representation theorem. As in the previous paragraph, the elements of $\mathcal{P}(X)$ are also bounded linear functionals on $C(X)$, with dual norm equal to 1. It is easy to see that $\mathcal{P}(X)$ is a convex set in the dual $C(X)^*$ of $C(X)$, and one can also check that $\mathcal{P}(X)$ is a closed set with respect to the weak* topology on $C(X)^*$. It follows that $\mathcal{P}(X)$ is compact with respect to the weak* topology on $C(X)^*$, by the Banach–Alaoglu theorem.

Let ϕ be a continuous mapping from X into itself, and put

$$T(f) = f \circ \phi \tag{25.4}$$

for each $f \in C(X)$. Thus

$$\|T(f)\| \leq \|f\| \tag{25.5}$$

for every $f \in C(X)$, so that T is a bounded linear mapping from $C(X)$ into itself, with operator norm less than or equal to 1. More precisely, the operator norm of T is equal to 1, because $T(\mathbf{1}_X) = \mathbf{1}_X$.

The corresponding dual mapping T^* is defined on $C(X)^*$ as usual by

$$T^*(\lambda) = \lambda \circ T \tag{25.6}$$

for each $\lambda \in C(X)^*$, so that

$$(T^*(\lambda))(f) = \lambda(T(f)) = \lambda(f \circ \phi) \tag{25.7}$$

for every $\lambda \in C(X)^*$ and $f \in C(X)$. If λ is a nonnegative linear functional on $C(X)$, then $T^*(\lambda)$ is nonnegative too, because $T(f) \geq 0$ on X when $f \geq 0$ on X. It follows easily that $T^*(\mathcal{P}(X)) \subseteq \mathcal{P}(X)$.

Note that T^* is continuous with respect to the weak* topology on $C(X)^*$, because it is the dual of a bounded linear mapping on $C(X)$. If we take $V = C(X)^*$ with the weak* topology and $E = \mathcal{P}(X)$, then we get that T^* has a fixed point in $\mathcal{P}(X)$. Alternatively, we can use the linearity of T^* to get approximate fixed points in $\mathcal{P}(X)$ more directly, as follows.

Let λ be an arbitrary element of $\mathcal{P}(X)$, and consider

$$\lambda_n = \frac{1}{n+1} \sum_{j=0}^{n} (T^*)^j(\lambda) \tag{25.8}$$

for each nonnegative integer n. Here $(T^*)^j$ denotes the jth power of T^* as a linear mapping on $C(X)^*$, which is the same as the dual of the jth power T^j of T on $C(X)$. As usual, we interpret T^j and $(T^*)^j$ as being the identity mappings on $C(X)$ and $C(X)^*$, respectively, when $j = 0$. Observe that $\lambda_n \in \mathcal{P}(X)$ for each n, since $(T^*)^j(\lambda) \in \mathcal{P}(X)$ for each j.

By construction,

$$\begin{aligned} T^*(\lambda_n) - \lambda_n &= \frac{1}{n+1} \sum_{j=1}^{n+1} (T^*)^j(\lambda) - \frac{1}{n+1} \sum_{j=0}^{n} (T^*)^j(\lambda) \\ &= \frac{1}{n+1} \left((T^*)^{n+1}(\lambda) - \lambda\right) \end{aligned} \tag{25.9}$$

for each n. Let $\|\mu\|_*$ denote the dual norm on $C(X)^*$ corresponding to the supremum norm on $C(X)$, so that

$$\|T^*(\mu)\|_* \leq \|\mu\|_* \tag{25.10}$$

for every $\mu \in C(X)^*$, by (25.5). Thus

$$\begin{aligned} \|T^*(\lambda_n) - \lambda_n\|_* &\leq \frac{1}{n+1} \left(\|(T^*)^{n+1}(\lambda)\|_* + \|\lambda\|_*\right) \\ &\leq \frac{2}{n+1} \|\lambda\|_* = \frac{2}{n+1} \end{aligned} \tag{25.11}$$

for each n, since $\|\lambda\|_* = 1$ when $\lambda \in \mathcal{P}(X)$.

To get a fixed point of T^* in $\mathcal{P}(X)$, one can use the compactness of $\mathcal{P}(X)$ with respect to the weak* topology on $C(X)^*$, as follows. Let A_l be the closure of the set of λ_n with $n \geq l$ with respect to the weak* topology on $C(X)^*$ for each $l \geq 0$. Thus $A_l \subseteq \mathcal{P}(X)$, $A_{l+1} \subseteq A_l$, and $A_l \neq \emptyset$ for every l. This implies that $\bigcap_{l=0}^\infty A_l \neq \emptyset$, because $\mathcal{P}(X)$ is compact and A_l is a closed subset of $\mathcal{P}(X)$ for each l, with respect to the weak* topology on $C(X)^*$. One can check that $T^*(\mu) = \mu$ for every $\mu \in \bigcap_{l=0}^\infty A_l$, using (25.11).

26 Composition mappings

Let X and Y be nonempty compact Hausdorff topological spaces, and let $C(X)$ and $C(Y)$ be the corresponding spaces of continuous real or complex-valued functions on X and Y, respectively. Also let $\|\cdot\|_{C(X)}$ and $\|\cdot\|_{C(Y)}$ denote the supremum norms on $C(X)$ and $C(Y)$, respectively, and let $\mathbf{1}_X$ and $\mathbf{1}_Y$ be the constant functions equal to 1 on X and Y. If ϕ is a continuous mapping from X into Y, then

$$T(f) = f \circ \phi \tag{26.1}$$

defines a bounded linear mapping from $C(Y)$ into $C(X)$. More precisely,

$$\|T(f)\|_{C(X)} \leq \|f\|_{C(Y)} \tag{26.2}$$

for every $f \in C(Y)$, and $T(\mathbf{1}_Y) = \mathbf{1}_X$, so that the operator norm of T is equal to 1 with respect to the supremum norms on $C(X)$ and $C(Y)$.

Suppose for the moment that X is a subset of Y, equipped with the induced topology. Note that X is a closed subset of Y, since X is compact and Y is Hausdorff, and conversely that closed subsets of Y are compact. In this case, $T(f)$ is simply the restriction of $f \in C(Y)$ to X, and the Tietze extension theorem implies that T maps $C(Y)$ onto $C(X)$, because Y is normal and X is a closed set in Y. More precisely, if f is a continuous function on X, then there is an extension of f to a continuous function on Y whose supremum norm on Y is equal to the supremum norm of f on X. Thus T behaves like a quotient mapping in terms of norms, and $C(X)$ can be identified with the quotient of $C(Y)$ by the closed linear subspace of continuous functions on Y that are equal to 0 at every point in X.

Now suppose that ϕ maps X onto Y. In this case, T is an isometric embedding of $C(Y)$ into $C(X)$, in the sense that

$$\|T(f)\|_{C(X)} = \|f\|_{C(Y)} \tag{26.3}$$

for every $f \in C(Y)$. Note that a function f on Y is continuous if and only if $f \circ \phi$ is continuous on X. The "only if" part is trivial, and to check the "if" part, it suffices to show that $f^{-1}(E)$ is a closed set in Y when E is a closed set in the range of f and $f \circ \phi$ is continuous. The continuity of $f \circ \phi$ implies that $(f \circ \phi)^{-1}(E) = \phi^{-1}(f^{-1}(E))$ is a closed set in X for every closed set E in the range of f, and so the main point is that $A \subseteq Y$ is a closed set when $\phi^{-1}(A)$ is a closed set in X. If $\phi^{-1}(A)$ is a closed set in X, then $\phi^{-1}(A)$ is also a compact set in X, because X is compact. Thus $\phi(\phi^{-1}(A))$ is a compact set in Y, since ϕ is continuous, which implies that $\phi(\phi^{-1}(A))$ is a closed set in Y, because Y is Hausdorff. Of course, $A = \phi(\phi^{-1}(A))$ when $\phi(X) = Y$, so that A is a closed set in Y, as desired.

Let ϕ be any continuous mapping from X into Y again, and put $Z = \phi(X)$. Also let ϕ_1 be the same as ϕ but considered as a continuous mapping from X onto Z with the topology induced from the one on Y, and let ϕ_2 is the obvious inclusion mapping from Z into Y, so that $\phi = \phi_2 \circ \phi_1$. If $T_1 : C(Z) \to C(X)$ and $T_2 : C(Y) \to C(Z)$ are the composition mappings corresponding to ϕ_1 and ϕ_2, respectively, then it is easy to see that $T = T_1 \circ T_2$. Of course, ϕ_1 and ϕ_2 are exactly as in the special cases described in the previous paragraphs.

27 Non-compact spaces

Let X and Y be nonempty topological spaces, not necessarily compact, and let $C_b(X)$ and $C_b(Y)$ be the corresponding spaces of bounded continuous real or complex-valued functions on X and Y, respectively. Also let $\|\cdot\|_{C_b(X)}$ and $\|\cdot\|_{C_b(Y)}$ be the supremum norms on $C_b(X)$ and $C_b(Y)$, and let $\mathbf{1}_X$ and $\mathbf{1}_Y$ be the constant functions equal to 1 on X and Y, as before. If ϕ is a continuous mapping from X into Y, then

$$T(f) = f \circ \phi \tag{27.1}$$

again defines a bounded linear mapping from $C_b(X)$ into $C_b(Y)$, with

$$\|T(f)\|_{C_b(X)} \le \|f\|_{C_b(Y)} \tag{27.2}$$

for every $f \in C_b(Y)$, and $T(\mathbf{1}_Y) = \mathbf{1}_X$. Of course, continuous real or complex-valued functions on compact spaces are automatically bounded, so that this includes the situation discussed in the previous section.

If X is a subset of Y with the topology induced from the one on Y, then $T(f)$ is simply the restriction of $f \in C_b(Y)$ to X. In particular, if Y is normal and X is a closed set in Y, then the Tietze extension theorem implies that T maps $C_b(Y)$ onto $C_b(X)$. If f is any bounded continuous function on Y and R is a nonnegative real number, then consider the function f_R on Y defined by

$$\begin{aligned} f_R(y) &= f(y) && \text{when } |f(y)| \le R, \text{ and} \\ &= R\,\frac{f(y)}{|f(y)|} && \text{when } |f(y)| > R. \end{aligned} \tag{27.3}$$

It is easy to see that f_R is a continuous function on Y for each $R \ge 0$, which is the same as the composition of f with a certain continuous function on $\mathbf{R}$ or $\mathbf{C}$, as appropriate. If we take R to be the supremum of $|f|$ over X, then f_R is a bounded continuous function on Y whose restriction to X is the same as the restriction of f to X, and the supremum norm of f_R on Y is the same as the supremum norm of the restriction to X. This implies that T behaves like a quotient mapping from $C_b(Y)$ onto its image in $C_b(X)$ in terms of norms, even when the image is a proper linear subspace of $C_b(X)$. Using this, one can show that the image of $C_b(Y)$ in $C_b(X)$ is a closed linear subspace with respect to the supremum norm, since $C_b(Y)$ is complete with respect to the supremum norm.

If ϕ is a continuous mapping from X onto a dense subset of Y, then

$$\|T(f)\|_{C_b(X)} = \|f\|_{C_b(Y)} \tag{27.4}$$

for every $f \in C_b(Y)$, so that T is an isometric embedding of $C_b(Y)$ into $C_b(X)$ with respect to the supremum norm. In particular, if X is compact and Y is Hausdorff, then $\phi(X)$ is a compact and hence closed set in Y, and $\phi(X) = Y$ when $\phi(X)$ is also dense in Y. Because $C_b(Y)$ is complete with respect to the supremum norm, (27.4) implies that $T(C_b(Y))$ is a closed linear subspace of $C_b(X)$, for any X and Y.

If ϕ is any continuous mapping from X into Y, then let Z be a subset of Y that contains $\phi(X)$ and is contained in the closure $\overline{\phi(X)}$ of $\phi(X)$ in Y. As before, ϕ can be expressed as $\phi_2 \circ \phi_1$, where ϕ_1 is the same as ϕ but considered as a continuous mapping from X into Z with the topology induced by the one on Y, and where ϕ_2 is the obvious inclusion mapping from Z into Y. This implies that $T = T_1 \circ T_2$, where $T_1 : C_b(Z) \to C_b(X)$ and $T_2 : C_b(Y) \to C_b(Z)$ are the composition mappings corresponding to ϕ_1 and ϕ_2, respectively.

28 Another openness condition

Let V and W be vector spaces, both real or both complex, and equipped with norms $\|v\|_V$ and $\|w\|_W$, respectively. Let T be a bounded linear mapping from V into W, and suppose that there are positive real numbers k, η such that $\eta < 1$ and for each $w \in W$ with $\|w\|_W \le 1$ there is a $v \in V$ with $\|v\|_V \le k$ and $\|w - T(v)\|_W \le \eta$. Equivalently, this means that for each $w \in W$ there is a $v \in V$ such that

$$\|v\|_V \le k\,\|w\|_W \tag{28.1}$$

and

$$\|w - T(v)\|_W \le \eta\,\|w\|_W. \tag{28.2}$$

Let $w \in W$ be given, and let v_1 be an element of V that satisfies (28.1) and (28.2). Applying this to $w - T(v_1)$ instead of w, we get an element v_2 of V with

$$\|v_2\|_V \le k\,\|w - T(v_1)\|_W \le k\,\eta\,\|w\|_W \tag{28.3}$$

and

$$\|w - T(v_1) - T(v_2)\|_W \le \eta\|w - T(v_1)\|_W \le \eta^2\,\|w\|_W. \tag{28.4}$$

Similarly, if $v_1, \ldots, v_n \in V$ have already been chosen in this way, then we can apply this to $w - \sum_{j=1}^n T(v_j)$ to get an element v_{n+1} of V such that

$$\|v_{n+1}\|_V \le k\,\Big\|w - \sum_{j=1}^n T(v_j)\Big\|_W \tag{28.5}$$

and

$$\left\| w - \sum_{j=1}^{n} T(v_j) - T(v_{n+1}) \right\|_W \le \eta \left\| w - \sum_{j=1}^{n} T(v_j) \right\|_W. \tag{28.6}$$

It follows that

$$\left\| w - \sum_{j=1}^{n} T(v_j) \right\|_W \le \eta^n \, \|w\|_W \tag{28.7}$$

for each n, and hence that

$$\|v_n\|_V \le k \, \eta^{n-1} \, \|w\|_W. \tag{28.8}$$

In particular,

$$\sum_{j=1}^{\infty} \|v_j\|_V \le \frac{k}{1-\eta} \, \|w\|_W. \tag{28.9}$$

If V is complete, then $\sum_{j=1}^{\infty} v_j$ converges in V, and satisfies

$$\left\| \sum_{j=1}^{\infty} v_j \right\|_V \le \frac{k}{1-\eta} \, \|w\|_W \tag{28.10}$$

and $T\Big(\sum_{j=1}^{\infty} v_j \Big) = w$. Otherwise, if V is not complete, then we get that

$$\overline{B}_W(0, r) \subseteq \overline{T(B_V(0, r\,k/(1-\eta)))} \tag{28.11}$$

for each $r > 0$. The main difference between this and the discussion in Section 16 is that here we use a single $\eta \in (0, 1)$, instead of arbitrarily small $\eta > 0$.

As in [22], the proof of Tietze's extension theorem can be considered as an example of this type of situation. Let Y be a normal topological space, and let X be a nonempty closed subset of Y. Also let $C_b(X)$ and $C_b(Y)$ be the vector spaces or bounded continuous real-valued functions on X and Y, respectively, with the corresponding supremum norms $\|\cdot\|_{C_b(X)}$ and $\|\cdot\|_{C_b(Y)}$. Suppose that $f \in C_b(X)$ satisfies $\|f\|_{C_b(X)} \le 1$, so that f takes values in $[-1, 1]$. Note that

$$A = \{x \in X : f(x) \le -1/3\} \quad \text{and} \quad B = \{x \in X : f(x) \ge 1/3\} \tag{28.12}$$

are disjoint closed subsets of X, and hence of Y. By Urysohn's lemma, there is a continuous real-valued function g on Y such that $g(x) = -1/3$ for every $x \in A$, $g(x) = 1/3$ for every $x \in B$, and $|g(y)| \le 1/3$ for every $y \in Y$. It is easy to see that

$$|f(x) - g(x)| \le 2/3 \tag{28.13}$$

for every $x \in X$, by considering separately the cases where $x \in A$, $x \in B$, and $x \in X \backslash (A \cup B)$. This puts us in the previous situation with $k = 1/3$ and $\eta = 2/3$, and where $T : C_b(Y) \to C_b(X)$ sends a continuous function on Y to its restriction to X. Of course, $C_b(Y)$ is complete with respect to the supremum norm, and so it follows that T maps $C_b(Y)$ onto $C_b(X)$ under these conditions, as desired.

References

[1] F. Albiac and N. Kalton, *Topics in Banach Space Theory*, Springer-Verlag, 2006.

[2] G. Bachman and L. Narici, *Functional Analysis*, Dover, 2000.

[3] L. Baggett, *Functional Analysis: A Primer*, Dekker, 1992.

[4] R. Bass, *Real Analysis for Graduate Students: Measure and Integration Theory*, 2011. http://homepages.uconn.edu/~rib02005/real.html

[5] R. Beals, *Topics in Operator Theory*, The University of Chicago Press, 1971.

[6] R. Beals, *Advanced Mathematical Analysis*, Springer-Verlag, 1973.

[7] R. Beals, *Analysis: An Introduction*, Cambridge University Press, 2004.

[8] B. Beauzamy, *Introduction to Banach Spaces and their Geometry*, 2nd edition, North-Holland, 1985.

[9] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis*, American Mathematical Society, 2000.

[10] R. Bhatia, *Notes on Functional Analysis*, Hindustan Book Agency, 2009.

[11] H. Brezis, *Functional Analysis, Sobolev Spaces, and Partial Differential Equations*, Springer-Verlag, 2011.

[12] J. Cerdà, *Linear Functional Analysis*, American Mathematical Society, Real Sociedad Matemática Española, 2010.

[13] N. Carothers, *A Short Course on Banach Space Theory*, Cambridge University Press, 2005.

[14] J. Conway, *A Course in Functional Analysis*, 2nd edition, Springer-Verlag, 1990.

[15] R. Edwards, *Functional Analysis*, Dover, 1995.

[16] Y. Eidelman, V. Milman, and A. Tsolomitis, *Functional Analysis: An Introduction*, American Mathematical Society, 2004.

[17] G. Folland, *Real Analysis: Modern Techniques and their Applications*, 2nd edition, Wiley, 1999.

[18] G. Folland, *A Guide to Advanced Real Analysis*, Mathematical Association of America, 2009.

[19] T. Gamelin and R. Greene, *Introduction to Topology*, 2nd edition, Dover, 1999.

[20] R. Goldberg, *Methods of Real Analysis*, 2nd edition, Wiley, 1976.

[21] S. Goldberg, *Unbounded Linear Operators*, Dover, 2006.

[22] S. Grabiner, *The Tietze extension theorem and the open mapping theorem*, American Mathematical Monthly **93** (1986), 190–191.

[23] V. Hansen, *Functional Analysis: Entering Hilbert Space*, World Scientific, 2006.

[24] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, 1975.

[25] S. Kesavan, *Functional Analysis*, Hindustan Book Agency, 2009.

[26] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.

[27] A. Knapp, *Advanced Real Analysis*, Birkhäuser, 2005.

[28] S. Krantz, *A Handbook of Real Variables*, Birkhäuser, 2004.

[29] S. Krantz, *Real Analysis and Foundations*, 2nd edition, Chapman & Hall / CRC, 2005.

[30] S. Krantz, *A Guide to Real Variables*, Mathematical Association of America, 2009.

[31] S. Krantz, *A Guide to Topology*, Mathematical Association of America, 2009.

[32] S. Krantz, *Essentials of Topology with Applications*, CRC press, 2010.

[33] S. Lang, *Real and Functional Analysis*, 3rd edition, Springer-Verlag, 1993.

[34] S. Lang, *Undergraduate Analysis*, 2nd edition, Springer-Verlag, 1997.

[35] P. Lax, *Functional Analysis*, Wiley, 2002.

[36] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces*, Lecture Notes in Mathematics **388**, Springer-Verlag, 1973.

[37] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces I: Sequence Spaces*, Springer-Verlag, 1977.

[38] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces II: Function Spaces*, Springer-Verlag, 1979.

[39] B. MacCluer, *Elementary Functional Analysis*, Springer-Verlag, 2009.

[40] R. Megginson, *An Introduction to Banach Space Theory*, Springer-Verlag, 1998.

[41] T. Morrison, *Functional Analysis: An Introduction to Banach Space Theory*, Wiley, 2001.

[42] L. Nachbin, *Introduction to Functional Analysis: Banach Spaces and Differential Calculus*, translated from the Portuguese by R. Aron, Dekker, 1981.

[43] L. Nirenberg, *Topics in Nonlinear Functional Analysis*, Chapter 6 by E. Zehnder, notes by R. Artino, New York University, American Mathematical Society, 2001.

[44] S. Promislow, *A First Course in Functional Analysis*, Wiley, 2008.

[45] H. Royden, *Real Analysis*, 3rd edition, Macmillan, 1988.

[46] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.

[47] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, 1987.

[48] W. Rudin, *Functional Analysis*, 2nd edition, McGraw-Hill, 1991.

[49] B. Rynne and M. Youngson, *Linear Functional Analysis*, 2nd edition, Springer-Verlag, 2008.

[50] K. Saxe, *Beginning Functional Analysis*, Springer-Verlag, 2002.

[51] M. Schechter, *Principles of Functional Analysis*, 2nd edition, American Mathematical Society, 2002.

[52] E. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.

[53] E. Stein and R. Shakarchi, *Functional Analysis: Introduction to Further Topics in Analysis*, Princeton University Press, 2011.

[54] C. Swartz, *An Introduction to Functional Analysis*, Dekker, 2002.

[55] C. Swartz, *Elementary Functional Analysis*, World Scientific, 2009.

[56] P. Wojtaszczyk, *Banach Spaces for Analysts*, Cambridge University Press, 1991.

[57] K. Yosida, *Functional Analysis*, Springer-Verlag, 1995.

[58] R. Zimmer, *Essential Results of Functional Analysis*, University of Chicago Press, 1990.