Functional Analysis

Matti Vuorinen

Department of Mathematics and Statistics
University of Turku

1. Introduction

This graduate course is an introduction to functional analysis based on

B.P. Rynne - M.A. Youngson: Linear Functional Analysis, 2nd ed. 2001, Springer

Prerequisites: Topology of metric spaces.
Idea: Linear algebra for function spaces: norm, convergence, subspaces, projection,
Goal is to build a unified framework for the theory of diff./integral equations.
Norms chosen according to application (e.g. approximation theory)

Definition 1.1

A vector space over a field $\mathbb{F}$ *is a non-empty set* V *together with two functions, one from* $V \times V$ *to* V *and the other from* $\mathbb{F} \times V$ *to* V *, denoted by* $x + y$ *and* αx *respectively, for all* $x, y \in V$ *and* $\alpha \in \mathbb{F}$*, such that, for any* $\alpha, \beta \in \mathbb{F}$*, and any* $x, y, z \in V$*,*

(a) $x + y = y + x, x + (y + z) = (x + y) + z$;

(b) *there exists a unique* $0 \in V$ *(independent of* x*) such that* $x + 0 = x$;

(c) *there exists a unique* $-x \in V$ *such that* $x + (-x) = 0$;

(d) $1x = x, \alpha(\beta(x)) = (\alpha\beta)x$;

(e) $\alpha(x + y) = \alpha x + \alpha y, (\alpha + \beta)x = \alpha x + \beta x$.

Def. 1.1 cont'd

If $\mathbb{F} = \mathbb{R}$ (respectively,$\mathbb{F} = \mathbb{C}$) then V is a real (respectively, complex) vector space. Elements of $\mathbb{F}$ are called scalars, while elements of V are called vectors. The operation $x + y$ is called vector addition, while the operation αx is called scalar multiplication.

Definition 1.2

Let V be a vector space. A non-empty set $U \subset V$ is a linear subspace of V if U is itself a vector space (with the same vector addition and scalar multiplication as in V). This is equivalent to the condition that $\alpha x + \beta y \in U$, for all $\alpha, \beta \in \mathbb{F}$ and $x, y \in U$ (which is called the subspace test).

Definition 1.3

Let V be a vector space, let $v = \{v_1, ..., v_k\} \subset V, k \geq 1$, be a finite set and let $A \subset V$ be an arbitrary non-empty set.

(a) *A linear combination of the elements of v is any vector of the form $x = \alpha_1 v_1 + \cdots + \alpha_k v_k \in V$, for any set of scalars $\alpha_1, \ldots, \alpha_k$.*

(b) *v is linearly independent if the following implication holds: $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0 \Leftrightarrow \alpha_1 = \cdots = \alpha_k = 0$.*

(c) *A is linearly independent if every finite subset of A is linearly independent. If A is not linearly independent then it is linearly dependent.*

Def. 1.3 cont'd

(d) The span of A (denoted SpA) is the set of all linear combinations of all finite subsets of A. This set is a linear subspace of V. Equivalently, SpA is the intersection of the set of all linear subspaces of V which contain A. Thus, SpA is the smallest linear subspace of V containing A (in the sense that if $A \subset B \subset V$ and B is a linear subspace of V then $SpA \subset B$).

(e) If v is linearly independent and $Spv = V$, then v is called a basis for V. It can be shown that if V has such a (finite) basis then all bases of V have the same number of elements. If this number is k then V is said to be k-dimensional (or, more generally, finite-dimensional), and we write $\dim V = k$.

Def. 1.3 cont'd

If V does not have such a finite basis it is said to be infinitedimensional.

(f) If v is a basis for V then any $x \in V$ can be written as a linear combination of the above form, with a unique set of scalars $\alpha_j, j = 1, \ldots, k$. These scalars (which clearly depend on x) are called the components of x with respect to the basis v.

(g) The set $\mathbb{F}^k$ is a vector space over $\mathbb{F}$ and the set of vectors $e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_k = (0, 0, 0, \ldots, 1)$, is a basis for $\mathbb{F}^k$. This notation will be used throughout the book, and this basis will be called the standard basis for $\mathbb{F}^k$.

Definition 1.4

Let V, W be vector spaces over $\mathbb{F}^k$. The Cartesian product $V \times W$ is a vector space with the following vector space operations. For any $\alpha \in \mathbb{F}$. and any $(x_j, y_j) \in V \times W, j = 1, 2$, let $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \alpha(x_1, y_1) = (\alpha x_1, \alpha y_1)$ (using the corresponding vector space operations in V and W).

Definition 1.5

Let S be a set and let V be a vector space over $\mathbb{F}$. We denote the set of functions $f : S \to V$ by $F(S, V)$. For any $\alpha \in \mathbb{F}$ and any $f, g \in F(S, V)$, we define functions $f + g$ and αf in $F(S, V)$ by $(f + g)(x) = f(x) + g(x)$, $(\alpha f)(x) = \alpha f(x)$, for all $x \in S$ (using the vector space operations in V). With these definitions the set $F(S, V)$ is a vector space over $\mathbb{F}$.

Example 1.6

If S is the set of integers $\{1, \ldots, k\}$ then the set $F(S, \mathbb{F})$ can be identified with the space $\mathbb{F}^k$ (by identifying an element $x \in \mathbb{F}^k$ with the function $f \in F(S, \mathbb{F})$ defined by $f(j) = x_j, 1 \leq j \leq k$).

Definition 1.7

Let V, W be vector spaces over the same scalar field $\mathbb{F}$. A function $T : V \to W$ is called a linear transformation (or mapping) if, for all $\alpha, \beta \in \mathbb{F}$ and $x, y \in V$,

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y).$$

The set of all linear transformations $T : V \to W$ will be denoted by $L(V, W)$. With the scalar multiplication and vector addition defined in Definition 1.1 the set $L(V, W)$ is a vector space (it is a subspace of $F(V, W)$). When $V = W$ we abbreviate $L(V, V)$ to $L(V)$.

Lemma 1.8

Let V, W, X be vector spaces and $T \in L(V, W)$, $S \in L(W, X)$. Then the composition $S \circ T \in L(V, X)$.

Lemma 1.9

Let V be a vector space, $R, S, T \in L(V)$, and $\alpha \in \mathbb{F}$. Then:

(a) $R \circ (S \circ T) = (R \circ S) \circ T$;

(b) $R \circ (S + T) = R \circ S + R \circ T$;

(c) $(S + T) \circ R = S \circ R + T \circ R$;

(d) $I_V \circ T = T \circ I_V = T$ *where* $I_V(x) = x, \forall x \in V$;

(e) $(\alpha S) \circ T = \alpha(S \circ T) = S \circ (\alpha T)$.

Lemma 1.10

Let V, W be vector spaces and $T \in L(V, W)$.

(a) $T(0) = 0$.

(b) *If U is a linear subspace of V then the set $T(U)$ is a linear subspace of W and $\dim T(U) \leq \dim U$ (as either finite numbers or ∞).*

(c) *If U is a linear subspace of W then the set $\{x \in V : T(x) \in U\}$ is a linear subspace of V.*

Definition 1.11

Let V, W be vector spaces and $T \in L(V, W)$.

(a) *The image of T (often known as the range of T) is the subspace $\mathrm{Im}\, T = T(V)$; the rank of T is the number $r(T) = \dim(\mathrm{Im}\, T)$.*

(b) *The kernel of T (often known as the null-space of T) is the subspace $\mathrm{Ker}\, T = \{x \in V : T(x) = 0\}$; the nullity of T is the number $n(T) = \dim(\mathrm{Ker}\, T)$. The rank and nullity, $r(T), n(T)$, may have the value ∞.*

(c) *T has finite rank if $r(T)$ is finite.*

(d) *T is one-to-one if, for any $y \in W$, the equation $T(x) = y$ has at most one solution x.*

Def. 1.11 cont'd

(e) T is onto if, for any $y \in W$, the equation $T(x) = y$ has at least one solution x.

(f) T is bijective if, for any $y \in W$, the equation $T(x) = y$ has exactly one solution x (that is, T is both one-to-one and onto).

Lemma 1.12

Let V, W be vector spaces and $T \in L(V, W)$.

(a) *T is one-to-one if and only if the equation $T(x) = 0$ has only the solution $x = 0$. This is equivalent to $\mathrm{Ker}\, T = \{0\}$ or $n(T) = 0$.*

(b) *T is onto if and only if $\mathrm{Im}\, T = W$. If $\dim W$ is finite this is equivalent to $r(T) = \dim W$.*

(c) *$T \in L(V, W)$ is bijective if and only if there exists a unique transformation $S \in L(W, V)$ which is bijective and $S \circ T = I_V$ and $T \circ S = I_W$.*

Lemma 1.12 cont'd

If V is k-dimensional then $n(T) + r(T) = k$ (in particular, $r(T)$ is necessarily finite, irrespective of whether W is finitedimensional). Hence, if W is also k-dimensional then T is bijective if and only if $n(T) = 0$.

Definition 1.13

Let V be a vector space and $T \in L(V)$. A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of T if the equation $T(x) = \lambda x$ has a non-zero solution $x \in V$, and any such non-zero solution is an eigenvector. The subspace $\mathrm{Ker}(T - \lambda I) \subset V$ is called the eigenspace (corresponding to λ) and the multiplicity of λ is the number $m_\lambda = n(T - \lambda I)$.

Lemma 1.14

Let V be a vector space and let $T \in L(V)$. Let $\{\lambda_1, \dots, \lambda_k\}$ be a set of distinct eigenvalues of T, and for each $1 \leq j \leq k$ let x_j be an eigenvector corresponding to λ_j . Then the set $\{x_1, \dots, x_k\}$ is linearly independent.

Theorem 1.15

(a) *The mapping* $T \in M_v^u(T)$ *is a bijective linear transformation from* $L(U, V)$ *to* $M_{mn}(F)$, *that is, if* $S, T \in L(U, V)$ *and* $\alpha \in \mathbb{F}$, *then*

$$M_v^u(\alpha T) = \alpha M_v^u(T), M_v^u(S + T) = M_v^u(S) + M_v^u(T).$$

(b) *If* $T \in L(U, V), S \in L(V, W)$ *(where* W *is l-dimensional, with basis w) then (again using standard matrix multiplication here).*

$$M_w^u(ST) = M_w^v(S) M_v^u(T)$$

Lemma 1.16

Let u be the standard basis of $\mathbb{F}^n$ and let v be the standard basis of $\mathbb{F}^m$. Let $C \in M_{mn}(\mathbb{F})$ and $T \in L(\mathbb{F}^n, \mathbb{F}^m)$. Then,

(a) $M^u{}_v(T_C) = C$;

(b) $T_B = T$ *(where* $B = M^u{}_v(T)$*).*

Definition 1.17

A metric on a set M is a function $d : M \times M \to \mathbb{R}$ with the following properties. For all $x, y, z \in M$,

(a) $d(x, y) \geq 0$;

(b) $d(x, y) = 0 \Leftrightarrow x = y$;

(c) $d(x, y) = d(y, x)$;

(d) $d(x, z) \leq d(x, y) + d(y, z)$ *(the triangle inequality).*

If d is a metric on M, then the pair (M, d) is called a metric space.

Example 1.18

For any integer $k \geq 1$, the function $d : \mathbb{F}^k \times \mathbb{F}^k \to \mathbb{R}$ defined by

$$d(x, y) = \left(\sum_{j=1}^{k} |x_j - y_j|^2 \right)^{1/2}$$

is a metric on the set $\mathbb{F}^k$. This metric will be called the standard metric on $\mathbb{F}^k$ and, unless otherwise stated, $\mathbb{F}^k$ will be regarded as a metric space with this metric. An example of an alternative metric on $\mathbb{F}^k$ is the function $d_1 : \mathbb{F}^k \times \mathbb{F}^k \to \mathbb{R}$ defined by

$$d_1(x, y) = \sum_{j=1}^{k} |x_j - y_j|.$$

Definition 1.19

Let (M, d) be a metric space and let N be a subset of M. Define $d_N : N \times N \to \mathbb{R}$ by $d_N(x, y) = d(x, y)$ for all $x, y \in N$ (that is, d_N is the restriction of d to the subset N). Then d_N is a metric on N, called the metric induced on N by d

Example 1.20

Using the definition of a sequence as a function from $\mathbb{N}$ to $\mathbb{F}$ we see that the space $F(\mathbb{N}, \mathbb{F})$ (see Definition 1.5) can be identified with the space consisting of all sequences in $\mathbb{F}$ (compare this with Example 1.6).

Definition 1.21

A sequence $\{x_n\}$ *in a metric space* (M, d) *converges to* $x \in M$ *(or the sequence* $\{x_n\}$ *is convergent) if, for every* $\varepsilon > 0$*, there exists* $N \in \mathbb{N}$ *such that* $d(x, x_n) < \varepsilon$*, for all* $n \geq N$*. As usual, we write*

$$\lim_{n\to\infty} x_n = x$$

or $x_n \to x$*. A sequence* $\{x_n\}$ *in* (M, d) *is a Cauchy sequence if, for every* $\varepsilon > 0$*, there exists* $N \in \mathbb{N}$ *such that* $d(x_m, x_n) < \varepsilon$*, for all* $m, n \geq N$*.*

Theorem 1.22

Suppose that $\{x_n\}$ *is a convergent sequence in a metric space* (M, d). *Then:*

(a) *the limit* $x = \lim_{n\to\infty} x_n$ *is unique;*

(b) *any subsequence of* $\{x_n\}$ *also converges to* x;

(c) $\{x_n\}$ *is a Cauchy sequence.*

Definition 1.23

Let (M, d) *be a metric space. For any* $x \in M$ *and any number* $r > 0$, *the set* $B_x(r) = \{y \in M : d(x, y) < r\}$ *will be called the open ball with centre* x *and radius* r. *If* $r = 1$ *the ball* $B_x(1)$ *is said to be an open unit ball. The set* $\{y \in M : d(x, y) \leq r\}$ *will be called the closed ball with centre* x *and radius* r. *If* $r = 1$ *this set will be called a closed unit ball.*

Definition 1.24

Let (M, d) be a metric space and let $A \subset M$.

(a) *A is bounded if there is a number $b > 0$ such that $d(x, y) < b$ for all $x, y \in A$.*

(b) *A is open if, for each point $x \in A$, there is an $\varepsilon > 0$ such that $B_x(\varepsilon) \subset A$.*

(c) *A is closed if the set $M \setminus A$ is open.*

(d) *A point $x \in M$ is a closure point of A if, for every $\varepsilon > 0$, there is a point $y \in A$ with $d(x, y) < \varepsilon$ (equivalently, if there exists a sequence $\{y_n\} \subset A$ such that $y_n \to x$).*

(e) *The closure of A, denoted by $\overline{A}$ or A, is the set of all closure points of A.*

(f) *A is dense (in M) if $\overline{A} = M$.*

Theorem 1.25

Let (M, d) be a metric space and let $A \subset M$.

(a) *$\overline{A}$ is closed and is equal to the intersection of the collection of all closed subsets of M which contain A (so $\overline{A}$ is the smallest closed set containing A).*

(b) *A is closed if and only if $\overline{A} = A$.*

(c) *A is closed if and only if, whenever $\{x_n\}$ is a sequence in A which converges to an element $x \in M$, then $x \in A$.*

(d) *$x \in \overline{A}$ if and only if*

$$\inf\{d(x, y) : y \in A\} = 0.$$

Theorem 1.25 cont'd

(e) For any $x \in M$ and $r > 0$, the "open" and "closed" balls in Definition 1.23 are open and closed in the sense of Definition 1.24. Furthermore, $\overline{B}_x(r) \subset \{y \in M : d(x,y) \leq r\}$, but these sets need not be equal in general (however, for most of the spaces considered in this book these sets are equal, see Exercise 2.10).

(f) A is dense if and only if, for any element $x \in M$ and any number $\varepsilon > 0$, there exists a point $y \in A$ with $d(x,y) < \varepsilon$ (equivalently, for any element $x \in M$ there exists a sequence $\{y_n\} \subset A$ such that $y_n \to x$).

Example 1.26

Let $M = \mathbb{R}$, with the standard metric, and let $N = (0, 1] \subset M$. If $A = (0, 1)$ then the closure of A in N is equal to N (so A is dense in N), but the closure of A in M is $[0, 1]$.

Definition 1.27

Let (M, d_M) and (N, d_N) be metric spaces and let $f : M \to N$ be a function.

(a) *f is continuous at a point $x \in M$ if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for $y \in M$, $d_M(x, y) < \delta \Rightarrow d_N(f(x), f(y)) < \varepsilon$.*

(b) *f is continuous (on M) if it is continuous at each point of M.*

(c) *f is uniformly continuous (on M) if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $x, y \in M$,*

$$d_M(x, y) < \delta \Rightarrow d_N(f(x), f(y)) < \varepsilon$$

(that is, the number δ can be chosen independently of $x, y \in M$).

Theorem 1.28

Suppose that $(M, d_M), (N, d_N)$, *are metric spaces and that* $f : M \to N$. *Then:*

(a) f *is continuous at* $x \in M$ *if and only if, for every sequence* $\{x_n\}$ *in* (M, d_M) *with* $x_n \to x$, *the sequence* $\{f(x_n)\}$ *in* (N, d_N) *satisfies* $f(x_n) \to f(x)$;

(b) f *is continuous on* M *if and only if either of the following conditions holds:*
(i) for any open set $A \subset N$, *the set* $f^{-1}(A) \subset M$ *is open;*
(ii) for any closed set $A \subset N$, *the set* $f^{-1}(A) \subset M$ *is closed*

Corollary 1.29

Suppose that $(M, d_M), (N, d_N)$, are metric spaces, A is a dense subset of M and $f, g : M \to N$ are continuous functions with the property that $f(x) = g(x)$ for all $x \in A$. Then $f = g$ (that is, $f(x) = g(x)$ for all $x \in M$).

Definition 1.30

A metric space (M, d) is complete if every Cauchy sequence in (M, d) is convergent. A set $A \subset M$ is complete (in (M, d)) if every Cauchy sequence lying in A converges to an element of A.

Theorem 1.31

For each $k \geq 1$, the space $\mathbb{F}^k$ with the standard metric is complete

Theorem 1.32

(Baire's Category Theorem) If (M, d) is a complete metric space and $M = \cup_{j=1}^{\infty} A_j$, where each $A_j \subset M, j = 1, 2, ...,$ is closed, then at least one of the sets A_j contains an open ball.

Definition 1.33

Let (M, d) be a metric space. A set $A \subset M$ is compact if every sequence $\{x_n\}$ in A contains a subsequence that converges to an element of A. A set $A \subset M$ is relatively compact if the closure $\bar{A}$ is compact. If the set M itself is compact then we say that (M, d) is a compact metric space.

Note. Compactness can also be defined in terms of "open coverings", and this definition is more appropriate in more general topological spaces, but in metric spaces both definitions are equivalent, and the above sequential definition is the only one that will be used in what follows.

Theorem 1.34

Suppose that (M, d) is a metric space and $A \subset M$. Then:

(a) *if A is complete then it is closed;*

(b) *if M is complete then A is complete if and only if it is closed;*

(c) *if A is compact then it is closed and bounded;*

(d) *(Bolzano-Weierstrass theorem) every closed, bounded subset of $\mathbb{F}^k$ is compact.*

Theorem 1.35

Suppose that (M, d) *is a compact metric space and* $f : M \to \mathbb{F}$ *is continuous. Then there exists a constant* $b > 0$ *such that* $|f(x)| \leq b$ *for all* $x \in M$ *(we say that* f *is bounded). In particular, if* $\mathbb{F} = \mathbb{R}$ *then the numbers* $\sup\{f(x) : x \in M\}$ *and* $\inf\{f(x) : x \in M\}$, *exist and are finite. Furthermore, there exist points* $x_s, x_i \in M$ *such that*

$$f(x_s) = \sup\{f(x) : x \in M\}, f(x_i) = \inf\{f(x) : x \in M\}.$$

Definition 1.36

Let (M, d) be a compact metric space. The set of continuous functions $f : M \to \mathbb{F}$ will be denoted by $C_{\mathbb{F}}(M)$. We define a metric on $C_{\mathbb{F}}(M)$ by

$$d(f, g) = \sup\{|f(x) - g(x)| : x \in M\}$$

(it can easily be verified that for any $f, g \in C_{\mathbb{F}}(M)$, the function $|f - g|$ is continuous so $d(f, g)$ is well-defined, by Theorem 1.35, and that d is a metric on $C_{\mathbb{F}}(M)$). This metric will be called the uniform metric and, unless otherwise stated, $C_{\mathbb{F}}(M)$ will always be assumed to have this metric.

Notation. Most properties of the space $C_{\mathbb{F}}(M)$ hold equally well in both the real and complex cases so, except where it is important to distinguish between these cases, we will omit the subscript and simply write $C(M)$. A similar convention will be adopted below for other spaces with both real and complex versions. Also, when M is a bounded interval $[a, b] \subset \mathbb{R}$ we write $C[a, b]$.

Definition 1.37

Suppose that (M, d) is a compact metric space and $\{f_n\}$ is a sequence in $C(M)$, and let $f : M \to \mathbb{F}$ be a function.

(a) *$\{f_n\}$ converges pointwise to f if $|f_n(x) - f(x)| \to 0$ for all $x \in M$.*

(b) *$\{f_n\}$ converges uniformly to f if $\sup\{|f_n(x) - f(x)| : x \in M\} \to 0$.*

Theorem 1.38

The metric space $C(M)$ is complete.

Theorem 1.39

(The Stone-Weierstrass Theorem) For any compact set $M \subset \mathbb{R}$, the set of polynomials $\mathcal{P}_{\mathbb{R}}$ is dense in $C_{\mathbb{R}}(M)$.

Definition 1.40

A set X is countable if it contains either a finite number of elements or infinitely many elements and can be written in the form $X = \{x_n : n \in \mathbb{N}\}$; in the latter case X is said to be countably infinite. A metric space (M, d) is separable if it contains a countable, dense subset. The empty set is regarded as separable.

Example 1.41

The space $\mathbb{R}$ is separable since the set of rational numbers is countably infinite (see Topology) and dense in $\mathbb{R}$.

Theorem 1.42

Suppose that (M, d) is a metric space and $A \subset M$.

(a) *If A is compact then it is separable.*

(b) *If A is separable and $B \subset A$ then B is separable.*

Definition 1.43

A σ-algebra (also known as a σ-field) is a class Σ of subsets of a set X with the properties:

(a) $\emptyset, X \in \Sigma$;

(b) $S \in \Sigma \Rightarrow X \setminus S \in \Sigma$;

(c) $S_n \in \Sigma, n = 1, 2, \cdots \Rightarrow \cup_{n=1}^{\infty} S_n \in \Sigma$.

A set $S \in \Sigma$ is said to be measurable.

Definition 1.44

Let X be a set and let Σ be a σ-algebra of subsets of X. A function $\mu : \Sigma \to \overline{\mathbb{R}}^+$ is a measure if it has the properties:

(a) $\mu(\emptyset) = 0$;

(b) *μ is countably additive, that is, if $S_j \in \Sigma, j = 1, 2, ...,$ are pairwise disjoint sets then*

$$\mu(\cup_{j=1}^{\infty} S_j) = \sum_{j=1}^{\infty} \mu(S_j).$$

The triple (X, Σ, μ) is called a measure space.

Definition 1.45

Let (X, Σ, μ) be a measure space. A set $S \in \Sigma$ with $\mu(S) = 0$ is said to have measure zero (or is a null set). A given property $P(x)$ of points $x \in X$ is said to hold almost everywhere if the set $\{x : P(x) \,\text{is false}\}$ has measure zero; alternatively, the property P is said to hold for almost every $x \in X$. The abbreviation a.e. will denote either of these terms.

Example 1.46

(Counting Measure) Let $X = \mathbb{N}$, let Σ_c be the class of all subsets of $\mathbb{N}$ and, for any $S \subset \mathbb{N}$, define $\mu_c(S)$ to be the number of elements of S. Then is Σ_c a σ-algebra and μ_c is a measure on Σ_c. This measure is called counting measure on $\mathbb{N}$. The only set of measure zero in this measure space is the empty set.

Example 1.47

(Lebesgue Measure) There is a σ-algebra Σ_L in $\mathbb{R}$ and a measure μ_L on Σ_L such that any finite interval $I = [a, b] \in \Sigma_L$ and $\mu_L(I) = \ell(I)$. The sets of measure zero in this space are exactly those sets A with the following property: for any $\varepsilon > 0$ there exists a sequence of intervals $I_j \subset \mathbb{R}, j = 1, 2, \ldots,$ such that

$$A \subset \cup_{j=1}^{\infty} I_j \text{ and } \sum_{j=1}^{\infty} \ell(I_j) < \varepsilon.$$

This measure is called Lebesgue measure and the sets in σ_L are said to be Lebesgue measurable.

Example 1.48

(Counting Measure) Suppose that $(X, \Sigma, \mu) = (\mathbb{N}, \Sigma_c, \mu_c)$ *(see Example 1.47). Any function* $f : \mathbb{N} \to \mathbb{F}$ *can be regarded as an* $\mathbb{F}$*-valued sequence* $\{a_n\}$ *(with* $a_n = f(n), n \geq 1$*), and since all subsets of* $\mathbb{N}$ *are measurable, every such sequence* $\{a_n\}$ *can be regarded as a measurable function. It follows from the above definitions that the sequence* $\{a_n\}$ *is integrable (with respect to* μ_c*) if and only if* $\sum_{n=1}^{\infty} |a_n| < \infty$*, and then the integral of* $\{a_n\}$ *is simply the sum* $\sum_{n=1}^{\infty} |a_n|$*. Instead of the general notation* $L^1(\mathbb{N})$*, the space of such sequences will be denoted by* ℓ^1 *(or* $\ell^1_{\mathbb{F}}$ *).*

Definition 1.49

(Lebesgue integral) Let $(X, \Sigma, \mu) = (\mathbb{R}^k, \Sigma_L, \mu_L)$ *, for some* $k \geq 1$. *If* $f \in L^1(\mathbb{R}^k)$ *(or* $f \in L^1(S)$, *with* $S \in \Sigma_L$*) then* f *is said to be Lebesgue integrable.*

Theorem 1.50

If $I = [a, b] \subset \mathbb{R}$ *is a bounded interval and* $f : I \to \mathbb{R}$ *is bounded and Riemann integrable on* I*, then* f *is Lebesgue integrable on* I*, and the values of the two integrals of* f *coincide. In particular, continuous functions on* I *are Lebesgue integrable.*

Theorem 1.51

Let (X, Σ, μ) be a measure space and let $f \in L^1(X)$.

(a) *If $f(x) = 0$ a.e., then $f \in L^1(X)$ and $\int_X f d\mu = 0$.*

(b) *If $\alpha \in \mathbb{R}$ and $f, g \in L^1(X)$ then the functions $f + g$ and αf (see Definition 1.5) belong to $L^1(X)$ and*

$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu, \int_X \alpha f d\mu = \alpha \int_X f d\mu.$$

In particular, $L^1(X)$ is a vector space.

Theorem 1.51 cont'd

(c) If $f, g \in L^1(X)$ and $f(x) \le g(x)$ for all $x \in X$, then

$$\int_X f d\mu \le \int_X g d\mu.$$

If, in addition, $f(x) < g(x)$ for all $x \in S$, with $\mu(S) > 0$, then

$$\int_X f d\mu < \int_X g d\mu.$$

Definition 1.52

Suppose that f is a measurable function and there exists a number b such that $f(x) \leq b$ a.e. Then we can define the essential supremum of f to be

$$\text{ess sup} f = \inf\{b : f(x) \leq b \text{ a.e.}\}.$$

It is a simple (but not completely trivial) consequence of this definition that $f(x) \leq \text{ess sup} f$ a.e. The essential infimum of f can be defined similarly. A measurable function f is said to be essentially bounded if there exists a number b such that $|f(x)| \leq b$ a.e.

Definition 1.53

Define the spaces

$$\mathcal{L}^p(X) = \{f : f \text{ is measurable and } (\int_X |f|^p d\mu)^{1/p} < \infty\}, 1 \leq p < \infty;$$

$$\mathcal{L}^\infty(X) = \{f : f \text{ is measurable and } \operatorname{ess\,sup} |f| < \infty\}.$$

We also define the corresponding sets $L^p(X)$ by identifying functions in $\mathcal{L}^p(X)$ which are a.e. equal and considering the corresponding spaces of equivalence classes (in practice, we again simply refer to representative functions of these equivalence classes rather than the classes themselves).

Theorem 1.54

Suppose that f and g are measurable functions. Then the following inequalities hold (infinite values are allowed). Minkowski's inequality (for $1 \leq p < \infty$):

$$\left(\int_X |f+g|^p d\mu\right)^{1/p} \leq \left(\int_X |f|^p d\mu\right)^{1/p} + \left(\int_X |g|^p d\mu\right)^{1/p}$$

$$\operatorname{ess\,sup} |f+g| \leq \operatorname{ess\,sup} |f| + \operatorname{ess\,sup} |g|.$$

Hölder's inequality (for $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$) :

$$\int_X |fg| d\mu. \leq \left(\int_X |f|^p d\mu\right)^{1/p} \left(\int_X |g|^q d\mu\right)^{1/q}$$

$$\int |fg| d\mu \leq \operatorname{ess\,sup} |f| \int |g| d\mu.$$

Corollary 1.55

Suppose that $1 \leq p \leq \infty$.

(a) $L^p(X)$ *is a vector space (essentially, this follows from Minkowski's inequality together with simple properties of the integral).*

(b) *The function*

$$d_p(f,g) = \left(\int_X |f-g|^p d\mu \right)^{1/p}, 1 \leq p < \infty,$$

$$d_\infty(f,g) = \operatorname{ess\,sup} |f-g|,$$

is a metric on $L^p(X)$ *(condition (b) in Definition 1.7 follows from properties (a) and (c) in Theorem 1.51, together with the construction of the spaces* $L^p(X)$*, while Minkowski's inequality shows that* d_p *satisfies the triangle inequality).*

Example 1.56

(Counting Measure) Suppose that $1 \le p \le \infty$. *In the special case where* $(X, \Sigma, \mu) = (\mathbb{N}, \Sigma_c, \mu_c)$, *the space* $L^p(\mathbb{N})$ *consists of the set of sequences* $\{a_n\}$ *in* $\mathbb{F}$ *with the property that*

$$\left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p} < \infty \,, \text{for } 1 \le p < \infty$$

$$\sup\{|a_n| : n \in \mathbb{N}\} < \infty, \ \text{for } p = \infty.$$

These spaces will be denoted by ℓ^p *(or* $\ell^p_{\mathbb{F}}$ *). Note that since there are no sets of measure zero in this measure space, there is no question of taking equivalence classes here. By Corollary 1.55, the spaces* ℓ^p *are both vector spaces and metric spaces. The standard metric on* ℓ^p *is defined analogously to the above expressions in the obvious manner.*

Corollary 1.57

Minkowski's inequality (for $1 \leq p < \infty$):

$$\left(\sum_{j=1}^{\infty} |x_j + y_j|^p \right)^{1/p} \leq \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} + \left(\sum_{j=1}^{\infty} |y_j|^p \right)^{1/p} .$$

Hölder's inequality (for $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$):

$$\sum_{j=1}^{\infty} |x_j y_j| \leq \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \left(\sum_{j=1}^{\infty} |y_j|^q \right)^{1/q}$$

Here, k and the values of the sums may be ∞.

Corollary 1.58

$$\sum_{j=1}^{\infty} |x_j y_j| \leq \left(\sum_{j=1}^{\infty} |x_j|^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} |y_j|^2 \right)^{1/2} .$$

Definition 1.59

Let $\tilde{e}_1 = (1, 0, 0, ...)$, $\tilde{e}_2 = (0, 1, 0, ...)$, $\ldots$. *For any* $n \in \mathbb{N}$ *the sequence* $\tilde{e}_n \in \ell^p$ *for all* $1 \leq p < \infty$.

Theorem 1.60

Suppose that $1 \leq p \leq \infty$. *Then the metric space* $L^p(X)$ *is complete. In particular, the sequence space* ℓ^p *is complete.*

Theorem 1.61

Suppose that $[a, b]$ *is a bounded interval and* $1 \leq p < \infty$*. Then the set* $C[a, b]$ *is dense in* $L^p[a, b]$.

2. Normed spaces

Definition 2.1

(a) *Let* X *be a vector space over* $\mathbb{F}$. *A norm on* X *is a function* $||\cdot|| : X \to \mathbb{R}$ *such that for all* $x, y \in X$ *and* $\alpha \in \mathbb{F}$,
(i) $||x|| \geq 0$;
(ii) $||x|| = 0$ *if and only if* $x = 0$;
(iii) $||\alpha x|| = |\alpha|||x||$;
(iv) $||x + y|| \leq ||x|| + ||y||$.

(b) *A vector space* X *on which there is a norm is called a normed vector space or just a normed space.*

(c) *If* X *is a normed space, a unit vector in* X *is a vector* x *such that* $||x|| = 1$.

Example 2.2

The function $||\cdot||:\mathbb{F}^n\to\mathbb{R}$ *defined by*

$$||(x_1,...,x_n)|| = \left(\sum_{j=1}^{n}|x_j|^2\right)^{1/2}$$

is a norm on $\mathbb{F}^n$ *called the standard norm on* $\mathbb{F}^n$*.*

Example 2.3

Let X be a finite-dimensional vector space over $\mathbb{F}$ with basis $\{e_1, e_2, \ldots, e_n\}$. Any $x \in X$ can be written as $x = \sum_{j=1}^{n} \lambda_j e_j$ for unique $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$. Then the function $||\cdot|| : X \to \mathbb{R}$ defined by

$$||x|| = \left(\sum_{j=1}^{n} |\lambda_j|^2 \right)^{1/2}$$

is a norm on X.

Example 2.4

Let M be a compact metric space and let $C_{\mathbb{F}}(M)$ be the vector space of continuous, $\mathbb{F}$-valued functions defined on M. Then the function $||\cdot|| : C_{\mathbb{F}}(M) \to \mathbb{R}$ defined by

$$||f|| = \sup\{|f(x)| : x \in M\}$$

is a norm on $C_{\mathbb{F}}(M)$ called the standard norm on $C_{\mathbb{F}}(M)$.

Example 2.5

Let (X, Σ, μ) *be a measure space.*

(a) *If* $1 \leq p < \infty$ *then*

$$||f||_p = \left(\int_X |f|^p d\mu\right)^{1/p}$$

is a norm on $L^p(X)$ *called the standard norm on* $L^p(X)$.

(b) $||f||_\infty = \operatorname{ess\,sup}\{|f(x)| : x \in X\}$ *is a norm on* $L^\infty(X)$ *called the standard norm on* $L^\infty(X)$.

Example 2.6

(a) *If* $1 \leq p < \infty$ *then* $||\{x_n\}||_p = (\sum_{n=1}^{\infty} |x_n|^p)^{1/p}$ *is a norm on* ℓ^p *called the standard norm on* ℓ^p*.*

(b) $||\{x_n\}||_\infty = \sup\{|x_n| : n \in \mathbb{N}\}$ *is a norm on* ℓ^∞ *called the standard norm on* ℓ^∞*.*

Example 2.7

Let X *be a vector space with a norm* $||\cdot||$ *and let* S *be a linear subspace of* X*. Let* $||\cdot||_S$ *be the restriction of* $||\cdot||$ *to* S*. Then* $||\cdot||_S$ *is a norm on* S*.*

Example 2.8

Let X and Y be vector spaces over $\mathbb{F}$ and let $Z = X \times Y$ be the Cartesian product of X and Y. This is a vector space by Definition 1.4. If $||\cdot||_1$ is a norm on X and $||\cdot||_2$ is a norm on Y then $||(x,y)|| = ||x||_1 + ||y||_2$ defines a norm on Z.

Lemma 2.9

Let X be a vector space with norm $||\cdot||$.If $d : X \times X \to \mathbb{R}$ is defined by $d(x,y) = ||x - y||$ then (X,d) is a metric space.

Notation. If X is a vector space with norm $||\cdot||$ and d is the metric defined by $d(x,y) = ||x - y||$ then d is called the metric associated with $||\cdot||$.

Whenever we use a metric or a metric space concept, for example, convergence, continuity or completeness, in a normed space then we will always use the metric associated with the norm even if this is not explicitly stated. The metrics associated with the standard norms are already familiar.

Example 2.10

The metrics associated with the standard norms on the following spaces are the standard metrics.

(a) $\mathbb{F}^n$;

(b) $C_{\mathbb{F}}(M)$ *where* M *is a compact metric space;*

(c) $L^p(X)$ *for* $1 \leq p < \infty$ *where* (X, Σ, μ) *is a measure space;*

(d) $L^\infty(X)$ *where* (X, Σ, μ) *is a measure space.*

Theorem 2.11

Let X be a vector space over $\mathbb{F}$ with norm $||\cdot||$. Let $\{x_n\}$ and $\{y_n\}$ be sequences in X which converge to x, y in X respectively and let $\{\alpha_n\}$ be a sequence in $\mathbb{F}$ which converges to α in $\mathbb{F}$. Then:

(a) $|||x|| - ||y||| \leq ||x - y||$;

(b) $\lim_{n\to\infty} ||x_n|| = ||x||$;

(c) $\lim_{n\to\infty}(x_n + y_n) = x + y$;

(d) $\lim_{n\to\infty} \alpha_n x_n = \alpha x$.

Definition 2.12

Let X be a vector space and let $||\cdot||_1$ and $||\cdot||_2$ be two norms on X. The norm $||\cdot||_2$ is equivalent to the norm $||\cdot||_1$ if there exists $M, m > 0$ such that for all $x \in X$

$$m||x||_1 \leq ||x||_2 \leq M||x||_1.$$

Lemma 2.13

Let X be a vector space and let $||\cdot||_1$,$||\cdot||_2$, $||\cdot||_3$ be three norms on X. Let $||\cdot||_2$ be equivalent to $||\cdot||_1$ and let $||\cdot||_3$ be equivalent to $||\cdot||_2$.

(a) $||\cdot||_1$ *is equivalent to* $||\cdot||_2$.

(b) $||\cdot||_3$ *is equivalent to* $||\cdot||_1$.

Lemma 2.14

Let X be a vector space and let $||\cdot||$ and $||\cdot||_1$ be norms on X. Let d and d_1 be the metrics defined by $d(x,y) = ||\cdot||$ and $d_1(x,y) = ||\cdot||_1$. Suppose that there exists $K > 0$ such that $||x|| \leq K||x||_1$ for all $x \in X$. Let $\{x_n\}$ be a sequence in X.

(a) *If $\{x_n\}$ converges to x in the metric space (X, d_1) then $\{x_n\}$ converges to x in the metric space (X, d).*

(b) *If $\{x_n\}$ is Cauchy in the metric space (X, d_1) then $\{x_n\}$ is Cauchy in the metric space (X, d).*

Corollary 2.15

Let X be a vector space and let $||\cdot||$ and $||\cdot||_1$ be equivalent norms on X. Let d and d_1 be the metrics defined by $d(x,y) = ||x-y||$ and $d_1(x,y) = ||x-y||_1$. Let $\{x_n\}$ be a sequence in X.

(a) *$\{x_n\}$ converges to x in the metric space (X,d) if and only if $\{x_n\}$ converges to x in the metric space (X,d_1).*

(b) *$\{x_n\}$ is Cauchy in the metric space (X,d) if and only if $\{x_n\}$ is Cauchy in the metric space (X,d_1).*

(c) *(X,d) is complete if and only if (X,d_1) is complete.*

Theorem 2.16

Let X be a finite-dimensional vector space with norm $||\cdot||$ and let $\{e_1, e_2, \ldots, e_n\}$ be a basis for X. Another norm on X was defined in Example 2.3 by

$$||\sum_{j=1}^{n} \lambda_j e_j||_1 = \left(\sum_{j=1}^{n} |\lambda_j|^2\right)^{1/2}.$$

The norms $||\cdot||$ and $||\cdot||_1$ are equivalent.

Corollary 2.17

If $||\cdot||$ and $||\cdot||_2$ are any two norms on a finite-dimensional vector space X then they are equivalent.

Lemma 2.18

Let X be a finite-dimensional vector space over $\mathbb{F}$ and let $\{e_1, e_2, \ldots, e_n\}$ be a basis for X. If $||\cdot|| : X \to \mathbb{R}$ is the norm on X defined in Theorem 2.16 then X is a complete metric space.

Corollary 2.19

If $||\cdot||$ is any norm on a finite-dimensional space X then X is a complete metric space.

Corollary 2.20

If Y is a finite-dimensional subspace of a normed vector space X, then Y is closed.

Example 2.21

Let

$$S = \{\{x_n\} \in \ell^\infty : \text{there exists } N \in \mathbb{N} \text{ such that } x_n = 0 \text{ for } n \geq N\},$$

so that S is the linear subspace of ℓ^∞ consisting of sequences having only finitely many non-zero terms. Then S is not closed.

Lemma 2.22

If X is a normed vector space and S is a linear subspace of X then $\overline{S}$ is a linear subspace of X.

Definition 2.23

Let X be a normed vector space and let E be any non-empty subset of X. The closed linear span of E, denoted by $\overline{Sp}\, E$, is the intersection of all the closed linear subspaces of X which contain E.

Lemma 2.24

Let X be a normed space and let E be any non-empty subset of X.

(a) *$\overline{Sp}\, E$ is a closed linear subspace of X which contains E.*

(b) *$\overline{Sp}\, E = \overline{Sp\, E}$, that is, $\overline{Sp}\, E$ is the closure of $Sp\, E$.*

Theorem 2.25

(Riesz' Lemma) Suppose that X is a normed vector space, Y is a closed linear subspace of X such that $Y \neq X$ and α is a real number such that $0 < \alpha < 1$. Then there exists $x_\alpha \in X$ such that $||x_\alpha|| = 1$ and $||x_\alpha - y|| > \alpha$ for all $y \in Y$.

Theorem 2.26

If X is an infinite-dimensional normed vector space then neither $D = \{x \in X : ||x|| \leq 1\}$ nor $K = \{x \in X : ||x|| = 1\}$ is compact.

Definition 2.27

A Banach space is a normed vector space which is complete under the metric associated with the norm.

Theorem 2.28

(a) *Any finite-dimensional normed vector space is a Banach space.*

(b) *If X is a compact metric space then $C_{\mathbb{F}}(X)$ is a Banach space.*

(c) *If (X, Σ, μ) is a measure space then $L^p(X)$ is a Banach space for $1 \leq p \leq \infty$.*

(d) *ℓ^p is a Banach space for $1 \leq p \leq \infty$.*

(e) *If X is a Banach space and Y is a linear subspace of X then Y is a Banach space if and only if Y is closed in X.*

Definition 2.29

Let X be a normed space and let $\{x_k\}$ be a sequence in X. For each positive integer n let $s_n = \sum_{k=1}^{n} x_k$ be the nth partial sum of the sequence. The series $\sum_{k=1}^{\infty} x_k$ is said to converge if $\lim_{n\to\infty} s_n$ exists in X and, if so, we define

$$\sum_{k=1}^{\infty} x_k = \lim_{n\to\infty} s_n .$$

Theorem 2.30

Let X be a Banach space and let $\{x_n\}$ be a sequence in X. If the series $\sum_{k=1}^{\infty} ||x_k||$ converges then the series $\sum_{k=1}^{\infty} x_k$ converges.

Inner product spaces

Definition 3.1

Let X be a real vector space. An inner product on X is a function $(\cdot,\cdot) : X \times X \to \mathbb{R}$ such that for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{R}$,

(a) $(x, x) \geq 0$;

(b) $(x, x) = 0$ *if and only if* $x = 0$;

(c) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$;

(d) $(x, y) = (y, x)$.

Example 3.2

The function $(\cdot,\cdot) : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ defined by $(x, y) = \sum_{n=1}^{k} x_n y_n$, is an inner product on $\mathbb{R}^k$. This inner product will be called the standard inner product on $\mathbb{R}^k$.

Definition 3.3

Let X be a complex vector space. An inner product on X is a function $(\cdot,\cdot) : X \times X \to \mathbb{C}$ such that for all $x, y, z \in X$, $\alpha, \beta \in \mathbb{C}$,

(a) $(x, x) \in \mathbb{R}$ *and* $(x, x) \geq 0$;

(b) $(x, x) = 0$ *if and only if* $x = 0$;

(c) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$;

(d) $(x, y) = \overline{(y, x)}$.

Example 3.4

The function $(\cdot,\cdot) : \mathbb{C}^k \times \mathbb{C}^k \to \mathbb{C}$ defined by $(x, y) = \sum_{n=1}^{k} x_n \overline{y}_n$, is an inner product on $\mathbb{C}^k$. This inner product will be called the standard inner product on $\mathbb{C}^k$.

Definition 3.5

A real or complex vector space X with an inner product $(\cdot,\cdot)$ is called an inner product space.

Example 3.6

Let X be a k-dimensional vector space with basis $\{e_1,\ldots,e_k\}$. Let $x,y\in X$ have the representation $x=\sum_{n=1}^{k}\lambda_n e_n$, $y=\sum_{n=1}^{k}\mu_n e_n$. The function $(\cdot,\cdot):X\times X\to\mathbb{F}$ defined by $(x,y)=\sum_{n=1}^{k}\lambda_n\overline{\mu}_n$, is an inner product on X.

Example 3.7

If $f,g\in L^2(X)$ then $fg\in L^1(X)$ and the function $(\cdot,\cdot):L^2(X)\times L^2(X)\to\mathbb{F}$ defined by $(f,g)=\int_X f\overline{g}\,d\mu$ is an inner product on $L^2(X)$. This inner product will be called the standard inner product on $L^2(X)$.

Example 3.8

If $a = \{a_n\}, b = \{b_n\} \in \ell^2$, then the sequence $\{a_n b_n\} \in \ell^1$ and the function $(\cdot,\cdot) : \mathbb{F}^k \times \mathbb{F}^k \to \mathbb{F}$ defined by $(a,b) = \sum_{n=1}^{\infty} a_n \overline{b}_n$ is an inner product on ℓ^2. This inner product will be called the standard inner product on ℓ^2.

Example 3.9

Let X be an inner product space with inner product $(\cdot,\cdot)$ and let S be a linear subspace of X. Let $(\cdot,\cdot)_S$ be the restriction of $(\cdot,\cdot)$ to S. Then $(\cdot,\cdot)_S$ is an inner product on S.

Example 3.10

Let X and Y be inner product spaces with inner products $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$, respectively, and let $Z = X \times Y$ be the Cartesian product space (see Definition 1.4). Then the function $(\cdot,\cdot) : Z \times Z \to \mathbb{F}$ defined by $((u,v),(x,y)) = (u,x)_1 + (v,y)_2$ is an inner product on Z.

Remark 3.11

We should note that, although the definitions in Examples 2.8 and 3.10 are natural, the norm induced on Z by the above inner product has the form $\sqrt{||x||_1^2 + ||y||_2^2}$ (where $||\cdot||_1, ||\cdot||_2$ are the norms induced by the inner products $(\cdot,\cdot)_1, (\cdot,\cdot)_2$), whereas the norm defined on Z in Example 2.8 has the form $||x||_1 + ||y||_2$. These two norms are not equal, but they are equivalent so in discussing analytic properties it makes no difference which one is used. However, the induced norm is somewhat less convenient to manipulate due to the square root term. Thus, in dealing with Cartesian product spaces one generally uses the norm in Example 2.8 if only norms are involved, but one must use the induced norm if inner products are also involved.

Lemma 3.12

Let X be an inner product space, $x, y, z \in X$ and $\alpha, \beta \in \mathbb{F}$. Then,

(a) $(0, y) = (x, 0) = 0$;

(b) $(x, \alpha y + \beta z) = \overline{\alpha}(x, y) + \overline{\beta}(x, z)$;

(c) $(\alpha x + \beta\, y, \alpha\, x + \beta\, y) = |\alpha|^2(x, x) + \alpha\overline{\beta}(x, y) + \beta\overline{\alpha}(y, x) + |\beta|^2(y, y)$.

Lemma 3.13

Let X be an inner product space and let $x, y \in X$. Then:

(a) $|(x, y)|^2 \leq (x, x)(y, y), x, y \in X$;

(b) *the function $|| \cdot || : X \to \mathbb{R}$ defined by $||x|| = (x, x)^{1/2}$, is a norm on X.*

Lemma 3.14

Let X be an inner product space with inner product $(\cdot, \cdot)$. Then for all $u, v, x, y \in X$:

(a) $(u + v, x + y) - (u - v, x - y) = 2(u, y) + 2(v, x)$;

(b) $4(u, y) =$
$(u+v, x+y)-(u-v, x-y)+i(u+iv, x+iy)-i(u-iv, x-iy)$
(for complex X).

Theorem 3.15

Let X be an inner product space with inner product $(\cdot,\cdot)$ and induced norm $||\cdot||$. Then for all $x, y \in X$:

(a)
$$||x+y||^2 + ||x-y||^2 = 2\left(||x||^2 + ||y||^2\right)$$
(the parallelogram rule);

(b) *if X is real then*
$$4(x,y) = ||x+y||^2 - ||x-y||^2;$$

(c) *if X is complex then*
$$4(x,y) = ||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2$$
(the polarization identity).

Example 3.16

The standard norm on the space $C[0,1]$ is not induced by an inner product.

Lemma 3.17

Let X be an inner product space and suppose that $\{x_n\}$ and $\{y_n\}$ are convergent sequences in X, with $\lim_{n\to\infty} x_n = x, \lim_{n\to\infty} y_n = y$. Then

$$\lim_{n\to\infty} (x_n, y_n) = (x, y).$$

Definition 3.18

Let X be an inner product space. The vectors $x, y \in X$ are said to be orthogonal if $(x, y) = 0$.

Definition 3.19

Let X be an inner product space. The set $\{e_1, \ldots, e_k\} \subset X$ is said to be orthonormal if $||e_n|| = 1$ for $1 \leq n \leq k$, and $(e_m, e_n) = 0$ for all $1 \leq m, n \leq k$ with $m \neq n$.

Lemma 3.20

(a) *An orthonormal set* $\{e_1, \ldots, e_k\}$ *in any inner product space* X *is linearly independent. In particular, if* X *is* k*-dimensional then the set* $\{e_1, \ldots, e_k\}$ *is a basis for* X *and any vector* $x \in X$ *can be expressed in the form*

$$x = \sum_{n=1}^{k} (x, e_n) e_n$$

(in this case $\{e_1, \ldots, e_k\}$ *is usually called an orthonormal basis and the numbers* (x, e_n) *are the components of* x *with respect to this basis).*

(b) *Let* $\{v_1, \ldots, v_k\}$ *be a linearly independent subset of an inner product space* X*, and let* $S = Sp\{v_1, \ldots, v_k\}$*. Then there is an orthonormal basis* $\{e_1, \ldots, e_k\}$ *for* S*.*

Remark 3.21

The inductive construction of the basis in part (b) of Lemma 3.20, using the formulae

$$b_{k+1} = v_{k+1} - \sum_{n=1}^{k} (v_{k+1}, e_n) e_n, \; e_{k+1} = \frac{b_{k+1}}{||b_{k+1}||},$$

is called the Gram–Schmidt algorithm, and is described in more detail in Linear algebra.

Theorem 3.22

Let X be a k-dimensional inner product space and let $\{e_1, \ldots, e_k\}$ be an orthonormal basis for X. Then, for any numbers $\alpha_n \in \mathbb{F}$, $n = 1, \ldots, k$,

$$||\sum_{n=1}^{k} \alpha_n e_n||^2 = \sum_{n=1}^{k} |\alpha_n|^2.$$

Definition 3.23

An inner product space which is complete with respect to the metric associated with the norm induced by the inner product is called a Hilbert space.

Example 3.24

(a) *Every finite-dimensional inner product space is a Hilbert space.*

(b) $L^2(X)$ *with the standard inner product is a Hilbert space.*

(c) ℓ^2 *with the standard inner product is a Hilbert space.*

Lemma 3.25

If H is a Hilbert space and $Y \subset H$ is a linear subspace, then Y is a Hilbert space if and only if Y is closed in H.

Definition 3.26

Let X be an inner product space and let A be a subset of X. The orthogonal complement of A is the set

$$A^{\perp} = \{x \in X : (x, a) = 0 \text{ for all } a \in A\}.$$

Thus

Example 3.27

If $X = \mathbb{R}^3$ *and* $A = \{(a_1, a_2, 0) : a_1, a_2 \in \mathbb{R}\}$, *then*

$$A^{\perp} = \{(0, 0, x_3) : x_3 \in \mathbb{R}\} = \{(0, 0, x_3) : x_3 \in \mathbb{R}\}.$$

Example 3.28

Suppose that X *is a* k*-dimensional inner product space, and* $\{e_1, \ldots, e_k\}$ *is an orthonormal basis for* X*. If* $A = Sp\{e_1, \ldots, e_p\}$*, for some* $1 \leq p < k$*, then* $A^{\perp} = Sp\{e_{p+1}, \ldots, e_k\}$*.*

Lemma 3.29

If X is an inner product space and $A \subset X$ then:

(a) $0 \in A^{\perp}$.

(b) *If* $0 \in A$ *then* $A \cap A^{\perp} = \{0\}$, *otherwise* $A \cap A^{\perp} = \emptyset$ *.*

(c) $0^{\perp} = X$; $X^{\perp} = \{0\}$.

(d) *If A contains an open ball $B_a(r)$, for some $a \in X$ and some positive $r > 0$, then $A^{\perp} = \{0\}$; in particular, if A is a non-empty open set then $A^{\perp} = \{0\}$.*

(e) *If $B \subset A$ then $A^{\perp} \subset B^{\perp}$. .*

(f) *$A^{\perp}$ is a closed linear subspace of X.*

(g) $A \subset (A^{\perp})^{\perp}$.

Lemma 3.30

Let Y be a linear subspace of an inner product space X. Then

$$x \in Y^{\perp} \Leftrightarrow ||x - y|| \geq ||x||, \forall y \in Y.$$

Definition 3.31

A subset A of a vector space X is convex if, for all $x, y \in A$ and all $\alpha \in [0, 1]$, $\alpha x + (1 - \alpha)y \in A$.

Theorem 3.32

Let A be a non-empty, closed, convex subset of a Hilbert space H and let $p \in H$. Then there exists a unique $q \in A$ such that

$$||p - q|| = \inf\{||p - a|| : a \in A\}.$$

Remark 3.33

Theorem 3.32 shows that if A is a non-empty, closed, convex subset of a Hilbert space H and p is a point in H, then there is a unique point q in A which is the closest point in A to p. In finite dimensions, even if the set A is not convex the existence of the point q can be proved in a similar manner (using the compactness of closed bounded sets to give the necessary convergent sequence). However, in this case the point q need not be unique (for example, let A be a circle in the plane and p be its centre, then q can be any point on A). In infinite dimensions, closed bounded sets need not be compact (see Theorem 2.26), so the existence question is more difficult and q may not exist if A is not convex.

Theorem 3.34

Let Y be a closed linear subspace of a Hilbert space H. For any $x \in H$, there exists a unique $y \in Y$ and $z \in Y^{\perp}$ such that $x = y + z$. Also,

$$||x||^2 = ||y||^2 + ||z||^2.$$

Notation. Suppose that Y is a closed linear subspace of a Hilbert space H and $x \in H$. The decomposition $x = y + z$, with $y \in Y$ and $z \in Y^{\perp}$, will be called the orthogonal decomposition of x with respect to Y.

Corollary 3.35

If Y is a closed linear subspace of a Hilbert space H then $Y^{\perp\perp} = Y$.

Corollary 3.36

If Y is any linear subspace of a Hilbert space H then $Y^{\perp\perp} = \overline{Y}$.

Definition 3.37

Let X be an inner product space. A sequence $\{e_n\} \subset X$ is said to be an orthonormal sequence if $||e_n|| = 1$ for all $n \in \mathbb{N}$, and $(e_m, e_n) = 0$ for all $m, n \in \mathbb{N}$ with $m \neq n$.

Example 3.38

The sequence $\{\tilde{e}_n\}$ (see Definition 1.59) is an orthonormal sequence in ℓ^2. Note that each of the elements of this sequence (in ℓ^2) is itself a sequence (in $\mathbb{F}$). This can be a source of confusion, so it is important to keep track of what space a sequence lies in.

Example 3.39

The set of functions $\{e_n\}$, *where* $e_n(x) = (2\pi)^{-1/2} e^{inx}$ *for* $n \in \mathbb{Z}$, *is an orthonormal sequence in the space* $L^2_{\mathbb{C}}[-\pi, \pi]$.

Theorem 3.40

Any infinite-dimensional inner product space X *contains an orthonormal sequence.*

Lemma 3.41

(Bessel's Inequality) Let X be an inner product space and let $\{e_n\}$ be an orthonormal sequence in X. For any $x \in X$ the (real) series $\sum_{n=1}^{\infty} |(x, e_n)|^2$ converges and

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq ||x||^2.$$

Theorem 3.42

Let H be a Hilbert space and let $\{e_n\}$ be an orthonormal sequence in H. Let $\{\alpha_n\}$ be a sequence in $\mathbb{F}$. Then the series

$$\sum_{n=1}^{\infty} \alpha_n e_n$$

converges if and only if $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$. If this holds, then

$$||\sum_{n=1}^{\infty} \alpha_n e_n||^2 = \sum_{n=1}^{\infty} |\alpha_n|^2.$$

Remark 3.43

The result of Theorem 3.42 can be rephrased as: the series

$$\sum_{n=1}^{\infty} \alpha_n e_n$$

converges if and only if the sequence $\{\alpha_n\} \in \ell^2$.

Corollary 3.44

Let H *be a Hilbert space and let* $\{e_n\}$ *be an orthonormal sequence in* H. *For any* $x \in H$ *the series* $\sum_{n=1}^{\infty}(x, e_n)e_n$ *converges.*

Example 3.45

In $\mathbb{R}^3$, consider the orthonormal set $\{\hat{e}_1, \hat{e}_2\}$, and let $x = (3, 0, 4)$, say. Then $(x, \hat{e}_1)\hat{e}_1 + (x, \hat{e}_2)\hat{e}_2 \neq x$.

Example 3.46

Let $\{e_n\}$ be an orthonormal sequence in a Hilbert space H, and let S be the subsequence $S = \{e_{2n}\}, n \in \mathbb{N}$ (that is, S consists of just the even terms in the sequence $\{e_n\}$). Then S is an orthonormal sequence in H with infinitely many elements, but, for instance, $e_1 \neq \sum_{n=1}^{\infty} \alpha_{2n} e_{2n}$, for any numbers α_{2n}.

Theorem 3.47

Let H be a Hilbert space and let $\{e_n\}$ be an orthonormal sequence in H. The following conditions are equivalent:

(a) $\{e_n : n \in \mathbb{N}\}^{\perp} = \{0\}$;

(b) $\overline{Sp}\{e_n : n \in \mathbb{N}\} = H$;

(c) $||x||^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 \quad \text{for all } x \in H$;

(d) $x = \sum_{n=1}^{\infty} (x, e_n) e_n \quad \text{for all } x \in H$;

Remark 3.48

The linear span $(Sp\{e_n\})$ *of the set* $\{e_n\}$ *consists of all possible finite linear combinations of the vectors in* $\{e_n\}$*, that is, all possible finite sums in the above expansion. However, for the expansion to hold for all x H it is necessary to also consider infinite sums in the expansion. This corresponds to considering the closed linear span* $(Sp\{e_n\})$*. In finite dimensions the linear span is necessarily closed, and so equals the closed linear span, so it is not necessary to distinguish between these concepts in finite-dimensional linear algebra.*

Definition 3.49

Let H *be a Hilbert space and let* $\{e_n\}$ *be an orthonormal sequence in* H*. Then* $\{e_n\}$ *is called an orthonormal basis for* H *if any of the conditions in Theorem 3.47 hold.*

Remark 3.50

Some books call an orthonormal basis $\{e_n\}$ *a complete orthonormal sequence - the sequence is "complete" in the sense that there are enough vectors in it to span the space (as in Theorem 3.47). We prefer not to use the term "complete" in this sense to avoid confusion with the previous use of "complete" to describe spaces in which all Cauchy sequences converge.*

Example 3.51

The orthonormal sequence $\{e_n\}$ *in* ℓ^2 *given in Example 3.38 is an orthonormal basis. This basis will be called the standard orthonormal basis in* ℓ^2.

Theorem 3.52

(a) *Finite dimensional normed vector spaces are separable.*
(b) *An infinite-dimensional Hilbert space H is separable if and only if it has an orthonormal basis.*

Example 3.53

The Hilbert space ℓ^2 is separable.

Theorem 3.54

The set of functions

$$C = \left\{ c_0(x) = (1/\pi)^{1/2}, c_n(x) = (2/\pi)^{1/2} \cos nx : n \in \mathbb{N} \right\}$$

is an orthonormal basis in $L^2[0,\pi]$.

Corollary 3.55

The space $L^2[0, \pi]$ *is separable.*

Theorem 3.56

The set of functions

$$S = \left\{ s_n(x) = (2/\pi)^{1/2} \sin nx : n \in \mathbb{N} \right\}$$

is an orthonormal basis in $L^2[0, \pi]$.

Corollary 3.57

The sets of functions

$$E = \{e_n(x) = (2\pi)^{-1/2} e^{inx} : n \in \mathbb{Z}\}, \quad F = \{2^{-1/2}c_0, 2^{-1/2}c_n, 2^{-1/2}s_n$$

are orthonormal bases in the space $L^2_{\mathbb{C}}[-\pi,\pi]$. *The set* F *is also an orthonormal basis in the space* $L^2_{\mathbb{R}}[-\pi,\pi]$ *(the set* E *is clearly not appropriate for the space* $L^2_{\mathbb{R}}[-\pi,\pi]$ *since the functions in* E *are complex).*

4. Linear Operators

Lemma 4.1

Let X and Y be normed linear spaces and let $T : X \to Y$ be a linear transformation. The following are equivalent:

(a) *T is uniformly continuous;*

(b) *T is continuous;*

(c) *T is continuous at 0;*

(d) *there exists a positive real number k such that $||T(x)|| \leq k$ whenever $x \in X$ and $||x|| \leq 1$;*

(e) *there exists a positive real number k such that $||T(x)|| \leq k||x||$ for all $x \in X$.*

Example 4.2

The linear transformation $T : C_{\mathbb{F}}[0,1] \to \mathbb{F}$ *defined by* $T(f) = f(0)$ *is continuous.*

Lemma 4.3

If $\{c_n\} \in \ell^\infty$ *and* $\{x_n\} \in \ell^p$, *where* $1 \leq p < \infty$, *then* $\{c_n x_n\} \in \ell^p$ *and*

$$\sum_{n=1}^{\infty} |c_n x_n|^p \leq ||\{c_n\}||_\infty^p \sum_{n=1}^{\infty} |x_n|^p.$$

Example 4.4

If $\{c_n\} \in \ell^\infty$, then the linear transformation $T : \ell^1 \to \mathbb{F}$ defined by

$$T(\{x_n\}) = \sum_{n=1}^{\infty} c_n x_n$$

is continuous.

Example 4.5

If $\{c_n\} \in \ell^\infty$, then the linear transformation $T : \ell^2 \to \ell^2$ defined by

$$T(\{x_n\}) = \{c_n x_n\}$$

is continuous.

Definition 4.6

Let X and Y be normed linear spaces and let $T : X \to Y$ be a linear transformation. T is said to be bounded if there exists a positive real number k such that $||T(x)|| \leq k||x||$ for all $x \in X$.

Notation. Let X and Y be normed linear spaces. The set of all continuous linear transformations from X to Y is denoted by $B(X, Y)$. Elements of $B(X, Y)$ are also called bounded linear operators or linear operators or sometimes just operators.

Example 4.7

Let $a, b \in \mathbb{R}$, *let* $k : [a, b] \times [a, b] \to \mathbb{C}$ *be continuous and let* $M = \sup\{|k(s, t)| : (s, t) \in [a, b] \times [a, b]\}$. *(a) If* $g \in C[a, b]$, *then* $f : [a, b] \to \mathbb{C}$ *defined by*

$$f(s) = \int_a^b k(s, t)\, g(t)\, dt$$

is in $C[a, b]$. *(b) If the linear transformation* $K : C[a, b] \to C[a, b]$ *is defined by*

$$(K(g))(s) = \int_a^b k(s, t)\, g(t)\, dt$$

then $K \in B(C[a, b], C[a, b])$ *and*

$$||K(g)|| \leq M(b - a)||g||.$$

Example 4.8

Let $\mathcal{P}$ be the linear subspace of $C_{\mathbb{C}}[0,1]$ consisting of all polynomial functions. If $T : \mathcal{P} \to \mathcal{P}$ is the linear transformation defined by $T(p) = p'$, where p' is the derivative of p, then T is not continuous.

Theorem 4.9

Let X be a finite-dimensional normed space, let Y be any normed linear space and let $T : X \to Y$ be a linear transformation. Then T is continuous.

Example 4.10

Let $\mathcal{P}$ be the linear subspace of $C_{\mathbb{C}}[0,1]$ consisting of all polynomial functions. If $T : \mathcal{P} \to \mathbb{C}$ is the linear transformation defined by

$$T(p) = p'(1),$$

where p' is the derivative of p, then T is not continuous.

Lemma 4.11

If X and Y are normed linear spaces and $T : X \to Y$ is a continuous linear transformation then $\mathrm{Ker}(T)$ is closed.

Definition 4.12

If X and Y are normed spaces and $T : X \to Y$ is a linear transformation, the graph of T is the linear subspace $\mathcal{G}(T)$ of $X \times Y$ defined by

$$\mathcal{G}(T) = \{(x, Tx) : x \in X\}.$$

Lemma 4.13

If X and Y are normed spaces and $T : X \to Y$ is a continuous linear transformation then $\mathcal{G}(T)$ is closed.

Lemma 4.14

Let X and Y be normed linear spaces and let $S, T \in B(X, Y)$ with $||S(x)|| \leq k_1||x||$ and $||T(x)|| \leq k_2||x||$ for all $x \in X$. Let $\lambda \in \mathbb{F}$. Then

(a) $||(S + T)(x)|| \leq (k_1 + k_2)||x||$ *for all* $x \in X$;

(b) $||(\lambda S)(x)|| \leq |\lambda|k_1||x||$ *for all* $x \in X$;

(c) *$B(X, Y)$ is a linear subspace of $L(X, Y)$ and so $B(X, Y)$ is a vector space.*

Lemma 4.15

Let X and Y be normed spaces. If $||E|| : B(X, Y) \to \mathbb{R}$ is defined by

$$||T|| = \sup\{||T(x)|| : ||x|| \leq 1\}$$

then $||E||$ is a norm on $B(X, Y)$.

Definition 4.16

Let X and Y be normed linear spaces and let $T \in B(X, Y)$. The norm of T is defined by

$$||T|| = \sup\{||T(x)|| : ||x|| \leq 1\}.$$

Definition 4.17

Let $\mathbb{F}^p$ have the standard norm and let A be a $m \times n$ matrix with entries in $\mathbb{F}$. If $T : \mathbb{F}^n \to \mathbb{F}^m$ is the bounded linear transformation defined by $T(x) = Ax$ then the norm of the matrix A is defined by $||A|| = ||T||$.

Example 4.18

If $T : C_{\mathbb{F}}[0, 1] \to \mathbb{F}$ is the bounded linear operator defined by $T(f) = f(0)$ then $||T|| = 1$.

Theorem 4.19

Let X be a normed linear space and let W be a dense subspace of X. Let Y be a Banach space and let $S \in B(W, Y)$.

(a) *If $x \in X$ and $\{x_n\}$ and $\{y_n\}$ are sequences in W such that $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = x$ then $\{S(x_n)\}$ and $\{S(y_n)\}$ both converge and $\lim_{n\to\infty} S(x_n) = \lim_{n\to\infty} S(y_n)$.*

(b) *There exists $T \in B(X, Y)$ such that $||T|| = ||S||$ and $Tx = Sx$ for all $x \in W$.*

Definition 4.20

Let X and Y be normed linear spaces and let $T \in L(X, Y)$. If $||T(x)|| = ||x||$ for all $x \in X$ then T is called an isometry.

Example 4.21

If X is a normed space and I is the identity linear transformation on X then I is an isometry.

Example 4.22

(a) *If $x = (x_1, x_2, x_3, \dots) \in \ell^2$ then $y = (0, x_1, x_2, x_3, \dots) \in \ell^2$.*

(b) *The linear transformation $S : \ell^2 \to \ell^2$ defined by $S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, x_3, \dots)$ is an isometry.*

Lemma 4.23

Let X and Y be normed linear spaces and let $T \in L(X, Y)$. If T is an isometry then T is bounded and $||T|| = 1$.

Definition 4.24

If X and Y are normed linear spaces and T is an isometry from X onto Y then T is called an isometric isomorphism and X and Y are called isometrically isomorphic.

Theorem 4.25

Let H be an infinite-dimensional Hilbert space over $\mathbb{F}$ with an orthonormal basis $\{e_n\}$. Then there is an isometry T of H onto $\ell^2_{\mathbb{F}}$ such that $T(e_n) = \tilde{e}_n$ for all $n \in \mathbb{N}$.

Corollary 4.26

Any infinite-dimensional, separable Hilbert space H over $\mathbb{F}$ is isometrically isomorphic to $\ell^2_{\mathbb{F}}$.

Theorem 4.27

If X is a normed linear space and Y is a Banach space then $B(X, Y)$ is a Banach space.

Definition 4.28

Let X be a normed space. Linear transformations from X to $\mathbb{F}$ are called linear functionals. The space $B(X, \mathbb{F})$ is called the dual space of X and denoted X'.

Corollary 4.29

If X is a normed vector space then X' is a Banach space.

Lemma 4.30

If X, Y and Z are normed linear spaces and $T \in B(X, Y)$ and $S \in B(Y, Z)$ then $S \circ T \in B(X, Z)$ and

$$||S \circ T|| \leq ||S||||T||.$$

Definition 4.31

Let X, Y and Z be normed linear spaces and $T \in B(X, Y)$ and $S \in B(Y, Z)$. The composition $S \circ T$ of S and T will be denoted by ST and called the product of S and T.

Notation. If X is a normed linear space the set $B(X, X)$ of all bounded linear operators from X to X will be denoted by $B(X)$.

Theorem 4.32

Let X be a normed linear space.

(a) *$B(X)$ is an algebra with identity and hence a ring with identity.*

(b) *If $\{T_n\}$ and $\{S_n\}$ are sequences in $B(X)$ such that $\lim_{n\to\infty} T_n = T$ and $\lim_{n\to\infty} S_n = S$, then $\lim_{n\to\infty} S_n T_n = S T$.*

Notation. Let X be a normed space and let $T \in B(X)$.

(a) TT will be denoted by T^2, TTT will be denoted by T^3, and more generally the product of T with itself n times will be denoted by T^n.

(b) If $a_0, a_1, \ldots, a_n \in \mathbb{F}$ and $p : \mathbb{F} \to \mathbb{F}$ is the polynomial defined by $p(x) = a_0 + a_1 x + \cdots + a_n x_n$, then we define $p(T)$ by $p(T) = a_0 I + a_1 T + \cdots + a_n T^n$.

Lemma 4.33

Let X be a normed linear space and let $T \in B(X)$. If p and q are polynomials and $\lambda, \mu \in \mathbb{C}$, then

(a) $(\lambda p + \mu q)(T) = \lambda p(T) + \mu q(T)$;

(b) $(pq)(T) = p(T)q(T)$.

Definition 4.34

Let X, Y be normed linear spaces. An operator $T \in B(X, Y)$ is said to be invertible if there exists $S \in B(Y, X)$ such that $ST = I_X$, $TS = I_Y$, in which case S is the inverse of T and is denoted by T^{-1}.

Lemma 4.35

If X, Y, Z are normed linear spaces and $T_1 \in B(X, Y)$, $T_2 \in B(Y, Z)$ are invertible, then:

(a) T_1^{-1} *is invertible with inverse* T_1;

(b) $T_2 T_1$ *is invertible with inverse* $T_1^{-1} T_2^{-1}$.

Remark 4.36

If $X = Y$ we have seen that $B(X)$ has additional algebraic properties compared with the space $B(X, Y)$. In particular, if $T \in B(X)$ then powers $T^n, n = 1, 2, \ldots,$ are well-defined. Similarly, if T is invertible, then inverse powers $T^{-n} = (T^{-1})^n, n = 1, 2, \ldots,$ are well-defined.

Definition 4.37

Let X, Y be normed linear spaces. If an invertible operator $T \in B(X, Y)$ exists then X, Y are isomorphic, and T is an isomorphism (between X and Y).

Lemma 4.38

If the normed linear spaces X, Y, are isomorphic, then:

(a) $\dim X < \infty$ *if and only if* $\dim Y < \infty$, *in which case* $\dim X = \dim Y$;

(b) X *is separable if and only if* Y *is separable;*

(c) X *is complete (i.e., Banach) if and only if* Y *is complete (i.e., Banach).*

Example 4.39

For any $h \in C[0,1]$ *let* $T_h \in B(L^2[0,1])$ *be defined by*

$$(T_h u)(t) = h(t)u(t), u \in L^2[0,1].$$

If $f \in C[0,1]$ *is defined by* $f(t) = 1 + t$, *then* T_f *is invertible.*

Theorem 4.40

Let X *be a Banach space. If* $T \in B(X)$ *is an operator with* $||T|| < 1$ *then* $I - T$ *is invertible and the inverse is given by*

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

Notation. The series in Theorem 4.40 is sometimes called the Neumann series.

Example 4.41

Let $A \in \mathbb{C}$ *and let* $k : [a, b] \times [a, b] \to \mathbb{R}$, $a, b \in \mathbb{R}$, *be defined by* $k(x, y) = A \sin(x - y)$. *Show that if* $|A| < 1$ *then for any* $f \in C[a, b]$ *there exists* $g \in C[a, b]$ *such that*

$$g(x) = f(x) + \int_0^1 k(x, y) g(y) \, dy.$$

Corollary 4.42

Let X, Y *be Banach spaces. The set* $\mathcal{A}$ *of invertible operators in* $B(X, Y)$ *is open.*

Theorem 4.43

(Open Mapping Theorem) Suppose that X *and* Y *are Banach spaces and* $T \in B(X, Y)$ *is surjective. Let*

$$L = \{T(x) : x \in X \text{ and } ||x|| \leq 1\},$$

with closure $\overline{L}$*. Then:*

(a) *there exists* $r > 0$ *such that* $\{y \in Y : ||y|| \leq r\} \subset \overline{L}$;

(b) $\{y \in Y : ||y|| \leq r/2\} \subset L$;

(c) *if, in addition,* T *is one-to-one then* T *is invertible.*

Corollary 4.44

(Closed Graph Theorem) If X *and* Y *are Banach space and* T *is a linear transformation from* X *into* Y *such that* $\mathcal{G}(T)$*, the graph of* T*, is closed, then* T *is continuous.*

Corollary 4.45

(Banach's Isomorphism Theorem) If X, Y *are Banach spaces and* $T \in B(X, Y)$ *is bijective, then* T *is invertible.*

Lemma 4.46

If X, Y *are normed linear spaces and* $T \in B(X, Y)$ *is invertible then, for all* $x \in X$, $||Tx|| \geq ||T^{-1}||^{-1}||x||$.

Lemma 4.47

Suppose that X *is a Banach space,* Y *is a normed space and* $T \in B(X, Y)$*. If there exists* $\alpha > 0$ *such that* $||Tx|| \geq \alpha||x||$ *for all* $x \in X$*, then* $\mathrm{Im}(T)$ *is closed.*

Theorem 4.48

Suppose that X, Y are Banach spaces, and $T \in B(X, Y)$. The following are equivalent:

(a) *T is invertible;*

(b) *$\mathrm{Im}(T)$ is dense in Y and there exists $\alpha > 0$ such that $||Tx|| \geq \alpha ||x||$ for all $x \in X$.*

Corollary 4.49

Suppose that X, Y are Banach spaces, and $T \in B(X, Y)$. The operator T is not invertible if and only if $\mathrm{Im}(T)$ is not dense in Y or there exists a sequence $\{x_n\}$ in X with $||x_n|| = 1$ for all $n \in \mathbb{N}$ but $\lim_{n\to\infty} T(x_n) = 0$.

Example 4.50

For any $h \in C[0,1]$, *let* $T_h \in B(L^2[0,1])$ *be defined as above. If* $f \in C[0,1]$ *is defined by* $f(t) = t$, *then* T_f *is not invertible.*

Example 4.51

For any $h \in C[0,1]$, *let* $T_h \in B(L^2[0,1])$ *be defined as above. If* $f \in C[0,1]$ *is defined by* $f(t) = 1 + t$, *then* T_f *is invertible.*

Theorem 4.52

(Uniform Boundedness Principle) Let U, X *be Banach spaces. Suppose that* S *is a non-empty set and, for each* $s \in S$, $T_s \in B(U, X)$. *If, for each* $u \in U$, *the set* $\{||T_s(u)|| : s \in S\}$ *is bounded then the set* $\{||T_s|| : s \in S\}$ *is bounded.*

Corollary 4.53

Let U, X be Banach spaces and $T_n \in B(U, X)$, $n = 1, 2, \dots$. Suppose that, $\lim_{n\to\infty} T_n u$ exists, for each $u \in U$, and define $Tu = \lim_{n\to\infty} T_n u$. Then $T \in B(U, X)$.

The End