

Sums and averages

Stephen Semmes
Rice University

Abstract

These informal notes are concerned with sums and averages in various situations in analysis.

Contents

1	Real and complex numbers	3
2	Cesaro means	4
3	Admissible series	5
4	Abel summability	6
5	Cauchy products	8
6	Norms on vector spaces	9
7	Inner product spaces	10
8	Convexity	11
9	A few estimates	12
10	Operator norms	13
11	Linear mappings	14
12	Convergence	15
13	Completeness	16
14	The supremum norm	17
15	Algebras	18
16	Banach algebras	19

17 Invertibility	20
18 Submultiplicative sequences	21
19 Invertibility, 2	21
20 Spectrum	22
21 Averages in normed algebras	23
22 Invertibility, 3	24
23 The open mapping theorem	25
24 The uniform boundedness principle	26
25 Strong operator convergence	27
26 Convergence of averages	27
27 Hilbert spaces	28
28 Unitary transformations	30
29 Rotations	31
30 Fourier series	31
31 Measure-preserving transformations	32
32 Sequence spaces	33
33 Maximal functions	34
34 L^p estimates	35
35 Maximal functions, 2	36
36 Discrete maximal functions	37
37 Another variant	37
38 Transference	38
39 Almost-everywhere convergence	40
40 Multiplication operators	42
41 Doubling spaces	43

42 Ultrametrics	43
43 Sequence spaces, 2	44
44 Snowflake transforms	44
45 Invariant measures	45
References	47

1 Real and complex numbers

Let \mathbf{R} be the field of real numbers. The *absolute value* of a real number x is denoted $|x|$ and defined to be equal to x when $x \geq 0$ and to $-x$ when $x \leq 0$. Thus $|x| \geq 0$,

$$(1.1) \quad |x + y| \leq |x| + |y|,$$

and

$$(1.2) \quad |xy| = |x| |y|$$

for every $x, y \in \mathbf{R}$.

If A is a nonempty set of real numbers with an upper bound, then there is a unique real number which is the least upper bound or supremum of A , denoted $\sup A$. Similarly, a nonempty set A of real numbers with a lower bound has a greatest lower bound or infimum, denoted $\inf A$. It is sometimes convenient to put $\sup A = +\infty$ when A has no upper bound, or $\inf A = -\infty$ when A has no lower bound. Normally we shall only be concerned with infima of sets of nonnegative real numbers here, which are also nonnegative.

Let \mathbf{C} be the field of complex numbers. A complex number z can be expressed as $x + yi$, where $x, y \in \mathbf{R}$ and $i^2 = -1$. In this case, x and y are known as the real and imaginary parts of z , and are denoted $\operatorname{Re} z$, $\operatorname{Im} z$, respectively. The *complex conjugate* \bar{z} of z is defined by

$$(1.3) \quad \bar{z} = x - yi.$$

In particular, $\overline{\bar{z}} = z$ and

$$(1.4) \quad \operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}.$$

Observe that

$$(1.5) \quad \overline{z + w} = \bar{z} + \bar{w}$$

and

$$(1.6) \quad \overline{z\bar{w}} = \bar{z}w$$

for every $z, w \in \mathbf{C}$. The *modulus* of $z = x + yi$ is the nonnegative real number given by

$$(1.7) \quad |z| = (x^2 + y^2)^{1/2}.$$

Equivalently,

$$(1.8) \quad |z|^2 = z \bar{z},$$

and hence

$$(1.9) \quad |z w| = |z| |w|$$

for every $z, w \in \mathbf{C}$.

Of course, $|\operatorname{Re} z|, |\operatorname{Im} z| \leq |z|$ for every $z \in \mathbf{C}$. If $z, w \in \mathbf{C}$, then

$$(1.10) \quad \begin{aligned} |z + w|^2 &= (z + w)(\bar{z} + \bar{w}) \\ &= |z|^2 + z \bar{w} + \bar{z} w + |w|^2 \\ &= |z|^2 + 2 \operatorname{Re} z \bar{w} + |w|^2, \end{aligned}$$

$$(1.11)$$

since $\overline{z \bar{w}} = \bar{z} w$. This implies that

$$(1.12) \quad |z + w|^2 \leq |z|^2 + 2|z||w| + |w|^2 = (|z| + |w|)^2,$$

and therefore

$$(1.13) \quad |z + w| \leq |z| + |w|.$$

2 Cesaro means

As usual, a sequence $\{z_j\}_{j=0}^{\infty}$ of complex numbers converges to $z \in \mathbf{C}$ if for every $\epsilon > 0$ there is a nonnegative integer L such that

$$(2.1) \quad |z_j - z| < \epsilon$$

for every $j \geq L$. In this case, one can show that the sequence of averages

$$(2.2) \quad \zeta_n = \frac{z_1 + \cdots + z_n}{n+1}$$

also converges to z as $n \rightarrow \infty$. However, there are also sequences $\{z_j\}_{j=0}^{\infty}$ of complex numbers that do not converge, but for which the corresponding sequence $\{\zeta_n\}_{n=0}^{\infty}$ of averages does converge. For example, if $z_j = (-1)^j$, then $\zeta_n = 0$ when n is odd and $\zeta_n = 1/(n+1)$ when n is even, and $\lim_{n \rightarrow \infty} \zeta_n = 0$.

If z is any complex number, then

$$(2.3) \quad (z - 1) \sum_{j=0}^n z^j = z^{n+1} - 1,$$

where $z^j = 1$ for every $z \in \mathbf{C}$ when $j = 0$. Hence

$$(2.4) \quad \sum_{j=0}^n z^j = \frac{z^{n+1} - 1}{z - 1}$$

when $z \neq 1$. It follows that

$$(2.5) \quad \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{j=0}^n z^j = 0$$

when $|z| = 1$ and $z \neq 1$. This extends the case of $z = -1$ described in the previous paragraph.

Let $\sum_{j=0}^{\infty} a_j$ be an infinite series of complex numbers, and consider the sequence of partial sums

$$(2.6) \quad b_l = \sum_{j=0}^l a_j.$$

By definition, $\sum_{j=0}^{\infty} a_j$ converges if $\{b_l\}_{l=0}^{\infty}$ converges, in which event

$$(2.7) \quad \sum_{j=0}^{\infty} a_j = \lim_{l \rightarrow \infty} b_l.$$

The average β_n of b_0, \dots, b_n is also given by

$$(2.8) \quad \beta_n = \sum_{j=0}^n \frac{n+1-j}{n+1} a_j.$$

The series $\sum_{j=0}^{\infty} a_j$ is said to be *Cesaro summable* if $\{\beta_n\}_{n=0}^{\infty}$ converges.

Consider the case of a geometric series

$$(2.9) \quad \sum_{j=0}^{\infty} a^j,$$

where $a \in \mathbf{C}$ and $a^j = 1$ when $j = 0$ again. If $|a| < 1$, then $\lim_{j \rightarrow \infty} a^j = 0$, and the corresponding geometric series converges with

$$(2.10) \quad \sum_{j=0}^{\infty} a^j = \frac{1}{1-a}$$

by the previous computations. If $|a| \geq 1$, then $|a^j| = |a|^j \geq 1$ for every j , and the geometric series does not converge in the conventional sense. It is Cesaro summable with sum $1/(1-a)$ when $|a| = 1$ and $a \neq 1$, by the earlier computations for the partial sums applied twice to estimate their averages too.

3 Admissible series

Let $\sum_{j=0}^{\infty} a_j$ be an infinite series of complex numbers. If

$$(3.1) \quad \sum_{j=0}^{\infty} a_j t^j$$

converges for some real number $t \leq 1$, then

$$(3.2) \quad \lim_{j \rightarrow \infty} a_j t^j = 0,$$

which implies that $\{a_j t^j\}_{j=0}^{\infty}$ is a bounded sequence. Conversely, if $\{a_j t^j\}_{j=0}^{\infty}$ is bounded, then

$$(3.3) \quad \sum_{j=0}^{\infty} |a_j| r^j$$

converges for $0 \leq r < t$, by comparison with a convergent geometric series.

Let us say that an infinite series $\sum_{j=0}^{\infty} a_j$ of complex numbers is *admissible* if any of the previous conditions holds for every positive real number $r < 1$ or $t < 1$, as appropriate, so that each of the other conditions also holds for every $0 \leq r < 1$ or $0 \leq t < 1$. This is the same as saying that

$$(3.4) \quad f(z) = \sum_{j=0}^{\infty} a_j z^j$$

has radius of convergence greater than or equal to 1. If $b_l = \sum_{j=0}^l a_j$ are the partial sums of $\sum_{j=0}^{\infty} a_j$, then $\sum_{l=0}^{\infty} b_l$ is admissible too. Indeed, $a_j = O(R^j)$ implies that $b_l = O(R^l)$ for each $R > 1$.

Put $b_{-1} = 0$, so that

$$(3.5) \quad \sum_{j=0}^n a_j z^j = \sum_{j=0}^n (b_j - b_{j-1}) z^j = \sum_{j=0}^n b_j z^j - \sum_{j=0}^n b_{j-1} z^j.$$

Since

$$(3.6) \quad \sum_{j=0}^n b_{j-1} z^j = \sum_{j=1}^n b_{j-1} z^j = \sum_{j=0}^{n-1} b_j z^{j+1},$$

we get that

$$(3.7) \quad \sum_{j=0}^n a_j z^j = \sum_{j=0}^{n-1} b_j (z^j - z^{j+1}) + b_n z^n = (1 - z) \sum_{j=0}^{n-1} b_j z^j + b_n z^n.$$

If $|z| < 1$, then admissibility of $\sum_{n=0}^{\infty} b_n$ implies that $\lim_{n \rightarrow \infty} b_n z^n = 0$, and

$$(3.8) \quad f(z) = (1 - z) \sum_{j=0}^{\infty} b_j z^j.$$

4 Abel summability

If $\sum_{j=0}^{\infty} a_j$ converges, then it is well known that

$$(4.1) \quad \lim_{r \rightarrow 1^-} \sum_{j=0}^{\infty} a_j r^j$$

exists and is equal to $\sum_{j=0}^{\infty} a_j$. An admissible series $\sum_{j=0}^{\infty} a_j$ of complex numbers is said to be *Abel summable* when this limit exists.

For example, $\sum_{j=0}^{\infty} a^j$ is admissible for any complex number a with $|a| = 1$, and

$$(4.2) \quad \sum_{j=0}^{\infty} a^j z^j = \frac{1}{1 - a z}$$

for every $z \in \mathbf{C}$ with $|z| < 1$, which implies that $\sum_{j=0}^{\infty} a^j$ is Abel summable to $1/(1 - a)$ when $a \neq 1$. Similarly, $\sum_{j=0}^{\infty} (j + 1) a^j$ is admissible when $|a| = 1$, and

$$(4.3) \quad \sum_{j=0}^{\infty} (j + 1) a^j z^j = \frac{1}{(1 - a z)^2}$$

for $|z| < 1$, so that $\sum_{j=0}^{\infty} j a^j$ is Abel summable to $1/(1 - a)^2$ when $a \neq 1$.

Let $\{b_j\}_{j=0}^{\infty}$ be a sequence of complex numbers, and consider

$$(4.4) \quad \beta_n = \frac{1}{n + 1} \sum_{j=0}^n b_j.$$

If $\{\beta_n\}_{n=0}^{\infty}$ converges, then

$$(4.5) \quad \frac{n}{n + 1} \beta_{n-1} = \frac{1}{n + 1} \sum_{j=0}^{n-1} b_j$$

converges to the same value. This implies that

$$(4.6) \quad \frac{b_n}{n + 1} = \beta_n - \frac{n}{n + 1} \beta_{n-1} \rightarrow 0$$

as $n \rightarrow \infty$. If $\sum_{j=0}^{\infty} a_j$ is a Cesaro summable series of complex numbers, then one can apply this to $b_n = \sum_{j=0}^n a_j$ to get that

$$(4.7) \quad \frac{a_n}{n + 1} = \frac{b_n - b_{n-1}}{n + 1} \rightarrow 0$$

as $n \rightarrow \infty$. In particular, $\sum_{j=0}^{\infty} (j + 1) a^j$ is not Cesaro summable for any $a \in \mathbf{C}$ with $|a| = 1$.

If $\sum_{j=0}^{\infty} a_j$ is Cesaro summable, then $\sum_{j=0}^{\infty} a_j$ is an admissible series, since $a_n = O(n)$. It is well known that $\sum_{j=0}^{\infty} a_j$ is Abel summable in this case, and with the same sum. For if $b_l = \sum_{j=0}^l a_j$ are the partial sums and $c_n = \sum_{l=0}^n b_l$ are their partial sums, then

$$(4.8) \quad \sum_{j=0}^{\infty} a_j z^j = (1 - z) \sum_{j=0}^{\infty} b_j z^j = (1 - z)^2 \sum_{j=0}^{\infty} c_j z^j$$

when $|z| < 1$, as in the preceding section. Equivalently,

$$(4.9) \quad \sum_{j=0}^{\infty} a_j z^j = (1-z)^2 \sum_{j=0}^{\infty} (j+1) \beta_j z^j,$$

when $|z| < 1$, where $\beta_j = c_j/(j+1)$ is the average of b_0, \dots, b_j , as before. Of course,

$$(4.10) \quad \sum_{j=0}^{\infty} (j+1) z^j = \frac{1}{(1-z)^2}$$

is the derivative of the usual geometric series. If the β_j 's were constant, then the desired conclusion would follow immediately. If $\beta_j \rightarrow 0$ as $j \rightarrow \infty$, then one can check that $\sum_{j=0}^{\infty} a_j$ is Abel summable with sum equal to 0. If $\sum_{j=0}^{\infty} a_j$ is Cesaro summable, so that $\{\beta_j\}_{j=0}^{\infty}$ converges, then one can combine these two cases to show that $\sum_{j=0}^{\infty} a_j$ is Abel summable with the same sum.

5 Cauchy products

If $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ are infinite series of complex numbers, then their *Cauchy product* is the infinite series $\sum_{n=0}^{\infty} c_n$ whose terms are given by

$$(5.1) \quad c_n = \sum_{j=0}^n a_j b_{n-j}.$$

Formally,

$$(5.2) \quad \sum_{n=0}^{\infty} c_n = \left(\sum_{j=0}^{\infty} a_j \right) \left(\sum_{l=0}^{\infty} b_l \right),$$

and in particular this holds when $a_j = b_l = 0$ for all but finitely many j and l . Moreover, (5.2) holds when $a_j = 0$ for all but finitely many j or $b_l = 0$ for all but finitely many l . Note that $\sum_{n=0}^{\infty} c_n z^n$ is the Cauchy product of $\sum_{j=0}^{\infty} a_j z^j$ and $\sum_{l=0}^{\infty} b_l z^l$ for every $z \in \mathbf{C}$.

If $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge absolutely, which means that $\sum_{j=0}^{\infty} |a_j|$ and $\sum_{l=0}^{\infty} |b_l|$ converge, then $\sum_{n=0}^{\infty} c_n$ converges absolutely, and (5.2) holds. In connection with this, observe that

$$(5.3) \quad |c_n| \leq \sum_{j=0}^n |a_j| |b_{n-j}|,$$

where the sum on the right side of the inequality corresponds exactly to the Cauchy product of $\sum_{j=0}^{\infty} |a_j|$ and $\sum_{l=0}^{\infty} |b_l|$. One can also show that $\sum_{n=0}^{\infty} c_n$ converges and satisfies (5.2) when $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge and at least one of these two series converges absolutely.

If $a_j, b_j = O(R^j)$ for some $R > 0$, then $c_n = O(n R^n)$. This implies that $\sum_{n=0}^{\infty} c_n$ is admissible when $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ are admissible. In this case,

the corresponding power series $\sum_{j=0}^{\infty} a_j z^j$, $\sum_{l=0}^{\infty} b_l z^l$, and $\sum_{n=0}^{\infty} c_n z^n$ converge absolutely for every $z \in \mathbf{C}$ with $|z| < 1$, and satisfy

$$(5.4) \quad \sum_{n=0}^{\infty} c_n z^n = \left(\sum_{j=0}^{\infty} a_j z^j \right) \left(\sum_{l=0}^{\infty} b_l z^l \right).$$

If $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ are Abel summable, then $\sum_{n=0}^{\infty} c_n$ is Abel summable, and their Abel sums satisfy (5.2). This follows from the analogous statement for the corresponding power series on the unit disk, as in the previous paragraph.

6 Norms on vector spaces

Let V be a real or complex vector space. A real-valued function N on V is said to be a *norm* if $N(v) \geq 0$ for every $v \in V$, $N(v) = 0$ if and only if $v = 0$,

$$(6.1) \quad N(tv) = |t|N(v)$$

for every $v \in V$ and real or complex number t , as appropriate, and

$$(6.2) \quad N(v+w) \leq N(v) + N(w)$$

for every $v, w \in V$. Thus the absolute value and modulus determine norms on \mathbf{R} and \mathbf{C} as one-dimensional vector spaces, respectively.

Let n be a positive integer, and consider the spaces \mathbf{R}^n and \mathbf{C}^n of n -tuples of real and complex numbers. As usual, these are vector spaces with respect to coordinatewise addition and multiplication. Consider

$$(6.3) \quad \|v\|_1 = |v_1| + \cdots + |v_n|$$

and

$$(6.4) \quad \|v\|_{\infty} = \max(|v_1|, \dots, |v_n|)$$

for $v = (v_1, \dots, v_n)$ in \mathbf{R}^n or \mathbf{C}^n . It is easy to see that $\|v\|_1$ and $\|v\|_{\infty}$ are norms on \mathbf{R}^n and \mathbf{C}^n .

The standard Euclidean norm on \mathbf{R}^n and \mathbf{C}^n is defined by

$$(6.5) \quad \|v\|_2 = (|v_1|^2 + \cdots + |v_n|^2)^{1/2}.$$

This clearly satisfies the positivity and homogeneity requirements of a norm. The triangle inequality will be discussed in the next two sections.

Observe that

$$(6.6) \quad \|v\|_{\infty} \leq \|v\|_1$$

and

$$(6.7) \quad \|v\|_{\infty} \leq \|v\|_2$$

for every v in \mathbf{R}^n or \mathbf{C}^n . One can also check that

$$(6.8) \quad \|v\|_2 \leq \|v\|_1$$

for every v in \mathbf{R}^n or \mathbf{C}^n , using the first inequality. More precisely,

$$(6.9) \quad \|v\|_2^2 \leq \|v\|_1 \|v\|_{\infty} \leq \|v\|_1^2,$$

where the first step follows directly from the definitions.

7 Inner product spaces

Let V be a real or complex vector space again. An *inner product* on V is a real or complex-valued function $\langle v, w \rangle$ defined for $v, w \in V$ with the following properties. First, $\langle v, w \rangle$ is linear as a function of v for each fixed $w \in V$. Second,

$$(7.1) \quad \langle w, v \rangle = \langle v, w \rangle$$

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

$$(7.2) \quad \langle w, v \rangle = \overline{\langle v, w \rangle}$$

for every $v, w \in V$ in the complex case. This implies that $\langle v, w \rangle$ is linear in w in the real case, and that it is conjugate-linear in the complex case. This also implies that $\langle v, v \rangle \in \mathbf{R}$ for every $v \in V$ in the complex case. The third condition asks that

$$(7.3) \quad \langle v, v \rangle > 0$$

for every $v \in V$ with $v \neq 0$. Of course, $\langle 0, 0 \rangle = 0$ by linearity.

For each $v \in V$, let $\|v\|$ be the nonnegative real number defined by

$$(7.4) \quad \|v\| = \langle v, v \rangle^{1/2}.$$

The *Cauchy–Schwarz inequality* states that

$$(7.5) \quad |\langle v, w \rangle| \leq \|v\| \|w\|$$

for every $v, w \in V$. To show this, one can start with

$$(7.6) \quad \langle v + tw, v + tw \rangle \geq 0$$

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

$$(7.7) \quad 2|t| |\langle v, w \rangle| \leq \|v\|^2 + |t|^2 \|w\|^2,$$

by expanding the inner product and collecting terms, and choosing the sign of t in the real case or $t/|t|$ in the complex case to get the absolute value or modulus of the inner product on the left. The Cauchy–Schwarz inequality follows by taking $|t| = \|v\|/\|w\|$ when $v, w \neq 0$.

Using the Cauchy–Schwarz inequality, one gets that

$$(7.8) \quad \begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &\leq \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

Hence

$$(7.9) \quad \|v + w\| \leq \|v\| + \|w\|$$

for every $v, w \in V$. Thus $\|v\|$ is a norm on V , since the positivity and homogeneity conditions are clearly satisfied.

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

$$(7.10) \quad \langle v, w \rangle = \sum_{j=1}^n v_j w_j$$

for $v = (v_1, \dots, v_n), w = (w_1, \dots, w_n) \in \mathbf{R}^n$, and

$$(7.11) \quad \langle v, w \rangle = \sum_{j=1}^n v_j \overline{w_j}$$

in the complex case. The corresponding norm

$$(7.12) \quad \|v\| = \left(\sum_{j=1}^n |v_j|^2 \right)^{1/2}$$

is the same as the standard Euclidean norm $\|v\|_2$.

8 Convexity

A set E in a real or complex vector space V is said to be *convex* if

$$(8.1) \quad t v + (1 - t) w \in E$$

for every $v, w \in E$ and real number t such that $0 < t < 1$. For example, if N is a norm on V , then the closed unit ball

$$(8.2) \quad B = \{v \in V : N(v) \leq 1\}$$

is convex.

Conversely, if N is a real-valued function on V which satisfies the positivity and homogeneity conditions of a norm, and if the corresponding closed unit ball B is convex, then one can show that N satisfies the triangle inequality and hence is a norm. For if \hat{v}, \hat{w} are nonzero vectors in V , then we can apply (8.1) with $v = \hat{v}/N(\hat{v}), w = \hat{w}/N(\hat{w})$,

$$(8.3) \quad t = \frac{N(\hat{v})}{N(\hat{v}) + N(\hat{w})},$$

and $E = B$ to get that

$$(8.4) \quad \frac{\hat{v} + \hat{w}}{N(\hat{v}) + N(\hat{w})} \in B,$$

which says exactly that

$$(8.5) \quad N(\hat{v} + \hat{w}) \leq N(\hat{v}) + N(\hat{w}),$$

as desired.

A real-valued function ϕ on the real line is convex if

$$(8.6) \quad \phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y)$$

for every $x, y, t \in \mathbf{R}$ with $0 < t < 1$. For example,

$$(8.7) \quad \phi(x) = |x|^p$$

is convex when $p \geq 1$.

If $p \geq 1$ and $v \in \mathbf{R}^n$ or \mathbf{C}^n , then put

$$(8.8) \quad \|v\|_p = \left(\sum_{j=1}^n |v_j|^p \right)^{1/p}.$$

This satisfies the positivity and homogeneity requirements of a norm, and one can use the convexity of $\phi(x) = |x|^p$ to show that the corresponding unit ball is convex, and hence that $\|v\|_p$ is a norm.

9 A few estimates

Observe that

$$(9.1) \quad \|v\|_\infty \leq \|v\|_p$$

for every $v \in \mathbf{R}^n, \mathbf{C}$ and $p \geq 1$. If $1 \leq p < q < \infty$, then

$$(9.2) \quad \|v\|_q^q \leq \|v\|_p^p \|v\|_\infty^{q-p} \leq \|v\|_p^q,$$

which implies that

$$(9.3) \quad \|v\|_q \leq \|v\|_p.$$

In the other direction, it is easy to see that

$$(9.4) \quad \|v\|_p \leq n^{1/p} \|v\|_\infty$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n and $p \geq 1$.

If $1 \leq p < q < \infty$, then

$$(9.5) \quad \|v\|_p \leq n^{1/p-1/q} \|v\|_q$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n . Equivalently,

$$(9.6) \quad \left(\frac{1}{n} \sum_{j=1}^n |v_j|^p \right)^{1/p} \leq \left(\frac{1}{n} \sum_{j=1}^n |v_j|^q \right)^{1/q}.$$

More precisely, this is the same as

$$(9.7) \quad \left(\frac{1}{n} \sum_{j=1}^n |v_j|^p \right)^{q/p} \leq \frac{1}{n} \sum_{j=1}^n |v_j|^q,$$

which can be derived from the convexity of $\phi(x) = |x|^{q/p}$ on the real line.

If N is any norm on \mathbf{R}^n or \mathbf{C}^n , then

$$(9.8) \quad N(v) \leq A \|v\|_2$$

for some $A > 0$ and every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. This can be verified using the triangle inequality to estimate $N(v)$ in terms of the norms of the standard basis vectors. The triangle inequality also implies that

$$(9.9) \quad N(v) - N(w), N(w) - N(v) \leq N(v - w)$$

for every $v, w \in \mathbf{R}^n$ or \mathbf{C}^n , and hence that

$$(9.10) \quad |N(v) - N(w)| \leq N(v - w) \leq A \|v - w\|_2.$$

This shows that N is continuous with respect to the standard topology on \mathbf{R}^n or \mathbf{C}^n .

The unit sphere in \mathbf{R}^n or \mathbf{C}^n with respect to the standard Euclidean norm consists of the vectors v such that $\|v\|_2 = 1$. By compactness, the continuous function N attains its minimum on the unit sphere, which is positive. Thus

$$(9.11) \quad N(v) \geq a$$

for some $a > 0$ and every $v \in \mathbf{R}^n$ or \mathbf{C}^n with $\|v\|_2 = 1$. It follows that

$$(9.12) \quad a \|v\|_2 \leq N(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , by homogeneity.

10 Operator norms

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

$$(10.1) \quad \|T(v)\|_W \leq A \|v\|_V$$

for every $v \in V$. If V is \mathbf{R}^n or \mathbf{C}^n and $\|v\|_V = \|v\|_p$ for some $p, 1 \leq p \leq \infty$, then every linear mapping $T : V \rightarrow W$ is bounded, as one can see by expressing any $v \in V$ as a linear combination of the standard basis vectors. This also works for any norm on \mathbf{R}^n or \mathbf{C}^n , since any norm is equivalent to the p -norms, as in the previous section. The same statement holds as well for any norm on any finite-dimensional vector space V , because V is then isomorphic to \mathbf{R}^n or \mathbf{C}^n for some n .

The *operator norm* of a bounded linear mapping $T : V \rightarrow W$ is defined by

$$(10.2) \quad \|T\|_{op} = \sup\{\|T(v)\|_W : v \in V, \|v\|_V \leq 1\}.$$

Equivalently, (10.1) holds with $A = \|T\|_{op}$, and $\|T\|_{op}$ is the smallest nonnegative real number with this property. One can check that the operator norm is a norm on the vector space of bounded linear mappings from V into W .

Suppose that V_1 , V_2 , and V_3 are vector spaces, all real or all complex, equipped with norms as before, and that $T_1 : V_1 \rightarrow V_2$ and $T_2 : V_2 \rightarrow V_3$ are bounded linear mappings. It is easy to see that the composition $T_2 \circ T_1$, defined by

$$(10.3) \quad (T_2 \circ T_1)(v) = T_2(T_1(v)), \quad v \in V_1,$$

is a bounded linear mapping from V_1 into V_3 , and that

$$(10.4) \quad \|T_2 \circ T_1\|_{op,13} \leq \|T_1\|_{op,12} \|T_2\|_{op,23},$$

where the subscripts indicate the spaces involved in the operator norms. For any normed vector space V , the identity operator I defined by $I(v) = v$ for each $v \in V$ is a bounded linear mapping from V into itself, and satisfies

$$(10.5) \quad \|I\|_{op} = 1,$$

using the same norm on V as both the domain and range.

Suppose that V is \mathbf{R}^n or \mathbf{C}^n equipped with the norm

$$(10.6) \quad \|v\|_1 = \sum_{j=1}^n |v_j|,$$

and let e_1, \dots, e_n be the standard basis of V , so that the j th coordinate of e_j is equal to 1 and the other coordinates are 0. In this case,

$$(10.7) \quad \|T\|_{op} = \max\{\|T(e_1)\|_W, \dots, \|T(e_n)\|_W\}$$

for any linear mapping $T : V \rightarrow W$. If V is any vector space with any norm $\|\cdot\|_V$ and the norm $\|\cdot\|_W$ on W is associated to an inner product $\langle \cdot, \cdot \rangle_W$, then

$$(10.8) \quad \|T\|_{op} = \sup\{|\langle T(v), w \rangle_W| : v \in V, \|v\|_V \leq 1, w \in W, \|w\|_W \leq 1\}.$$

Indeed, for each $z \in W$,

$$(10.9) \quad \|z\|_W = \sup\{|\langle z, w \rangle_W| : w \in W, \|w\|_W \leq 1\},$$

since the inner product is bounded by the norm because of the Cauchy–Schwarz inequality, and equality occurs with $w = z/\|z\|_W$ when $z \neq 0$ and with any w when $z = 0$.

11 Linear mappings

Let T be a linear mapping from a vector space V into itself, and let T^j be the composition of j factors of T for each positive integer j . Thus $T^1 = T$, $T^2 = T \circ T$, etc., and it is convenient to take T^0 to be the identity mapping I on V . As in the case of real and complex numbers, one can consider infinite series of the form

$$(11.1) \quad \sum_{j=0}^{\infty} T^j,$$

and limits of sequences of averages of the form

$$(11.2) \quad \frac{I + T + \cdots + T^n}{n}.$$

One can also consider these expressions applied to individual vectors in V .

More precisely, if V has finite dimension, then limits of sequences of vectors in V can be defined in terms of the corresponding sequences of coefficients with respect to a basis of V , and limits of linear mappings can be defined in terms of the entries of the corresponding matrices. By standard arguments, convergence is independent of the particular choice of basis of V . Convergence can also be defined with respect to a norm on any vector space, as in the next section, and is equivalent to using a basis when V has finite dimension.

Suppose for instance that $v \in V$ is an eigenvector of T with eigenvalue $\lambda \in \mathbf{R}$ or \mathbf{C} , as appropriate, so that

$$(11.3) \quad T(v) = \lambda v.$$

For each j ,

$$(11.4) \quad T^j(v) = \lambda^j v,$$

and we are back to sequences and series of real and complex numbers. If V has finite dimension and there is a basis of V consisting of eigenvectors of T , then T^j is diagonalized by the same basis for each j , and the previous sequences and series of linear mappings are reduced to sequences and series of complex numbers.

Remember that any linear mapping T on a finite-dimensional complex vector space of positive dimension has a nonzero eigenvector, as a consequence of the fundamental theorem of algebra. If T is not diagonalizable, then the behavior of T^j can still be analyzed in terms of the Jordan canonical form.

12 Convergence

Let V be a real or complex vector space equipped with a norm $\|\cdot\|$. A sequence $\{v_j\}_j$ of vectors in V is said to *converge* to $v \in V$ if for every $\epsilon > 0$ there is an L such that

$$(12.1) \quad \|v_j - v\| < \epsilon$$

for every $j \geq L$. This is the same as the usual definition of convergence of a sequence of real or complex numbers when $V = \mathbf{R}$ or \mathbf{C} and the norm is given by the absolute value or modulus. As usual, the limit of a sequence is unique when it exists.

If $\{v_j\}_j, \{w_j\}_j$ are sequences of vectors in V that converge to $v, w \in V$, respectively, then the sequence of sums $v_j + w_j$ converges to $v + w$. Similarly, if $\{v_j\}_j$ converges to v in V , and $\{t_j\}_j$ is a sequence of real or complex numbers that converges to $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then $\{t_j v_j\}_j$ converges to $t v$ in V . These statements can be verified using standard arguments. An infinite series $\sum_{j=0}^{\infty} a_j$ with terms in V converges if the corresponding sequence of partial sums

$b_n = \sum_{j=0}^n a_j$ converges in V . If $\sum_{j=0}^{\infty} a_j$ converges, then $\{a_j\}_j$ converges to 0 in V , as in the case of real or complex numbers.

If $V = \mathbf{R}^n$ or \mathbf{C}^n with norm $\|v\|_p$ for some p , $1 \leq p \leq \infty$, then a sequence $\{v_j\}_j$ converges to $v \in V$ if and only if the corresponding n sequences of coordinates of the v_j 's converge to the coordinates of v as sequences of real or complex numbers. This also works for any norm on \mathbf{R}^n or \mathbf{C}^n , since all norms on these spaces are equivalent. There are analogous statements for any finite-dimensional vector space V and any basis in V , using a linear mapping that sends the given basis to the standard basis in \mathbf{R}^n or \mathbf{C}^n , as appropriate.

Suppose that V and W are vector spaces, both real or both complex, and equipped with norms. The space of bounded linear mappings from V into W is also a vector space, and convergence of sequences and series in this space can be defined in terms of the operator norm.

13 Completeness

Let V be a real or complex vector space with a norm $\|\cdot\|$. A sequence $\{v_j\}_j$ of vectors in V is said to be a *Cauchy sequence* if for every $\epsilon > 0$ there is an L such that

$$(13.1) \quad \|v_j - v_l\| < \epsilon$$

for every $j, l \geq L$. Convergent sequences are Cauchy sequences, and V is said to be *complete* if every Cauchy sequence in V converges to an element of V . A complete vector space with respect to a norm is known as a *Banach space*, and it is a *Hilbert space* if the norm is determined by an inner product. It is well known that \mathbf{R} and \mathbf{C} are complete with respect to the usual absolute value and modulus. A sequence in \mathbf{R}^n or \mathbf{C}^n is a Cauchy sequence with respect to a p -norm if and only if the n sequences of its coordinates are Cauchy sequences in \mathbf{R} or \mathbf{C} , as appropriate. It follows that \mathbf{R}^n and \mathbf{C}^n are complete with respect to the p -norms, since \mathbf{R} and \mathbf{C} are complete. Hence \mathbf{R}^n and \mathbf{C}^n are complete with respect to any norm, by equivalence of norms. This implies in turn that finite-dimensional vector spaces are always complete.

Let $\sum_{j=0}^{\infty} a_j$ be an infinite series with terms in V . This series converges *absolutely* if

$$(13.2) \quad \sum_{j=0}^{\infty} \|a_j\|$$

converges as an infinite series of nonnegative real numbers. The sequence of partial sums of an absolutely convergent series is a Cauchy sequence, just as for absolutely convergent series of real or complex numbers. If V is a Banach space, then it follows that every absolutely convergent series of vectors in V converges in V .

Conversely, if every absolutely convergence series in V converges, then V is complete. For suppose that $\{v_j\}_j$ is a Cauchy sequence in V , and let $\{v_{j_l}\}_{l=0}^{\infty}$ be a subsequence of $\{v_j\}_j$ such that

$$(13.3) \quad \|v_{j_l} - v_{j_{l+1}}\| \leq 2^{-l}$$

for each l . Thus $\sum_{l=0}^{\infty}(v_{j_l} - v_{j_{l+1}})$ converges absolutely in V , and hence converges in V by hypothesis. This implies that $\{v_{j_l}\}_{l=0}^{\infty}$ converges in V , since

$$(13.4) \quad \sum_{l=0}^n (v_{j_l} - v_{j_{l+1}}) = v_{j_0} - v_{j_{n+1}}$$

for each n . Therefore $\{v_j\}_j$ converges in V , because a Cauchy sequence with a convergent subsequence converges to the same limit.

Let V and W be vector spaces, both real or both complex, and equipped with norms. If W is complete, then the vector space of bounded linear mappings from V into W is complete with respect to the operator norm. For if $\{T_j\}_j$ is a Cauchy sequence of bounded linear mappings from V into W , then $\{T_j(v)\}_j$ is a Cauchy sequence in W for each $v \in V$. Because W is complete, $\{T_j(v)\}_j$ converges in W , and its limit determines a linear mapping T from V into W . One can check that T is a bounded linear mapping from V into W , and that $\{T_j\}_j$ converges to T in the operator norm.

14 The supremum norm

A continuous real or complex-valued function f on a topological space X is said to be *bounded* if there is a nonnegative real number A such that

$$(14.1) \quad |f(x)| \leq A$$

for every $x \in X$. In this case, the *supremum norm* of f is defined by

$$(14.2) \quad \|f\|_{sup} = \sup\{|f(x)| : x \in X\}.$$

This is a norm on the vector space of bounded real or complex-valued continuous functions on X .

Convergence of a sequence of bounded continuous functions on X with respect to the supremum norm is the same as uniform convergence. If $\{f_j(x)\}_j$ is a Cauchy sequence of bounded continuous functions on X with respect to the supremum norm, then $\{f_j(x)\}_j$ is a Cauchy sequence of real or complex numbers for each $x \in X$, as appropriate. Hence $\{f_j\}_j$ converges pointwise to a function f on X , and one can use the Cauchy condition with respect to the supremum norm to show that $\{f_j\}_j$ converges uniformly to f . This implies that f is bounded and continuous on X , by well-known results about uniform convergence. Thus $\{f_j\}_j$ converges to f in the supremum norm, and the space of bounded continuous functions on X is complete with respect to the supremum norm.

Let V be a real or complex vector space equipped with a norm $\|v\|_V$. A continuous function f on X with values in V is said to be *bounded* if $\|f(x)\|_V$ is bounded as a real-valued function on X , in which event the supremum norm of f with respect to $\|v\|_V$ is defined by

$$(14.3) \quad \|f\|_{sup,V} = \sup\{\|f(x)\|_V : x \in X\}.$$

This is a norm on the vector space $\mathcal{C}_b(X, V)$ of bounded continuous functions on X with values in V . If V is complete with respect to $\|v\|_V$, then $\mathcal{C}_b(X, V)$ is complete with respect to $\|f\|_{sup, V}$, by an argument like the one in the previous paragraph.

For example, if $v \in V$ and ϕ is a bounded continuous real or complex-valued function on X , depending on whether V is a real or complex vector space, then $\Phi(x) = \phi(x)v$ is a bounded continuous function on X with values in V , and

$$(14.4) \quad \|\Phi\|_{sup, V} = \|\phi\|_{sup} \|v\|_V.$$

If f is bounded continuous V -valued function on X , then $\|f(x)\|_V$ is a bounded continuous real-valued function on X whose supremum norm is $\|f\|_{sup, V}$. Of course, every continuous function on X with values in V is bounded when X is compact.

15 Algebras

Let \mathcal{A} be an associative algebra over the real or complex numbers. Thus \mathcal{A} is a vector space over the real or complex numbers equipped with a binary operation of multiplication ab which is associative in the sense that

$$(15.1) \quad (ab)c = a(bc)$$

for every $a, b, c \in \mathcal{A}$. More precisely, multiplication is asked to be a bilinear mapping, which means that $a \mapsto ab$ is a linear mapping for each $b \in \mathcal{A}$, and $b \mapsto ab$ is linear for each $a \in \mathcal{A}$. A pair of elements a, b of \mathcal{A} *commute* with each other if

$$(15.2) \quad ab = ba,$$

and \mathcal{A} is a commutative algebra if this holds for every $a, b \in \mathcal{A}$.

It will be convenient to suppose also that there be a nonzero multiplicative identity element e in \mathcal{A} , which is to say that

$$(15.3) \quad ae = ea = a$$

for every $a \in \mathcal{A}$. An element a of \mathcal{A} is said to be *invertible* if there is a $b \in \mathcal{A}$ such that

$$(15.4) \quad ab = ba = e.$$

In this case, the inverse b of a is unique, and denoted a^{-1} . If a is invertible, and $c \in \mathcal{A}$ commutes with a , then c commutes with a^{-1} as well, since

$$(15.5) \quad a^{-1}c = a^{-1}ca a^{-1} = a^{-1}aca^{-1} = ca^{-1}.$$

If a_1, a_2 are invertible elements of \mathcal{A} , then their product $a_1 a_2$ is invertible too, and is given by

$$(15.6) \quad (a_1 a_2)^{-1} = a_2^{-1} a_1^{-1}.$$

Conversely, if a_1, a_2 are commuting elements of \mathcal{A} whose product $a_1 a_2$ is invertible, then a_1 and a_2 are each invertible. In this case, $a_1 a_2$ and hence its inverse commute with a_1 and a_2 , and

$$(15.7) \quad a_1^{-1} = (a_1 a_2)^{-1} a_2, \quad a_2^{-1} = a_1 (a_1 a_2)^{-1}.$$

For example, the real or complex-valued continuous functions on a topological space X form a commutative algebra with respect to pointwise multiplication of functions. The constant function equal to 1 at every element of X is the multiplicative identity element in this algebra, and a continuous function f on X is invertible in this algebra if and only if $f(x) \neq 0$ for every $x \in X$, in which event the inverse of f is given by $1/f(x)$. The bounded continuous function on X also form an algebra which is a subalgebra of the algebra of all continuous functions on X . In order for a bounded continuous function f to be invertible in this subalgebra, it is necessary that there be an $\eta > 0$ such that

$$(15.8) \quad |f(x)| \geq \eta$$

for every $x \in X$, so that $1/f(x)$ is bounded on X . If X is compact, then these two algebras are the same.

If V is a vector space, then the linear mappings on V form an algebra with composition as multiplication. The identity operator I on V is the multiplicative identity element of this algebra, and a linear mapping T on V is invertible in the algebra if and only if it is a one-to-one mapping of V onto itself. If V is equipped with a norm, then the bounded linear mappings on V with respect to this norm form an algebra which is a subalgebra of the algebra of all linear mappings on V . Invertibility of a bounded linear mapping T on V in this subalgebra means that the inverse mapping T^{-1} is also bounded. If V has finite dimension, then these two algebras are the same.

16 Banach algebras

Let \mathcal{A} be an associative algebra with nonzero multiplicative identity element e over the real or complex numbers, and suppose that \mathcal{A} is equipped with a norm $\|a\|$. If

$$(16.1) \quad \|a b\| \leq \|a\| \|b\|$$

for every $a, b \in \mathcal{A}$ and

$$(16.2) \quad \|e\| = 1,$$

then $(\mathcal{A}, \|a\|)$ is said to be a *normed algebra*. In particular, this implies that the product of a pair of convergent sequences in \mathcal{A} converges to the product of the limits of the sequences, just as for products of convergent sequences of real or complex numbers.

If \mathcal{A} is complete with respect to $\|a\|$, then \mathcal{A} is said to be a *Banach algebra*. For example, the algebra of bounded continuous real or complex-valued functions on a topological space is a Banach algebra.

Let V be a vector space over the real or complex numbers equipped with a norm. The algebra of bounded linear mappings on V is a normed algebra with respect to the operator norm. If V is complete, then the algebra of bounded linear operators on V is a Banach algebra.

17 Invertibility

Let $(\mathcal{A}, \|a\|)$ be a normed algebra with nonzero multiplicative identity element e . For each $a \in \mathcal{A}$ and positive integer j , let a^j be the product of j factors of a , so that $a^1 = a$, $a^2 = a a$, etc., with $a^0 = e$. Thus

$$(17.1) \quad \|a^j\| \leq \|a\|^j$$

for each j . If $\|a\| < 1$, then $\{a^j\}_{j=0}^{\infty}$ converges to 0 in \mathcal{A} ,

$$(17.2) \quad \sum_{j=0}^{\infty} \|a^j\| \leq \sum_{j=0}^{\infty} \|a\|^j = \frac{1}{1 - \|a\|},$$

and hence $\sum_{j=0}^{\infty} a^j$ converges absolutely in \mathcal{A} .

If $\|a\| < 1$ and \mathcal{A} is a Banach algebra, then it follows that $\sum_{j=0}^{\infty} a^j$ converges in \mathcal{A} . For each $n \geq 0$,

$$(17.3) \quad (e - a) \left(\sum_{j=0}^n a^j \right) = \left(\sum_{j=0}^n a^j \right) (e - a) = e - a^{n+1},$$

which implies that

$$(17.4) \quad (e - a) \left(\sum_{j=0}^{\infty} a^j \right) = \left(\sum_{j=0}^{\infty} a^j \right) (e - a) = e,$$

since $a^{n+1} \rightarrow 0$ as $n \rightarrow \infty$. Therefore, $e - a$ is invertible, with

$$(17.5) \quad (e - a)^{-1} = \sum_{j=0}^{\infty} a^j.$$

This is a fundamental property of Banach algebras.

Let x be an invertible element of \mathcal{A} , so that

$$(17.6) \quad 1 = \|x x^{-1}\| \leq \|x\| \|x^{-1}\|.$$

If $y \in \mathcal{A}$ satisfies

$$(17.7) \quad \|x - y\| < \frac{1}{\|x^{-1}\|},$$

then

$$(17.8) \quad y = x - (x - y) = x(e - x^{-1}(x - y))$$

is invertible by the remarks of the previous paragraph. In particular, the invertible elements of \mathcal{A} form an open set.

If $\|a\| < 1$, then

$$(17.9) \quad \|(e - a)^{-1} - e\| = \left\| \sum_{j=1}^{\infty} a^j \right\| \leq \sum_{j=1}^{\infty} \|a\|^j = \frac{\|a\|}{1 - \|a\|}.$$

Hence $x \mapsto x^{-1}$ is a continuous mapping on the set of invertible elements of \mathcal{A} .

18 Submultiplicative sequences

A sequence $\{r_j\}_{j=1}^{\infty}$ of nonnegative real numbers is said to be *submultiplicative* if

$$(18.1) \quad r_{j+l} \leq r_j r_l$$

for every $j, l \geq 1$. In this case,

$$(18.2) \quad \lim_{n \rightarrow \infty} r_n^{1/n} = \inf_{n \geq 1} r_n^{1/n},$$

where the existence of the limit is part of the conclusion. To see this, observe that

$$(18.3) \quad r_{j+n+l} \leq (r_n)^j (r_1)^l$$

and hence

$$(18.4) \quad (r_{j+n+l})^{1/(j+n+l)} \leq [(r_n)^{1/n}]^{jn/(j+n+l)} (r_1)^{l/(j+n+l)}$$

for every $j, l, n \geq 1$. If n is fixed, then the right side can be approximated by $(r_n)^{1/n}$ for j sufficiently large and $0 \leq l < n$.

19 Invertibility, 2

For each element x of a Banach algebra \mathcal{A} , $r_j = \|x^j\|$ is a submultiplicative sequence of nonnegative real numbers. Put

$$(19.1) \quad \rho(x) = \lim_{n \rightarrow \infty} \|x^n\|^{1/n} = \inf_{n \geq 1} \|x^n\|^{1/n} \leq \|x\|,$$

which is known as the *spectral radius* of x , at least in the case of complex Banach algebras. Thus

$$(19.2) \quad \rho(tx) = |t| \rho(x)$$

for each real or complex number t , as appropriate.

If $\rho(x) < 1$, then $\sum_{j=0}^{\infty} \|x^j\|$ converges, and hence $\sum_{j=0}^{\infty} x^j$ converges in \mathcal{A} . As before, the sum is equal to the inverse of $e - x$ under these conditions. Note that $\rho(x) < 1$ if and only if $\|x^n\|^{1/n} < 1$ for some n , which is the same as $\|x^n\| < 1$. Invertibility of $e - x$ can also be obtained from the invertibility of $e - x^n$, since the latter is the product of the former and $e + x + \cdots + x^{n-1}$.

If $\rho(x) \leq 1$ and t is a real or complex number, such that $|t| < 1$, as appropriate, then consider

$$(19.3) \quad \sum_{j=0}^{\infty} t^j x^j = (e - tx)^{-1}.$$

If $e - x$ is invertible, then $(e - tx)^{-1}$ extends continuously to a neighborhood of $t = 1$. Conversely, if there is a sequence $\{t_j\}_j$ such that $|t_j| < 1$ for each j , $\{t_j\}_j$ converges to 1, and $(e - t_j x)^{-1}$ converges in \mathcal{A} , then $e - x$ is invertible and the limit of $(e - t_j x)^{-1}$ is equal to its inverse.

Suppose that \mathcal{A} is the algebra of bounded continuous real or complex-valued functions on a topological space X , with the supremum norm. In this case,

$$(19.4) \quad \|f^n\|_{sup} = \|f\|_{sup}^n$$

for each $n \geq 1$ and f , and hence $\rho(f) = \|f\|_{sup}$.

20 Spectrum

Let T be a linear mapping on a real or complex vector space V . To say that a real or complex number λ , as appropriate, is an eigenvalue of T means exactly that $T - \lambda I$ has nontrivial kernel. In particular, $T - \lambda I$ is not invertible. Conversely, if V has finite dimension and $T - \lambda I$ is not invertible, then $T - \lambda I$ has nontrivial kernel, and λ is an eigenvalue of T .

Suppose that V is equipped with a norm $\|\cdot\|_V$, and that T is a bounded linear operator on V . If v is a nonzero eigenvector of T with eigenvalue λ , then

$$(20.1) \quad |\lambda| \|v\|_V = \|T(v)\|_V \leq \|T\|_{op} \|v\|_V$$

implies that

$$(20.2) \quad |\lambda| \leq \|T\|_{op}.$$

Similarly, λ^n is an eigenvalue of T^n for each positive integer n , and hence

$$(20.3) \quad |\lambda|^n \leq \|T^n\|_{op}.$$

Thus

$$(20.4) \quad |\lambda| \leq \|T^n\|_{op}^{1/n}$$

for each n , and therefore $|\lambda| \leq \rho(T)$.

Let $(\mathcal{A}, \|\cdot\|)$ be a Banach algebra with nonzero multiplicative identity element e , let x be an element of \mathcal{A} , and let λ be a real or complex number, as appropriate. If $|\lambda| > \|x\|$, then

$$(20.5) \quad \lambda e - x = \lambda(e - \lambda^{-1}x)$$

is invertible. The same conclusion holds when $|\lambda| > \rho(x)$. Equivalently,

$$(20.6) \quad |\lambda| \leq \rho(x)$$

when $\lambda e - x$ is not invertible in \mathcal{A} .

The set

$$(20.7) \quad \sigma(x) = \{\lambda : \lambda e - x \text{ is not invertible in } \mathcal{A}\}$$

is known as the *spectrum* of x in \mathcal{A} , especially in the complex case. This is a closed set in \mathbf{R} or \mathbf{C} , as appropriate, since the complementary *resolvent* set of λ such that $\lambda e - x$ is invertible is an open set when \mathcal{A} is a Banach algebra. If \mathcal{A} is a complex Banach algebra, then a famous theorem states that $\sigma(x) \neq \emptyset$ for every $x \in \mathcal{A}$. Basically, if $\sigma(x) = \emptyset$, then

$$(20.8) \quad (\lambda e - x)^{-1}$$

would be a nonconstant holomorphic \mathcal{A} -valued function on the complex plane that tends to 0 as $|\lambda| \rightarrow \infty$, a contradiction. This is still a holomorphic \mathcal{A} -valued function on the complement of $\sigma(x)$ in the complex plane for any $x \in \mathcal{A}$. Another famous theorem states that

$$(20.9) \quad \rho(x) = \max\{|\lambda| : \lambda \in \sigma(x)\}$$

when \mathcal{A} is a complex Banach algebra, as a consequence of the convergence of

$$(20.10) \quad \sum_{j=0}^{\infty} \alpha^j x^j$$

when α is a nonzero complex number such that $|\lambda| < 1/|\alpha|$ for every $\lambda \in \sigma(x)$.

21 Averages in normed algebras

Let $(\mathcal{A}, \|\cdot\|)$ be a normed algebra with nonzero multiplicative identity element e . If $x \in \mathcal{A}$ and $\|x\| < 1$, then

$$(21.1) \quad \left\| \sum_{j=0}^n x^j \right\| \leq \sum_{j=0}^n \|x^j\| \leq \sum_{j=0}^n \|x\|^j \leq \frac{1}{1 - \|x\|}$$

for each n . Hence

$$(21.2) \quad \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{j=0}^n x^j = 0.$$

Suppose now that $\|x\| = 1$. Thus

$$(21.3) \quad \left\| \frac{1}{n+1} \sum_{j=0}^n x^j \right\| \leq \frac{1}{n+1} \sum_{j=0}^n \|x^j\| \leq \frac{1}{n+1} \sum_{j=0}^n \|x\|^j = 1$$

for each n . Of course,

$$(21.4) \quad \frac{1}{n+1} \sum_{j=0}^n x^j = e$$

for each n when $x = e$.

For any $x \in \mathcal{A}$ and $n \geq 1$,

$$(21.5) \quad (e - x) \sum_{j=0}^n x^j = e - x^{n+1}.$$

This implies that

$$(21.6) \quad \sum_{j=0}^n x^j = (e - x)^{-1} (e - x^{n+1})$$

when $e - x$ is invertible in \mathcal{A} . If $\|x\| = 1$ and $e - x$ is invertible, then

$$(21.7) \quad \left\| \sum_{j=0}^n x^j \right\| \leq \|(e - x)^{-1}\| \|e - x^{n+1}\| \leq 2 \|(e - x)^{-1}\|,$$

so that (21.2) holds in this case too.

If $e - x$ is invertible, then

$$(21.8) \quad \begin{aligned} \sum_{l=0}^n \sum_{j=0}^l x^j &= \sum_{l=0}^n (e - x)^{-1} (e - x^{l+1}) \\ &= (n + 1) (e - x)^{-1} - (e - x)^{-2} (x - x^{n+2}) \end{aligned}$$

for each n . If in addition $\|x\| = 1$, then it follows that

$$(21.9) \quad \lim_{n \rightarrow \infty} \frac{1}{n + 1} \sum_{l=0}^n \sum_{j=0}^l x^j = (e - x)^{-1}.$$

22 Invertibility, 3

Let \mathcal{A} be a Banach algebra with nonzero multiplicative identity element e . Suppose that $x \in \mathcal{A}$ has the property that the sums

$$(22.1) \quad \sum_{j=0}^n x^j$$

are uniformly bounded in \mathcal{A} . In particular, $x^0 = e, x, x^2, \dots$ is a bounded sequence in \mathcal{A} , so that

$$(22.2) \quad \sum_{j=0}^{\infty} r^j x^j$$

converges absolutely when $0 \leq r < 1$. Using summation by parts, we get that (22.2) is equal to

$$(22.3) \quad (1 - r) \sum_{l=0}^{\infty} r^l \left(\sum_{j=0}^l x^j \right).$$

Thus the boundedness of (22.1) implies the boundedness of (22.2). Hence the inverse of $e - rx$ has bounded norm when $0 \leq r < 1$, since it is given by (22.2). It follows that $e - x$ is invertible, by the results of Section 17.

23 The open mapping theorem

Let $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ be Banach spaces, both real or both complex, and let T be a bounded linear mapping from V into W . If T maps V onto W , then Banach's open mapping theorem says that T sends open subsets of V to open subsets of W . In particular, if T is a one-to-one mapping of V onto W , then it follows that T^{-1} is a bounded linear mapping from W onto V .

For each $r > 0$, put

$$(23.1) \quad B_V(r) = \{v \in V : \|v\|_V < r\},$$

and let $B_W(r)$ be defined in the same way. Because T is linear, it suffices to show that there is an $r > 0$ such that

$$(23.2) \quad T(B_V(1)) \supseteq B_W(r).$$

The hypothesis that T map V onto W implies that

$$(23.3) \quad \bigcup_{n=1}^{\infty} T(B_V(n)) = W.$$

Hence W is the union of the closure $\overline{T(B_V(n))}$ of $T(B_V(n))$, $n \geq 1$, and it follows from the Baire category theorem that $\overline{T(B_V(n))}$ contains a nonempty open set for some n . Using linearity, one can show that there is an $r_1 > 0$ such that

$$(23.4) \quad \overline{T(B_V(1))} \supseteq B_W(r_1).$$

Let us use completeness of V and linearity of T to show that

$$(23.5) \quad T(B_V(2)) \supseteq B_W(r_1).$$

Let $w \in B_W(r_1)$ be given. By (23.4), there is a $v_0 \in B_V(1)$ such that

$$(23.6) \quad \|w - T(v_0)\|_W < \frac{r_1}{2}.$$

Applying the same argument to $2(w - T(v_0))$, we get that there is a $v_1 \in B_V(1/2)$ such that

$$(23.7) \quad \|w - T(v_0) - T(v_1)\|_W < \frac{r_1}{4}.$$

Repeating the process, we get $v_0, v_1, v_2, \dots \in V$ such that

$$(23.8) \quad \|v_j\|_V < 2^{-j}$$

for each j , and

$$(23.9) \quad \left\| w - \sum_{j=0}^l T(v_j) \right\| < 2^{-l-1} r_1$$

for each l . Thus

$$(23.10) \quad \sum_{j=0}^{\infty} \|v_j\|_V < 2,$$

and so $\sum_{j=0}^{\infty} v_j$ converges in V and satisfies

$$(23.11) \quad \left\| \sum_{j=0}^{\infty} v_j \right\|_V < 2.$$

Moreover,

$$(23.12) \quad T\left(\sum_{j=0}^{\infty} v_j\right) = w,$$

which implies (23.5).

24 The uniform boundedness principle

Let V and W be vector spaces, both real or both complex, equipped with norms $\|\cdot\|_V$ and $\|\cdot\|_W$, respectively. Let T_1, T_2, \dots be a sequence of bounded linear mappings from V into W such that $\{T_j(v)\}_{j=1}^{\infty}$ is a bounded sequence in W for each $v \in V$. If V is complete, then a theorem of Banach and Steinhaus implies that the operator norms of the T_j 's are bounded.

As a variant of this, let f_1, f_2, \dots be a sequence of nonnegative real-valued continuous functions on a metric space M such that $\{f_j(x)\}_{j=1}^{\infty}$ is bounded for each $x \in M$. If M is complete, then there is a nonempty open set in M on which the f_j 's are uniformly bounded. To see this, consider

$$(24.1) \quad E_n = \{x \in M : f_j(x) \leq n \text{ for each } j\},$$

which is a closed set in M for each n by continuity. The hypothesis of pointwise boundedness means exactly that

$$(24.2) \quad \bigcup_{n=1}^{\infty} E_n = M,$$

and the Baire category theorem implies that E_n contains a nonempty open set for some n , as desired.

Let us apply this to $M = V$ and $f_j(v) = \|T_j(v)\|_W$. Note that $\|T_j(v)\|_W$ is a continuous function on V , since $T_j : V \rightarrow W$ is a bounded linear mapping. Because of linearity, uniform boundedness of $\|T_j(v)\|_W$ on a nonempty open set in V implies that the operator norms of the T_j 's are bounded.

In practice, we are especially interested in situations where $\{T_j(v)\}_{j=1}^{\infty}$ converges in W for each $v \in V$.

25 Strong operator convergence

Let V and W be vector spaces, both real or both complex, with norms $\|\cdot\|_V$ and $\|\cdot\|_W$. A sequence T_1, T_2, \dots of bounded linear mappings from V into W is said to converge *strongly* to a linear mapping T from V into W if the following two conditions are satisfied. First, the operator norms of the T_j 's are uniformly bounded, so that there is an $A \geq 0$ such that

$$(25.1) \quad \|T_j\|_{op} \leq A$$

for each j . Second,

$$(25.2) \quad \lim_{j \rightarrow \infty} T_j(v) = T(v)$$

in W for each $v \in V$. It follows that T is a bounded linear mapping from V into W , with

$$(25.3) \quad \|T\|_{op} \leq A.$$

Of course, convergence in the operator norm implies strong convergence. If V is complete, then the pointwise convergence of a sequence of bounded linear mappings on V implies the boundedness of their operator norms, as in the previous section. Conversely, if the operator norms of the T_j 's are uniformly bounded, and if T is a bounded linear mapping from V into W , then convergence of $T_j(v)$ to $T(v)$ for every v in a dense set in V implies the same property for every $v \in V$. Similarly, if the operator norms of the T_j 's are uniformly bounded, and if $\{T_j(v)\}_{j=1}^{\infty}$ is a Cauchy sequence in W for each v in a dense set in V , then $\{T_j(v)\}_{j=1}^{\infty}$ is a Cauchy sequence in W for every $v \in V$. If W is complete, then it follows that $\{T_j\}_{j=1}^{\infty}$ converges strongly to a bounded linear mapping from V into W .

26 Convergence of averages

Let V be a vector space with a norm $\|\cdot\|$, and let T be a bounded linear operator on V with $\|T\|_{op} \leq 1$. Thus

$$(26.1) \quad \left\| \frac{1}{n+1} \sum_{j=0}^n T^j \right\|_{op} \leq \frac{1}{n+1} \sum_{j=0}^n \|T^j\|_{op} \leq \frac{1}{n+1} \sum_{j=0}^n \|T\|_{op}^j \leq 1$$

for each n .

For $v \in V$, let us consider the convergence in V of

$$(26.2) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(v).$$

If $T(v) = v$, then $T^j(v) = v$ for each j , and (26.2) is equal to v for every n .

If $v = T(u) - u$ for some $u \in V$, then

$$(26.3) \quad \sum_{j=0}^n T^j(v) = T^{n+1}(u) - u$$

for each n . Hence (26.2) tends to 0 as $n \rightarrow \infty$, since $T^{n+1}(u)$ is bounded. The same conclusion holds when v is in the closure of the set of $T(u) - u$, $u \in V$.

Observe that

$$(26.4) \quad T\left(\frac{1}{n+1} \sum_{j=0}^n T^j(v)\right) - \frac{1}{n+1} \sum_{j=0}^n T^j(v) = \frac{1}{n+1}(T^{n+1}(v) - v)$$

converges to 0 as $n \rightarrow \infty$ for every $v \in V$. If (26.2) converges for some $v \in V$, then it follows that the limit is an eigenvector of T with eigenvalue 1.

27 Hilbert spaces

Let $(V, \langle \cdot, \cdot \rangle)$ be a Hilbert space, and let E be a nonempty closed convex set in V . Let $v \in V$ be given, and consider

$$(27.1) \quad r = \inf\{\|v - w\| : w \in E\}.$$

Let w_1, w_2, \dots be a sequence of elements of E such that

$$(27.2) \quad \lim_{j \rightarrow \infty} \|v - w_j\| = r.$$

Because E is convex, $(w_j + w_l)/2 \in E$ for every j, l , and hence

$$(27.3) \quad \left\|v - \frac{w_j + w_l}{2}\right\| \geq r.$$

The parallelogram law

$$(27.4) \quad \|a - b\|^2 + \|a + b\|^2 = 2\|a\|^2 + 2\|b\|^2$$

with $a = v - w_j$, $b = v - w_l$ implies that

$$(27.5) \quad \|w_j - w_l\|^2 + \left\|2v - w_j - w_l\right\|^2 = 2\|v - w_j\|^2 + 2\|v - w_l\|^2.$$

It follows that

$$(27.6) \quad \lim_{j, l \rightarrow \infty} \|w_j - w_l\| = 0,$$

which is to say that $\{w_j\}_{j=1}^{\infty}$ is a Cauchy sequence in V . This sequence converges, because V is complete, and its limit w is an element of E and satisfies

$$(27.7) \quad \|v - w\| = r.$$

Let us apply this to a closed linear subspace W of V . For each $v \in V$, the preceding argument implies that there is a $w \in W$ such that

$$(27.8) \quad \|v - w\| \leq \|v - w - z\|$$

for every $z \in W$. By standard computations, this implies in turn that

$$(27.9) \quad \langle v - w, z \rangle = 0$$

for every $z \in W$, in the same way that the minimum of a function is attained at a critical point. Conversely, the latter condition implies that

$$(27.10) \quad \|v - w - z\|^2 = \|v - w\|^2 + \|z\|^2 \geq \|v - w\|^2$$

for every $z \in W$. If $u \in W$ also satisfies $\langle v - u, z \rangle = 0$ for every $z \in W$, then

$$(27.11) \quad \|u - w\|^2 = \langle u - w, u - w \rangle = \langle u - v, u - w \rangle + \langle v - w, u - w \rangle = 0,$$

since $u - w \in W$, and hence $u = w$.

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

$$(27.12) \quad W^\perp = \{y \in V : \langle y, z \rangle = 0 \text{ for every } z \in W\}.$$

The previous arguments show that every element of V can be expressed in a unique way as the sum of an element of W and an element of W^\perp . Let us check that

$$(27.13) \quad W = (W^\perp)^\perp.$$

Every element of W is contained in $(W^\perp)^\perp$ by definition, and so it suffices to check that $x \in (W^\perp)^\perp$ is in W . If $x = w + y$ for some $w \in W \subseteq (W^\perp)^\perp$ and $y \in W^\perp$, then it follows that $y \in (W^\perp)^\perp$, so that $\langle y, y \rangle = 0$, $y = 0$, and $x = w$. If W is any linear subspace of V , then W^\perp can be defined in the same way, and is a closed linear subspace of V . Of course, the closure \overline{W} of W is a closed linear subspace of V , and one can check that

$$(27.14) \quad \overline{W}^\perp = W^\perp.$$

Therefore $\overline{W} = (\overline{W}^\perp)^\perp = (W^\perp)^\perp$.

If W is a closed linear subspace of V , then the *orthogonal projection* $P = P_W$ of V onto W is the linear mapping on V defined by

$$(27.15) \quad P(v) = w$$

for $v = w + y$ with $w \in W$, $y \in W^\perp$. Thus

$$(27.16) \quad \|P(v)\| \leq \|v\|$$

for every $v \in V$, since

$$(27.17) \quad \|v\|^2 = \|w\|^2 + \|y\|^2$$

in this case. In particular, $\|P\|_{op} = 1$ except when $W = \{0\}$ and $P = 0$.

28 Unitary transformations

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. A one-to-one linear mapping T from V onto itself is said to be *unitary* if

$$(28.1) \quad \langle T(v), T(w) \rangle = \langle v, w \rangle$$

for every $v, w \in V$. This may also be described as an orthogonal transformation in the real case. If T is unitary, then

$$(28.2) \quad \|T(v)\| = \|v\|$$

for every $v \in V$, by taking $v = w$ in the previous equation. Conversely, a linear mapping T of V onto itself which preserves norms is one-to-one and satisfies (28.1), by a polarization argument. Unitary transformations are obviously bounded, with operator norm 1. Note that the inverse of a unitary transformation is also unitary.

Suppose now that V is complete, which is to say that V is a Hilbert space. Consider the linear subspace

$$(28.3) \quad W = \{T(u) - u : u \in V\}.$$

Thus $T(W) = W$, since $T(V) = V$. Because T is unitary,

$$(28.4) \quad \langle y, T(u) - u \rangle = \langle T^{-1}(y) - y, u \rangle$$

for every $u, y \in V$. It follows that W^\perp consists of the $y \in V$ such that $T^{-1}(y) = y$, which is equivalent to $T(y) = y$.

As in the previous section, every $v \in V$ can be expressed in a unique way as $w + y$, where $w \in \overline{W}$, $y \in W^\perp$, and hence $T(y) = y$. This implies that

$$(28.5) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(v) = \frac{1}{n+1} \sum_{j=0}^n T^j(w) + y$$

converges to y in V as $n \rightarrow \infty$, as in Section 26.

Let T_W be the restriction of T to \overline{W} , and let I_W be the identity mapping on \overline{W} . If $I_W - T_W$ has a bounded inverse on \overline{W} , then

$$(28.6) \quad \frac{1}{n+1} \sum_{j=0}^n (T_W)^j$$

converges to 0 in the operator norm on \overline{W} as $n \rightarrow \infty$, as in Section 21. In this case,

$$(28.7) \quad \frac{1}{n+1} \sum_{j=0}^n T^j$$

converges to the orthogonal projection of V onto W^\perp in the operator norm.

29 Rotations

Put

$$(29.1) \quad \mathbf{T} = \{\alpha \in \mathbf{C} : |\alpha| = 1\}.$$

For each $\alpha \in \mathbf{T}$, let R_α be the linear mapping on the vector space of continuous complex-valued functions on \mathbf{T} defined by

$$(29.2) \quad R_\alpha(f)(z) = f(\alpha z)$$

for each $z \in \mathbf{T}$. This is an isometry of the space of continuous functions on \mathbf{T} onto itself with respect to the supremum norm.

If $f(z) = z^l$ for some integer l , then

$$(29.3) \quad R_\alpha(f) = \alpha^l f,$$

so that f is an eigenvector of R_α with eigenvalue α^l . Note that $f(z)$ can also be expressed as \bar{z}^{-l} , since $|z| = 1$. It follows that

$$(29.4) \quad \frac{1}{n+1} \sum_{j=0}^n R_\alpha^j(f)$$

converges uniformly on \mathbf{T} as $n \rightarrow \infty$ when f is one of these eigenfunctions, or a finite linear combination of these eigenfunctions.

It is well known that the finite linear combinations of the functions z^l , $l \in \mathbf{Z}$, are dense in the space of continuous functions on \mathbf{T} with respect to the supremum norm. This implies that (29.4) converges uniformly on \mathbf{T} for every continuous function f on \mathbf{T} .

One can also consider the Lebesgue spaces $L^p(\mathbf{T})$ of measurable complex-valued functions f on \mathbf{T} such that $|f|^p$ is integrable with respect to Lebesgue measure on \mathbf{T} , $1 \leq p < \infty$. For each $\alpha \in \mathbf{T}$, R_α determines an isometric linear mapping of $L^p(\mathbf{T})$ onto itself, which is a unitary mapping when $p = 2$. The sum (29.4) also converges in the L^p norm for every $f \in L^p(\mathbf{T})$ when $1 \leq p < \infty$, since it converges for a dense class of functions.

30 Fourier series

If f is a continuous complex-valued function on the unit circle \mathbf{T} and $\alpha \in \mathbf{T}$, then

$$(30.1) \quad \int_{\mathbf{T}} f(\alpha z) |dz| = \int_{\mathbf{T}} f(z) |dz|.$$

Here $|dz|$ denotes the element of arc length integration on \mathbf{T} , which is invariant under rotations. This implies that

$$(30.2) \quad \int_{\mathbf{T}} z^l |dz| = 0$$

for every nonzero integer l .

For any pair of continuous functions f_1, f_2 on \mathbf{T} , put

$$(30.3) \quad \langle f_1, f_2 \rangle_{\mathbf{T}} = \frac{1}{2\pi} \int_{\mathbf{T}} f_1(z) \overline{f_2(z)} |dz|.$$

This defines an inner product on the vector space of complex-valued continuous functions on \mathbf{T} , which extends to the Lebesgue space $L^2(\mathbf{T})$ of square-integrable functions on \mathbf{T} , so that the latter becomes a Hilbert space.

Observe that

$$(30.4) \quad \begin{aligned} \langle z^j, z^l \rangle_{\mathbf{T}} &= 0 && \text{when } j \neq l \\ &= 1 && \text{when } j = l. \end{aligned}$$

Thus the functions $z^l, l \in \mathbf{Z}$, are orthonormal with respect to this inner product.

The Fourier coefficients of a function f on \mathbf{T} are defined by

$$(30.5) \quad \widehat{f}(j) = \langle f, z^j \rangle_{\mathbf{T}} = \frac{1}{2\pi} \int_{\mathbf{T}} f(z) z^{-j} |dz|,$$

and the Fourier series associated to f is

$$(30.6) \quad \sum_{j=-\infty}^{\infty} \widehat{f}(j) z^j.$$

It is well known that the Abel sums of the Fourier series of a continuous function f on \mathbf{T} , given by

$$(30.7) \quad \sum_{j=-\infty}^{\infty} \widehat{f}(j) r^{|j|} z^j, \quad 0 < r < 1,$$

converge to $f(z)$ uniformly on \mathbf{T} as $r \rightarrow 1$, because these sums can be expressed as averages of f concentrated near z on \mathbf{T} . Similarly, the Cesaro means

$$(30.8) \quad \frac{1}{n+1} \sum_{l=0}^n \sum_{j=-l}^l \widehat{f}(j) z^j$$

converge to $f(z)$ uniformly on \mathbf{T} as $n \rightarrow \infty$.

31 Measure-preserving transformations

Let X be a set with a σ -algebra of measurable subsets and a positive measure μ . Let ϕ be a one-to-one mapping of X onto itself such that $E \subseteq X$ is measurable if and only if $\phi(E)$ is measurable, and

$$(31.1) \quad \mu(\phi(E)) = \mu(E)$$

for every measurable set $E \subseteq X$. Consider the linear mapping

$$(31.2) \quad T(f) = f \circ \phi$$

acting on measurable functions f on X . If f is an integrable function on X , then $T(f)$ is too, and

$$(31.3) \quad \int_X T(f) d\mu = \int_X f d\mu.$$

Similarly, T determines an isometry of $L^p(X)$ onto itself for each p , $1 \leq p \leq \infty$.

In particular, T defines a unitary mapping on the Hilbert space $L^2(X)$. As in Section 28,

$$(31.4) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(f)$$

converges in $L^2(X)$ for each $f \in L^2(X)$. For any p ,

$$(31.5) \quad \left\| \frac{1}{n+1} \sum_{j=0}^n T^j(f) \right\|_p \leq \frac{1}{n+1} \sum_{j=0}^n \|T^j(f)\|_p = \|f\|_p$$

for each $f \in L^p(X)$ and $n \geq 0$, since T is an isometry on $L^p(X)$. If f is in $L^1(X)$ and $L^\infty(X)$, then $f \in L^p(X)$ for $1 < p < \infty$, and the previous statements imply that the averages converge in $L^p(X)$ for $1 < p < \infty$. Because $L^1(X) \cap L^\infty(X)$ is dense in $L^p(X)$ for $1 < p < \infty$, it follows that the averages converge in $L^p(X)$ for every $f \in L^p(X)$ when $1 < p < \infty$.

If $\mu(X) < \infty$, then $L^2(X)$ is already a dense linear subspace of $L^1(X)$. In this case, convergence in $L^2(X)$ implies convergence in $L^1(X)$, and so the averages converge in $L^1(X)$ for every $f \in L^2(X)$. Hence the averages converge in $L^1(X)$ for every $f \in L^1(X)$.

Suppose that X is the set \mathbf{Z} of integers, and that μ is counting measure on \mathbf{Z} . If $\phi(l) = l + 1$ for each n and f is a function equal to 1 at one point and to 0 elsewhere, then the averages have L^1 norm equal to 1 for each n , but converge to 0 uniformly on X .

32 Sequence spaces

Let A be a finite set with at least two elements, and let X be the set of doubly-infinite sequences $x = \{x_j\}_{j=-\infty}^{\infty}$ with $x_j \in A$ for each j . Thus X is a compact Hausdorff space with respect to the product topology defined using the discrete topology on A . Let $\phi : X \rightarrow X$ be the shift mapping such that the j th term of $\phi(x)$ is equal to x_{j+1} . This is a homeomorphism from X onto itself. Suppose that to each $a \in A$ is assigned a nonnegative real number $w(a)$ such that

$$(32.1) \quad \sum_{a \in A} w(a) = 1.$$

This defines a probability measure on A , which leads to a probability measure on X . This probability measure on X is preserved by the shift mapping ϕ .

33 Maximal functions

Let f be a locally-integrable function on the real line. Consider the Hardy–Littlewood maximal function f^* associated to f , defined by

$$(33.1) \quad f^*(x) = \sup_{x \in I} \frac{1}{|I|} \int_I |f(y)| dy,$$

where the supremum is taken over all open intervals $I = (a, b)$ in \mathbf{R} that contain x , and where $|I|$ denotes the length $b - a$ of I . Thus $f^*(x) \geq 0$ by construction, and $f^*(x) = +\infty$ is certainly possible. The mapping $f \mapsto f^*$ is sublinear in the sense that

$$(33.2) \quad (f_1 + f_2)^* \leq f_1^* + f_2^*$$

for any locally-integrable functions f on \mathbf{R} , and

$$(33.3) \quad (tf)^* = |t|f^*$$

for any locally-integrable function f on \mathbf{R} and constant t . If f is bounded, then f^* is bounded, and

$$(33.4) \quad \|f^*\|_\infty \leq \|f\|_\infty.$$

If f is any locally-integrable function on \mathbf{R} , then f^* is lower semicontinuous. This means that

$$(33.5) \quad E_\lambda = \{x \in \mathbf{R} : f^*(x) > \lambda\}$$

is an open set for each $\lambda \geq 0$. For if $f^*(x) > \lambda$ for some $x \in \mathbf{R}$ and $\lambda \geq 0$, then

$$(33.6) \quad \frac{1}{|I|} \int_I |f(y)| dy > \lambda$$

for some open interval I in \mathbf{R} that contains x . This implies in turn that $f^* > \lambda$ on I , and hence that $I \subseteq E_\lambda$, as desired.

If f is an integrable function on \mathbf{R} , then

$$(33.7) \quad |E_\lambda| \leq \frac{2}{\lambda} \int_{\mathbf{R}} |f(y)| dy$$

for every $\lambda > 0$, where $|E_\lambda|$ is the Lebesgue measure of E_λ . In particular, $f^* < \infty$ almost everywhere on \mathbf{R} . To prove this estimate, it suffices to show that

$$(33.8) \quad |K| \leq \frac{2}{\lambda} \int_{\mathbf{R}} |f(y)| dy$$

for every compact set $K \subseteq E_\lambda$.

By compactness, there are finitely many open intervals I_1, \dots, I_n such that

$$(33.9) \quad \frac{1}{|I_i|} \int_{I_i} |f(y)| dy > \lambda$$

for each l , and

$$(33.10) \quad K \subseteq \bigcup_{l=1}^n I_l.$$

We may also suppose that each $x \in \mathbf{R}$ is contained in at most two of the I_l 's. For if a point is contained in three of these intervals, then one of the intervals is contained in the union of the other two, and may be dropped from the collection. This implies that

$$(33.11) \quad \begin{aligned} |K| \leq \sum_l |I_l| &\leq \sum_l \frac{1}{\lambda} \int_{I_l} |f(y)| dy \\ &\leq \frac{2}{\lambda} \int_{\bigcup_l I_l} |f(y)| dy \\ &\leq \frac{2}{\lambda} \int_{\mathbf{R}} |f(y)| dy, \end{aligned}$$

as desired.

34 L^p estimates

Let ϕ be a nonnegative measurable function on the real line, and let p be a real number with $p \geq 1$. Observe that

$$(34.1) \quad \begin{aligned} \int_{\mathbf{R}} \phi(x)^p dx &= \int_{\mathbf{R}} \int_0^{\phi(x)} p \lambda^{p-1} d\lambda dx \\ &= \int_0^\infty p \lambda^{p-1} |\{x \in \mathbf{R} : \phi(x) > \lambda\}| d\lambda. \end{aligned}$$

We would like to apply this to $\phi = f^*$, where $f \in L^p(\mathbf{R})$ and $p > 1$.

For each $\lambda > 0$, let f_λ be the function defined by

$$(34.2) \quad \begin{aligned} f_\lambda(x) &= f(x) \quad \text{when } |f(x)| \leq \frac{\lambda}{2} \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Thus

$$(34.3) \quad f^*(x) \leq f_\lambda^*(x) + (f - f_\lambda)^*(x) \leq \frac{\lambda}{2} + (f - f_\lambda)^*(x)^*$$

for every $x \in \mathbf{R}$, by subadditivity of f^* and (33.4), which implies that

$$(34.4) \quad \{x \in \mathbf{R} : f^*(x) > \lambda\} \subseteq \left\{x \in \mathbf{R} : (f - f_\lambda)^*(x) > \frac{\lambda}{2}\right\}.$$

Combining this with (33.7) applied to $f - f_\lambda$, we get that

$$(34.5) \quad |\{x \in \mathbf{R} : f^*(x) > \lambda\}| \leq \frac{4}{\lambda} \int_{\{y \in \mathbf{R} : |f(y)| > \lambda/2\}} |f(y)| dy$$

for every $\lambda > 0$. Hence

$$(34.6) \quad \int_{\mathbf{R}} f^*(x)^p dx \leq \int_0^\infty 4p \lambda^{p-2} \int_{\{y \in \mathbf{R}: |f(y)| > \lambda/2\}} |f(y)| dy d\lambda.$$

Interchanging the order of integration, we get that

$$(34.7) \quad \int_{\mathbf{R}} f^*(x)^p dx \leq \int_{\mathbf{R}} \int_0^{2|f(y)|} 4p \lambda^{p-2} |f(y)| d\lambda dy.$$

If $p > 1$, then it follows that

$$(34.8) \quad \int_{\mathbf{R}} f^*(x)^p dx \leq \frac{4p 2^{p-1}}{p-1} \int_{\mathbf{R}} |f(y)|^p dy.$$

If f is not equal to 0 almost everywhere on \mathbf{R} , then one can check that $f^*(x)$ is at least a positive constant multiple of $1/|x|$ when $|x|$ is sufficiently large. This implies that f^* is not integrable on \mathbf{R} , although it can still be locally integrable.

35 Maximal functions, 2

Let $(M, d(x, y))$ be a metric space. For each $x \in M$ and $r > 0$, the open ball with center x and radius r is defined by

$$(35.1) \quad B(x, r) = \{y \in M : d(x, y) < r\}.$$

If $B = B(x, r)$ is an open ball and $a > 0$, then aB denotes the ball $B(x, ar)$.

Suppose that B_1, B_2, \dots is a finite sequence of open balls in M , or an infinite sequence of balls whose radii converge to 0. Let β_1 be one of these balls with maximum radius, and let β_2 be another one of these balls which is disjoint from β_1 and has maximal radius, if there is one. Continuing in this manner, we get a finite or infinite sequence of pairwise disjoint balls β_1, β_2, \dots chosen from B_1, B_2, \dots such that for each B_i there is a β_j which intersects B_i and has radius at least that of B_i . Hence $B_i \subseteq 3\beta_j$, and so

$$(35.2) \quad \bigcup_i B_i \subseteq \bigcup_j 3\beta_j.$$

A positive Borel measure μ on M is said to be *doubling* if the measure of each open ball in M is positive and finite, and if there is a $C > 0$ such that

$$(35.3) \quad \mu(2B) \leq C \mu(B)$$

for every open ball B in M . If f is a locally integrable function on M with respect to μ , then the corresponding maximal function f^* is defined by

$$(35.4) \quad f^*(x) = \sup_{x \in B} \frac{1}{\mu(B)} \int_B |f| d\mu,$$

where the supremum is taken over all open balls B in M with $x \in B$. As in [7, 8], one can show that f^* satisfies properties like those in the previous sections, using the covering lemma described in the preceding paragraph in particular.

For example, this applies to Lebesgue measure on \mathbf{R}^n equipped with the standard metric. As in [32], one can also consider \mathbf{R}^{n+1} equipped with a metric associated to parabolic dilations. In addition to nonstandard dilations, there can be more complicated geometric structure related to nilpotent Lie groups. Many common fractals are equipped with natural doubling measures.

36 Discrete maximal functions

If f is a function on the set \mathbf{Z} of integers, then the corresponding maximal function is defined by

$$(36.1) \quad f^*(l) = \sup_{a \leq l \leq b} \frac{1}{b-a+1} \sum_{j=a}^b |f(j)|,$$

where more precisely the supremum is taken over all integers a, b with $a \leq l \leq b$, and may be infinite. As before, $f \mapsto f^*$ is sublinear, f^* is bounded when f is bounded, and satisfies

$$(36.2) \quad \sup_{l \in \mathbf{Z}} f^*(l) \leq \sup_{l \in \mathbf{Z}} |f(l)|.$$

If we identify functions on \mathbf{Z} with functions on \mathbf{R} that are constant on intervals of the form $[l, l+1)$ for $l \in \mathbf{Z}$ and have the same values at the integers, then this maximal function is obviously less than or equal to the previous one on \mathbf{R} . Using this observation, or arguments analogous to the earlier ones, it is easy to see that

$$(36.3) \quad |\{j \in \mathbf{Z} : f^*(j) > \lambda\}| \leq \frac{2}{\lambda} \sum_{j=-\infty}^{\infty} |f(j)|$$

for every $\lambda > 0$, where $|E|$ is the number of elements of $E \subseteq \mathbf{Z}$, and

$$(36.4) \quad \sum_{j=-\infty}^{\infty} f^*(j)^p \leq \frac{4p2^{p-1}}{p-1} \sum_{j=-\infty}^{\infty} |f(j)|^p$$

when $1 < p < \infty$.

37 Another variant

Let X be a set with a positive measure μ , and consider $X \times \mathbf{Z}$ with the measure which is the product of μ and counting measure. If f is a measurable function on $X \times \mathbf{Z}$, which is to say that $f(x, l)$ is measurable in x for each l , then consider the maximal function f^* defined by

$$(37.1) \quad f^*(x, l) = \sup_{a \leq l \leq b} \frac{1}{b-a+1} \sum_{j=a}^b |f(x, j)|,$$

which is the maximal function of $f(x, l)$ in l for each $x \in X$. Thus $f \mapsto f^*$ is sublinear, f^* is bounded when f is bounded, and

$$(37.2) \quad \|f\|_\infty \leq \|f\|_\infty.$$

Moreover,

$$(37.3) \quad \sum_{j=-\infty}^{\infty} \mu(\{x \in X : f^*(x, j) > \lambda\}) \leq \frac{2}{\lambda} \sum_{j=-\infty}^{\infty} \int_X |f(x, j)| d\mu(x)$$

for each $\lambda > 0$, and

$$(37.4) \quad \sum_{j=-\infty}^{\infty} \int_X f^*(x, j)^p d\mu(x) \leq \frac{4p2^{p-1}}{p-1} \sum_{j=-\infty}^{\infty} \int_X |f(x, j)|^p d\mu(x)$$

when $1 < p < \infty$.

38 Transference

Let X be a set with a positive measure μ , let ϕ be a one-to-one measure-preserving mapping of X onto itself, and let

$$(38.1) \quad T(f) = f \circ \phi$$

be the corresponding linear transformation on measurable functions on X . Consider the maximal function

$$(38.2) \quad A^*(f)(x) = \sup_{k, l \geq 0} \frac{1}{k+l+1} \sum_{j=-k}^l |T^j(f)(x)|,$$

where the supremum is taken over nonnegative integers k, l . It is also convenient to consider the approximations to this defined by

$$(38.3) \quad A_n^*(f)(x) = \max_{0 \leq k, l \leq n} \frac{1}{k+l+1} \sum_{j=-k}^l |T_j(f)(x)|.$$

Thus $A_n^*(f)(x)$ is monotone increasing in n , and tends to $A^*(f)(x)$ as $n \rightarrow \infty$. As usual, A^* and A_n^* are sublinear operators that send bounded functions to bounded functions, with

$$(38.4) \quad \|A_n^*(f)\|_\infty \leq \|A^*(f)\|_\infty \leq \|f\|_\infty.$$

By construction, they also commute with T , which is to say that

$$(38.5) \quad A_n^*(T(f)) = T(A_n^*(f)), \quad A^*(T(f)) = T(A^*(f)).$$

In particular,

$$(38.6) \quad \mu(\{x \in X : A_n^*(T^l(f))(x) > \lambda\}) = \mu(\{x \in X : A_n^*(f)(x) > \lambda\})$$

for each n , l , and λ , and

$$(38.7) \quad \int_X A_n^*(T^l(f))^p d\mu = \int_X A_n^*(f)^p d\mu$$

for each n , l , and p .

Let N be a large positive integer. If f is a measurable function on X , then let $F(x, j)$ be the function on $X \times \mathbf{Z}$ defined by

$$(38.8) \quad F(x, j) = T^j(f)(x)$$

when $0 \leq j \leq N$, and $F(x, j) = 0$ otherwise. Because ϕ is measure-preserving,

$$(38.9) \quad \sum_{j=-\infty}^{\infty} \int_X |F(x, j)|^p d\mu(x) = (N+1) \int_X |f|^p d\mu$$

for each p . Let $F^*(x, l)$ be the maximal function in the second variable, as in the previous section. If $n \leq l \leq N - n$, then

$$(38.10) \quad A_n^*(T^l(f))(x) \leq F^*(x, l).$$

It follows that

$$(38.11) \quad \begin{aligned} (N - 2n) \mu(\{x \in X : A_n^*(f)(x) > \lambda\}) \\ \leq \sum_{j=-\infty}^{\infty} \mu(\{x \in X : F^*(x, j) > \lambda\}) \end{aligned}$$

for each λ , and hence

$$(38.12) \quad \begin{aligned} (N - 2n) \mu(\{x \in X : A_n^*(f)(x) > \lambda\}) \\ \leq \frac{2}{\lambda} \sum_{j=-\infty}^{\infty} \int_X |F(x, j)| d\mu(x) \\ = (N + 1) \frac{2}{\lambda} \int_X |f| d\mu. \end{aligned}$$

Equivalently,

$$(38.13) \quad \mu(\{x \in X : A_n^*(f)(x) > \lambda\}) \leq \frac{(N + 1)}{(N - 2n)} \frac{2}{\lambda} \int_X |f| d\mu$$

for every $N > 2n$, and so

$$(38.14) \quad \mu(\{x \in X : A_n^*(f)(x) > \lambda\}) \leq \frac{2}{\lambda} \int_X |f| d\mu,$$

which implies that

$$(38.15) \quad \mu(\{x \in X : A^*(f)(x) > \lambda\}) \leq \frac{2}{\lambda} \int_X |f| d\mu$$

for every λ .

Similarly, for each p ,

$$(38.16) \quad (N - 2n) \int_X A_n^*(f)^p d\mu \leq \sum_{j=-\infty}^{\infty} \int_X F^*(x, j)^p d\mu(x).$$

If $1 < p < \infty$, then

$$(38.17) \quad \begin{aligned} (N - 2n) \int_X A_n^*(f)^p d\mu &\leq \frac{4p2^{p-1}}{p-1} \sum_{j=-\infty}^{\infty} \int_X |F(x, j)|^p d\mu(x) \\ &= (N + 1) \frac{4p2^{p-1}}{p-1} \int_X |f|^p d\mu. \end{aligned}$$

Thus

$$(38.18) \quad \int_X A_n^*(f)^p d\mu \leq \frac{(N + 1)}{(N - 2n)} \frac{4p2^{p-1}}{p-1} \int_X |f|^p d\mu$$

when $N > 2n$, which implies that

$$(38.19) \quad \int_X A_n^*(f)^p d\mu \leq \frac{4p2^{p-1}}{p-1} \int_X |f|^p d\mu,$$

and hence

$$(38.20) \quad \int_X A^*(f)^p d\mu \leq \frac{4p2^{p-1}}{p-1} \int_X |f|^p d\mu.$$

39 Almost-everywhere convergence

A standard application of maximal function estimates is to the existence of pointwise limits almost everywhere. For example, if f is a locally integrable function on the real line, then

$$(39.1) \quad \frac{1}{|I|} \int_I |f(y) - f(x)| dy \rightarrow 0$$

for almost every $x \in \mathbf{R}$, where the limit is taken for open intervals I with $x \in I$ as $|I| \rightarrow 0$. One may as well suppose that f is integrable here, or even has compact support, since the problem is local. The limit is trivial when f is continuous, and any integrable function can be approximated in the L^1 norm by a continuous function. Using maximal function estimates, one can show that

the approximations by continuous functions also work for the pointwise limits almost everywhere.

One can also show that maximal functions associated to the Abel and Cesaro sums of the Fourier series of an integrable function are bounded by the analogue of the Hardy–Littlewood maximal function on the circle. Hence these maximal functions satisfy the same sort of estimates as before. This can be used to show that the Abel and Cesaro sums converge to the function almost everywhere, which could be derived from the limit described in the previous paragraph as well, and which is basically the same type of statement anyway.

Suppose now that X is a set with a positive measure μ , ϕ is a measure-preserving transformation on X , and that $T(f) = f \circ \phi$ for measurable functions f on X . If $f \in L^p(X)$, $1 < p < \infty$, then

$$(39.2) \quad \int_X \sum_{n=0}^{\infty} \frac{|T^n(f)(x)|^p}{(n+1)^p} d\mu(x) = \sum_{n=0}^{\infty} \frac{1}{(n+1)^p} \int_X |T^n(f)(x)|^p d\mu \\ = \sum_{n=0}^{\infty} \frac{1}{(n+1)^p} \int_X |f|^p d\mu < \infty.$$

Hence

$$(39.3) \quad \sum_{n=0}^{\infty} \frac{|T^n(f)(x)|^p}{(n+1)^p} < \infty$$

for almost every $x \in X$, and thus

$$(39.4) \quad \lim_{n \rightarrow \infty} \frac{T^n(f)(x)}{n+1} = 0$$

for almost every $x \in X$. Of course,

$$(39.5) \quad \frac{\|T^n(f)\|_{\infty}}{n+1} = \frac{\|f\|_{\infty}}{n+1} \rightarrow 0$$

as $n \rightarrow \infty$ when f is bounded X .

For each $x \in X$, consider the sequence of averages

$$(39.6) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(f)(x).$$

If $T(f) = f$, then

$$(39.7) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(f)(x) = f(x)$$

for each n , and converge as $n \rightarrow \infty$ trivially. If $f = T(b) - b$ for some $b \in L^p(X)$, $1 < p \leq \infty$, then

$$(39.8) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(f)(x) = \frac{T^{n+1}(b)(x) - b(x)}{n+1}$$

converges to 0 almost everywhere on X as $n \rightarrow \infty$, by the remarks in the preceding paragraph. As in Section 28, the functions of the form

$$(39.9) \quad a + (T(b) - b), \quad a, b \in L^2(X), T(a) = a,$$

are dense in $L^2(X)$, since T defines a unitary transformation on $L^2(X)$. Thus the averages (39.6) converge almost everywhere on X for a dense class of functions in $L^2(X)$. Using the boundedness of the maximal function on $L^2(X)$, one can show that the averages converge almost everywhere for any $f \in L^2(X)$. This also holds for $f \in L^p(X)$ when $1 \leq p < \infty$, because $L^p(X) \cap L^2(X)$ is dense in $L^p(X)$, and using the maximal function estimates on $L^p(X)$.

40 Multiplication operators

Let (X, μ) be a measure space, and let b be a bounded measurable function on X . The corresponding multiplication operator is defined by

$$(40.1) \quad T(f) = bf.$$

This determines a bounded linear operator on $L^p(X)$ for $1 \leq p \leq \infty$, with operator norm equal to the essential supremum norm of b for each p .

In particular, if $\{b_j\}_j$ is a sequence of bounded measurable functions on X that converges to b in the L^∞ norm, then the corresponding sequence $\{T_j\}_j$ of multiplication operators converges to T in the operator norm on $L^p(X)$ for each p , $1 \leq p \leq \infty$. If instead $\{b_j\}_j$ is a sequence of bounded measurable functions with uniformly bounded L^∞ norms that converges to b pointwise almost everywhere on X , then the dominated convergence theorem implies that $\{T_j(f)\}_j$ converges to $T(f)$ in the L^p norm for every $f \in L^p(X)$ when $1 \leq p < \infty$.

For example, if $\|b\|_\infty < 1$, then b^j converges to 0 in the L^∞ norm as $j \rightarrow \infty$. If $\|b\|_\infty = 1$ and $|b(x)| < 1$ for almost every $x \in X$, then $\|b^j\|_\infty = 1$ for each j and $\lim_{j \rightarrow \infty} b(x)^j = 0$ almost everywhere on X .

Suppose that $\|b\|_\infty \leq 1$, and consider

$$(40.2) \quad a_n = \frac{1}{n+1} \sum_{j=0}^n b^j.$$

Thus

$$(40.3) \quad \frac{1}{n+1} \sum_{j=0}^n T^j(f) = a_n f.$$

Clearly $\|a_n\|_\infty \leq 1$ for each n . As in Section 2, $\{a_n\}_n$ converges pointwise to the function that is equal to 1 when b is equal to 1 and to 0 elsewhere. Hence the sequence of averages of the T_j 's converges strongly to the operator of multiplication by the characteristic function of the set where $b = 1$ on $L^p(X)$ when $1 \leq p < \infty$.

41 Doubling spaces

A metric space $(M, d(x, y))$ is said to be *doubling* if there is a $C > 0$ such that every ball in M of radius r is contained in the union of $\leq C$ balls of radius $r/2$. Thus every ball of radius r can also be covered by $\leq C^2$ balls of radius $r/4$, etc. Equivalently, M is doubling if there is a $C' > 0$ such that every set in M of diameter t can be covered by $\leq C'$ subsets of diameter $\leq t/2$.

For example, it is easy to see that finite-dimensional Euclidean spaces are doubling. In this case, it suffices to consider coverings of the unit ball, because a covering of any other ball can be reduced to the unit ball using translations and dilations. A similar argument works on nilpotent Lie groups under suitable conditions. If $(M, d(x, y))$ is a metric space which is doubling and $E \subseteq M$, then E is doubling with respect to the restriction of $d(x, y)$ to E . In particular, subsets of finite-dimensional Euclidean spaces are doubling with respect to the induced metric.

Let M be a metric space with a doubling measure μ , and let B be a ball of radius r in M . Suppose that $A \subseteq B$ has the property that $d(x, y) \geq r/2$ for every $x, y \in A$ with $x \neq y$. Thus the balls $B(x, r/2)$, $x \in A$, are pairwise disjoint, and $B(x, r/2) \subseteq (3/2)B$ for every $x \in B$, and hence

$$(41.1) \quad \sum_{x \in A} \mu(B(x, r/2)) \leq \mu((3/2)B).$$

This implies an upper bound for the number of elements of A in terms of the doubling constant for μ , since

$$(41.2) \quad B \subseteq B(x, 2r)$$

for each $x \in B$, so that $\mu(B)$ is less than or equal to a constant multiple of $\mu(B(x, r/2))$. If A is a maximal set of this type, then for each $z \in B$ there is an $x \in A$ such that $d(x, z) \leq r/2$, which means that

$$(41.3) \quad B \subseteq \bigcup_{x \in A} B(x, r/2),$$

from which it follows that M is doubling.

Remember that a set E in a metric space M is said to be *totally bounded* if for each $\epsilon > 0$, E can be covered by finitely many balls of radius ϵ . Totally bounded subsets of any metric space are automatically bounded, and bounded subsets of a doubling metric space are totally bounded. If M is complete, then $K \subseteq M$ is compact if and only if K is closed and totally bounded. If M is also doubling, then it follows that $K \subseteq M$ is compact if and only if K is closed and bounded.

42 Ultrametrics

A metric $d(x, y)$ on a set M is said to be an *ultrametric* if

$$(42.1) \quad d(x, z) \leq \max(d(x, y), d(y, z))$$

for every $x, y, z \in M$, which is stronger than the usual triangle inequality. Some examples of ultrametric spaces are given in the next section.

Let $(M, d(x, y))$ be an ultrametric space. If $x, y \in M$ and $r, t > 0$, then either

$$(42.2) \quad B(x, r) \subseteq B(y, t),$$

or

$$(42.3) \quad B(y, t) \subseteq B(x, r),$$

or

$$(42.4) \quad B(x, r) \cap B(y, t) = \emptyset.$$

More precisely, the first alternative holds when $d(x, y)$ and r are less than or equal to t , the second alternative holds when $d(x, y)$ and t are less than or equal to r , and the third alternative holds when r and t are strictly less than $d(x, y)$.

Suppose that \mathcal{B} is a collection of open balls in M . The remarks of the previous paragraph imply that the maximal elements of \mathcal{B} are pairwise disjoint. Under suitable conditions, every element of \mathcal{B} is contained in a maximal element.

This is a very simple covering argument for ultrametric spaces. Doubling conditions are still important, even if they are not required for a covering lemma.

43 Sequence spaces, 2

Let A be a set with at least two elements, and let M be the set of sequences $x = \{x_j\}_{j=1}^{\infty}$ with $x_j \in A$ for each j . Fix a positive real number $\rho < 1$, and put

$$(43.1) \quad d(x, y) = \rho^n$$

when $x, y \in M$, $x \neq y$, and n is the largest nonnegative integer such that $x_j = y_j$ for $j \leq n$, and otherwise put $d(x, y) = 0$. This defines an ultrametric on M for which the corresponding topology is the product topology, when M is considered as the product of infinitely many copies of A equipped with the discrete topology. Let us restrict our attention to the case where A has only finitely many elements, so that M is compact. It is easy to see that M is doubling in this case. Let $w(a)$ be a positive real number for each $a \in A$ whose sum is equal to 1. This defines a probability measure on A which leads to a probability measure μ on M , and one can check that μ is a doubling measure on M .

44 Snowflake transforms

If a, x, y are nonnegative real numbers and $0 < a < 1$, then

$$(44.1) \quad (x + y)^a \leq x^a + y^a.$$

This is because

$$(44.2) \quad x + y \leq (x^a + y^a) \max(x, y)^{1-a}$$

and

$$(44.3) \quad \max(x, y) \leq (x^a + y^a)^{1/a}$$

imply that $x + y \leq (x^a + y^a)^{1/a}$.

If $(M, d(x, y))$ is a metric space and $0 < a < 1$, then it follows that $d(x, y)^a$ is also a metric on M . If $d(x, y)$ is an ultrametric on M , then $d(x, y)^a$ is an ultrametric on M for every $a > 0$. Note that $d(x, y)^a$ determines the same topology on M as $d(x, y)$ in both situations.

For each $p \in M$ and $r > 0$, the open ball in M centered at p with radius r with respect to $d(x, y)$ is the same as the open ball centered at p with radius r^a with respect to $d(x, y)^a$. Using this fact, it is easy to see that M is doubling with respect to $d(x, y)$ if and only if it is doubling with respect to $d(x, y)^a$. Similarly, a positive Borel measure μ on X is doubling with respect to $d(x, y)$ if and only if it is doubling with respect to $d(x, y)^a$. Of course, the doubling constants for $d(x, y)^a$ are not normally the same as for $d(x, y)$.

The diameter of a set $E \subseteq M$ with respect to $d(x, y)^a$ is equal to the a th power of the diameter of E with respect to $d(x, y)$. It is easy to see that the t -dimensional Hausdorff measure of $E \subseteq M$ with respect to $d(x, y)$ is equal to the (t/a) -dimensional Hausdorff measure of E with respect to $d(x, y)^a$, using the preceding observation and the definition of Hausdorff measures. In particular, the Hausdorff dimension of E with respect to $d(x, y)^a$ is equal to the Hausdorff dimension of E with respect to $d(x, y)$ divided by a .

45 Invariant measures

Let X be a compact Hausdorff topological space, and let $\mathcal{C}(X)$ be the space of continuous real-valued functions on X equipped with the supremum norm. A linear functional λ on $\mathcal{C}(X)$ is said to be *bounded* if it is bounded as a linear mapping into \mathbf{R} as a one-dimensional vector space with the absolute value as norm. Thus λ is bounded if there is an $A \geq 0$ such that

$$(45.1) \quad |\lambda(f)| \leq A \|f\|_{sup}$$

for every $f \in \mathcal{C}(X)$, in which case the dual norm $\|\lambda\|_*$ of λ is the smallest value of A for which this condition holds. This defines a norm on the vector space $\mathcal{C}(X)^*$ of bounded linear functionals on $\mathcal{C}(X)$, which is the same as the operator norm of λ as a linear mapping from $\mathcal{C}(X)$ into \mathbf{R} .

A linear functional λ on $\mathcal{C}(X)$ is said to be *nonnegative* if

$$(45.2) \quad \lambda(f) \geq 0$$

for every $f \in \mathcal{C}(X)$ which is nonnegative in the sense that $f(x) \geq 0$ for every $x \in X$. If λ is a nonnegative linear functional on $\mathcal{C}(X)$, then λ is bounded, and

$$(45.3) \quad \|\lambda\|_* = \lambda(1),$$

where the right side means λ applied to the constant function equal 1 on X . Every nonnegative linear functional on $\mathcal{C}(X)$ can be expressed by integration

with respect to a positive Borel measure on X , by the Riesz representation theorem.

A set $U \subseteq \mathcal{C}(X)^*$ is an open set in the weak* topology if for every $\lambda \in U$ there are finitely many continuous functions f_1, \dots, f_n on X and positive real numbers r_1, \dots, r_n such that

$$(45.4) \quad \{\lambda' \in \mathcal{C}(X)^* : |\lambda'(f_1) - \lambda(f_1)| < r_1, \dots, |\lambda'(f_n) - \lambda(f_n)| < r_n\} \subseteq U.$$

The Banach–Alaoglu theorem implies that

$$(45.5) \quad B^* = \{\lambda \in \mathcal{C}(X)^* : \|\lambda\|_* \leq 1\}$$

is compact with respect to this topology. If the topology on X is determined by a metric, then $\mathcal{C}(X)$ is separable, and one can show that the topology on B^* induced by the weak* topology is determined by a metric.

Let ϕ be a homeomorphism of X onto itself, and let

$$(45.6) \quad T(f) = f \circ \phi$$

be the corresponding linear operator on $\mathcal{C}(X)$. This leads to a dual linear operator T^* on $\mathcal{C}(X)^*$, defined by

$$(45.7) \quad T^*(\lambda)(f) = \lambda(T(f)).$$

Note that $T(f) \geq 0$ when $f \geq 0$, and hence $T^*(\lambda)$ is nonnegative when λ is nonnegative. Similarly, T preserves the supremum norm on $\mathcal{C}(X)$, and so T^* preserves the dual norm on $\mathcal{C}(X)^*$. Of course, T automatically sends constant functions to themselves on X .

Consider

$$(45.8) \quad \lambda_n = \frac{1}{n+1} \sum_{j=0}^n (T^*)^j(\lambda)$$

for each nonnegative integer n and $\lambda \in \mathcal{C}(X)^*$. Observe that

$$(45.9) \quad T^*(\lambda_n) - \lambda_n = \frac{1}{n+1} ((T^*)^{n+1}(\lambda) - \lambda),$$

which implies that

$$(45.10) \quad \|T^*(\lambda_n) - \lambda_n\|_* \leq \frac{2\|\lambda\|_*}{n+1}.$$

In particular,

$$(45.11) \quad \lim_{n \rightarrow \infty} \|T^*(\lambda_n) - \lambda_n\|_* = 0$$

Let E_l be the closure of $\lambda_l, \lambda_{l+1}, \dots$ in the weak* topology. By compactness,

$$(45.12) \quad E = \bigcap_{l=0}^{\infty} E_l \neq \emptyset.$$

If the topology on X is determined by a metric, then there is a subsequence of $\lambda_0, \lambda_1, \dots$ that converges in the weak* topology, and whose limit is in E . Every element of E is invariant under T^* , by the remarks of the previous paragraph.

If λ is nonnegative, then λ_n is nonnegative for each n , and every element of E is nonnegative. In this case, $\lambda_n(1) = \lambda(1)$ for each n , and every element of E has the same value when applied to the constant function 1. It follows that λ_n and every element of E has the same norm as λ when $\lambda \geq 0$. Using this, one can check that there is a nonnegative linear functional on $\mathcal{C}(X)$ with norm 1 which is invariant under T^* , and which corresponds to an invariant probability measure on X .

References

- [1] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, American Mathematical Society, 1997.
- [2] L. Ambrosio and P. Tilli, *Topics on Analysis in Metric Spaces*, Oxford University Press, 2004.
- [3] W. Arveson, *A Short Course on Spectral Theory*, Springer-Verlag, 2002.
- [4] R. Beals, *Topics in Operator Theory*, University of Chicago Press, 1971.
- [5] R. Beals, *Analysis: An Introduction*, Cambridge University Press, 2004.
- [6] S. Bochner and K. Chandrasekharan, *Fourier Transforms*, Princeton University Press, 1949.
- [7] R. Coifman and G. Weiss, *Analyse Harmonique Non-Commutative sur certains Espaces Homogènes*, Lecture Notes in Mathematics **242**, Springer-Verlag, 1971.
- [8] R. Coifman and G. Weiss, *Transference Methods in Analysis*, American Mathematical Society, 1976.
- [9] R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis*, Bulletin of the American Mathematical Society **83** (1977), 569–645.
- [10] J. Conway, *A Course in Functional Analysis*, 2nd edition, Springer-Verlag, 1990.
- [11] G. David and S. Semmes, *Fracture Fractals and Broken Dreams: Self-Similar Geometry through Metric and Measure*, Oxford University Press, 1997.
- [12] R. DeVore and R. Sharpley, *Maximal Functions Measuring Smoothness*, Memoirs of the American Mathematical Society **293**, 1984.

- [13] J. Duoandikoetxea, *Fourier Analysis*, translated and revised from the 1985 Spanish original by D. Cruz-Urbe, American Mathematical Society, 2001.
- [14] Y. Eidelman, V. Milman, and A. Tsolomitis, *Functional Analysis: An Introduction*, American Mathematical Society, 2004.
- [15] K. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, 1986.
- [16] K. Falconer, *Techniques in Fractal Geometry*, Wiley, 1997.
- [17] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 2nd edition, Wiley, 2003.
- [18] H. Federer, *Geometric Measure Theory*, Springer-Verlag, 1969.
- [19] G. Folland, *Real Analysis*, 2nd edition, Wiley, 1999.
- [20] G. Folland and E. Stein, *Hardy Spaces on Homogeneous Groups*, Princeton University Press, 1982.
- [21] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.
- [22] J. García-Cuerva and J. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland, 1984.
- [23] R. Goldberg, *Methods of Real Analysis*, 2nd edition, Wiley, 1976.
- [24] P. Halmos, *Lectures on Ergodic Theory*, Chelsea, 1960.
- [25] P. Halmos, *Finite-Dimensional Vector Spaces*, Springer-Verlag, 1974.
- [26] P. Halmos, *A Hilbert Space Problem Book*, 2nd edition, Springer-Verlag, 1982.
- [27] P. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, 2nd edition, AMS Chelsea Publishing, 1998.
- [28] V. Hansen, *Functional Analysis: Entering Hilbert Space*, World Scientific, 2006.
- [29] J. Heinonen, *Lectures on Analysis on Metric Spaces*, Springer-Verlag, 2001.
- [30] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, 1975.
- [31] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton University Press, 1941.
- [32] B. Jones, *A Class of Singular Integrals*, American Journal of Mathematics **86** (1964), 441–462.

- [33] F. Jones, *Lebesgue Integration on Euclidean Spaces*, Jones and Bartlett, 1993.
- [34] J.-L. Journé, *Calderón–Zygmund Operators, Pseudodifferential Operators, and the Cauchy Integral of Calderón*, Lecture Notes in Mathematics **994**, Springer-Verlag, 1983.
- [35] Y. Katznelson, *An Introduction to Harmonic Analysis*, 3rd edition, Cambridge University Press, 2004.
- [36] J. Kigami, *Analysis on Fractals*, Cambridge University Press, 2001.
- [37] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.
- [38] A. Knapp, *Advanced Real Analysis*, Birkhäuser, 2005.
- [39] S. Krantz, *A Panorama of Harmonic Analysis*, Mathematical Association of America, 1999.
- [40] S. Krantz, *Function Theory of Several Complex Variables*, 2nd edition, AMS Chelsea Publishing, 2001.
- [41] S. Krantz, *Real Analysis and Foundations*, 2nd edition, Chapman & Hall / CRC, 2005.
- [42] S. Krantz and H. Parks, *The Geometry of Domains in Space*, Birkhäuser, 1999.
- [43] S. Lang, *Real and Functional Analysis*, 3rd edition, Springer-Verlag, 1993.
- [44] P. Lax, *Functional Analysis*, Wiley, 2002.
- [45] R. Macías and C. Segovia, *Lipschitz functions on spaces of homogeneous type*, Advances in Mathematics **33** (1979), 257–270.
- [46] R. Macías and C. Segovia, *A decomposition into atoms of distributions on spaces of homogeneous type*, Advances in Mathematics **33** (1979), 271–309.
- [47] R. Mañé, *Ergodic Theory and Differentiable Dynamics*, translated from the Portuguese by S. Levy, Springer-Verlag, 1987.
- [48] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, Cambridge University Press, 1995.
- [49] M. Nadkarni, *Basic Ergodic Theory*, 2nd edition, Birkhäuser, 1998.
- [50] P. Nicholls, *The Ergodic Theory of Discrete Groups*, Cambridge University Press, 1989.
- [51] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1989.
- [52] M. Pollicott, *Lectures on Ergodic Theory and Pesin Theory on Compact Manifolds*, Cambridge University Press, 1993.

- [53] M. Pollicott and M. Yuri, *Dynamical Systems and Ergodic Theory*, Cambridge University Press, 1998.
- [54] S. Promislow, *A First Course in Functional Analysis*, Wiley, 2008.
- [55] C. Rickart, *General Theory of Banach Algebras*, van Nostrand, 1960.
- [56] C. Rogers, *Hausdorff Measures*, with a foreword by K. Falconer, Cambridge University Press, 1998.
- [57] H. Royden, *Real Analysis*, 3rd edition, Macmillan, 1988.
- [58] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.
- [59] W. Rudin, *Function Theory in the Unit Ball of \mathbf{C}^n* , Springer-Verlag, 1980.
- [60] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, 1987.
- [61] W. Rudin, *Functional Analysis*, 2nd edition, McGraw-Hill, 1991.
- [62] D. Ruelle, *Dynamical Zeta Functions for Piecewise Monotone Maps of the Interval*, American Mathematical Society, 1994.
- [63] B. Rynne and M. Youngson, *Linear Functional Analysis*, 2nd edition, Springer-Verlag, 2008.
- [64] C. Sadosky, *Interpolation of Operators and Singular Integrals: An Introduction to Harmonic Analysis*, Dekker, 1979.
- [65] K. Saxe, *Beginning Functional Analysis*, Springer-Verlag, 2002.
- [66] M. Schechter, *Principles of Functional Analysis*, 2nd edition, American Mathematical Society, 2002.
- [67] C. Silva, *Invitation to Ergodic Theory*, American Mathematical Society, 2008.
- [68] Y. Sinai, *Introduction to Ergodic Theory*, translated by V. Scheffer, Princeton University Press, 1976.
- [69] Y. Sinai, *Topics in Ergodic Theory*, Princeton University Press, 1994.
- [70] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970.
- [71] E. Stein, *Topics in Harmonic Analysis Related to the Littlewood–Paley Theory*, Princeton University Press, 1970.
- [72] E. Stein, *Boundary Behavior of Holomorphic Functions of Several Complex Variables*, Princeton University Press, 1972.

- [73] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, with the assistance of T. Murphy, Princeton University Press, 1993.
- [74] E. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.
- [75] E. Stein and R. Shakarchi, *Complex Analysis*, Princeton University Press, 2003.
- [76] E. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.
- [77] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.
- [78] R. Strichartz, *The Way of Analysis*, Jones and Bartlett, 1995.
- [79] R. Strichartz, *Differential Equations on Fractals*, Princeton University Press, 2006.
- [80] K. Stromberg, *Introduction to Classical Real Analysis*, Wadsworth, 1981.
- [81] D. Stroock, *A Concise Introduction to the Theory of Integration*, 3rd edition, Birkhäuser, 1999.
- [82] M. Taibleson, *Fourier Analysis on Local Fields*, Princeton University Press, 1975.
- [83] A. Torchinsky, *Real Variables*, Addison-Wesley, 1988.
- [84] A. Torchinsky, *Real-Variable Methods in Harmonic Analysis*, Dover, 2004.
- [85] N. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and Geometry on Groups*, Cambridge University Press, 1992.
- [86] P. Walters, *Ergodic Theory — Introductory Lectures*, Lecture Notes in Mathematics **458**, Springer-Verlag, 1975.
- [87] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, 1982.
- [88] R. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Dekker, 1977.
- [89] R. Zimmer, *Ergodic Theory and Semisimple Groups*, Birkhäuser, 1984.
- [90] A. Zygmund, *Trigonometric Series*, Volumes I and II, 3rd edition, with a foreword by R. Fefferman, Cambridge University Press, 2002.