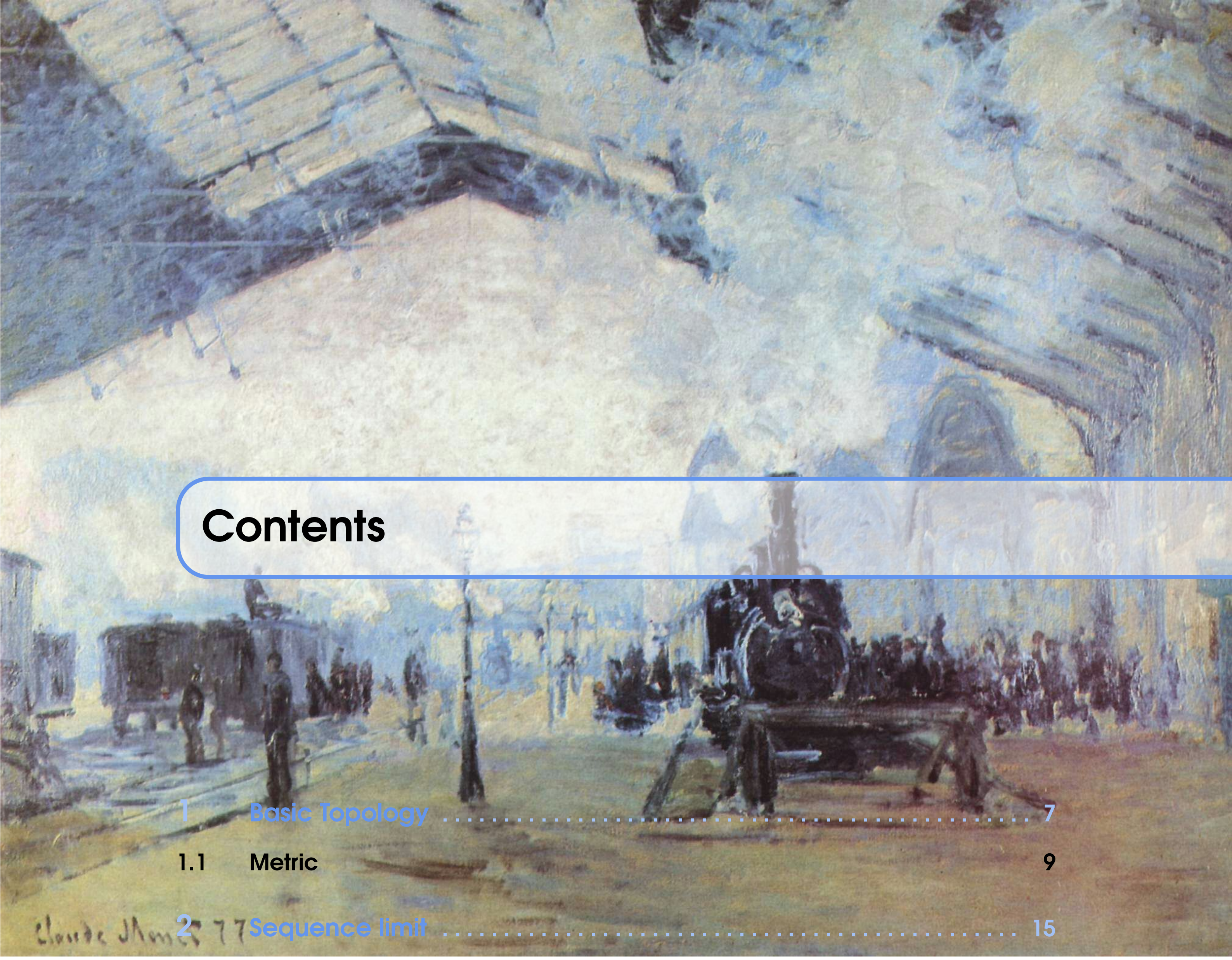
Analysis

Xin Wen



Contents

1. Basic Topology

Definition 1.0.1 For any positive integer n, let $\mathbb{N}_n$ be the set whose elements are the integers $1,2,\ldots,n$. let $\mathbb{N}$ be the set consisting of all positive integers. For any set A, we say:

1. A is finite if $A \sim N_n$ for some n (the empty set is also considered to be finite).
2. A is infinite if A is not finite.
3. A is countable if $A \sim N$.
4. A is uncountable if A is neither finite nor countable.
5. A is at most countable if A is finite or countable.

Proposition 1.0.1
1. A finite set cannot be equivalent to one of its proper subsets.
2. A is infinite if A is equivalent to one of its proper subsets.
3. Every infinite subset of a countable set A is countable.
4. No uncountable set can be a subset of a countable set.

Theorem 1.0.2 Let $\{E_n\}, n = 1,2,3,\ldots,$ be a sequence of countable sets, and put

$$S = \bigcup_{n=1}^{\infty} E_n$$

Then S is countable.

Proof. Let every set E_n be arranged in a sequence $\{x_{ni}\}, k = 1,2,3,\ldots$ and consider the infinite array

$$\begin{array}{cccc} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

in which the elements of E_n form the n th row. The array contains all elements of S. As indicated by the arrows, these elements can be arranged in a sequence

$$x_{11}; x_{21}, x_{12}; x_{31}, x_{22}, x_{13}; x_{41}, x_{32}, x_{23}, x_{14} : \ldots$$

If any two of the sets E_n have elements in common, these will appear more than once in last equation. Hence there is a subset T of the set of all positive integers such that $S \sim T$, which shows that S is at most countable. since $E_1 \subset S$, and E_1 is infinite, S is infinite, and thus countable. ■

Theorem 1.0.3 Let A be a countable set, and let B_n be the set of all n -tuples, $(a_1, \ldots, a_n)$, where $a_k \in A (k = 1, \ldots, n)$, and the elements $a_1, \ldots, a_n$ need not be distinct. Then B_n is countable.

Proof. That B_1 is countable is evident, since $B_1 = A$. Suppose B_{n-1} is countable $(n = 2,3,4,\ldots)$. The clements of B_n are of the form (18) $(b,a) \quad (b \in B_{n-1}, a \in A)$ For every fixed b, the set of pairs (b,a) is equivalent to A, and hence countable. Thus B_n is the union of a countable set of countable sets. By Theorem 2.12, B_n is countable. The theorem follows by induction.

■

Theorem 1.0.4 Let A be the set of all sequences whose elements are the digits 0 and 1. This set A is uncountable. The elements of A are sequences like $1,0,0,1,0,1,1,1,\ldots$

Proof. Let E be a countable subset of A, and let E consist of the sequences $s_1, s_2, s_3, \ldots$. We construct a sequence s as follows. If the n th digit in s_n is 1, we let the n^{th} digit of s be 0, and vice versa. Then the sequence s differs from every member of E in at least one place; hence $s \notin E$. But clearly $s \in A$, so that E is a proper subset of A We have shown that every countable subset of A is a proper subset of A. It follows that A is uncountable (for otherwise A would be a proper subset of A, which is absurd). ■

R The set of all real numbers is uncountable.

1.1 Metric

Definition 1.1.1 A set X, whose elements we shall call points, is said to be a $metricspace$ if with any two points p and q of X there is associated a real number $d(p,q)$, called the distance from p to q, such that

1. $d(p,q) > 0$ if $p \neq q; d(p,p) = 0$
2. $d(p,q) = d(q,p)$
3. $d(p,q) \leq d(p,r) + d(r,q)$, for any $r \in X$

Any function with these three properties is called a distance function, or a metric.

Definition 1.1.2 If $a_i < b_i$ for $i = 1,\dots,k$, the set of all points $\mathbf{x} = (x_1,\dots,x_k)$ in R^k whose coordinates satisfy the inequalities $a_i \leq x_i \leq b_i (1 \leq i \leq k)$ is called a **k-cell.** Thus a 1-cell is an interval, a 2-cell is a rectangle, etc.

Definition 1.1.3 We call a set $E \subset R^k$ **convex** if

$$\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$$

whenever $\mathbf{x} \in E, \mathbf{y} \in E$, and $0 < \lambda < 1$ For example, balls are convex, and k-cells are convex.

Definition 1.1.4 Let X be a metric space. All points and sets mentioned below are understood to be elements and subsets of X.

1. A neighborhood of a point p is a set $N_r(p)$ consisting of all points q such that $d(p,q) < r$. The number r is called the radius of $N_r(p)$
2. A point p is a limit point of the set E if every neighborhood of p contains a point $q \neq p$ such that $q \in E$
3. If $p \in E$ and p is not a limit point of E, then p is called an isolated point of E
4. E is closed if every limit point of E is a point of E
5. A point p is an interior point of E if there is a neighborhood N of p such that $N \subset E$
6. E is open if every point of E is an interior point of E
7. The complement of E (denoted by E^c) is the set of all points $p \in X$ such that $p \notin E$
8. E is perfect if E is closed and if every point of E is a limit point of E
9. E is bounded if there is a real number M and a point $q \in X$ such that $d(p,q) < M$ for all $p \in E$
10. E is dense in X if every point of X is a limit point of E, or a point of E(or both)

Theorem 1.1.1 Every neighborhood is an open set.

Proof. Consider a neighborhood $E = N_r(p)$, and let q be any point of E. Then there is a positive real number h such that

$$d(p,q) = r - h$$

For all points s such that $d(q,s) < h$, we have then

$$d(p,s) \leq d(p,q) + d(q,s) < r - h + h = r$$

so that $s \in E$. Thus q is an interior point of E ■

Theorem 1.1.2
1. If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E
2. A finite point set has no limit points.

Definition 1.1.5 If X is a metric space, if $E \subset X$, and if E' denotes the set of all limit points of E in X, then the closure of E is the set $\bar{E} = E \cup E'$

Theorem 1.1.3 If X is a metric space and $E \subset X$, then
1. $\bar{E}$ is closed,
2. $E = \bar{E}$ if and only if E is closed,
3. $\bar{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$

By (1) and (3), $\bar{E}$ is the smallest closed subset of X that contains E.

Proof.
1. If $p \in X$ and $p \notin \bar{E}$ then p is neither a point of E nor a limit point of E Hence p has a neighborhood which does not intersect E. The complement of $\bar{E}$ is the refore open. Hence $\bar{E}$ is closed.
2. If $E = \bar{E}, (a)$ implies that E is closed. If E is closed, then $E' \subset E$ [by Definitions $2.18(d)$ and 2.26], hence $\bar{E} = E$
3. If F is closed and $F \supset E$, then $F \supset F'$, hence $F \Rightarrow E'$. Thus $F \supset E$

■

R 只要包涵 E 的闭集一定包含 E 的闭包

Theorem 1.1.4 Let E be a nonempty set of real numbers which is bounded above. Let $y = \sup E$. Then $y \in \bar{E}$. Hence $y \in E$ if E is closed.

Proof. If $y \in E$ then $y \in E$. Assume $y \notin E$. For every $h > 0$ there exists then a point $x \in E$ such that $y - h < x < y$, for otherwise $y - h$ would be an upper bound of E. Thus y is a limit point of E. Hence $y \in \bar{E}$

■

Theorem 1.1.5 Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X

Proof. Suppose E is open relative to Y. To each $p \in E$ there is a positive number r_p such that the conditions $d(p,q) < r_p, q \in Y$ imply that $q \in E$ Let V_p be the set of ail $q \in X$ such that $d(p,q) < r_p$, and define

$$G = \bigcup_{p \in E} V_p$$

Then G is an open subset of X, by Theorems 2.19 and 2.24 since $p \in V_p$ for all $p \in E$, it is clear that $E \subset G \cap Y$ By our choice of V_p, we have $V_p \cap Y \subset E$ for every $p \in E$, so that $G \cap Y \subset E$. Thus $E = G \cap Y$, and one half of the theorem is proved. Conversely, if G is open in X and $E = G \cap Y$, every $p \in E$ has a neighborhood $V_p \subset G$. Then $V_p \cap Y \subset E$, so that E is open relative to Y

■

Definition 1.1.6 — Compact. By an open cover of a set E in a metric space X we mean a collection $\{G_a\}$ of open subsets of X such that $E \subset U_a G_a$

Definition 1.1.7 A subset K of a metric space X is said to be compact if every open cover of K contains a finite subcover.

More explicitly, the requirement is that if $\{G_a\}$ is an open cover of K, then there are finitely many indices $\alpha_1, \ldots, \alpha_n$ such that

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$$

R every finite set is compact.

Theorem 1.1.6 Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.

Theorem 1.1.7 Compact subsets of metric spaces are closed.

Proof. Let K be a compact subset of a metric space X. We shall prove that the complement of K is an open subset of X. Suppose $p \in X, p \notin K$. If $q \in K$, let V_q and W_q be neighborhoods of p and q, respectively, of radius less than $\frac{1}{2}d(p,q)$ [see Definition 2.18(a)] . since K is compact, there are finitely many points $q_1, \ldots, q_n$ in K such that $K \subset W_{q_1} \cup \cdots \cup W_{q_n} = W$ If $V = V_{q_1} \cap \cdots \cap V_{q_n}$, then V is a neighborhood of p which does not intersect W. Hence $V \subset K^c$, so that p is an interior point of K^c. The theorem follows. ■

Theorem 1.1.8 Closed subsets of compact sets are compact.

Theorem 1.1.9 If F is closed and K is compact, then $F \cap K$ is compact.

Theorem 1.1.10 If $\{K_x\}$ is a collection of **compact** subsets of a metric space X such that the intersection of every finite subcollection of $\{K_a\}$ is nonempty, then $\cap K_\alpha$ is nonempty.

Proof. Fix a member K_1 of $\{K_a\}$ and put $G_a = K_a^c$. Assume that no point of K_1 belongs to every K_α. Then the sets G_α form an open cover of K_1 and since K_1 is compact, there are finitely many indices $\alpha_1, \ldots, \alpha_n$ such that $K_1 \subset G_{a_1} \cup \cdots \cup G_{a_n}$. But this means that

$$K_1 \cap K_{a_1} \cap \cdots \cap K_{a_n}$$

is empty, in contradiction to our hypothesis. ■

Corollary 1.1.11 If $\{K_n\}$ is a sequence of nonempty compact sets such that $K_n \supset K_{n+1}$ $(n = 1, 2, 3, \ldots)$, then $\cap_1^\infty K_n$ is not empty.

Theorem 1.1.12 If E is an infinite subset of a compact set K, then E has a limit point in K.

Proof. If no point of K were a limit point of E, then each $q \in K$ would have a neighborhood V_q which contains at most one point of E (namely. q, if $q \in E$). It is clear that no finite subcollection of $\{V_q\}$ can cover E and the same is true of K, since $E \subset K$. This contradicts the compactness of K

■

Corollary 1.1.13 If $\{I_n\}$ is a sequence of intervals in R^1, such that $I_n \supset I_{n+1}$ $(n = 1, 2, 3, \ldots)$, then $\bigcap_1^\infty I_n$ is not empty.

Theorem 1.1.14 Let k be a positive integer. If I_n is a sequence of $k - cells$ such that $I_n \supset I_{n+1} (n = I, 2, 3, ...)$, then $\bigcap_1^\infty I_n$ is not empty.

Theorem 1.1.15 Every k-cell is compact.

Proof. Let I be a k -cell, consisting of all points $x = (x_1, \ldots, x_k)$ such that $a_j \leq x_j \leq b_j (1 \leq j \leq k)$. Put

$$\delta = \left\{ \sum_i^k (b_j - a_j)^2 \right\}^{1/2}$$

Then $\rfloor \mathbf{x}-\mathbf{y}| \leq \delta$, if $\mathbf{x} \in I, \mathbf{y} \in I$ Suppose, to get a contradiction, that there exists an open cover $\{G_a\}$ of I which contains no finite subcover of I. Put $c_j=\left(a_j+b_j\right)/2$. The intervals $\left[a_j, c_j\right]$ and $\left[c_j, b_j\right]$ then determine $2^k k$ -cells Q_i whose union is I At least one of these sets Q_i, call it I_1, cannot be covered by any finite subcollection of $\{G_\alpha\}$ (otherwise I could be so covered). We next subdivide I_1 and continue the process. We obtain a sequence $\{I_n\}$ with the following properties:

1. $I \supset I_1 \supset I_2 \supset I_3 \supset \cdots$
2. I_n is not covered by any finite subcollection of $\{G_a\}$
3. if $\mathbf{x} \in I_n$ and $\mathbf{y} \in I_n$, then $|\mathbf{x}-\mathbf{y}| \leq 2^{-n} \delta$

By (a) and Theorem 2.39, there is a point x^* which lies in every I_n For some $\alpha, \mathbf{x}^* \in G_\alpha$. since G_α is open, there exists $r>0$ such that $|\mathbf{y}-\mathbf{x}^*|<r$ implies that $\mathbf{y} \in G_\alpha$. If n is so large that $2^{-n} \delta<r$ (there is such an n, for otherwise $2^n \leq \delta / r$ for all positive integers n, which is absurd since R is archimedean), then (c) implies that $I_n \subset G_\alpha$, which contradicts (b). This completes the proof. ■

Theorem 1.1.16 If a set E in R^k has one of the following three properties, then it has the other two:

1. E is closed and bounded.
2. E is compact.
3. Every infinite subset of E has a limit point in E.

Proof. If (a) holds, then $E \subset I$ for some k -cell I, and (b) follows from Theorems 2.40 and 2.35. Theorem 2.37 shows that (b) implies (c). It remains to be shown that (c) implies (a) If E is not bounded, then E contains points $\mathbf{x}_n$ with

$$|\mathbf{x}_n|>n \quad (n=1,2,3, \ldots)$$

The set S consisting of these points x_n is infinite and clearly has no limit point in R^k, hence has none in E. Thus (c) implies that E is bounded. If E is not closed, then there is a point $x_0 \in R^k$ which is a limit point of E but not a point of E. For $n=1,2,3, \ldots$, there are points $\mathbf{x}_n \in E$ such that $|\mathbf{x}_n-\mathbf{x}_0|<1 / n$. Let S be the set of these points $\mathbf{x}_n$. Then S is infinite (otherwise $|\mathbf{x}_n-\mathbf{x}_0|$ would have a constant positive value, for infinitely many n), S has x_0 as a limit point, and S has no other limit point in R^k. For if $\mathbf{y} \in R^k, \mathbf{y} \neq \mathbf{x}_0$, then

$$\begin{aligned} |\mathbf{x}_n-\mathbf{y}| & \geq |\mathbf{x}_0-\mathbf{y}|-|\mathbf{x}_n-\mathbf{x}_0| \\ & \geq |\mathbf{x}_0-\mathbf{y}|-\frac{1}{n} \geq \frac{1}{2}|\mathbf{x}_0-\mathbf{y}| \end{aligned}$$

for all but finitely many n; this shows that y is not a limit point of S (Theorem 2.20).

Thus S has no limit point in E; hence E must be closed if (c) holds.

■

R (2) and (3) are equivalent in any metric space but that (1) does not, in general, imply (2) and (3).

Theorem 1.1.17 Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.

Definition 1.1.8 — Separated. Two subsets A and B of a metric space X are said to be separated if both $A \cap \bar{B}$ and $\bar{A} \cap B$ are empty, i.e., if no point of A lies in the closure of B and no point of B lies in the closure of A. A set $E \subset X$ is said to be connected if E is not a union of two nonempty separated sets.

R Separated sets are of course disjoint, but disjoint sets need not be separated. For example, the interval [0,1] and the segment (1,2) are not separated, since 1 is a limit point of $(1,2)$. However, the segments (0,1) and (1,2) are separated.

The connected subsets of the line have a particularly simple structure:

Definition 1.1.9 — connected. A subset E of the real line R^1 is connected if and only if it has the following property: If $x \in E, y \in E$, and $x < z < y$, then $z \in E$.

Proof. If there exist $x \in E, y \in E$, and some $z \in (x,y)$ such that $z \notin E$, then

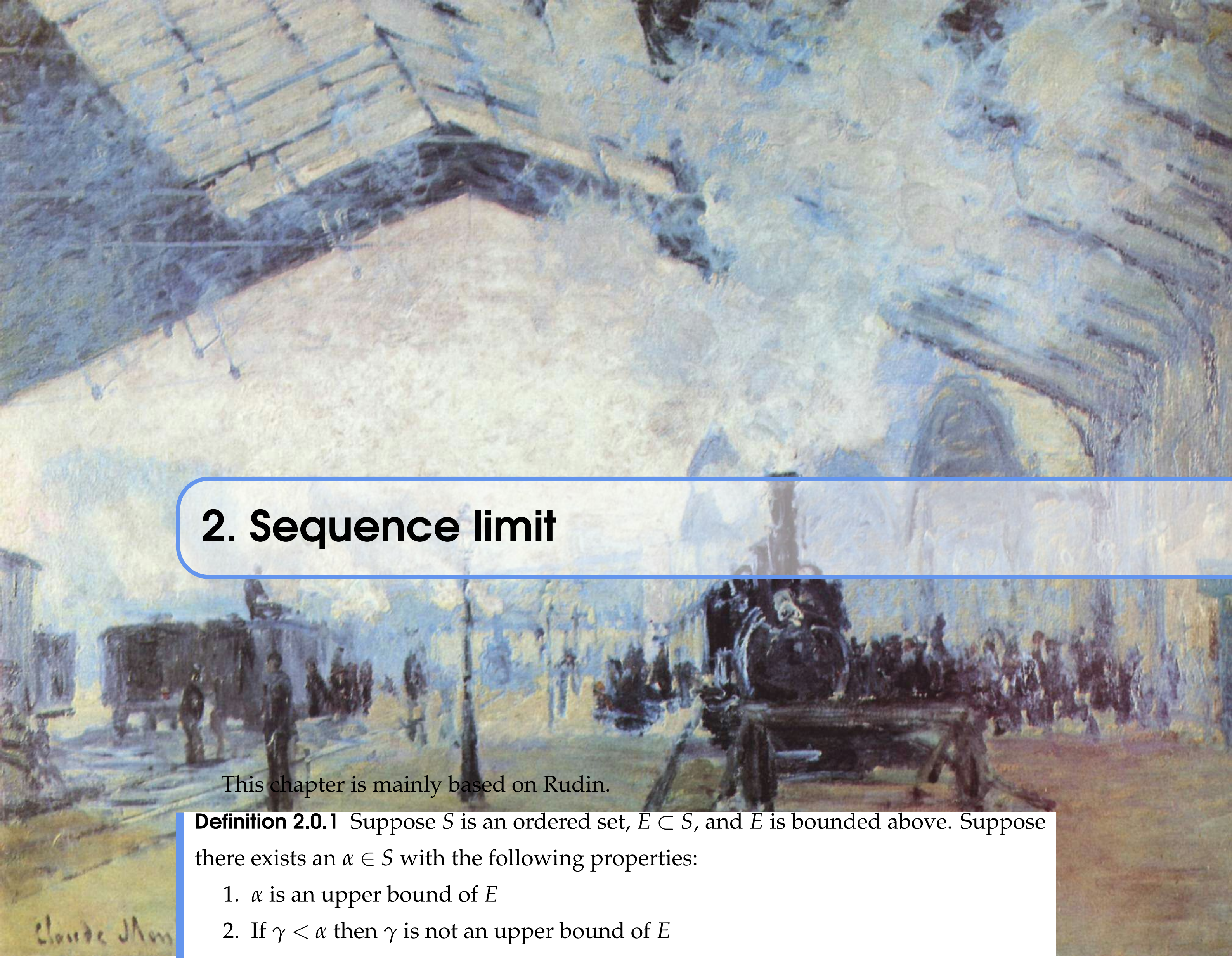
$E = A_z \cup B_z$ where

$$A_z = E \cap (-\infty, z), \quad B_z = E \cap (z, \infty)$$

since $x \in A_z$ and $y \in B_z, A$ and B are nonempty. since $A_z \subset (-\infty, z)$ and $B_z \subset (z, \infty)$, they are separated. Hence E is not connected. To prove the converse, suppose E is not connected. Then there are nonempty separated sets A and B such that $A \cup B = E$. Pick $x \in A, y \in B$ and assume (without loss of generality) that $x < y$. Define

$$z = \sup(A \cap [x,y])$$

By Theorem 2.28, $z \in \bar{A}$; hence $z \notin B$. In particular, $x \leq z < y$ If $z \notin A$, it follows that $x < z < y$ and $z \notin E$ If $z \in A$, then $z \notin \bar{B}$, hence there exists z_1 such that $z < z_1 < y$ and $z_1 \notin B$. Then $x < z_1 < y$ and $z_1 \notin E$ ∎



2. Sequence limit

This chapter is mainly based on Rudin.

Definition 2.0.1 Suppose S is an ordered set, $E \subset S$, and E is bounded above. Suppose there exists an $\alpha \in S$ with the following properties:

1. α is an upper bound of E
2. If $\gamma < \alpha$ then γ is not an upper bound of E

Then α is called the **least upper bound** of E [that there is at most one such α is clear from (ii)] or the **supremum** of E, and we write

$$\alpha = \sup E$$

The greatest lower bound, or **infimum**, of a set E which is bounded below is defined in the same manner: The statement $\alpha = \inf E$ means that α is a lower bound of E and that no β with $\beta > \alpha$ is a lower bound of E

R If $a = \sup E$ exists, then a may or may not be a member of E.

Definition 2.0.2 An ordered set S is said to have the least-upper-bound property if the following is true : if $E \subset S$, E is not empty, and E is bounded above, then $\sup E$ exists in S.

Theorem 2.0.1 Suppose S is an ordered set with the least-upper-bound property, $B \subset S$, B is not empty, and B is bounded below. Let L be the set of all lower bounds of B.

Then

$$\alpha = \sup L$$

exists in S, and $\alpha = \inf B$. In particular, $\inf B$ exists in S.

Proof. Since B is bounded below, L is not empty. since L consists of exactly those $y \in S$ which satisfy the inequality $y \leq x$ for every $x \in B$, we see that every $x \in B$ is an upper bound of L. Thus L is bounded above. Our hypothesis about S implies therefore that L has a supremum in S; call it α. Then following to prove $\alpha = \inf B$.

If $\gamma < \alpha$ then (see Definition)γ is not an upper bound of L hence $\gamma \notin B$. It follows that $\alpha \leq x$ for every $x \in B$. Thus $\alpha \in L$ If $\alpha < \beta$ then $\beta \notin L$, since α is an upper bound of L We have shown that $\alpha \in L$ but $\beta \notin L$ if $\beta > \alpha$. In other words, α is a lower bound of B, but β is not if $\beta > \alpha$. This means that $\alpha = \inf B$. ■

Theorem 2.0.2 — Archimedean property of $\mathbb{R}$. If $x \in \mathbb{R}, y \in \mathbb{R}$, and $x > 0$, then there is a positive integer n such that $nx > y$.

If $x \in \mathbb{R}, y \in \mathbb{R}$, and $x > y$ then there exists a $p \in \mathbb{Q}$ such that $x < p < y$.

Definition 2.0.3 — converge. A sequence $\{p_n\}$ in a metric space X is said to **converge** if there is a point $p \in X$ with the following property: For every $\varepsilon > 0$ there is an integer N such that $n \geq N$ implies that $d(p_n, p) < \varepsilon$. (Here d denotes the distance in X.) In this case we also say that $\{p_n\}$ converges to p, or that p is the limit of $\{p_n\}$, and we write $p_n \rightarrow p$, or

$$\lim_{n \to \infty} p_n = p$$

If $\{p_n\}$ does not converge, it is said to diverge.

Theorem 2.0.3 Let $\{p_n\}$ be a sequence in a metric space X.

1. $\{p_n\}$ converges to $p \in X$ if and only if every neighborhood of p contains all but finitely many of the terms of $\{p_n\}$
2. If $p \in X, p' \in X$, and if $\{p_n\}$ converges to p and to p', then $p' = p$
3. If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.
4. If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = \lim_{n \to \infty} p_n$

Definition 2.0.4 — Subsequence. Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3 < \cdots$. Then the sequence $\{p_n\}$ is called a subsequence of $\{p_n\}$. If $\{p_n\}$ converges, its limit is called a subsequential limit of $\{p_n\}$

It is clear that $\{p_n\}$ converges to p if and only if every subsequence of $\{p_n\}$ converges to p.

Theorem 2.0.4 1. If $\{p_n\}$ is a sequence in a compact metric space X, then some subsequence of $\{p_n\}$ converges to a point of X.
2. Every bounded sequence in R^k contains a convergent subsequence.

Proof. Let E be the range of $\{p_n\}$. If E is infinite, then E has a limit point $p \in X$. Choose n_1 so that $d(p, p_{n_1}) < 1$. Having chosen $n_1, \cdots, n_{i-1}$, then there is an integer $n_i > n_{i-1}$ such that $d(p, p_{n_i}) < 1/i$. Then $\{p_n\}$ converges to p. ■

Theorem 2.0.5 The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subset of X.

2.1 Cauchy sequence

Definition 2.1.1 — Cauchy sequence 1. A sequence $\{p_n\}$ in a metric space X is said to be a **Cauchy sequence** if for every $\varepsilon > 0$ there is an integer N such that $d(p_n, p_m) < \varepsilon$ if $n > N$ and $m > N$.

Definition 2.1.2 Let E be a subset of a metric space X, and let S be the set of all real numbers of the form $d(p,q)$, with $p \in E$ and $q \in E$. The sup of S is called the diameter of E.

Definition 2.1.3 — Cauchy sequence 2. If $\{p_n\}$ is a sequence in X and if E_N consists of the points $p_N, p_{N+1}, p_{N+2}, \ldots$ it is clear from the two preceding definitions that $\{p_n\}$ is a Cauchy sequence if and only if

$$\lim_{N\to\infty} \operatorname{diam} E_N = 0$$

Theorem 2.1.1 1. If $\bar{E}$ is the closure of a set E in a metric space X, then

$$\operatorname{diam} \bar{E} = \operatorname{diam} E$$

2. If K_n is a sequence of compact sets in X such that $K_n \supset K_{n+1}$ $(n = 1,2,3,\ldots)$ and if $\lim_{n\to\infty} \operatorname{diam} K_n = 0$ then $\bigcap_1^x K_n$ consists of exactly one point.

Proof. Just prove the first theorem. Since $E \subset \bar{E}$, it is clear that

$$\operatorname{diam} E \leq \operatorname{diam} \bar{E}$$

Fix $s > 0$, and choose $p \in \bar{E}, q \in \bar{E}$. By the definition of $\bar{E}$, there are points p', q', in E such that $d\left(p,p'\right)<\varepsilon, d\left(q,q'\right)<\varepsilon$. Hence

$$\begin{aligned} d(p,q) &\leq d\left(p,p'\right)+d\left(p'q'\right)+d\left(q',q\right) \\ &< 2\varepsilon+d\left(p',q'\right) \leq 2\varepsilon+\operatorname{diam} E \end{aligned}$$

It follows that

$$\operatorname{diam} \bar{E} \leq 2\varepsilon+\operatorname{diam} E$$

and since ε was arbitrary, (a) is proved.

■

Theorem 2.1.2

1. In any metric space X, every convergent sequence is a Cauchy sequence.
2. If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some points of X.
3. In $\mathbb{R}^k$, every Cauchy sequence converges.

Proof. Let $\{p_n\}$ be a Cauchy sequence in the compact space X. For $N=1,2,3,\ldots$, let E_N be the set consisting of $p_N, p_{N+1}, p_{N+2}, \ldots$ Then

$$\lim_{N\to\infty} \operatorname{diam} \bar{E}_N = 0$$

by Definition 3.9 and Theorem 3.10(a). Being a closed subset of the compact space X, each $\bar{E}_N$ is compact (Theorem 2.35). Also $E_N \supset E_{N+1}$ so that $\bar{E}_N \supset \bar{E}_{N+1}$ Theorem 3.10(b) shows now that there is a unique $p \in X$ which lies in every $\bar{E}_N$ Let $\varepsilon>0$ be given. By (3) there is an integer N_0 such that $\operatorname{diam} \bar{E}_N<\varepsilon$ if $N \geq N_0$. since $p \in \bar{E}_N$, it follows that $d(p,q)<\varepsilon$ for every $q \in \bar{E}_N$, hence for every $q \in E_N$. In other words, $d\left(p,p_n\right)<\varepsilon$ if $n \geq N_0$. This says precisely that $p_n \to p$ ■

R Compact 的空间里的 Cauchy 就是收敛的，Complete 是由 Cauchy 收敛定义的

Definition 2.1.4 A metric space in which every Cauchy sequence converges is said to be **complete**.

R All compact metric spaces and all Euclidean spaces are complete. And every closed subset E of a complete metric space X is complete.

Definition 2.1.5 A sequence $\{s_n\}$ of real numbers is said to be

1. monotonically increasing if $s_n \leq s_{n+1} (n=1,2,3,\ldots)$
2. monotonically decreasing if $s_n \geq s_{n+1} (n=1,2,3,\ldots)$

The class of monotonic sequences consists of the increasing and the decreasing sequences.

Theorem 2.1.3 Suppose $\{s_n\}$ is monotonic. Then $\{s_n\}$ converges if and only if it is bounded.

R 单调有界没有空间的限制条件

2.2 UPPER AND LOWER LIMITS

Definition 2.2.1 Let $\{s_n\}$ be a sequence of real numbers with the following property: For every real M there is an integer N such that $n > N$ implies $s_n \geq M$, denoted by $s_n \to \infty$.

Definition 2.2.2 Let $\{s_n\}$ be a sequence of real numbers. Let E be the set of numbers x (in the extended real number system) such that $s_{n_k} \to x$ for some subsequence $\{s_{n_k}\}$. This set E contains all **subsequential limits**, plus possibly the numbers $+\infty, -\infty$.

$$s^* = \sup E$$

$$s_* = \inf E$$

The numbers s^*, s_* are called **the upper and lower limits** of $\{s_n\}$; we use the notation

$$\limsup_{n\to\infty} s_n = s^*, \quad \liminf_{n\to\infty} s_n = s_*$$

Theorem 2.2.1 Let $\{s_n\}$ be a sequence of real numbers. Let E and s^* have the same meaning as in previous Definition. Then s^* has the following two properties:

1. $s^* \in E$
2. If $x > s^*$, there is an integer N such that $n \geq N$ implies $s_n < x$. Moreover, s^* is the only number with the properties (a) and (b). Of course, an analogous result is true for s_*

Theorem 2.2.2 If $s_n \leq t_n$ for $n \geq N$, where N is fixed, then

$$\liminf_{n\to\infty} s_n \leq \liminf_{n\to\infty} t_n$$

$$\limsup_{n\to\infty} s_n \leq \limsup_{n\to\infty} t_n$$

Proposition 2.2.3 1.

$$\limsup x_{n_k} \leq \limsup x_n$$

2.
$$\liminf x_{n_k} \geq \liminf x_n$$

3.
$$\liminf x_n \leq \lim x_{n_k} \leq \limsup x_n$$

4.
$$\exists \{x_{n_k}\}, \lim x_{n_k} = \limsup x_n$$

5.
$$\exists \{x_{n_k}\}, \lim x_{n_k} = \liminf x_n$$

6.
$$\overline{\lim_{n\to\infty}}\, x_n = \lim_{n\to\infty} \sup_{k=n} x_k = \inf_n \sup x_k \; \underline{\lim_{n\to\infty}}\, x_n = \lim_{n\to\infty} \inf_{n\to\infty} x_k = \sup_n \inf_{k\geqslant n} x_k$$

7.
$$\underline{\lim_{n\to\infty}}\, x_n \cdot \underline{\lim_{n\to\infty}}\, y_n \leqslant \underline{\lim_{n\to\infty}}\, x_n y_n \leqslant \underline{\lim_{n\to\infty}}\, x_n \cdot \overline{\lim_{n\to\infty}}\, y_n$$

8.
$$\underline{\lim_{n\to\infty}}\, x_n \cdot \overline{\lim_{n\to\infty}}\, y_n \leqslant \overline{\lim_{n\to\infty}}\, x_n y_n \leqslant \overline{\lim_{n\to\infty}}\, x_n \cdot \overline{\lim_{n\to\infty}}\, y_n$$

Theorem 2.2.4

$$\liminf_{n\to\infty} x_n + \liminf_{n\to\infty} y_n \leq \liminf_{n\to\infty} (x_n + y_n) \leq \liminf_{n\to\infty} x_n + \limsup_{n\to\infty} y_n \tag{2.1}$$

Theorem 2.2.5

$$\liminf_{n\to\infty} x_n + \limsup_{n\to\infty} y_n \leq \limsup_{n\to\infty} (x_n + y_n) \leq \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n \tag{2.2}$$

2.3 SOME SPECIAL SEQUENCES

The proofs will all be based on the following remark: If $0 \leq x_n \leq s_n$ for $n \geq N$ where N is some fixed number, and if $s_n \to 0$, then $x_n \to 0$

Theorem 2.3.1 1. If $p > 0$, then $\lim_{n\to\infty} \frac{1}{n^p} = 0$

2. If $p > 0$, then $\lim_{n\to\infty} \sqrt[n]{p} = 1$
3. $\lim_{x\to\infty} \sqrt[n]{n} = 1$
4. If $p > 0$ and α is real, then $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$
5. If$|x| < 1$, then $\lim x^n = 0$

Proof. Prove 4. (d) Let k be an integer such that $k > \alpha, k > 0$. For $n > 2k$

$$(1+p)^n > \binom{n}{k} p^k = \frac{n(n-1)\cdots(n-k+1)}{k!} p^k > \frac{n^k p^k}{2^k k!}$$

Hence

$$0 < \frac{n^a}{(1+p)^n} < \frac{2^k k!}{p^k} n^{x-k} \quad (n > 2k)$$

since $\alpha - k < 0, n^{\alpha-k} \to 0$, by (a)

■

2.4 Series

Definition 2.4.1 Given a sequence $\{a_n\}$, we use the notation

$$\sum_{n=p}^{q} a_n \quad (p \leq q)$$

to denote the sum $a_p + a_{p+1} + \cdots + a_q$.

With $\{a_n\}$ we associate a sequence $\{s_n\}$, where

$$s_n = \sum_{k=1}^{n} a_k$$

For $\{s_n\}$ we also use the symbolic expression

$$a_1 + a_2 + a_3 +$$

or, more concisely,

$$\sum_{n=1}^{\infty} a_n$$

The symbol we call an **infinite series**, or just a series.

The numbers s_n are called the **partial sums** of the series. If $\{s_n\}$ converges to s, we say that the **series converges**, and write

$$\sum_{n=1}^{\infty} a_n = s$$

The number s is called the sum of the series; but it should be clearly understood that s is the limit of a sequence of sums, and is not obtained simply by addition. If $\{s_n\}$ diverges, the series is said to diverge.

Theorem 2.4.1 — Cauchy criterion. Σa_n converges if and only if for every $\varepsilon > 0$ there is an integer N such that $|\sum_{k=n}^{m} a_k| \leq \varepsilon$ if $m \geq n \geq N$

Theorem 2.4.2 If $\sum_{n=0}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.(By m=n)

Theorem 2.4.3 A series of nonnegative1 terms converges if and only if its partial sums form a bounded sequence.

Theorem 2.4.4 — Comparsion Test.

1. If $|a_n| \leq c_n$ for $n \geq N_0$, where N_0 is some fixed integer, and if Σc_n converges, then Σa_n converges.
2. If $a_n \geq d_n \geq 0$ for $n \geq N_0$, and if Σd_n diverges, then Σa_n diverges.

Theorem 2.4.5 If $0 \leq x < 1$, then

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

If $x \geq 1$, the series diverges.

Proof. If $x \neq 1$

$$s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$$

■

Theorem 2.4.6 Suppose $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the series $\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$ converges.

Proof. By Theorem 3.24, it suffices to consider boundedness of the partial sums. Let

$$s_n = a_1 + a_2 + \cdots + a_n$$
$$t_k = a_1 + 2a_2 + \cdots + 2^k a_{2^k}$$

For $n < 2^k$

$$\begin{aligned} s_n &\leq a_1 + (a_2 + a_3) + \cdots + (a_{2k} + \cdots + a_{2k+1-1}) \\ &\leq a_1 + 2a_2 + \cdots + 2^k a_{2^*} \\ &= t_k \end{aligned}$$

so that

$$s_n \leq t_k$$

On the other hand, if $n > 2^k$

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \\ &\geq \frac{1}{2} a_1 + a_2 + 2a_4 + \cdots + 2^{k-1} a_{2^k} \\ &= \frac{1}{2} t_k \end{aligned}$$

so that

$$2s_n \geq t_k$$

The sequences $\{s_n\}$ and $\{t_k\}$ are either both bounded or both unbounded. This completes the proof. ■

Theorem 2.4.7 $\sum \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \leq 1$

Proof.

$$\sum_{k=0}^{\infty} 2^k \cdot \frac{1}{2^{k_p}} = \sum_{k=0}^{\infty} 2^{(1-p)k}$$

Now, $2^{1-p} < 1$ if and only if $1 - p < 0$, and the result follows by comparison with the geometric series (take $x = 2^{1-p}$ in Theorem 3.26). ■

Theorem 2.4.8 If $p > 1$

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$$

converges; if $p \leq 1$, the series diverges.

Proof.

$$\sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k (\log 2^k)^p} = \sum_{k=1}^{\infty} \frac{1}{(k \log 2)^p} = \frac{1}{(\log 2)^p} \sum_{k=1}^{\infty} \frac{1}{k^p}$$

■

Definition 2.4.2

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}$$

Here $n! = 1 \cdot 2 \cdot 3 \cdots n$ if $n \geq 1$, and $0! = 1$ Since $s_n = 1 + 1 + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \cdots + \frac{1}{1 \cdot 2 \cdots n}$ $< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} < 3$, the series converges,

Theorem 2.4.9

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e$$

Proof. Let

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}, \quad t_n = \left(1 + \frac{1}{n}\right)^n$$

By the binomial theorem,

$$t_n = 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \cdots$$
$$+ \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right)$$

Hence $t_n \leq s_n$, so that (14)

$$\limsup t_n \leq e$$

by Theorem 3.19. Next, if $n \geq m$

$$t_n \geq 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \cdots + \frac{1}{m!}\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{m-1}{n}\right)$$

Let $n \to \infty$, keeping m 'fixed. We get

$$\liminf_{n \to x} t_n \geq 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!}$$

so that

$$s_m \leq \liminf t_n$$

Letting $m \to \infty$, we finally get (I5)

$$e \leq \liminf_{n \to \infty} t_n$$

The theorcm follows from (14) and (15).

■

R

$$e - s_n = \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots$$
$$< \frac{1}{(n+1)!}\left\{1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots\right\} = \frac{1}{n!n}$$

(16)

$$0 < e - s_n < \frac{1}{n!n}$$

Thus s_{10}, for instance, approximates e with an error less than 10^{-7}. The inequality (16) is of theoretical interest as well, since it enables us to prove the irrationality of e very easily.

Theorem 2.4.10 e is irrational.

Proof. Suppose e is rational. Then $e = p/q$, where p and q are positive integers. By (16) (17)

$$0 < q!\left(e - s_q\right) < \frac{1}{q}$$

By our assumption, $q!e$ is an integer. Since

$$q!s_q = q!\left(1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{q!}\right)$$

is an integer, we see that $q!\left(e - s_q\right)$ is an integer. since $q \geq 1, (17)$ implies the existence of an integer between 0 and 1. We have thus reached a contradiction. ■

2.4.1 THE ROOT AND RATIO TESTS

Theorem 2.4.11 — Root Test. Giren Σa_n, put $\alpha = \lim_{n\to\infty} \sup_{n\to\infty} \sqrt[n]{|a_n|}$ Then

1. if $\alpha < 1, \Sigma a_n$ converges:
2. if $\alpha > 1, \Sigma a_n$ dicerges;
3. if $\alpha = 1$, the test gires no information.

R This theorem does not require $\sum a_n$ should exist, so it take limsup.

Proof. If $\alpha < 1$, we can choose β so that $\alpha < \beta < 1$, and an integer N such that

$$\sqrt[n]{|a_n|} < \beta$$

for $n \geq N$. That is, $n \geq N$ implies

$$|a_n| < \beta^n$$

since $0 < \beta < 1, \Sigma\beta^n$ converges. Convergence of Σa_n follows now from the comparison test.

If $\alpha > 1$, then, again by Theorem 3.17, there is a sequence $\{n_k\}$ such that

Hence $|a_n| > 1$ for infinitely many values of n, so that the condition $a_n \to 0$, necessary for convergence of Σa_n, does not hold (Theorem 3.23. To prove (c), we consider the series

$$\Sigma\frac{1}{n}, \Sigma\frac{1}{n^2}$$

For each of these series $\alpha = 1$, but the first diverges, the second converges. ■

Theorem 2.4.12 — Ratio Test. The series Σa_n, not requiring exists, so the following is about limsup and liminf.

1. converges if $\lim_{n\to\infty}\sup_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1$
2. diverges if $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ for $n \geq n_0$, where n_0 is some fixed integer.
3. diverges if $\gamma = \lim\inf\left|\frac{a_{n+1}}{a_n} > 1\right|$
4. no information if $\gamma \leq 1 < \lim\sup\left|\frac{a_{n+1}}{a_n} > 1\right|$

Proof. If condition (a) holds, we can find $\beta < 1$, and an integer N, such that

$$\left|\frac{a_{n+1}}{a_n}\right| < \beta$$

for $n \geq N$. In particular,

$$\begin{aligned}
&|a_{N+1}| < \beta\,|a_N| \\
&|a_{N+2}| < \beta\,|a_{N+1}| < \beta^2\,|a_N| \\
&\dots\dots\dots\dots\dots\dots \\
&|a_{N+p}| < \beta^p\,|a_N|
\end{aligned}$$

That is,

$$|a_n| < |a_N|\,\beta^{-N}\cdot\beta^n$$

for $n \geq N$, and (a) follows from the comparison test, since $\Sigma\beta^n$ converges. If $|a_{n+1}| \geq |a_n|$ for $n \geq n_0$, it is easily seen that the condition $a_n \to 0$ does not hold, and $(\ b\ )$ follows. ■

R The knowledge that $\lim a_{n+1}/a_n = 1$ implies nothing about the convergence of Σa_n. The series $\Sigma 1/n$ and $\Sigma 1/n^2$ demonstrate this.

Theorem 2.4.13 For any sequence $\{c_n\}$ of positive numbers,

$$\begin{aligned}
&\lim_{n\to\infty}\inf\tfrac{c_{n+1}}{c_n} \leq \lim\inf_{n\to\infty}\sqrt[n]{c_n} \\
&\lim_{n\to\infty}\sup\sqrt[n]{c_n} \leq \lim_{n\to\infty}\sup\tfrac{c_{n+1}}{c_n}
\end{aligned}$$

Proof. We shall prove the second inequality; the proof of the first is quite similar. Put

$$\alpha = \limsup_{n\to\infty}\frac{c_{n+1}}{c_n}$$

If $\alpha = +\infty$, there is nothing to prove. If α is finite, choose $\beta > \alpha$. There is an integer N such that

$$\frac{c_{n+1}}{c_n} \leq \beta$$

for $n \geq N$. In particular, for any $p > 0$

$$c_{N+k+1} \leq \beta c_{N+k} \quad (k = 0, 1, \ldots, p-1)$$

Multiplying these inequalities, we obtain

$$c_{N+p} \leq \beta^p c_N$$

or

$$c_n \leq c_N \beta^{-N} \cdot \beta^n \quad (n \geq N)$$

Hence

$$\sqrt[n]{c_n} \leq \sqrt[n]{c_N \beta^{-N}} \cdot \beta$$

so that (18)

$$\limsup \sqrt[n]{c_n} \leq \beta$$

Since (1 8) is true for every $\beta > \alpha$, we have

$$\limsup \sqrt[n]{c_n} \leq \alpha$$

■

2.4.2 Power series

Theorem 2.4.14 Given the power series $\Sigma c_n z^n$, *put*

$$\alpha = \lim_{n \to \infty} \sup \sqrt[n]{|c_n|}, \quad R = \frac{1}{\alpha}$$

(If $\alpha = 0, R = +\infty$; if $\alpha = +\infty, R = 0$.) Then $\Sigma c_n z^n$ converges if $|z| < R$, and diverges if $|z| > R$.

Proof. Put $a_n = c_n z^n$, and apply the root test:

$$\limsup_{n \to \infty} \sqrt[n]{|a_n|} = |z| \lim_{n \to \infty} \sup_{\sqrt{n}} \sqrt[n]{|c_n|} = \frac{|z|}{R}$$

Note: R is called the **radius of convergence** of $\Sigma c_n z^n$ ■

Theorem 2.4.15 Given two sequences $\{a_n\}, \{b_n\}$, put $A_n = \sum_{k=0}^{n} a_k$ if $n \geq 0$; put $A_{-1} = 0$. Then, if $0 \leq p \leq q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

Proof. $\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} (A_n - A_{n-1}) b_n = \sum_{n=p}^{q} A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1}$ ∎

Theorem 2.4.16 Suppose

1. the partial sums A_n of Σa_n form a bounded sequence;
2. $b_0 \geq b_1 \geq b_2 \geq \cdots$
3. $\lim_{n\to\infty} b_n = 0$

Then $\Sigma a_n b_n$ converges.

Proof. Choose M such that $|A_n| \leq M$ for all n. Given $\varepsilon > 0$, there is an integer N such that $b_N \leq (\varepsilon/2M)$. For $N \leq p \leq q$, we have

$$\begin{aligned} \left|\sum_{n=p}^{q} a_n b_n\right| &= \left|\sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p\right| \\ &\leq M \left|\sum_{n=p}^{q-1} (b_n - b_{n+1}) + b_q + b_p\right| \\ &= 2Mb_p \leq 2Mb_N \leq \varepsilon \end{aligned}$$

Convergence now follows from the Cauchy criterion. We note that the first inequality in the above chain depends of course on the fact that

$$b_n - b_{n+1} \geq 0$$

∎

Theorem 2.4.17 Suppose

1. $|c_1| \geq |c_2| \geq |c_3| \geq \cdots$
2. $c_{2m-1} \geq 0, c_{2m} \leq 0 \quad (m = 1,2,3,\ldots)$
3. $\lim_{n\to\infty} c_n = 0$

Then Σc_n converges.

Theorem 2.4.18 Suppose the radius of convergence of $\Sigma c_n z^n$ is 1, and suppose $c_0 \geq c_1 \geq c_2 \geq \cdots, \lim_{n\to\infty} c_n = 0$. Then $\Sigma c_n z^n$ converges at every point on the circle $|z| = 1$, except possibly at $z = 1$

Proof. Put $a_n = z^n, b_n = c_n$. The hypotheses of Theorem 3.42 are then satisfied, since

$$|A_n| = \left|\sum_{m=0}^{n} z^m\right| = \left|\frac{1 - z^{m+1}}{1 - z}\right| \leq \frac{2}{|1 - z|}$$

if $|z| = 1, z \neq 1$

∎

Definition 2.4.3 The series Σa_n is said to **converge absolutely** if the series $\Sigma |a_n|$ converges.

Theorem 2.4.19 If Σa_n converges absolutely, then Σa_n converges.

Proof. The assertion follows from the inequality

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{n=k}^{m} |a_k|$$

plus the Cauchy criterion. ■

Definition 2.4.4 Given Σa_n and Σb_n, we put

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} \quad (n = 0,1,2,\ldots)$$

and call Σc_n the *product* of the two given series.

Theorem 2.4.20 Suppose

1. $\sum_{n=0}^{\infty} a_n$ converges absolutely,
2. $\sum_{n=0}^{\infty} a_n = A$
3. $\sum_{n=0}^{\infty} b_n = B$
4. $c_n = \sum_{k=0}^{n} a_k b_{n-k} \quad (n = 0,1,2,\ldots)$

Then

$$\sum_{n=0}^{\infty} c_n = AB$$

That is, the product of two convergent series converges, and to the right value, if at least one of the two series converges absolutely.

Proof. Put $A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{k=0}^{n} b_k, \quad C_n = \sum_{k=0}^{n} c_k, \quad \beta_n = B_n - \Delta$ Then

$$\begin{aligned} C_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + \cdots + a_n B_0 \\ &= a_0 (B + \beta_n) + a_1 (B + \beta_{n-1}) + \cdots + a_n (B + \beta_0) \end{aligned}$$

$= A_n B + a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0$

Put

$$\gamma_n = a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0$$

We wish to show that $C_n \to AB$. since $A_n B \to AB$, it suffices to show that (21)

$$\lim_{n\to\infty} \gamma_n = 0$$

Put

$$\alpha = \sum_{n=0}^{\infty} |a_n|$$

[It is here that we use (a).] Let $\varepsilon > 0$ be given. By $(c), \beta_n \to 0$. Hence we can choose N such that $|\beta_n| \leq \varepsilon$ for $n \geq N$, in which case

$$\begin{aligned} |\gamma_n| &\leq |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + |\beta_{N+1} a_{n-N-1} + \cdots + \beta_n a_0| \\ &\leq |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + \varepsilon\alpha \end{aligned}$$

Keeping N fixed, and letting $n \to \infty$, we get

$$\lim_{n\to\infty} |\gamma_n| \leq \varepsilon\alpha$$

since $a_k \to 0$ as $k \to \infty$. since ε is arbitrary, (21) follows. ■

R Another question which may be asked is whether the series Σc_n, if convergent, must have the sum AB. Abel showed that the answer is in the affirmative.

Theorem 2.4.21 If the series $\Sigma a_n, \Sigma b_n, \Sigma c_n$ converge to A, B, C, and $c_n = a_0 b_n + \cdots + a_n b_0$, then $C = AB$

Definition 2.4.5 — REARRANGEMENTS. Let $\{k_n\}, n = 1,2,3,\ldots$, be a sequence in which every positive integer appears once and only once (that is, $\{k_n\}$ is a $1-1$ function from J to J, in the notation of Definition 2.2). Putting

$$a_n' = a_{k_n} \quad (n = 1,2,3,\ldots)$$

we say that $\Sigma a_n'$ is a rearrangement of Σa_n

Theorem 2.4.22 Let Σa_n be a series of real numbers which converges, but not absolutely. Suppose $-\infty \leq \alpha \leq \beta \leq \infty$ Then there exists a rearrangement $\Sigma a_n'$ with partial sums s_n' such that

$$\text{lim inf } s_n' = \alpha, \text{ lim sup } s_n' = \beta$$

Theorem 2.4.23 If Σa_n is a series of complex numbers which converges absolutely, then every rearrangement of Σa_n converges, and the" all converge to the same sum.

Proof. Let $\Sigma a_n'$ be a rearrangement, with partial sums s_n'. Given $\varepsilon > 0$ there exists an integer N such that $m \geq n \geq N$ implies (26) $\sum_{i=n}^{m} |a_i| \leq \varepsilon$ Now choose p such that the integers $1,2,\ldots,N$ are all contained in the set $k_1, k_2, \ldots, k_p$ (we use the notation of Definition 3.52). Then if $n > p$ the numbers $a_1, \ldots, a_N$ will cancel in the difference $s_n - s_n'$, so that $|s_n - s_n'| \leq \varepsilon$, by (26). Hence $\{s_n'\}$ converges to the same sum as $\{s_n\}$ ■

3. Continuous

Definition 3.0.1 Let X and Y be metric spaces. Suppose $E \subset X, f$ maps E into Y, and p is a limit point of E. We write $f(x) \to q$ as $x \to p$, or

$$\lim_{x \to p} f(x) = q$$

if there is a point $q \in Y$ with the following property: For every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), q) < \varepsilon$$

for all points $x \in E$ for which

$$0 < d_x(x, p) < \delta$$

The symbols d_x and d_y refer to the distances in X and Y, respectively.

Definition 3.0.2 Let X, Y, E, f, and p be as in Definition 3.0.1 . Then

$$\lim_{x \to p} f(x) = q$$

if and only if

$$\lim_{n \to \infty} f(p_n) = q$$

for every sequence $\{p_n\}$ in E such that

$$p_n \neq p, \quad \lim_{n \to \infty} p_n = p$$

Corollary 3.0.1 If f has a limit at p, this limit is unique.

Definition 3.0.3 — continuous. Suppose X and Y are metric spaces, $E \subset X, p \in E$, and f maps E into Y. Then f is said to be **continuous** at p if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \varepsilon$$

for all points $x \in E$ for which $d_x(x, p) < \delta$. If f is continuous at every point of E, then f is said to be continuous on E.

R It should be noted that f has to be defined at the point p in order to be continuous at p. (Compare this with Definition 3.0.1)

R Isolated point is continuous point.

R Assume also that p is a limit point of E. Then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$

Theorem 3.0.2 A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y.

Proof. Suppose f is continuous on X and V is an open set in Y. We have to show that every point of $f^{-1}(V)$ is an interior point of $f^{-1}(V)$. So, suppose $p \in X$ and $f(p) \in \mathcal{V}$. since V is open, there exists $\varepsilon > 0$ such that $y \in V$ if $d_Y(f(p), y) < \varepsilon$; and since f is continuous at p, there exists $\delta > 0$ such that $d_y(f(x), f(p)) < \varepsilon$ if $d_x(x, p) < \delta$. Thus $x \in f^{-1}(V)$ as soon as $d_X(x, p) < \delta$

Conversely, suppose $f^{-1}(V)$ is open in X for every open set V in Y Fix $p \in X$ and $\varepsilon > 0$, let V be the set of all $y \in Y$ such that $d_y(y, f(p)) < \varepsilon$ Then V is open; hence $f^{-1}(V)$ is open; hence there exists $\delta > 0$ such that $x \in f^{-1}(V)$ as soon as $d_x(p, x) < \delta$. But if $x \in f^{-1}(V)$, then $f(x) \in V$, so that $d_Y(f(x), f(p)) < \varepsilon$ This completes the proof. ■

Theorem 3.0.3 From the triangle inequality one sees easily that

$$||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}| \quad \left(\mathbf{x}, \mathbf{y} \in R^k\right)$$

Hence the mapping $x \to |x|$ is a continuous real function on R^k.

R 所以距离，范数都是连续的

Definition 3.0.4 — Bounded. A mapping f of a set E into R^k is said to be bounded if there is a real number M such that $|\mathbf{f}(x)| \leq M$ for all $x \in E$

Theorem 3.0.4 Suppose f is a continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact. (i) then compere metric space Y.

R We have used the relation $f\left(f^{-1}(E)\right) \subset E$. valid for $E \subset Y$. If $E \subset X$, then $f^{-1}(f(E)) \supset E$; equality need not hold in either case.

Theorem 3.0.5 If $f(x)$is a continuous mapping of a compact metric space X into R^k, then f(X) is closed and bounded. Thus, f is bounded.

Theorem 3.0.6 Suppose f is a continuous real function on a compact metric space X, and

$$M = \sup_{p\in X} f(p), m = \inf_{p\in X} f(p)$$

Then there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$

R The conclusion may also be stated as follows: There exist points p and q in X such that $f(q) \leq f(x) \leq f(p)$ for all $x \in X$: that is. f attains its maximum (at p) and its minimum (at q).

Theorem 3.0.7 4.17 Theorem Suppose f is a continuous $1-1$ mapping of a compact metric space X onto a metric space Y. Then the inverse mapping f^{-1} defined on Y by

$$f^{-1}(f(x)) = x \quad (x \in \mathrm{X})$$

is a continuous mapping of Y onto X

Proof. Applying Theorem 4.8 to f^{-1} in place of f, we see that it suffices to prove that $f(V)$ is an open set in Y for every open set V in X. Fix such a set $\mathcal{V}$ The complement V^c of V is closed in X, hence compact (Theorem 2.35); hence $f\left(V^c\right)$ is a compact subset of Y (Theorem 4.14) and so is closed in Y (Theorem 2.34). since f is one-to-one and onto, $f(V)$ is the complement of $f\left(V^c\right)$. Hence $f(V)$ is open.

■

Definition 3.0.5 — uniformly continuous. Let f be a mapping of a metric space X into a metric space Y. We say that f is uniformly continuous on X if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(p), f(q)) < \varepsilon$$

for all p and q in X for which $d_x(p,q) < \delta$

Theorem 3.0.8 Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is uniformly continuous on X.

Proof. Let $\varepsilon > 0$ be given. since f is continuous, we can associate to each point $p \in X$ a positive number $\phi(p)$ such that

$$q \in X, d_X(p,q) < \phi(p) \text{ implies } d_Y(f(p), f(q)) < \frac{\varepsilon}{2}$$

Let $J(p)$ be the set of all $q \in X$ for which (17)

$$d_X(p,q) < \frac{1}{2}\phi(p)$$

since $p \in J(p)$, the collection of all sets $J(p)$ is an open cover of X; and since X is compact, there is a finite set of points $p_1, \ldots, p_n$ in X, such that (18) $X \subset J(p_1) \cup \cdots \cup J(p_n)$ We put (19) $_2\min[\phi(p_1), \ldots, \phi(p_n)]$ Then $\delta > 0$. (This is one point where the finiteness of the covering, inherent in the definition of compactness. is essential. The minimum of a finite set of positive numbers is positive, whereas the inf of an infinite set of positive numbers may very well be 0.)

Now let q and p be points of X, such that $d_x(p,q) < \delta$. By (18). there is an integer $m, 1 \le m \le n$, such that $p \in J(p_m)$; hence (20) $d_x(p, p_m) < \frac{1}{2}\phi(p_m)$ and we also have $d_X(q, p_m) \le d_x(p,q) + d_x(p, p_m) < \delta + \frac{1}{2}\phi(p_m) \le \phi(p_m)$ Finally, (16) shows that therefore $d_Y(f(p), f(q)) \le d_Y(f(p), f(p_m)) + d_Y(f(q), f(p_m)) < \varepsilon$ This completes the proof. ■

■ **Example 3.1** Let E be a noncompact set in R^1. Then

1. there exists a contimuous function on E which is not bounded; $f(x) = \frac{1}{x - x_0} \quad (x \in E)$, x_0 is not the point of E.
2. there exists a continuous and bounded function on E which has no maximum. The function g given by

   $$g(x) = \frac{1}{1 + (x - x_0)^2} \quad (x \in E)$$

   is continuous on E, and is bounded, since $0 < g(x) < 1$. It is clear that

   $$\sup_{x \in E} g(x) = 1$$

   whereas $g(x) < 1$ for all $x \in E$. Thus g has no maximum on E

3. If, in addition, E is bounded, then there exists a continuous function on E which is not uniformly continuous.

Theorem 3.0.9 If $f(x)$is a continuous mapping of a metric space X into a metric the space Y, and if E is a connected subset of X, then $f(E)$ is connected.

Proof. Assume, on the contrary, that $f(E) = A \cup B$, where A and B are nonempty separated subsets of Y. Put $G = E \cap f^{-1}(A), H = E \cap f^{-1}(B)$ Then $E = G \cup H$, and neither G nor H is empty. since $A \subset A$ (the closure of A), we have $G \subset f^{-1}(\bar{A})$; the latter set is closed, since f is continuous: hence $\bar{G} \subset f^{-1}(\bar{A})$. It follows that $f(\bar{G}) \subset \bar{A}$ since $f(H) = B$ and $\bar{A} \cap B$ is empty, we conclude that $\bar{G} \cap H$ is empty. The same argument shows that $G \cap \bar{H}$ is empty. Thus G and H are separated. This is impossible if E is connected. ■

Theorem 3.0.10 Let f be a continuous real function on the interval $[a,b]$. If the the portion $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a,b)$ such that $f(x) = c$. A similar result holds, of course, if $f(a) > f(b)$.

Roughly speaking, the theorem says that a continuous real function assumes all intermediate values on an interval.

Proof. By Theorem 2.47, [a, b] is connected: hence Theorem 4.22 shows that $f([a,b])$ is a connected subset of R^1, and the assertion follows if we appeal once more to Theorem 2.47.

■

3.0.1 DISCONTINUITIES

Definition 3.0.6 Let f be defined on (a,b). Consider any point x such that $a \leq x < b$. We write

$$f(x+) = q$$

if $f(t_n) \to q$ as $n \to \infty$, for all sequences $\{t_n\}$ in (x,b) such that $t_n \to x$. To obtain the definition of $f(x-)$, for $a < x \leq b$, we restrict ourselves to sequences $\{t_n\}$ in (a,x) It is clear that any point x of (a,b),$\lim f(t)$ exists if and only if

$$f(x+) = f(x-) = \lim_{t \to x} f(t)$$

Definition 3.0.7 Let f be defined on (a,b). If f is discontinuous at a point x, and if $f(x+)$ and $f(x-)$ exist. then f is said to have a discontinuity of the first kind, or a simple discontinuity, at x. Otherwise the discontinuity is said to be of the second kind. There are two ways in which a function can have a simple discontinuity: either $f(x+) \neq f(x-)$ [in which case the value $f(x)$ is immaterial], or $f(x+) = f(x-) \neq f(x)$

Definition 3.0.8 — MONOTONIC FUNCTIONS. Let f be real on (a,b). Then f is said to be monotonically to increasing on (a,b) if $a<x<y<b$ implies $f(x) \leq f(y)$. If the last inequality is reversed, we obtain the definition of a monotonically decreasing function. The class of monotonic functions consists of both the increasing and the decreasing functions.

Theorem 3.0.11 Let f be monotonically increasing on (a,b). Then $f(x+)$ and $f(x-)$ exist at every point of x of (u,b). More precisely, $\sup_{a<t<x} f(t) = f(x-\cdot) \leq f(x) \leq f(x+) = \inf_{x<t<b} f(t)$. Furthermore, if $a<x<y<b$, then $f(x+) \leq f(y-)$. Analogous results evidently hold for monotonically decreasing functions.

Proof. By hypothesis, the set of numbers $f(t)$, where $a<t<x$, is bounded above by the number $f(x)$, and therefore has a least upper bound which we shall denote by A. Evidently $A \leq f(x)$. We have to show that $A = f(x-)$ Let $\varepsilon > 0$ be given. It follows from the definition of A as a least upper bound that there exists $\delta > 0$ such that $a < x-\delta < x$ and

$$A - \varepsilon < f(x-\delta) \leq A$$

since f is monotonic, we have

$$f(x-\delta) \leq f(t) \leq A \quad (x-\delta < t < x)$$

Combining (27) and (28), we see that

$$|f(t) - A| < \varepsilon \quad (x-\delta < t < x)$$

Hence $f(x-) = A$ The second half of (25) is proved in precisely the same way. Next, if $a<x<y<b$, we see from (25) that

$$f(x+) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t)$$

The last equality is obtained by applying (25) to (a,y) in place of (a,b) Similarly,

$$f(y-) = \sup_{a<t<y} f(t) = \sup_{x<t<y} f(t)$$

Comparison of (29) and (30) gives (26). ■

Theorem 3.0.12 Monotonic functions have no discontinuities of the second kind.

Theorem 3.0.13 Let f be monotonic on (a,b). Then the set of points of $(a,b) at$ which f is discontinuous is at most countable.

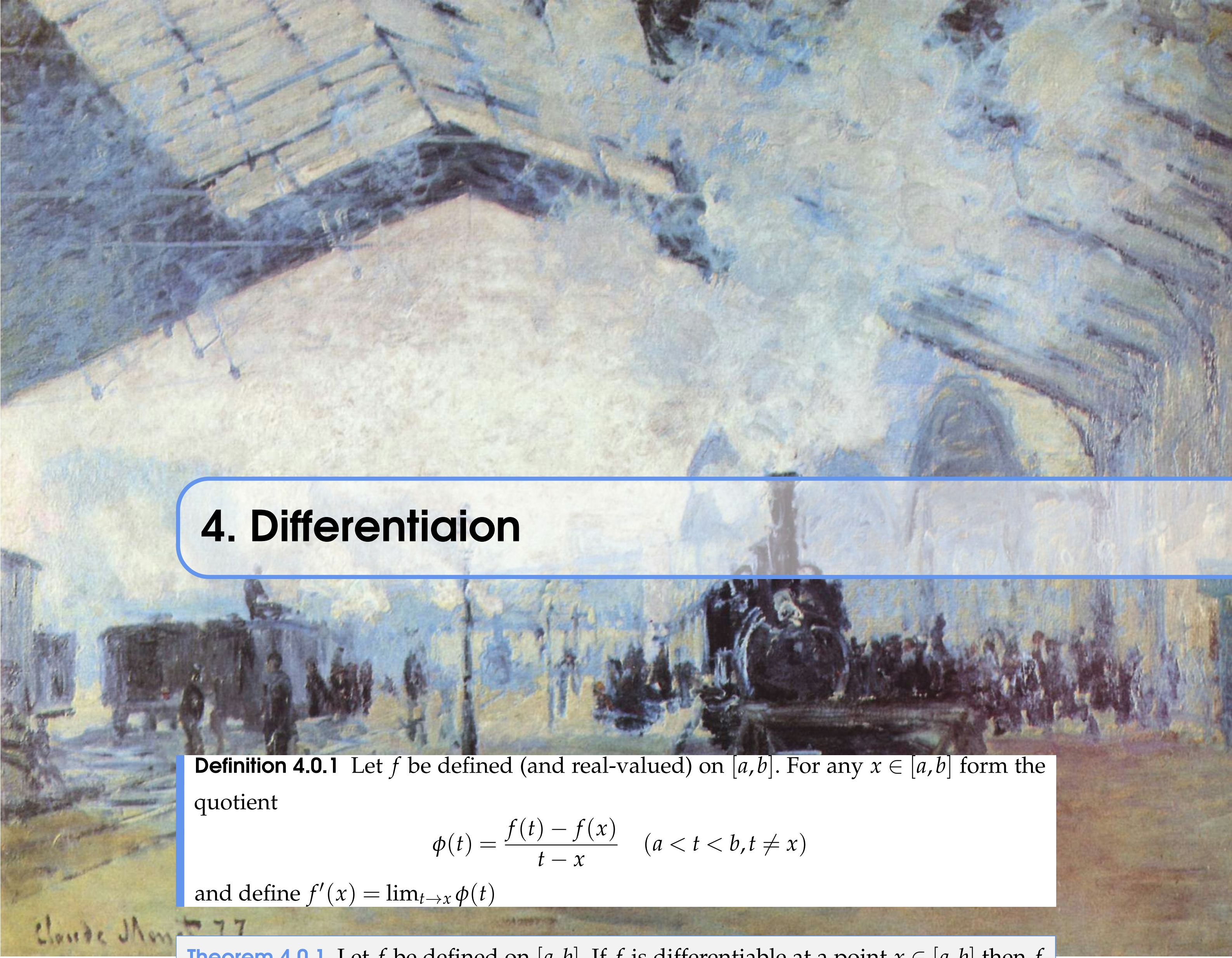
Proof. Suppose for the sake of definiteness, that f is increasing, and let E be the set of points at which f is discontinuous.

With every point x of E we associate a rational number $r(x)$ such that

$$f(x-) < r(x) < f(x+)$$

since $x_1 < x_2$ implies $f\left(x_1+\right) \leq f\left(x_2-\right)$, we see that $r\left(x_1\right) \neq r\left(x_2\right)$ if $x_1 \neq x_2$ We have thus established a $1-1$ correspondence between the set E and a subset of the set of rational numbers. The latter. as we know, is countable.

■



4. Differentiaion

Definition 4.0.1 Let f be defined (and real-valued) on $[a,b]$. For any $x \in [a,b]$ form the quotient

$$\phi(t) = \frac{f(t)-f(x)}{t-x} \quad (a<t<b, t \neq x)$$

and define $f'(x) = \lim_{t \to x} \phi(t)$

Theorem 4.0.1 Let f be defined on $[a,b]$. If f is differentiable at a point $x \in [a,b]$ then f is continuous at x

Definition 4.0.2 Let f be a real function defined on a metric space X. We say that f has a **local maximum** at a point $p \in X$ if there exists $\delta > 0$ such that $f(q) \leq f(p)$ for all $q \in X$ with $d(p,q) < \delta$

Theorem 4.0.2 Let f be defined on $[a,b]$; if f has a local maximum at a point $x \in (a,b)$, and if $f'(x)$ exists, then $f'(x) = 0$

Proof. If $x - \delta < t < x$, then

$$\frac{f(t)-f(x)}{t-x} \geq 0$$

Letting $t \to x$, we see that $f'(x) \geq 0$ ■

Theorem 4.0.3 — generalized mean-value theorem. If f and g are continuous real functions on $[a,b]$ which are differentiable in (a,b), then there is a point $x \in (a,b)$ at which

$$[f(b)-f(a)]g'(x) = [g(b)-g(a)]f'(x)$$

Note that differentiability is not required at the end points.

Proof. Put

$$h(t) = [f(b)-f(a)]g(t) - [g(b)-g(a)]f(t) \quad (a \le t \le b)$$

Then h is continuous on $[a,b], h$ is differentiable in (a,b), and (12)

$$h(a) = f(b)g(a) - f(a)g(b) = h(b)$$

To prove the theorem, we have to show that $h'(x) = 0$ for some $x \in (a,b)$ If h is constant, this holds for every $x \in (a,b)$. If $h(t) > h(a)$ for some $t \in (a,b)$, let x be a point on $[a,b]$ at which h attains its maximum (Theorem 4.16). By $(12), x \in (a,b)$, and Theorem 5.8 shows that $h'(x) = 0$ If $h(t) < h(a)$ for some $t \in (a,b)$, the same argument applies if we choose for x a point on $[a,b]$ where h attains its minimum.

■

Theorem 4.0.4 — Mean-value theorem. If f is a real continuous function on $[a,b]$ which is differentiable in (a,b), then there is a point $x \in (a,b)$ at which

$$f(b) - f(a) = (b-a)f'(x)$$

Theorem 4.0.5 Suppose f is a real differentiable function on $[a,b]$ and suppose $f'(a) < \lambda < f'(b)$. Then there is a point $x \in (a,b)$ such that $f'(x) = \lambda$

R 导函数的介质性

Corollary 4.0.6 If f is differentiable on $[a,b]$, then f' cannot have any simple discontinuities on $[a,b]$

Theorem 4.0.7 — L'shospital. Suppose f and g are real and differentiable in (a,b), and $g'(x) \neq 0$ for all $x \in (a,b)$, where $-\infty \le a < b \le +\infty$. Suppose

$$\frac{f'(x)}{g'(x)} \to A \text{ as } x \to a$$

If $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$ or if $g(x) \to +\infty$ as $x \to a$ then

$$\frac{f(x)}{g(x)} \to A \text{ as } x \to a$$

Proof. We first consider the case in which $-\infty \leq A < +\infty$. Choose a real number q such that $A < q$, and then choose r such that $A < r < q$, then there is a point $c \in (a,b)$ such that $a < x < c$ implies

$$\frac{f'(x)}{g'(x)} < r$$

If $a < x < y < c$, then Cauchy M-V theorem shows that there is a point $t \in (x,y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r$$

Suppose the first condition holds. Letting $x \to a$, we see that

$$\frac{f(y)}{g(y)} \leq r < q \quad (a < y < c)$$

Next, suppose the second holds. Keeping y fixed, we can choose a point $c_1 \in (a,y)$ such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < c_1$. Multiplying (18) by $[g(x) - g(y)]/g(x)$, we obtain

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)} (a < x < c_1)$$

If we let $x \to a$,then there is a point $c_2 \in (a,c_1)$ such that $\frac{f(x)}{g(x)} < q \quad (a < x < c_2)$ Summing up, (19) and (21) show that for any q, subject only to the condition $A < q$, there is a point c_2 such that $f(x)/g(x) < q$ if $a < x < c_2$ In the same manner, if $-\infty < A \leq +\infty$, and p is chosen so that $p < A$, we can find a point c_3 such that (22) $p < \frac{f(x)}{g(x)} \quad (a < x < c_3)$ and (16) follows from these two statements.

■

Theorem 4.0.8 — TAYLOR'S THEOREM. Suppose f is a real function on $[a,b], n$ is a positive integer, $f^{(n-1)}$ is continuous on $[a,b], f^{(n)}(t)$ exists for every $t \in (a,b)$. Let α, β be distinct points of $[a,b]$, and define

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k$$

Then there exists a point x between α and β such that $f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n$. For $n = 1$, this is just the mean value theorem. In general, the theorem shows that f can be approximated by a polynomial of degree $n-1$, and that (24) allows us to estimate the error, if we know bounds on $\left|f^{(n)}(x)\right|$

Proof. Proof Let M be the number defined by (25) $f(\beta) = P(\beta) + M(\beta - \alpha)^n$ and put (26) $g(t) = f(t) - P(t) - M(t-x)^n \quad (a \leq t \leq b)$ We have to show that $n!M = f^{(n)}(x)$ for some x between α and β. By (23) and (26) (27) $g^{(n)}(t) = f^{(n)}(t) - n!M \quad (a < t < b)$ Hence the proof will be complete if we can show that $g^{(n)}(x) = 0$ for some x between α and β since $P^{(k)}(x) = f^{(k)}(\alpha)$ for $k = 0, \ldots, n-1$. we have (28) $g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0$ Our choice of M shows that $g(\beta) = 0$, so that $g'(x_1) = 0$ for some x_1 between x and β. by the mean value theorem. since $g'(x) = 0$, we conclude similarly that $g''(x_2) = 0$ for some x_2 between α and x_1. After n steps we arrive at the conclusion that $g^{(n)}(x_n) = 0$ for some x_n between x and x_{n-1} that is. between x and β ■

4.0.1 DIFFERENTIATION OF VECTOR-VALUED FUNCTIONS

Theorem 4.0.9 Suppose **f** is a continuous mapping of [a.b] into R^k and **f** is differentiable in (a,b). Then there exists $x \in (a.b)$ such that

$$|\mathbf{f}(h) - \mathbf{f}(a)| \leq (b-a)\left|\mathbf{f}'(x)\right|$$

Proof. [1] Put $z = f(b) - f(a)$, and define

$$\varphi(t) = \mathbf{z} \cdot \mathbf{f}(t) \quad (a \leq t \leq b)$$

Then φ is a real-valued continuous function on $[a,b]$ which is differentiable in (a. b). The mean value theorem shows therefore that

$$\varphi(b) - \varphi(a) = (b-a)\varphi'(x) = (b-a)z \cdot \mathbf{f}'(x)$$

for some $x \in (a,b)$. On the other hand,

$$\varphi(h) - \varphi(a) = \mathbf{z} \cdot \mathbf{f}(b) - \mathbf{z} \cdot \mathbf{f}(a) = \mathbf{z} \cdot \mathbf{z} = |\mathbf{z}|^2$$

The Schwar7 inequality now gives

$$|\mathbf{z}|^2 = (b-a)\left|\mathbf{z} \cdot \mathbf{f}'(x)\right| \leq (b-u)|\mathbf{z}|\left|\mathbf{f}'(x)\right|$$

Hence $|\mathbf{z}| \leq (h-a)\left|\mathbf{f}'(x)\right|$, which is the desired conclusion. ■

5. THE RIEMANN-STIELTJES INTEGRAL

Definition 5.0.1 Let $[a,b]$ be a given interval. By a partition P of $[a,b]$ we mean a finite set of points $x_0, x_1, \dots, x_n$, where

$$a = x_0 \leq x_1 \leq \cdots \leq x_{n-1} \leq x_n = b$$

We write

$$\Delta x_i = x_i - x_{i-1} \quad (i = 1, \dots, n)$$

Now suppose f is a bounded real function defined on $[a,b]$. Corresponding to each partition P of $[a,b]$ we put

$$M_i = \sup f(x) \quad (x_{i-1} \leq x \leq x_i)$$
$$m_i = \inf f(x) \quad (x_{i-1} \leq x \leq x_i)$$
$$U(P,f) = \sum_{i=1}^{n} M_i \Delta x_i$$
$$L(P,f) = \sum_{i=1}^{n} m_i \Delta x_i$$

and finally

$$\overline{\int_a^b} f\,dx = \inf U(P,f)$$

$$\underline{\int_a^b} f dx = \sup L(P,f)$$

where the inf and the sup are taken over all partitions P of $[a,b]$. The left members of previous functions are called the **upper and lower Riemann integrals** of f over $[a,b]$, respectively

If the upper and lower integrals are equal, we say that f is Riemann integrable on $[a,b]$, we write $f \in \mathscr{R}$ (that is, $\mathscr{R}$ denotes the set of Riemann integrable functions), and we denote the common value by

$$\int_a^b f dx$$

or by

$$\int_a^b f(x) dx$$

This is the **Riemann integral** of f over $[a,b]$. since f is bounded, there exist two numbers, m and M, such that

$$m \leq f(x) \leq M \quad (a \leq x \leq b)$$

Hence, for every P

$$m(b-a) \leq L(P,f) \leq U(P,f) \leq M(b-a)$$

so that the numbers $L(P,f)$ and $U(P,f)$ form a bounded set. This shows that the upper and lower integrals are defined for every bounded function f.

Definition 5.0.2 — Riemann-Stieltjes integral. Let α be a monotonically increasing function on $[a,b]$ (since $\alpha(a)$ and $\alpha(b)$ are finite, it follows that α is bounded on $[a,b]$). Corresponding to each partition P of $[a,b]$, we write

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

It is clear that $\Delta\alpha_i \geq 0$. For any real function f which is bounded on $[a,b]$ we put

$$U(P,f,\alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$
$$L(P,f,\alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$$

where M_i, m_i have the same meaning as in Definition 6.1, and we define

$$\overline{\int_a^b} f d\alpha = \inf U(P,f,\alpha)$$

$$\int_{-a}^{b} f d\alpha = \sup L(P,f,\alpha)$$

the inf and sup again being taken over all partitions. If the left members of previous functions are equal, we denote their common value by

$$\int_{a}^{b} f d\alpha$$

or sometimes by

$$\int_{a}^{b} f(x) d\alpha(x)$$

This is the **Riemann-Stieltjes integral** (or simply the Stieltjes integral) of f with respect to α, over $[a,b]$, or we say that f is integrable with respect to α, in the Riemann sense, and write $f \in \mathscr{R}(\alpha)$

By taking $\alpha(x) = x$, the Riemann integral is seen to be a special case of the Riemann-Stieltjes integral. Let us mention explicitly, however, that in the general case α need not even be continuous.

R The integral depends on f, α, a and b, but not on the variable of integration, which may as well be omitted.

Definition 5.0.3 — refinement. We say that the partition P^* is a **refinement** of P if $P^* \supset P$ (that is, if every point of P is a point of P^*). Given two partitions, P_1 and P_2, we say that P^* is their **common refinement** if $P^* = P_1 \cup P_2$

Theorem 5.0.1 If P^* is a refinement of P, then

$$L(P,f,\alpha) \leq L(P^*,f,\alpha)$$

and

$$U(P^*,f,\alpha) \leq U(P,f,\alpha)$$

Proof. To prove the first equation, suppose first that P^* contains just one point more than P. Let this extra point be x^*, and suppose $x_{i-1} < x^* < x_i$, where x_{i-1} and x_i are two consecutive points of P. Put

$$\begin{aligned} w_1 &= \inf f(x) \quad (x_{i-1} \leq x \leq x^*) \\ w_2 &= \inf f(x) \quad (x^* \leq x \leq x_i) \end{aligned}$$

Clearly $w_1 \geq m_i$ and $w_2 \geq m_i$, where, as before,

$$m_i = \inf f(x) \quad (x_{i-1} \leq x \leq x_i)$$

Hence

$$\begin{aligned}
&L(P^*, f, \alpha) - L(P, f, \alpha) \\
&\quad = w_1 [\alpha(x^*) - \alpha(x_{i-1})] + w_2 [\alpha(x_i) - \alpha(x^*)] - m_i [\alpha(x_i) - \alpha(x_{i-1})] \\
&\quad = (w_1 - m_i) [\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i) [\alpha(x_i) - \alpha(x^*)] \geq 0
\end{aligned}$$

If P^* contains k points more than P, we repeat this reasoning k times. The proof of the other is analogous. ∎

Theorem 5.0.2 $\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha$ for any partition.

Proof. Let P^* be the common refinement of two partitions P_1 and P_2. By previous Theorem

$$L(P_1, f, \alpha) \leq L(P^*, f.\alpha) \leq U(P^*, f, \alpha) \leq U(P_2, f, \alpha)$$

Hence

$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$$

If P_2 is fixed and the sup is taken over all P_1, then

$$\underline{\int} f d\alpha \leq U(P_2, f, \alpha)$$

The theorem follows by taking the inf over all P_2 in another equation. ∎

Theorem 5.0.3 Theorem $f \in \mathscr{R}(\alpha)$ on $[a, b]$ if and only if for every $\varepsilon > 0$ there exists a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon \tag{5.1}$$

Proof For every P we have

$$L(P, f, \alpha) \leq \underline{\int} f d\alpha \leq \overline{\int} f d\alpha \leq U(P, f, \alpha)$$

Thus 5.1 implies

$$0 \leq \overline{\int} f d\alpha - \underline{\int} f d\alpha < \varepsilon$$

Hence, if (5.1) can be satisfied for every $\varepsilon > 0$, we have

$$\overline{\int} f d\alpha = \underline{\int} f d\alpha$$

that is, $f \in \mathscr{R}(\alpha)$ Conversely, suppose $f \in \mathscr{R}(\alpha)$, and let $\varepsilon > 0$ be given. Then there exist partitions P_1 and P_2 such that (14)

$$U(P_2, f, \alpha) - \int f d\alpha < \frac{\varepsilon}{2}$$

(15)

$$\int f d\alpha - L(P_1, f, \alpha) < \frac{\varepsilon}{2}$$

We choose P to be the common refinement of P_1 and P_2. Then shows that

$$U(P, f, \alpha) \leq U(P_2, f, \alpha) < \int f d\alpha + \frac{\varepsilon}{2} < L(P_1, f, \alpha) + \varepsilon \leq L(P, f, \alpha) + \varepsilon$$

so that (5.1) holds for this partition P

Theorem 5.0.4
1. If (5.1) holds for some P and some ε, then (5.1) holds (with the same ε) for every refinement of P
2. If (5.1) holds for $P = \{x_0, \ldots, x_n\}$ and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta\alpha_i < \varepsilon$$

3. If $f \in \mathbb{R}(x)$ and the hypotheses of (b) hold, then

$$\left| \sum_{i=1}^{n} f(t_i) \Delta\alpha_i - \int_a^b f d\alpha \right| < \varepsilon$$

Proof. Under the assumptions made in (b) both $f(s_i)$ and $f(t_i)$ lie in $[m_i, M_i]$, so that $|f(s_i) - f(t_i)| \leq M_i - m_i$. Thus

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta\alpha_i \leq U(P, f, \alpha) - L(P, f, \alpha)$$

which proves (b). The obvious inequalities

$$L(P, f, \alpha) \leq \sum f(t_i) \Delta\alpha_i \leq U(P, f, \alpha)$$

and

$$L(P, f, \alpha) \leq \int f d\alpha \leq U(P, f, \alpha)$$

prove (c)

■

Theorem 5.0.5 If f is continuous on $[a, b]$ then $f \in \mathscr{R}(\alpha)$ on $[a, b]$

Proof. Let $\varepsilon > 0$ be given. Choose $\eta > 0$ so that

$$[\alpha(b) - \alpha(a)]\eta < \varepsilon$$

Since f is uniformly continuous on $[a,b]$, there exists a $\delta > 0$ such that

$$|f(x) - f(t)| < \eta$$

if $x \in [a,b], t \in [a,b]$, and $|x - t| < \delta$ If P is any partition of $[a,b]$ such that $\Delta x_i < \delta$ for all i, then

$$M_i - n_i \leq \eta \quad (i - 1, \ldots, n)$$

and therefore

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{i=1}^{n} (M_i - m_i) \Delta x_i \\ &\leq \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \varepsilon \end{aligned}$$

$f \in \mathscr{R}(x)$

■

Theorem 5.0.6 If f is monotonic on $[a,b]$, and if α is continuous on $[a,b]$, then $f \in \mathscr{R}(\alpha)$. (We still assume, of course, that α is monotonic.)

Proof. Let $\varepsilon > 0$ be given. For any positive integer n, choose a partition such that

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n} \quad (i = 1, \ldots, n)$$

This is possible since α is continuous. We suppose that f is monotonically increasing (the proof is analogous in the other case). Then

$$M_i = f(x_i), \quad m_i = f(x_{i-1}) \quad (i = 1, \ldots, n)$$

so that

$$\begin{aligned} U(P,f,x) - L(P,f,\alpha) &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \\ &= \frac{\alpha(b) - \alpha(b)}{n} \cdot [f(b) - f(a)] < \varepsilon \end{aligned}$$

if n is taken large enough. $f \in \mathscr{R}(\alpha)$ ■

Theorem 5.0.7 Suppose f is bounded on $[a,b]$, f has only finitely many points of discontinuity on $[a,b]$, and α is continuous at every point at which f is discontinuous. Then $f \in \mathscr{R}(\alpha)$

Proof. Let $\varepsilon > 0$ be given. Put $M = \sup|f(x)|$, let E be the set of points at which f is discontinuous. since E is finite and α is continuous at every point of E, we can cover E by finitely many disjoint intervals $[u_j, v_j] \subset [a,b]$ such that the sum of the corresponding differences $\alpha(v_j) - \alpha(u_j)$ is less than ε. Furthermore, we can place these intervals in such a way that every point of $E \cap (a,b)$ lies in the interior of some $[u_j, v_j]$

Remove the segments (u_j, v_j) from $[a,b]$. The remaining set K is compact. Hence f is uniformly continuous on K, and there exists $\delta > 0$ such that $|f(s) - f(t)| < \varepsilon$ if $s \in K, t \in K, |s-t| < \delta$ Now form a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a,b]$, as follows: Each u_j occurs in P. Each v_j occurs in P. No point of any segment (u_j, v_j) occurs in P. If x_{i-1} is not one of the u_j, then $\Delta x_i < \delta$

Note that $M_i - m_i \leq 2M$ for every i, and that $M_i - m_i \leq \varepsilon$ unless x_{i-1} is one of the u_j. Hence, as in the proof of Theorem 6.8

$$U(P,f,x) - L(P,f,\alpha) \leq [\alpha(b) - \alpha(a)]\varepsilon + 2M\varepsilon$$

since ε is arbitrary, Theorem 5.0.3 shows that $f \in \mathscr{R}(\alpha)$ Note: If f and x have a common point of discontinuity, then f need not be in $\mathscr{R}(x)$.

■

Theorem 5.0.8 Suppose $f \in \mathscr{R}(\alpha)$ on $[a,b], m \leq f \leq M, \phi$ is continuous on $[m,M]$, and $h(x) = \phi(f(x))$ on $[a,b]$. Then $h \in \mathscr{R}(x)$ on $[a,b]$

Proof. Choose $\varepsilon > 0$. since ϕ is uniformly continuous on $[m,M]$, there exists $\delta > 0$ such that $\delta < c$ and $|\phi(s) - \phi(t)| < \varepsilon$ if $|s-t| \leq \delta$ and $s,t \in [m,M]$ since $f \in \mathscr{R}(x)$, there is a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a,b]$ such that

$$U(P,f,\alpha) - L(P,f,\alpha) < \delta^2$$

Let M_i, m_i have the same meaning as in Definition 6.1, and let M_i^*, m_i^* be the analogous numbers for h. Divide the numbers $1, \ldots, n$ into two classes: $i \in A$ if $M_i - m_i < \delta, i \in B$ if $M_i - m_i \geq \delta$ For $i \in A$, our choice of δ shows that $M_i^* - m_i^* \leq \varepsilon$ For $i \in B$, $\quad M_i^* - m_i^* \leq 2K$, where $K = \sup|\phi(t)|, m \leq t \leq M$. we have

$$\delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i \in B} (M_i - m_i) \Delta x_i < \delta^2$$

so that $\sum_{i \in B} \Delta\alpha_i < \delta$. It follows that

$$\begin{aligned} U(P,h,\alpha) - L(P,h,\alpha) &= \sum_{i \in A} (M_i^* - m_i^*) \Delta\alpha_i + \sum_{i \in B} (M_i^* - m_i^*) \Delta\alpha_i \\ &\leq \varepsilon[\alpha(b) - \alpha(a)] + 2K\delta < \varepsilon[\alpha(b) - \alpha(a) + 2K] \end{aligned}$$

since ε was arbitrary, Theorem 5.0.3 implies that $h \in \mathscr{R}(\alpha)$

■

5.0.1 PROPERTIES OF THE INTEGRAL

Proposition 5.0.9 1. If $f_1 \in \mathscr{R}(\alpha)$ and $f_2 \in \mathscr{R}(\alpha)$ on $[a,b]$, then

$$f_1 + f_2 \in \mathscr{R}(\alpha)$$

$cf \in \mathscr{R}(\alpha)$ for every constant c, and

$$\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha$$
$$\int_a^b cf d\alpha = c\int_a^b f d\alpha$$

2. If $f_1(x) \le f_2(x)$ on $[a,b]$, then

$$\int_a^b f_1 d\alpha \le \int_a^b f_2 d\alpha$$

3. If $f \in \mathscr{R}(\alpha)$ on $[a,b]$ and if $a < c < b$, then $f \in \mathscr{R}(\alpha)$ on $[a,c]$ and on $[c,b]$, and

$$\int_a^c f d\alpha + \int_c^b f d\alpha = \int_a^b f d\alpha$$

4. If $f \in \mathscr{R}(x)$ on $[a,b]$ and if $|f(x)| \le M$ on $[a,b]$, then

$$\left|\int_a^b f d\alpha\right| \le M[\alpha(b) - \alpha(a)]$$

5. If $f \in \mathscr{R}(\alpha_1)$ and $f \in \mathscr{R}(\alpha_2)$, then $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ and

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2$$

6. if $f \in \mathscr{R}(\alpha)$ and c is a positive constant, then $f \in \mathscr{R}(c\alpha)$ and

$$\int_a^b f d(c\alpha) = c\int_a^b f d\alpha$$

Proof. Proof If $f = f_1 + f_2$ and P is any partition of $[a,b]$, we have

$$L(P,f_1,\alpha) + L(P,f_2,\alpha) \le L(P,f,\alpha) \tag{5.2}$$
$$\le U(P,f,\alpha) \le U(P,f_1,\alpha) + U(P,f_2,\alpha) \tag{5.3}$$

If $f_1 \in \mathscr{R}(\alpha)$ and $f_2 \in \mathscr{R}(\alpha)$, let $\varepsilon > 0$ be given. There are partitions P_j $(j = 1,2)$ such that

$$U(P_j, f_j, \alpha) - L(P_j, f_j, \alpha) < \varepsilon$$

These inequalities persist if P_1 and P_2 are replaced by their common refinement P. Then (5.2) implies

$$U(P,f,\alpha) - L(P,f,\alpha) < 2\varepsilon$$

which proves that $f \in \mathscr{R}(\alpha)$ With this same P we have

$$U\left(P,f_j,\alpha\right) < \int f_j d\alpha + \varepsilon \quad (j = 1,2)$$

hence (5.2) implies

$$\int f dx \leq U(P,f,\alpha) < \int f_1 d\alpha + \int f_2 d\alpha + 2\varepsilon$$

since ε was arbitrary, we conclude that

$$\int f d\alpha \leq \int f_1 d\alpha + \int f_2 d\alpha \tag{5.4}$$

If we replace f_1 and f_2 in (5.4) by $-f_1$ and $-f_2$, the inequality is reversed, and the equality is proved.

■

Theorem 5.0.10 If $f \in \mathscr{R}(\alpha)$ and $g \in \mathscr{R}(\alpha)$ on $[a,b]$, then

1. $fg \in \mathscr{R}(x)$
2. $|f| \in \mathscr{R}(\alpha)$ and $\left|\int_a^b f d\alpha\right| \leq \int_a^b |f| d\alpha$

Proof. If we take $\phi(t) = t^2$, Theorem 5.0.8 shows that $f^2 \in \mathscr{R}(\alpha)$ if $f \in \mathscr{R}(x)$ The identity

$$4fg = (f+g)^2 - (f-g)^2$$

completes the proof of (a) If we take $\phi(t) = |t|$, Theorem 6.11 shows similarly that $|f|. \in \mathscr{R}(\alpha)$ Choose $c = \pm 1$, so that

$$c\int f d\alpha \geq 0$$

Then

$$\left|\int f d\alpha\right| = c\int f d\alpha = \int c f d\alpha \leq \int |f| d\alpha$$

since $cf \leq |f|$ ■

6.15 Theorem If $a < s < b$, f is bounded on $[a,b]$, f is continuous at s, and $\alpha(x) = I(x-s)$, then

$$\int_a^b f d\alpha = f(s)$$

Proof Consider partitions $P = \{x_0, x_1, x_2, x_3\}$, where $x_0 = a$, and $x_1 = s < x_2 < x_3 = b$. Then

$$U(P,f,\alpha) = M_2, \quad L(P,f,x) = m_2$$

since f is continuous at s, we see that M_2 and m_2 converge to $f(s)$ as $x_2 \to s$

Theorem 5.0.11 Suppose $c_n \geq 0$ for $1,2,3,\ldots,\Sigma c_n$ converges, $\{s_n\}$ is a sequence of distinct points in (a,b), and

$$\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n) \tag{5.5}$$

Let f be continuous on $[a,b]$. Then

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} c_n f(s_n) \tag{5.6}$$

Proof. The comparison test shows that the series (5.5) converges for every x. Its sum $\alpha(x)$ is evidently monotonic, and $\alpha(a) = 0, \alpha(b) = \Sigma c_n$. Let $\varepsilon > 0$ be given, and choose N so that

$$\sum_{N+1}^{\infty} c_n < \varepsilon$$

Put

$$\alpha_1(x) = \sum_{n=1}^{N} c_n I(x - s_n), \quad \alpha_2(x) = \sum_{N+1}^{\infty} c_n I(x - s_n)$$

$$\int_a^b f d\alpha_1 = \sum_{i=1}^{N} c_n f(s_n)$$

since $\alpha_2(b) - \alpha_2(a) < \varepsilon$

$$\left| \int_a^b f d\alpha_2 \right| \leq M\varepsilon$$

where $M = \sup|f(x)|$: since $\alpha = \alpha_1 + \alpha_2$, it follows that

$$\left| \int_a^b f d\alpha - \sum_{i=1}^{N} c_n f(s_n) \right| \leq M\varepsilon$$

If we let $N \to \infty$, we obtain (5.6) ∎

Theorem 5.0.12 Assume α increases monotonically and $\alpha' \in \mathscr{R}$ on $[a,b]$. Let f be a bounded real function on $[a,b]$ Then $f \in \mathscr{R}(\alpha)$ if and only if $f\alpha' \in \mathscr{R}$. In that case

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x)dx$$

Proof. Let $\varepsilon > 0$ be given and apply Theorem 5.0.3 to α' : There is a partition $P = \{x_0,\ldots,x_n\}$ of $[a,b]$ such that

$$U(P,\alpha') - L(P,\alpha') < \varepsilon \tag{5.7}$$

The mean value theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that

$$\Delta\alpha_i = \alpha'(t_i)\,\Delta x_i$$

for $i = 1, \ldots, n$. If $s_i \in [x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} \left|\alpha'(s_i) - \alpha'(t_i)\right| \Delta x_i < \varepsilon \tag{5.8}$$

by (5.7). Put $M = \sup|f(x)|$. since

$$\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\,\alpha'(t_i)\,\Delta x_i$$

it follows from (6.3) that

$$\left|\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i - \sum_{i=1}^{n} f(s_i)\,\alpha'(s_i)\,\Delta x_i\right| \leq M\varepsilon \tag{5.9}$$

In particular,

$$\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i \leq U\left(P, f\alpha'\right) + M\varepsilon$$

for all choices of $s_i \in [x_{i-1}, x_i]$, so that

$$U(P, f, \alpha) \leq U\left(P, f\alpha'\right) + M\varepsilon$$

The same argument leads from (6.4) to

$$U\left(P, f\alpha'\right) \leq U(P, f, \alpha) + M\varepsilon$$

Thus

$$\left|U(P, f, \alpha) - U\left(P, f\alpha'\right)\right| \leq M\varepsilon \tag{5.10}$$

Now note that (5.7) remains true if P is replaced by any refinement. Hence (5.10) also remains true. We conclude that

$$\left|\overline{\int_a^b} f d\alpha - \overline{\int_a^b} f(x)\alpha'(x)dx\right| \leq M\varepsilon$$

But ε is arbitrary. Hence

$$\overline{\int_a^b} f d\alpha = \overline{\int_a^b} f(x)\alpha'(x)dx$$

for any bounded f. The equality of the lower integrals follows from (6.4) in exactly the same way. The theorem follows.

$\blacksquare$

R The two preceding theorems illustrate the generality and flexibility which are inherent in the Stieltjes process of integration. If α is a pure step function, the integral reduces to a finite or infinite series. If α has an integrable derivative, the integral reduces to an ordinary Riemann integral. This makes it possible in many cases to study series and integrals simultaneously, rather than separately.

Theorem 5.0.13 — change of variable. Suppose φ is a strictly increasing continuous function that maps an interval $[A,B]$ onto $[a,b]$. Suppose α is monotonically increasing on $[a,b]$ and $f \in \mathscr{R}(\alpha)$ on $[a,b]$. Define β and g on $[A,B] by$

$$\beta(y) = \alpha(\varphi(y)), \quad g(y) = f(\varphi(y))$$

Then $g \in \mathscr{R}(\beta)$ and

$$\int_A^B g d\beta = \int_a^b f d\alpha$$

Proof. To each partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ corresponds a partition $Q = \{y_0, \ldots, y_n\}$ of $[A,B]$, so that $x_i = \varphi(y_i)$. All partitions of $[A,B]$ are obtained in this way. since the values taken by f on $[x_{i-1}, x_i]$ are exactly the same as those taken by g on $[y_{i-1}, y_i]$, we see that

$$U(Q,g,\beta) = U(P,f,\alpha), \quad L(Q,g,\beta) = L(P,f,\alpha)$$

since $f \in \mathscr{R}(\alpha), P$ can be chosen so that both $U(P,f,\alpha)$ and $L(P,f,\alpha)$ are close to $\int f d\alpha$. Hence (38), combined with Theorem 6.6, shows that $g \in \mathscr{R}(\beta)$. This completes the proof. ■

R Let us note the following special case: Take $\alpha(x) = x$. Then $\beta = \varphi$. Assume $\varphi' \in \mathscr{R}$ on $[A,B]$. We obtain

$$\int_a^b f(x)dx = \int_A^B f(\varphi(y))\varphi'(y)dy$$

5.0.2 INTEGRATION AND DIFFERENTIATION

Theorem 5.0.14 Let $f \in \mathscr{R}$ on $[a,b]$. For $a \leq x \leq b$, put

$$F(x) = \int_a^x f(t)dt$$

Then F is continuous on $[a,b]$; furthermore, if f is continuous at a point x_0 of $[a,b]$, then

F is differentiable at x_0, and

$$F'(x_0) = f(x_0)$$

Proof. Since $f \in \mathscr{R}, f$ is bounded. Suppose $|f(t)| \leq M$ for $a \leq t \leq b$ If $a \leq x < y \leq b$, then

$$|F(y) - F(x)| = \left| \int_x^y f(t)dt \right| \leq M(y - x)$$

by Theorem 5.0.9(c) and (d). Given $\varepsilon > 0$, we see that

$$|F(y) - F(x)| < \varepsilon$$

provided that $|y - x| < \varepsilon / M$. This proves continuity (and, in fact, uniform continuity) of F

Now suppose f is continuous at x_0. Given $\varepsilon > 0$, choose $\delta > 0$ such that

$$|f(t) - f(x_0)| < \varepsilon$$

if $|t - x_0| < \delta$, and $a \leq t \leq b$. Hence, if

$$x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta \quad \text{and} \quad a \leq s < t \leq b$$

we have, by Theorem 5.0.9 (d)

$$\left| \frac{F(t) - F(s)}{t - s} - f(x_0) \right| = \left| \frac{1}{t - s} \int_s^t [f(u) - f(x_0)]\, du \right| < \varepsilon$$

It follows that $F'(x_0) = f(x_0)$ ■

Theorem 5.0.15 — The fundamental theorem of calculus . If $f \in \mathscr{R}$ on $[a,b]$ and if there is a differentiable function F on $[a,b]$ such that $F' = f$, then

$$\int_a^b f(x)dx = F(b) - F(a)$$

Proof. Let $\varepsilon > 0$ be given. Choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ so that $U(P,f) - L(P,f) < \varepsilon$. The mean value theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that

$$F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$$

for $i = 1, \ldots, n$. Thus

$$\sum_{i=1}^n f(t_i)\Delta x_i = F(b) - F(a)$$

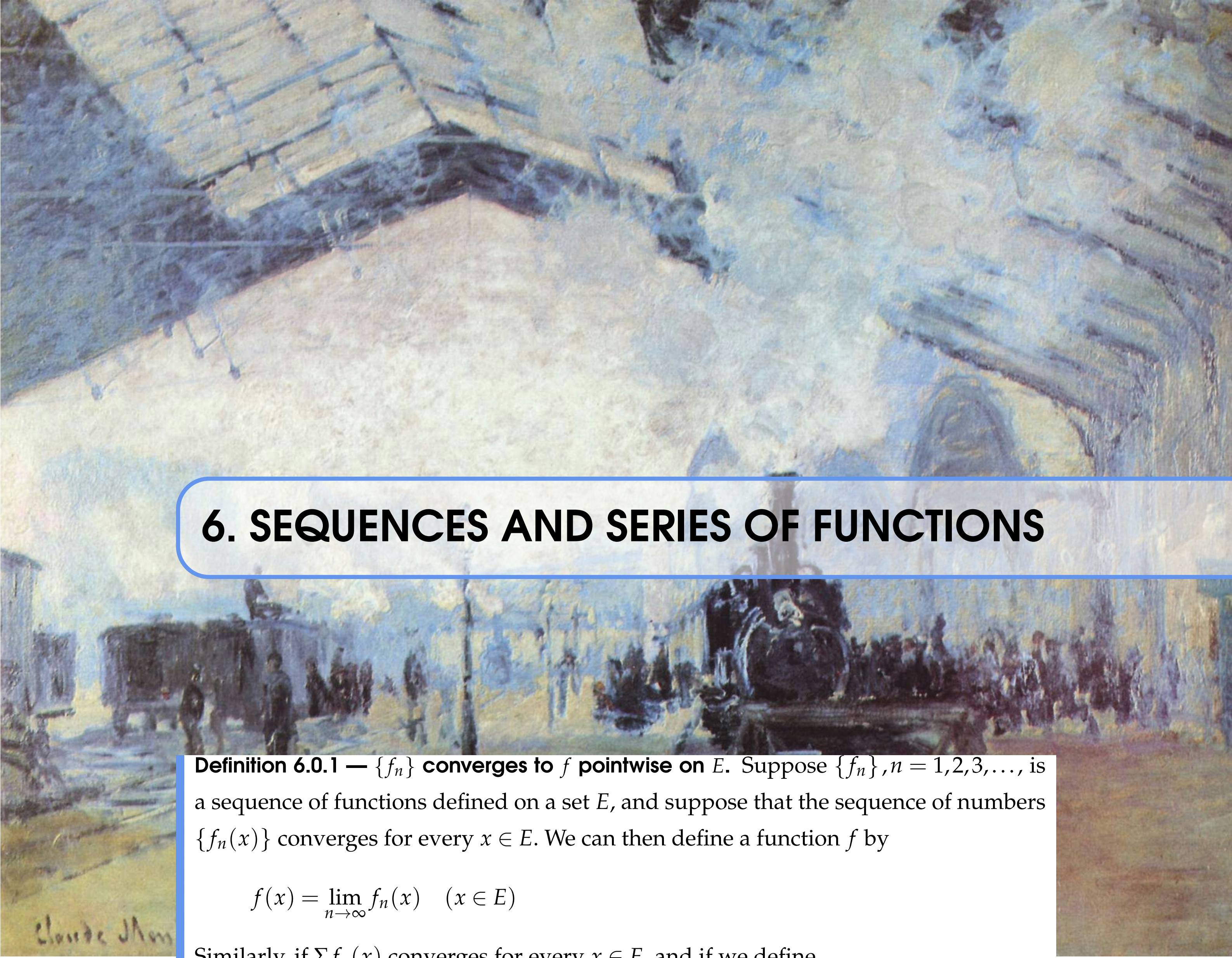
It now follows from Theorem 6.7(c) that

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| < \varepsilon$$

since this holds for every $\varepsilon > 0$, the proof is complete. ■

Theorem 5.0.16 — integration by parts. Suppose F and G are differentiable functions on $[a,b], F' = f \in \mathscr{R}$, and $G' = g \in \mathscr{R}$. Then

$$\int_a^b F(x)g(x)dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx$$



6. SEQUENCES AND SERIES OF FUNCTIONS

Definition 6.0.1 — $\{f_n\}$ converges to f pointwise on E. Suppose $\{f_n\}, n = 1,2,3,\ldots,$ is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in E$. We can then define a function f by

$$f(x) = \lim_{n\to\infty} f_n(x) \quad (x \in E)$$

Similarly, if $\Sigma f_n(x)$ converges for every $x \in E$, and if we define

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (x \in E)$$

the function f is called the sum of the series Σf_n

Definition 6.0.2 — converges uniformly. We say that a sequence of functions $\{f_n\}, n = 1,2,3,\ldots,$ converges uniformly on E to a function f if for every $\varepsilon > 0$ there is an integer N such that $n \geq N$ implies

$$|f_n(x) - f(x)| \leq \varepsilon \tag{6.1}$$

for all $x \in E$. It is clear that every uniformly convergent sequence is pointwise convergent. Quite explicitly, the difference between the two concepts is this: If $\{f_n\}$ converges pointwise on E, then there exists a function f such that, for every $\varepsilon > 0$, and for every $x \in E$, there is an integer N, depending on ε and on x, such that 6.1 holds if $n \geq N$; if $\{f_n\}$ converges uniformly on E, it is possible, for each $\varepsilon > 0$, to find one integer N which will do for all $x \in E$.

We say that the series $\Sigma f_n(x)$ converges uniformly on E if the sequence $\{s_n\}$ of partial sums defined by

$$\sum_{i=1}^{n} f_i(x) = s_n(x)$$

converges uniformly on E.

Theorem 6.0.1 — Cauchy criterion–converges uniformly. The sequence of functions $\{f_n\}$, defined on E, converges uniformly on E if and only if for every $\varepsilon > 0$ there exists an integer N such that $m \geq N$ $n \geq N, x \in E$ implies

$$|f_n(x) - f_m(x)| \leq \varepsilon$$

Proof. Suppose $\{f_n\}$ converges uniformly on E, and let f be the limit function. Then there is an integer N such that $n \geq N, x \in E$ implies

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2}$$

so that

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \leq \varepsilon$$

if $n \geq N, m \geq N, x \in E$

Conversely, suppose the Cauchy condition holds. By Theorem 3.11 the sequence $\{f_n(x)\}$ converges, for every x, to a limit which we may call $f(x)$. Thus the sequence $\{f_n\}$ converges on E, to f. We have to prove that the convergence is uniform.

Let $\varepsilon > 0$ be given, and choose N such that condition holds. Fix n, and let $m \to \infty$ in condition . Since $f_m(x) \to f(x)$ as $m \to \infty$, this gives

$$|f_n(x) - f(x)| \leq \varepsilon$$

for every $n \geq N$ and every $x \in E$, which completes the proof. ∎

Theorem 6.0.2 Suppose

$$\lim_{n\to\infty} f_n(x) = f(x) \quad (x \in E)$$

Put

$$M_n = \sup_{x \in E} |f_n(x) - f(x)|$$

Then $f_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n \to \infty$

Theorem 6.0.3 Suppose $\{f_n\}$ is a sequence of functions defined on E, and suppose

$$|f_n(x)| \leq M_n \quad (x \in E, n = 1,2,3,\dots)$$

Then Σf_n converges uniformly on E if ΣM_n converges.

Proof. By Cauchy criterion ■

Theorem 6.0.4 Suppose $f_n \to f$ uniformly on a set E in a metric space. Let x be a limit point of E, and suppose that

$$\lim_{t \to x} f_n(t) = A_n \quad (n = 1,2,3,\dots)$$

Then $\{A_n\}$ converges, and

$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n$$

In other words, the conclusion is that

$$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t)$$

Proof. Let $\varepsilon > 0$ be given. By the uniform convergence of $\{f_n\}$, there exists N such that $n \geq N, m \geq N, t \in E$ imply

$$|f_n(t) - f_m(t)| \leq \varepsilon$$

Letting $t \to x$, we obtain

$$|A_n - A_m| \leq \varepsilon$$

for $n \geq N, m \geq N$, so that $\{A_n\}$ is a Cauchy sequence and therefore converges, say to A Next,

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|$$

We first choose n such that

$$|f(t) - f_n(t)| \leq \frac{\varepsilon}{3}$$

for all $t \in E$ (this is possible by the uniform convergence), and such that

$$|A_n - A| \leq \frac{\varepsilon}{3}$$

Then, for this n, we choose a neighborhood V of x such that

$$|f_n(t) - A_n| \leq \frac{\varepsilon}{3}$$

if $t \in V \cap E, t \neq x$ Substituting the inequalities, we see that

$$|f(t) - A| \leq \varepsilon$$

provided $t \in V \cap E, t \neq x$.

■

Theorem 6.0.5 If $\{f_n\}$ is a sequence of continuous functions on E, and if $f_n \to f$ uniformly on E, then f is continuous on E

The converse is not true; that is, a sequence of continuous functions may converge to a continuous function, although the convergence is not uniform.

Theorem 6.0.6 Suppose K is compact, and

1. $\{f_n\}$ is a sequence of continuous functions on K
2. $\{f_n\}$ converges pointwise to a continuous function fon K
3. $f_n(x) \geq f_{n+1}(x)$ for all $x \in K, n = 1,2,3,\ldots$

Then $f_n \to f$ uniformly on K

Proof. Put $g_n = f_n - f$. Then g_n is continuous, $g_n \to 0$ pointwise, and $g_n \geq g_{n+1}$. We have to prove that $g_n \to 0$ uniformly on K

Let $\varepsilon > 0$ be given. Let K_n be the set of all $x \in K$ with $g_n(x) \geq \varepsilon$ since g_n is continuous, K_n is closed (Theorem 4.8), hence compact (Theorem 2.35). since $g_n \geq g_{n+1}$, we have $K_n \supset K_{n+1}$. Fix $x \in K$. since $g_n(x) \to 0$ we see that $x \notin K_n$ if n is sufficiently large. Thus $x \notin \bigcap K_n$. In other words, $\bigcap K_n$ is empty. Hence K_N is empty for some N (Theorem 2.36). It follows that $0 \leq g_n(x) < \varepsilon$ for all $x \in K$ and for all $n \geq N$. This proves the theorem. ■

■ **Example 6.1** Let us note that compactness is really needed here. For instance, if

$$f_n(x) = \frac{1}{nx+1} \quad (0 < x < 1; n = 1,2,3,\ldots)$$

then $f_n(x) \to 0$ monotonically in $(0,1)$, but the convergence is not uniform.

Definition 6.0.3 If X is a metric space, $\mathscr{C}(X)$ will denote the set of all complex-valued, continuous, bounded functions with domain X.

[Note that boundedness is redundant if X is compact. Thus $\mathscr{C}(X)$ consists of all complex continuous functions on X if X is compact.

We associate with each $f \in \mathscr{C}(X)$ its supremum norm

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

since f is assumed to be bounded, $\|f\| < \infty$. It is obvious that $\|f\| = 0$ only if $f(x) = 0$

for every $x \in X$, that is, only if $f = 0$. If $h = f + g$, then

$$|h(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|$$

for all $x \in X$; hence

$$\|f + g\| \leq \|f\| + \|g\|$$

Theorem 6.0.7 A sequence $\{f_n\}$ converges to f with respect to the metric of $\mathscr{C}(X)$ if and only if $f_n \to f$ uniformly on X

Definition 6.0.4 closed subsets of $\mathscr{C}(X)$ are sometimes called uniformly closed, the closure of a set $\mathscr{A} \subset \mathscr{C}(X)$ is called its uniform closure, and so on.

Theorem 6.0.8 The above metric makes $\mathscr{C}(X)$ into a complete metric space.

Proof. Let $\{f_n\}$ be a Cauchy sequence in $\mathscr{C}(X)$. This means that to each $\varepsilon > 0$ corresponds an N such that $\|f_n - f_m\| < \varepsilon$ if $n \geq N$ and $m \geq N$ It follows (by Theorem 7.8) that there is a function f with domain X to which $\{f_n\}$ converges uniformly. By Theorem 7.12, f is continuous. Moreover, f is bounded, since there is an n such that $|f(x) - f_n(x)| < 1$ for all $x \in X$, and f_n is bounded.

Thus $f \in \mathscr{C}(X)$, and since $f_n \to f$ uniformly on X, we have $\|f - f_n\| \to 0$ as $n \to \infty$ ■

6.0.1 UNIFORM CONVERGENCE AND INTEGRATION

Theorem 6.0.9 Let α be monotonically increasing on $[a,b]$. Suppose $f_n \in \mathscr{R}(\alpha)$ on $[a,b]$, for $n = 1,2,3,\ldots$, and suppose $f_n \to f$ uniformly on $[a,b]$. Then $f \in \mathscr{R}(\alpha)$ $on[a,b]$, and

$$\int_a^b f d\alpha = \lim_{n\to\infty} \int_a^b f_n d\alpha$$

(The existence of the limit is part of the conclusion.)

Proof. It suffices to prove this for real f_n. Put

$$\varepsilon_n = \sup|f_n(x) - f(x)|$$

the supremum being taken over $a \leq x \leq b$. Then

$$f_n - \varepsilon_n \leq f \leq f_n + \varepsilon_n$$

so that the upper and lower integrals of f satisfy

$$\int_a^b (f_n - \varepsilon_n)\, d\alpha \leq \underline{\int} f d\alpha \leq \overline{\int} f d\alpha \leq \int_a^b (f_n + \varepsilon_n)\, d\alpha \tag{6.2}$$

Hence

$$0 \leq \overline{\int} f d\alpha - \underline{\int} f d\alpha \leq 2\varepsilon_n[\alpha(b) - \alpha(a)]$$

since $\varepsilon_n \to 0$ as $n \to \infty$ (Theorem 6.0.3), the upper and lower integrals of f are equal. Thus $f \in \mathscr{R}(\alpha)$. Another application of (6.2) now yields

$$\left| \int_a^b f d\alpha - \int_a^b f_n d\alpha \right| \leq \varepsilon_n[\alpha(b) - \alpha(a)]$$

■

Corollary 6.0.10 If $f_n \in \mathscr{R}(\alpha)$ on $[a,b]$ and if

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (a \leq x \leq b)$$

the series converging uniformly on $[a,b]$, then

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n d\alpha$$

In other words, the series may be integrated term by term.

6.0.2 UNIFORM CONVERGENCE AND DIFFERENTIATION

Theorem 6.0.11 Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a,b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a,b]$. If $\{f_n'\}$ converges uniformly on $[a,b]$, then $\{f_n\}$ converges uniformly on $[a,b]$, to a function f, and

$$f'(x) = \lim_{n \to \infty} f_n'(x) \quad (a \leq x \leq b)$$

Proof. Let $\varepsilon > 0$ be given. Choose N such that $n \geq N, m \geq N$, implies

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$$

and

$$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)} \quad (a \leq t \leq b) \tag{6.3}$$

If we apply the mean value theorem to the function $f_n - f_m$,6.3 shows that

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leq \frac{|x-t|\varepsilon}{2(b-a)} \leq \frac{\varepsilon}{2} \tag{6.4}$$

for any x and t on $[a,b]$, if $n \geq N, m \geq N$. The inequality

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|$$

implies, by 6.4, that

$$|f_n(x) - f_m(x)| < \varepsilon \quad (a \le x \le b, n \ge N, m \ge N)$$

so that $\{f_n\}$ converges uniformly on $[a,b]$. Let

$$f(x) = \lim_{n\to\infty} f_n(x) \quad (a \le x \le b)$$

Let us now fix a point x on $[a,b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}, \quad \phi(t) = \frac{f(t) - f(x)}{t - x}$$

for $a \le t \le b, t \ne x$. Then

$$\lim_{t\to x} \phi_n(t) = f_n'(x) \quad (n = 1,2,3,\dots)$$

The first inequality in (6.4) shows that

$$|\phi_n(t) - \phi_m(t)| \le \frac{\varepsilon}{2(b-a)} \quad (n \ge N, m \ge N)$$

so that $\{\phi_n\}$ converges uniformly, for $t \ne x$. since $\{f_n\}$ converges to f, we conclude from that

$$\lim_{n\to\infty} \phi_n(t) = \phi(t)$$

uniformly for $a \le t \le b, t \ne x$ If we now apply Theorem 6.0.4 to $\{\phi_n\}$, show that

$$\lim_{t\to x} \phi(t) = \lim_{n\to\infty} f_n'(x)$$

and this is (27), by the definition of $\phi(t)$

■

R If the continuity of the functions f_n' is assumed in addition to the above hypotheses, then a much shorter proof can be based on Theorem 6.0.9 and the fundamental theorem of calculus.

■ **Example 6.2** There exists a real continuous function on the real line which is nowhere differentiable. Rudin 154

6.0.3 EQUICONTINUOUS FAMILIES OF FUNCTIONS

Definition 6.0.5 — pointwise bounded, uniformly bounded. Let $\{f_n\}$ be a sequence of functions defined on a set E. We say that $\{f_n\}$ is **pointwise bounded** on E if the sequence $\{f_n(x)\}$ is bounded for every $x \in E$, that is, if there exists a finite-valued function ϕ defined on E such that

$$|f_n(x)| < \phi(x) \quad (x \in E, n = 1,2,3,\dots)$$

We say that $\{f_n\}$ is **uniformly bounded** on E if there exists a number M such that

$$|f_n(x)| < M \quad (x \in E, n = 1,2,3,\dots)$$

Now if $\{f_n\}$ is pointwise bounded on E and E_1 is a countable subset of E it is always possible to find a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E_1$. This can be done by the diagonal process which is used in the proof of Theorem 7.23

However, even if $\{f_n\}$ is a uniformly bounded sequence of continuous functions on a compact set E, there need not exist a subsequence which converges pointwise on E.

R 函数列的点点有界--界是关于 x 的函数；函数列一致有界--界是常数 M

Definition 6.0.6 — equicontinuous . A family $\mathscr{F}$ of complex functions f defined on a set E in a metric space X is said to be **equicontinuous** on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon$$

whenever $d(x,y) < \delta, x \in E, y \in E$, and $f \in \mathscr{F}$. Here d denotes the metric of X It is clear that every member of an equicontinuous family is uniformly continuous.

Theorem 6.0.12 If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$

Proof. Proof Let $\{x_i\}, i = 1,2,3,\dots$, be the points of E, arranged in a sequence. since $\{f_n(x_1)\}$ is bounded, there exists a subsequence, which we shall denote by $\{f_{1,k}\}$, such that $\{f_{1,k}(x_1)\}$ converges as $k \to \infty$

Let us now consider sequences $S_1, S_2, S_3, \dots$, which we represent by the array

$$\begin{array}{lllllll} S_1: & f_{1,1} & f_{1,2} & f_{1,3} & f_{1,4} & \cdots & \cdots \\ S_2: & f_{2,1} & f_{2,2} & f_{2,3} & f_{2,4} & \cdots & \cdots \quad \cdots \\ S_3: & f_{3,1} & f_{3,2} & f_{3,3} & f_{3,4} & \cdots & \cdots\cdots \end{array}$$

and which have the following properties:

1. S_n is a subsequence of S_{n-1}, for $n = 2,3,4,\dots$
2. $\{f_{n,k}(x_n)\}$ converges, as $k \to \infty$ (the boundedness of $\{f_n(x_n)\}$ makes it possible to choose S_n in this way);
3. The order in which the functions appear is the same in each sequence ; i.e., if one function precedes another in S_1, they are in the same relation in every S_n, until one or the other is deleted. Hence, when going from one row in the above array to the next below, functions may move to the left but never to the right.

We now go down the diagonal of the array; i.e., we consider the sequence

$$S : f_{1,1} f_{2,2} f_{3,3} f_{4,4} \cdots$$

By (c), the sequence S (except possibly its first $n-1$ terms) is a subsequence of S_n, for $n = 1,2,3,\ldots,$ Hence (b) implies that $\{f_{n,n}(x_i)\}$ converges, as $n \to \infty$, for every $x_i \in E$ ■

Theorem 6.0.13 If K is a compact metric space, if $f_n \in \mathscr{C}(K)$ for $n = 1,2,3,\ldots,$ and if$\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K

Proof. Let $\varepsilon > 0$ be given. since $\{f_n\}$ converges uniformly, there is an integer N such that

$$\|f_n - f_N\| < \varepsilon \quad (n > N)$$

Since continuous functions are uniformly continuous on compact sets, there is a $\delta > 0$ such that

$$|f_i(x) - f_i(y)| < \varepsilon$$

if $1 \leq i \leq N$ and $d(x,y) < \delta$.

If $n > N$ and $d(x,y) < \delta$, it follows that

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\varepsilon$$

■

Theorem 6.0.14 If K is compact, if $f_n \in \mathscr{C}(K)$ for $n = 1,2,3,\ldots,$ and if $\{f_n\}$ is pointwise bounded and equicontimous on K, then

1. $\{f_n\}$ is uniformly bounded on K
2. $\{f_n\}$ contains a uniformly convergent subsequence.

Proof.
1. Let $\varepsilon > 0$ be given and choose $\delta > 0$, so that

   $$|f_n(x) - f_n(y)| < \varepsilon$$

   for all n, provided that $d(x,y) < \delta$. Since K is compact, there are finitely many points $p_1, \ldots, p$, in K such that to every $x \in K$ corresponds at least one p_i with $d(x, p_i) < \delta$. Since $\{f_n\}$ is pointwise bounded, there exist $M_i < \infty$ such that $|f_n(p_i)| < M_i$ for all n. If $M = \max(M_1, \ldots, M_r)$, then $|f(x)| < M + \varepsilon$ for every $x \in K$. This proves (a)
2. Let E be a countable dense subset of K. (For the existence of such a set E, see Exercise 25, Chap. 2.) Theorem 7.23 shows that $\{f_n\}$ has a subsequence $\{f_{n_i}\}$ such that $\{f_{n_i}(x)\}$ converges for every $x \in E$

Put $f_{n_1} = g_i$, to simplify the notation. We shall prove that $\{g_i\}$ converges uniformly on K

Let $\varepsilon > 0$, and pick $\delta > 0$ as in the beginning of this proof. Let $V(x,\delta)$ be the set of all $y \in K$ with $d(x,y) < \delta$. since E is dense in K, and K is compact, there are finitely many points $x_1,\ldots,x_m$ in E such that

$$K \subset V(x_1,\delta) \cup \cdots \cup V(x_m,\delta)$$

since $\{g_i(x)\}$ converges for every $x \in E$, there is an integer N such that

$$K \subset V(x_1,\delta) \cup \cdots \cup V(x_m,\delta) \tag{6.5}$$

whenever $i \geq N, j \geq N, 1 \leq s \leq m$ If $x \in K$,6.5 shows that $x \in V(x_s,\delta)$ for some s, so that

$$\left|g_i(x_s) - g_j(x_s)\right| < \varepsilon \tag{6.6}$$

for every i. If $i \geq N$ and $j \geq N$, it follows from (6.6) that

$$\begin{aligned}\left|g_i(x) - g_j(x)\right| &\leq \left|g_i(x) - g_i(x_s)\right| + \left|g_i(x_s) - g_j(x_s)\right| + \left|g_j(x_s) - g_j(x)\right| \\ &< 3\varepsilon\end{aligned}$$

■

6.0.4 THE STONE-WEIERSTRASS THEOREM

Theorem 6.0.15 — THE STONE-WEIERSTRASS THEOREM. If f is a continuous complex function on $[a,b]$, there exists a sequence of polynomials P_n such that

$$\lim_{n\to\infty} P_n(x) = f(x)$$

uniformly on $[a,b]$. If f is real, the P_n may be taken real.

Proof. Proof We may assume, without loss of generality, that $[a,b] = [0,1]$ We may also assume that $f(0) = f(1) = 0$. For if the theorem is proved for this case, consider

$$g(x) = f(x) - f(0) - x[f(1) - f(0)] \quad (0 \leq x \leq 1)$$

Here $g(0) = g(1) = 0$, and if g can be obtained as the limit of a uniformly convergent sequence of polynomials, it is clear that the same is true for f since $f - g$ is a polynomial.

Furthermore, we define $f(x)$ to be zero for x outside $[0,1]$. Then f is uniformly continuous on the whole line. We put

$$Q_n(x) = c_n\left(1 - x^2\right)^n \quad (n = 1,2,3,\ldots)$$

where c_n is chosen so that

$$\int_{-1}^{1} Q_n(x)dx = 1 \quad (n = 1,2,3,\dots) \tag{6.7}$$

We need some information about the order of magnitude of c_n. Since

$$\begin{aligned}\int_{-1}^{1} \left(1-x^2\right)^n dx = 2\int_{0}^{1} \left(1-x^2\right)^n dx \geq 2\int_{0}^{1/\sqrt{n}} \left(1-x^2\right)^n dx \\ \geq 2\int_{0}^{1/\sqrt{n}} \left(1-nx^2\right) dx \\ = \frac{4}{3\sqrt{n}} \\ > \frac{1}{\sqrt{n}}\end{aligned}$$

it follows from (6.7) that

$$c_n < \sqrt{n} \tag{6.8}$$

The inequality $\left(1-x^2\right)^n \geq 1 - nx^2$ which we used above is easily shown to be true by considering the function

$$\left(1-x^2\right)^n - 1 + nx^2$$

which is zero at $x = 0$ and whose derivative is positive in (0,1) For any $\delta > 0$,6.8 implies

$$Q_n(x) \leq \sqrt{n}\left(1-\delta^2\right)^n \quad (\delta \leq |x| \leq 1) \tag{6.9}$$

so that $Q_n \to 0$ uniformly in $\delta \leq |x| \leq 1$ Now set

$$P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt \quad (0 \leq x \leq 1) \tag{6.10}$$

Our assumptions about f show, by a simple change of variable, that

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_{0}^{1} f(t)Q_n(t-x)dt$$

and the last integral is clearly a polynomial in x. Thus $\{P_n\}$ is a sequence of polynomials, which are real if f is real. Given $\varepsilon > 0$, we choose $\delta > 0$ such that $|y-x| < \delta$ implies

$$|f(y) - f(x)| < \frac{\varepsilon}{2}$$

Let $M = \sup|f(x)|$. Using (6.7),(6.9), and the fact that $Q_n(x) \geq 0$, we see that for $0 \leq x \leq 1$ ■

Corollary 6.0.16 For every interval $[-a,a]$ there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and such that

$$\lim_{n\to\infty} P_n(x) = |x|$$

uniformly on $[-a,a]$

Proof. Proof By Theorem 6.0.15, there exists a sequence $\{P_n^*\}$ of real polynomials which converges to $|x|$ uniformly on $[-a,a]$. In particular, $P_n^*(0) \to 0$ as $n \to \infty$. The polynonials

$$P_n(x) = P_n^*(x) - P_n^*(0) \quad (n = 1,2,3,\ldots)$$

have desired properties. ■

Definition 6.0.7 A family $\mathscr{A}$ of complex functions defined on a set F is said to be an algehra if

1. $f + g \in \mathscr{A}$.
2. $fg \in \mathscr{A}$. and
3. $cf \in \mathscr{A}$ for all $\int \epsilon \mathscr{A}, g \in \mathscr{A}$ and for all complex constants c,

that is, if $\mathscr{A}$ is closed under addition, multiplication, and scalar multiplication. We shall also have to consider algebras of real functions; in this case, (iii) is of course only required to hold for all real c. If $\mathscr{A}$ has the property that $f \in \mathscr{A}$ whenever $f_n \in \mathscr{A} (n = 1,2,3,\ldots)$ and $f_n \to f$ uniformly on E, then $\mathscr{A}$ is said to be uniformly closed.

Let $\mathscr{B}$ be the set of all functions which are limits of unifurmly convergent sequences of members of $\mathscr{A}$. Then $\mathscr{B}$ is called the uniform closure of $\mathscr{A}$.

■ **Example 6.3** The set of all polynomials is an algebra, and the Weierstrass theorem may be stated by saying that the set of continuous functions on $[a,b]$ is the uniform closure of the set of polynomials on $[a,b]$

Theorem 6.0.17 Let $\mathscr{B}$ be the uniform closure of an algebra a of bounded functions. Then $\mathscr{B}$ is a uniformly closed algebra.

Proof. If $f \in \mathscr{B}$ and $g \in \mathscr{B}$, there exist uniformly convergent sequences $\{f_n\}, \{g_n\}$ such that $f_n \to f, g_n \to g$ and $f_n \in \mathscr{A}, g_n \in \mathscr{A}$. Since we are dealing with bounded functions, it is easy to show that

$$f_n + g_n \to f + g, \quad f_n q_n \to fg, \quad cf_n \to cf$$

where c is any constant, the convergence being uniform in each case. Hence $f + g \in \mathscr{B}, fg \in \mathscr{B}$, and $cf \in \mathscr{B}$, so that $\mathscr{B}$ is an algebra. And $\mathscr{B}$ is (uniformly) closed. ■

Definition 6.0.8 — separate points. Let $\mathscr{A}$ be a family of functions on a set E. Then $\mathscr{A}$ is said to **separate points** on E if to every pair of distincts point $x_1, x_2 \in E$ there corresponds a function $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$

If to each $x \in E$ there corresponds a function $g \in \mathscr{A}$ such that $g(x) \neq 0$ we say that $\mathscr{A}$ **vanishes at no point** of E

The algebra of all polynomials in one variable clearly has these properties on R^1. An example of an algebra which does not separate points is the set of all even polynomials, say on $[-1,1]$, since $f(-x) = f(x)$ for every even function f.

Theorem 6.0.18 Suppose $\mathscr{A}$ is an algebra of functions on a set E, $\mathscr{A}$ separates points on E, and $\mathscr{A}$ vanishes at no point of E. Suppose x_1, x_2 are distinct points of E, and c_1, c_2 are constants (real if $\mathscr{A}$ is a real algebra). Then $\mathscr{A}$ contains *a* function f such that

$$f(x_1) = c_1, \quad f(x_2) = c_2$$

Proof. The assumptions show that $\mathscr{A}$ contains functions g, h, and k such that

$$g(x_1) \neq g(x_2), \quad h(x_1) \neq 0, \quad k(x_2) \neq 0$$

Put

$$u = gk - g(x_1)k, \quad v = gh - g(x_2)h$$

Then $u \in \mathscr{A}, v \in \mathscr{A}, u(x_1) = v(x_2) = 0, u(x_2) \neq 0$, and $v(x_1) \neq 0$. Therefore

$$f = \frac{c_1 v}{v(x_1)} + \frac{c_2 u}{u(x_2)}$$

has the desired properties.

■

7.32 Theorem Let $\mathscr{A}$ be an algebra of real continuous functions on a compact set K. If $\mathscr{A}$ separates points on K and if $\mathscr{A}$ vanishes at no point of K, then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all real continuous functions on K We shall divide the proof into four steps.

Step 1: If $f \in \mathscr{B}$, then $|f| \in \mathscr{B}$

Proof. Let

$$a = \sup |f(x)| \quad (x \in K)$$

and let $\varepsilon > 0$ be given. By Corollary 6.0.16 there exist real numbers $c_1, \ldots, c_n$ such that

$$\left| \sum_{i=1}^{n} c_i y^i - |y| \right| < \varepsilon \quad (-a \leq y \leq a)$$

since $\mathscr{B}$ is an algebra, the function

$$g = \sum_{i=1}^{n} c_i f^i$$

is a member of $\mathscr{B}$. Then we have

$$|g(x) - |f(x)|| < \varepsilon \quad (x \in K)$$

since $\mathscr{B}$ is uniformly closed, this shows that $|f| \in \mathscr{B}$ ■

Step 2: If $f \in \mathscr{B}$ and $g \in \mathscr{B}$, then $\max(f,g) \in \mathscr{B}$ and $\min(f,g) \in \mathscr{B}$ By $\max(f,g)$ we mean the function h defined by

$$h(x) = \begin{cases} f(x) & \text{if } f(x) \geq g(x) \\ g(x) & \text{if } f(x) < g(x) \end{cases}$$

and $\min(f,g)$ is defined likewise.

Proof. Step 2 follows from step 1 and the identities

$$\begin{aligned} \max(f,g) &= \tfrac{f+g}{2} + \tfrac{|f-g|}{2} \\ \min(f,g) &= \tfrac{f+g}{2} - \tfrac{|f-g|}{2} \end{aligned}$$

By iteration, the result can of course be extended to any finite set of functions: If $f_1, \ldots, f_n \in \mathscr{B}$, then

$$\max(f_1, \ldots, f_n) \in \mathscr{B}$$

and

$$\min(f_1, \ldots, f_n) \in \mathscr{B}$$

■

Step 3: Given a real function f, continuous on K, a point $x \in K$, and $\varepsilon > 0$, there exists a function $g_x \in \mathscr{B}$ such that $g_x(x) = f(x)$ and

$$g_x(t) > f(t) - \varepsilon \quad (t \in K) \tag{6.11}$$

Proof. Since $\mathscr{A} \subset \mathscr{B}$ and $\mathscr{A}$ satisfies the hypotheses of Theorem 7.31 so does $\mathscr{B}$. Hence, for every $y \in K$, we can find a function $h_y \in \mathscr{B}$ such that

$$h_y(x) = f(x), \quad h_y(y) = f(y) \tag{6.12}$$

By the continuity of h_y there exists an open set J_y, containing y such that

$$h_y(t) > f(t) - \varepsilon \quad (t \in J_y)$$

since K is compact, there is a finite set of points $y_1,\ldots,y_n$ such that

$$K \subset J_{y_1} \cup \cdots \cup J_{y_n}$$

Put

$$g_x = \max\left(h_{y_1},\ldots,h_{y_n}\right)$$

By step 2,$g \in \mathscr{B}$, and the previous relations show that g_x has the other required properties. ■

Step 4: Given a real function f, continuous on K, and $\varepsilon > 0$, there exists a function $h \in \mathscr{B}$ such that

$$|h(x) - f(x)| < \varepsilon \quad (x \in K) \tag{6.13}$$

since $\mathscr{B}$ is uniformly closed, this statement is equivalent to the conclusion of the theorem.

Proof. Let us consider the functions g_x, for each $x \in K$, constructed in step 3. By the continuity of g_x, there exist open sets V_x containing x such that

$$g_x(t) < f(t) + \varepsilon \quad (t \in V_x) \tag{6.14}$$

Since K is compact, there exists a finite set of points $x_1,\ldots,x_n$ such that

$$K \subset V_{x_1} \cup \cdots \cup V_{x_m} \tag{6.15}$$

Put

$$h = \min\left(g_{x_1},\ldots,g_{x_m}\right)$$

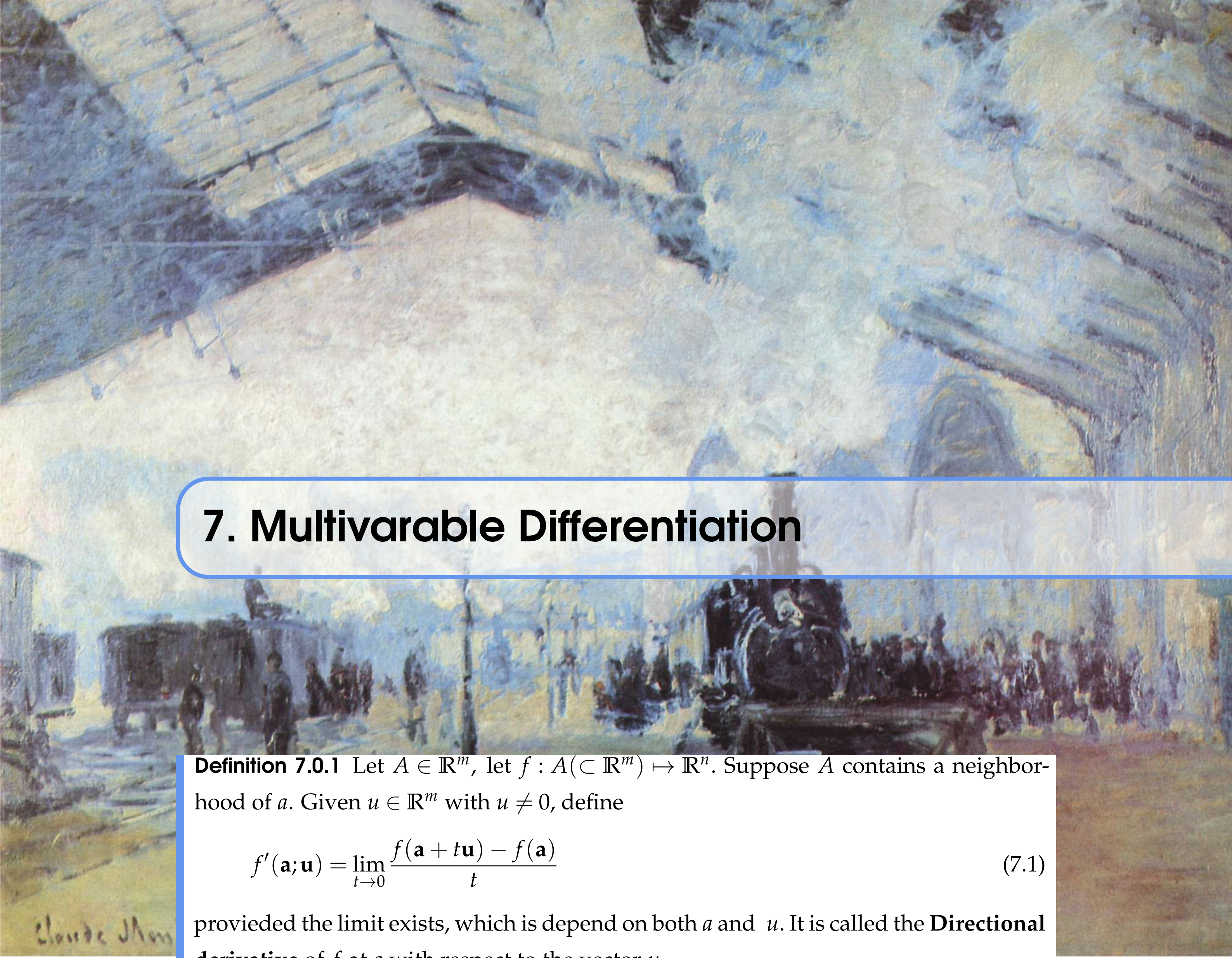
By step 2,$h \in \mathscr{B}$, and (6.11) implies

$$h(t) > f(t) - \varepsilon \quad (t \in K) \tag{6.16}$$

whereas (6.14) and (6.15) imply

$$h(t) < f(t) + \varepsilon \quad (t \in K) \tag{6.17}$$

Finally, (6.13) follows from (6.16) and (6.17) ■



7. Multivarable Differentiation

Definition 7.0.1 Let $A \in \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}^n$. Suppose A contains a neighborhood of a. Given $u \in \mathbb{R}^m$ with $u \neq 0$, define

$$f'(\mathbf{a};\mathbf{u}) = \lim_{t \to 0} \frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} \tag{7.1}$$

provieded the limit exists, which is depend on both a and u. It is called the **Directional derivative** of f at a with respect to the vector u.

Definition 7.0.2 Let $A \subset \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}^n$. Suppose A contains a neighborhood of a. We say that f is differentiable at a if there is an n by m matrix B such that

$$\frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - B \cdot \mathbf{h}}{|\mathbf{h}|} \to \mathbf{0} \quad \text{as} \quad \mathbf{h} \to \mathbf{0} \tag{7.2}$$

The matrix B, which is unique, is called the derivative of f at a ; it is denoted $Df(a)$.

Theorem 7.0.1 Let $A \subset \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}^n$. If f is differentiable at a, then all the directional derivatives of f at a exist, and

$$f'(\mathbf{a};\mathbf{u}) = \mathcal{D}f(\mathbf{a}) \cdot \mathbf{u} \tag{7.3}$$

(Let $\mathbf{h} = t\mathbf{u}$)

Theorem 7.0.2 Let $A \subset \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}^n$. If f is diff. at a, then f is conti. at a.

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = |\mathbf{h}| \left[\frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - B \cdot \mathbf{h}}{|\mathbf{h}|} \right] + B \cdot \mathbf{h}$$

Definition 7.0.3 Let $A \subset \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}$. We define the j^{th} partial derivative of f at a to be the directional derivative of f at a with respect to the vector e_j , provided this derivative exists; and we denote it by $D_i f(a)$. That is,

$$D_j f(\mathbf{a}) = \lim_{t \to 0} \left(f\left(\mathbf{a} + t\mathbf{e}_j\right) - f(\mathbf{a})\right)/t \tag{7.4}$$

R By def, partial derivative is special directional derivative.

Theorem 7.0.3 Let $A \subset \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}$. If f is differentiable at a, then

$$Df(\mathbf{a}) = \left[D_1 f(\mathbf{a}) \quad D_2 f(\mathbf{a}) \quad \cdots \quad D_m f(\mathbf{a}) \right] \tag{7.5}$$

That is, if $Df(a)$ exists, it is the row matrix whose entries are the partial derivatives of f at a.

$$D_j f(\mathbf{a}) = f'\left(\mathbf{a}; \mathbf{e}_j\right) = Df(\mathbf{a}) \cdot \mathbf{e}_j$$

Theorem 7.0.4 Let $A \subset \mathbb{R}^m$, let $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}^n$. Suppose A contains a neighborhood of a. Let $f_i : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}$ be the i^{th} component function of f, so that

$$f(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix}$$

(a) The function f is differentiable at a if and only if each component function f_i is differentiable at a. (b) If f is differentiable at a, then its derivative is the $n \times m$ matrix whose i^{th} row is the derivative of the function f_i.

R This theorem tells us that

$$f(\mathbf{a}) = \begin{bmatrix} Df_1(\mathbf{a}) \\ \vdots \\ Df_n(\mathbf{a}) \end{bmatrix}$$

so that $Df(a)$ is the matrix whose entry in row i and column j is $D_j f_i(a)$.

Definition 7.0.4 If the partial derivatives of the component functions f_i of f exist at a, then one can form the matrix that has $D_j f_i(a)$ as its entry in row i and column j. This matrix is called the **Jacobian matrix** of f. If f is differentiable at a, this matrix equals $\mathcal{D}f(a)$, which is the result of previous theorem. However, it is possible for the partial derivatives, and hence the Jacobian matrix, to exist, without it following that f is differentiable at a, which means J exists as long as partial derivatives exists.

■ **Example 7.1** Consider a differentiable function $g : \mathbb{R}^3 \mapsto \mathbb{R}$. Its derivative is the row matrix

$$Dg(\mathbf{x}) = \left[\begin{array}{lll} D_1 g(\mathbf{x}) & D_2 g(\mathbf{x}) & D_3 g(\mathbf{x}) \end{array}\right]$$

the directional derivative equals the matrix product $Dg(x)u$. And the vector field

$$\vec{\nabla} g = \operatorname{grad} g = (D_1 g)\,\mathbf{e}_1 + (D_2 g)\,\mathbf{e}_2 + (D_3 g)\,\mathbf{e}_3$$

is called the **gradient** of g. ■

Theorem 7.0.5 — partical exists and conti. $\Longrightarrow$f diff. Let A be open in $\mathbb{R}^m$. Suppose that the partial derivatives $D_j f_i(a)$ of the component functions of f exist at each point x of A and are continuous on A.(This condition is called continuously differentiable C^1) Then f is differentiable at each point of A.

Proof. We just need to prove each component functions of f is differentiable.Therefore we may restrict ourselves to the case of a real-valued function $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}$. Let a be a point of A. We are given that, for some ε, the partial derivatives $D_j f(x)$ exist and are continuous for $|x - a| < \varepsilon$. We next to show that f is differentiable at a.

Step 1: Let $\mathbf{h}$ be a point of $\mathbb{R}^m$ with $0 < |h| < \varepsilon$, and the components of h is $h_1, \cdots h_m$. We want to use one-dim mean-value theorem, so consider the following sequence of points of$\mathbb{R}^m$;

$$\begin{aligned}
\mathbf{p}_0 &= \mathbf{a} \\
\mathbf{p}_1 &= \mathbf{a} + h_1 \mathbf{e}_1 \\
\mathbf{p}_2 &= \mathbf{a} + h_1 \mathbf{e}_1 + h_2 \mathbf{e}_2 \\
&\dots \\
\mathbf{p}_m &= \mathbf{a} + h_1 \mathbf{e}_1 + \cdots + h_m \mathbf{e}_m = \mathbf{a} + \mathbf{h}
\end{aligned}$$

each p_i and p_{i+1} are just different with one component $h_{i+1} e_{i+1}$. The points p_i all belong to the (closed) cube C of radius $|h|$ centered at a. We need to deal with the difference $f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})$, we want to add some point between them and cut them into some pieces, then we can use M-V theorem to each piece, finally to control the $f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})$

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \sum_{j=1}^{m} \left[f\left(\mathbf{p}_j\right) - f\left(\mathbf{p}_{j-1}\right)\right]$$

Let j be fixed, and define

$$\phi(t) = f\left(\mathbf{p}_{j-1} + t\mathbf{e}_j\right)$$

Assume $h_j \neq 0$ for the moment. $t \in [0, h_j] = I, P_{j-1} + te_j \in [P_{j-1}, P_j]$. As t varies, only the j^{th} component of the point $P_{j-1} + te_j$ varies. Hence because $D_j f$ exists at each point of A, the function ϕ is differentiable on an open interval containing I. Applying the mean-value theorem to ϕ, we conclude that

$$\phi\left(h_j\right) - \phi(0) = \phi'\left(c_j\right) h_j$$

So

$$f\left(\mathbf{p}_j\right) - f\left(\mathbf{p}_{j-1}\right) = D_j f\left(\mathbf{q}_j\right) h_j$$

where q_j is the point $P_{j-1} + c_j e_j$ of the line segment from P_{j-1} to p_j, which lies in C. So now we have

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \sum_{j=1}^{m} D_j f\left(\mathbf{q}_j\right) h_j \tag{7.6}$$

where each point $q_j$lies in the cube C of radius $|h|$ centered at a.

Step 2: Let B be the matrix

$$B = [D_1 f(\mathbf{a}) \cdots D_m f(\mathbf{a})]$$

Then

$$B \cdot \mathbf{h} = \sum_{j=1}^{m} D_j f(\mathbf{a}) h_j$$

$$\frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - B \cdot \mathbf{h}}{|\mathbf{h}|} = \sum_{j=1}^{m} \frac{\left[D_j f\left(\mathbf{q}_j\right) - D_j f(\mathbf{a})\right] h_j}{|\mathbf{h}|}$$

then we let $h \to 0$. Since q_j lies in the cube C of radius $|h|$ centered at a, we have $q_j \to a$. Since the partials of f are continuous at a, the factors in brackets all go to zero. The factors $h_i / |h|$ are of course bounded in absolute value by 1. Hence the entire expression goes to zero, as desired.

■

Theorem 7.0.6 Let A be open in $\mathbb{R}^m$, $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}$ be a function of class C^2. Then for each $a \in A$,

$$D_k D_j f(\mathbf{a}) = D_j D_k f(\mathbf{a})$$

Theorem 7.0.7 — Mean-value theorem. Let A be open in $\mathbb{R}^m$, $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}$ be differentiable on A. If A contains the line segment with end points **a** and **a+ h**, then there is a point $\mathbf{c} = a + t_0\mathbf{h}$ with $0 < t_0 < 1$ of this line segment such that

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = Df(\mathbf{c}) \cdot \mathbf{h}$$

Theorem 7.0.8 Let A be open in $\mathbb{R}^m$, $f : A(\subset \mathbb{R}^m) \mapsto \mathbb{R}^n$ let $f(a) = b$. Suppose that g maps a neighborhood of $b$into $\mathbb{R}^n$, that $g(b) = a$, and

$$g(f(x)) = x$$

for all x in a neighborhood of a. If f is differentiable at a and if g is differentiable at b, then

$$Dg(\mathbf{b}) = [Df(\mathbf{a})]^{-1}$$

R The preceding theorem implies that if a differentiable function f is to have a differentiable inverse, it is necessary that the matrix Df be non-singular.

■ **Example 7.2** Con■sider the chain rule, for example. If

$$f : \mathbb{R}^m \to \mathbb{R}^n \quad \text{and} \quad g : \mathbb{R}^n \to \mathbb{R}$$

then the composite function $F = gof$ maps $\mathbb{R}^m$ into $\mathbb{R}$, and its derivative is given by the formula

$$DF(\mathbf{x}) = Dg(f(\mathbf{x})) \cdot Df(\mathbf{x})$$

$$[D_1F(\mathbf{x}) \cdots D_mF(\mathbf{x})]$$
$$= [D_1g(f(\mathbf{x})) \quad \cdots \quad D_ng(f(\mathbf{x}))] \begin{bmatrix} D_1f_1(\mathbf{x}) & \cdots & D_mf_1(\mathbf{x}) \\ \cdots & & \cdots \\ D_1f_n(\mathbf{x}) & \cdots & D_mf_n(\mathbf{x}) \end{bmatrix}$$

The formula for the j^{th} partial derivative of F is thus given by the equa■tion

$$D_jF(\mathbf{x}) = \sum_{k=1}^{n} D_kg(f(\mathbf{x}))D_jf_k(\mathbf{x}) = \sum_i D_ig(f(x))\dot{D}_jf_i$$

■

7.0.1 Inverse function and implict function

Theorem 7.0.9 Let A be open in $\mathbb{R}^n$, $f : A(\subset \mathbb{R}^n) \mapsto \mathbb{R}^n$ be of class C^1. If $Df(a)$ is non-singular for , then there exists an $\alpha > 0$ such that the inequality

$$|f(\mathbf{x}_0) - f(\mathbf{x}_1)| \geq \alpha |\mathbf{x}_0 - \mathbf{x}_1|$$

holds for allx_0, x_1 in some open cube $C(a;\varepsilon)$ centered at a. It follows that f is one-to-one on this open cube.

Theorem 7.0.10 Let A be open in $\mathbb{R}^n$, $f : A(\subset \mathbb{R}^n) \mapsto \mathbb{R}^n$ be of class C^r. let $B = f(A)$. If f is one-to-one on A and if $Df(x)$ is non-singular for $x \in A$, then the set B is open in $\mathbb{R}^n$ and the inverse function $g : B \mapsto A$ is of class C^r .

Theorem 7.0.11 — The inverse function theorem. Let A be open in $\mathbb{R}^n$,$f : A(\subset \mathbb{R}^n) \mapsto \mathbb{R}^n$ be of class C^r. If $Df(x)$ is non-singular at the point a of A, there is a neighborhood U of the point a such that f carries U in a one-to-one fashion onto an open set V of $\mathbb{R}^n$ and the inverse function is of class C^r .

Definition 7.0.5 Let A be open in $\mathbb{R}^m$,$f : A(\subset \mathbb{R}^n) \mapsto \mathbb{R}^n$ is differentiable . Let$f_1, \cdots, f_n$ be the component functions of f. We sometimes use the notation

$$Df = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_m)} = \partial f / \partial \mathbf{x}$$

Theorem 7.0.12 Let A be open in $\mathbb{R}^{k+n}$; let $f : A(\subset \mathbb{R}^{k+n}) \mapsto \mathbb{R}^n$ be differentiable. Write $f$in the form$f(x,y)$, for$x \in R^k$ and $y \in R^n$; then Df has the form

$$Df = \begin{bmatrix} \partial f / \partial \mathbf{x} & \partial f / \partial \mathbf{y} \end{bmatrix}$$

Suppose there is a differentiable function $g : B \mapsto \mathbb{R}^n$ defined on an open set B in $\mathbb{R}^k$, such that

$$f(\mathbf{x}, g(\mathbf{x})) = \mathbf{0}$$

for all $x \in B$. Then for $x \in B$,

$$\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, g(\mathbf{x})) + \frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, g(\mathbf{x})) \cdot Dg(\mathbf{x}) = \mathbf{0}$$

This equation implies that if the$n \times n$ matrix $\partial f / \partial y$ is non-singular at the point $(x, g(x))$, then

$$Dg(\mathbf{x}) = -\left[\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, g(\mathbf{x}))\right]^{-1} \cdot \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, g(\mathbf{x}))$$

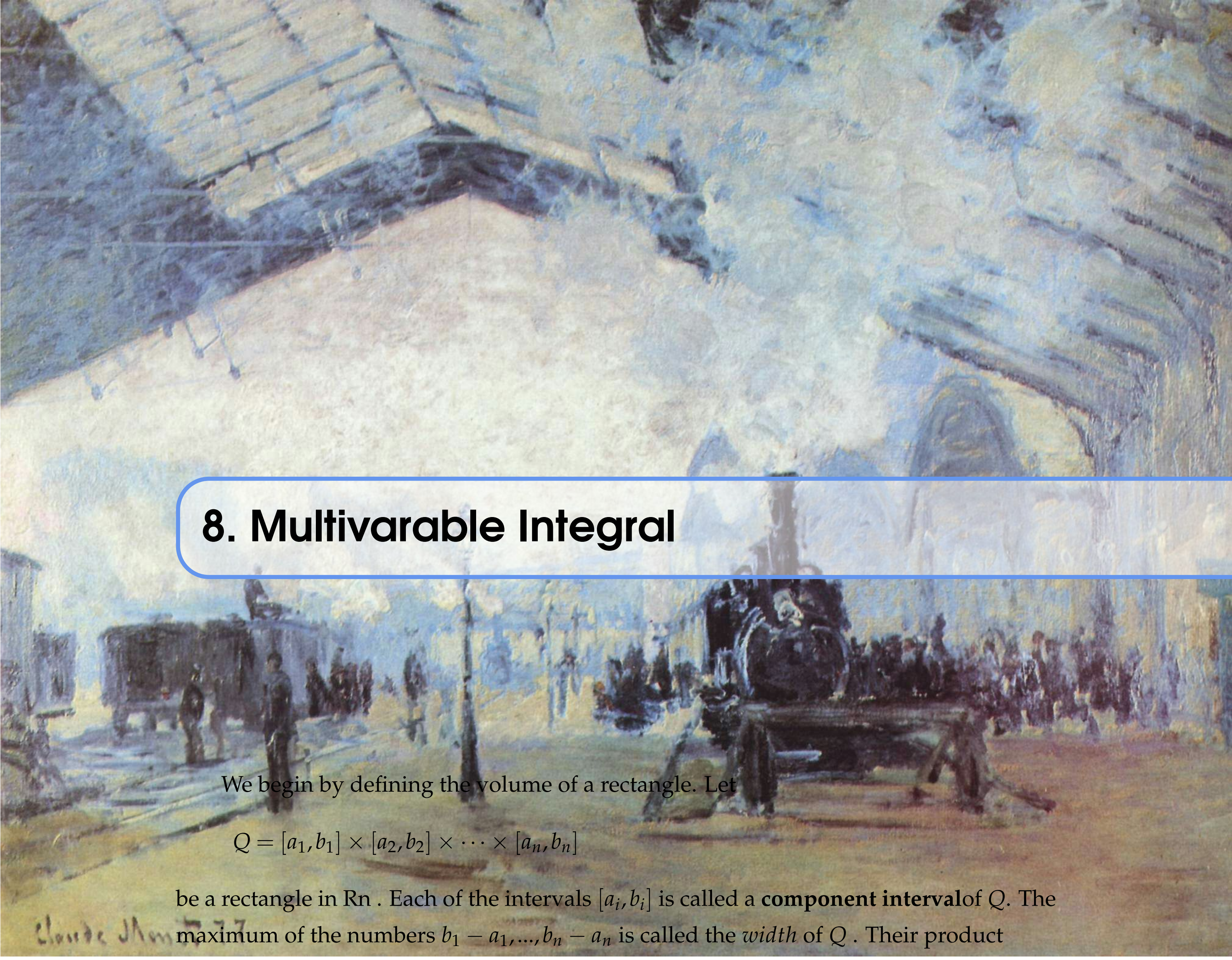
Theorem 7.0.13 — (Implicit function theorem). Let A be open in $\mathbb{R}^{k+n}$; let $f : A(\subset \mathbb{R}^{k+n}) \mapsto \mathbb{R}^n$ of class C^r. Write f in the form $f(x,y)$, for$x \in R^k$ and $y \in R^n$. Suppose that (a,b) is a point of A such that $f(a,b) = 0$ and

$$\det \frac{\partial f}{\partial \mathbf{y}}(\mathbf{a},\mathbf{b}) \neq 0$$

Then there is a neighborhood B of a in $\mathbb{R}^k$and a unique continuous function $g : B \mapsto \mathbb{R}^n$ such that $g(a) = b$and

$$f(\mathbf{x}, g(\mathbf{x})) = \mathbf{0}$$

for all $x \in B$. The function g is in fact of class C^r.



8. Multivarable Integral

We begin by defining the volume of a rectangle. Let

$$Q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

be a rectangle in Rn . Each of the intervals $[a_i, b_i]$ is called a **component interval**of Q. The maximum of the numbers $b_1 - a_1, ..., b_n - a_n$ is called the *width* of Q . Their product

$$v(Q) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$$

is called the **volume** of Q .

Definition 8.0.1 Given a closed interval$[a, b]$ of $\mathbb{R}$, a **partition** of $[a, b]$ is a finite collection P of points of $[a, b]$ that includes the points a and b.

$$a = t_0 < t_1 < \cdots < t_k = b$$

each of the intervals $[t_{i-1}, ti]$, for$i = 1, ..., k$, is called a subinterval deterÂ■mined by P , o f the interval$[a, b]$.

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

More generally, given a rectangle in $\mathbb{R}^n$, a partition P of Q is an n-tuple $(P_i, ..., P_n)$such that P_j is a partition of $[a_j, b_j]$ for each j. If for each j, I_j is one of the subintervals determined by P_j of the interval $[a_j, b_j]$, then the rectangle

$$R = I_1 \times \cdots \times I_n$$

is called a subrectangle determined by$P$, of the rectangle Q. The maxiÂ■mum width of these subrectangles is called the **mesh** of P.

Definition 8.0.2 Let Q be a rectangle in $\mathbb{R}^n$, let $f : Q \mapsto \mathbb{R}$, assume f is bounded. Let P be a partition of Q. For each subrectangle R determined by P, let

$$m_R(f) = \inf\{f(\mathbf{x})|\mathbf{x} \in R\}$$
$$M_R(f) = \sup\{f(\mathbf{x})|\mathbf{x} \in R\}$$

We define the **lower sum** and the **upper sum**, respectively, of **f**, determined by **P**, by the equations

$$L(f,P) = \sum_R m_R(f) \cdot v(R)$$
$$U(f,P) = \sum_R M_R(f) \cdot v(R)$$

where the summations extend over all subrectangles R determined by P.

9. Measures

Definition 9.0.1 — Algebras. Let X be an arbitrary set. A collection $\mathscr{A}$ of subsets of X is an algebra on X if

1. $X \in \mathscr{A}$
2. for each set A that belongs to $\mathscr{A}$, the set A^c belongs to $\mathscr{A}$
3. for each finite sequence $A_1, \ldots, A_n$ of sets that belong to $\mathscr{A}$, the set $\cup_{i=1}^{n} A_i$ belongs to $\mathscr{A}$, and
4. for each finite sequence $A_1, \ldots, A_n$ of sets that belong to $\mathscr{A}$, the set $\cap_{i=1}^{n} A_i$ belongs to $\mathscr{A}$

Of course, in conditions (b), (c), and (d), we have required that $\mathscr{A}$ **be closed under complementation, under the formation of finite unions, and under the formation of finite intersections.**

Definition 9.0.2 — σ -algebra. Let X be an arbitrary set. A collection $\mathscr{A}$ of subsets of X is a σ -algebra [1] on X if

1. $X \in \mathscr{A}$
2. for each set A that belongs to $\mathscr{A}$, the set A^c belongs to $\mathscr{A}$
3. for each infinite sequence $\{A_i\}$ of sets that belong to $\mathscr{A}$, the set $\cup_{i=1}^{\infty} A_i$ belongs to $\mathscr{A}$, and
4. for each infinite sequence $\{A_i\}$ of sets that belong to $\mathscr{A}$, the set $\cap_{i=1}^{\infty} A_i$ belongs to $\mathscr{A}$

Thus a σ -algebra on X is a family of subsets of X that contains X and is closed **under complementation, under the formation of countable unions, and under the**

formation of countable intersections. We can replace condition (a) with the requirement that $\mathscr{A}$ be nonempty.

Definition 9.0.3 — $\mathscr{A}$ -measurable. If $\mathscr{A}$ is a σ -algebra on the set X, it is sometimes convenient to call a subset of X $\mathscr{A}$ -measurable if it belongs to $\mathscr{A}$

■ **Example 9.1**
1. Let X be a set, and let $\mathscr{A}$ be the collection of all subsets of X. Then $\mathscr{A}$ is a σ -algebra on X
2. Let X be a set, and let $\mathscr{A} = \{\varnothing, X\}$. Then $\mathscr{A}$ is a σ -algebra on X
3. Let $\mathscr{A}$ be the collection of all subsets of $\mathbb{R}$ that are unions of finitely many intervals of the form $(a,b], (a,+\infty)$, or $(-\infty,b]$. It is easy to check that each set that belongs to $\mathscr{A}$ is the union of a finite disjoint collection of intervals of the types listed above, and then to check that $\mathscr{A}$ is an algebra on $\mathbb{R}$ (the empty set belongs to $\mathscr{A}$, since it is the union of the empty, and hence finite, collection of intervals). **The algebra $\mathscr{A}$ is not a σ -algebra**; for example, the bounded open subintervals of $\mathbb{R}$ are unions of sequences of sets in $\mathscr{A}$ but do not themselves belong to $\mathscr{A}$

Proposition 9.0.1 Let X be a set. Then the intersection of an arbitrary nonempty collection of σ -algebras on X is a σ -algebra on X.

Proposition 9.0.2 The union of a family of σ -algebras can fail to be a σ -algebra.

Corollary 9.0.3 Corollary 1.1.3. Let X be a set, and let $\mathscr{F}$ be a family of subsets of X. Then there is a smallest σ -algebra on X that includes $\mathscr{F}$

Definition 9.0.4 The smallest σ -algebra is called the σ -algebra generated by $\mathscr{F}$ and is often denoted by $\sigma(\mathscr{F})$

Definition 9.0.5 The Borel σ-algebra on $\mathbb{R}^d$ is the σ -algebra on $\mathbb{R}^d$ generated by the collection of open subsets of $\mathbb{R}^d$; it is denoted by $\mathscr{B}(\mathbb{R}^d)$. The Borel subsets of $\mathbb{R}^d$ are those that belong to $\mathscr{B}(\mathbb{R}^d)$. In case $d = 1$, one generally writes $\mathscr{B}(\mathbb{R})$ in place of $\mathscr{B}(\mathbb{R}^1)$

Proposition 9.0.4 The σ -algebra $\mathscr{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ is generated by each of the following collections of sets:
1. the collection of all closed subsets of $\mathbb{R}$
2. the collection of all subintervals of $\mathbb{R}$ of the form $(-\infty, b]$
3. the collection of all subintervals of $\mathbb{R}$ of the form $(a, b]$.

Proposition 9.0.5 The following properties of the σ -algebra $\mathscr{B}(\mathbb{R})$:
1. It contains virtually [2] every subset of $\mathbb{R}$ that is of interest in analysis.
2. It is small enough that it can be dealt with in a fairly constructive manner.

Proposition 9.0.6 Proposition 1.1.5. The σ -algebra $\mathscr{B}(\mathbb{R}^d)$ of Borel subsets of $\mathbb{R}^d$ is generated by each of the following collections of sets: (a) the collection of all closed subsets of $\mathbb{R}^d$; (b) the collection of all closed half-spaces in $\mathbb{R}^d$ that have the form $\{(x_1,\ldots,x_d) : x_i \leq b\}$ for some index i and some b in $\mathbb{R}$ (c) the collection of all rectangles in $\mathbb{R}^d$ that have the form

$$\{(x_1,\ldots,x_d) : a_i < x_i \leq b_i \text{ for } i = 1,\ldots,d\}$$

Definition 9.0.6 Let $\mathscr{G}$ be the family of all open subsets of $\mathbb{R}^d$, and let $\mathscr{F}$ be the family of all closed subsets of $\mathbb{R}^d$. (Of course$\mathscr{G}$ and $\mathscr{F}$ depend on the dimension d, and it would have been more precise to write $\mathscr{G}\left(\mathbb{R}^d\right)$ and $\mathscr{F}\left(\mathbb{R}^d\right)$.)

Let $\mathscr{G}_\delta$ be the collection of all **intersections** of sequences of sets in $\mathscr{G}$, and let $\mathscr{F}_\sigma$ be the collection of all **unions** of sequences of sets in $\mathscr{F}$. Sets in $\mathscr{G}_\delta$ are often called G_δ 's, and sets in $\mathscr{F}_\sigma$ are often called F_σ 's.

Theorem 9.0.7 Proposition 1.1.6. Each closed subset of $\mathbb{R}^d$ is a G_δ, and each open subset of $\mathbb{R}^d$ is an F_σ

Proof. Suppose that F is a closed subset of $\mathbb{R}^d$. We need to construct a sequence $\{U_n\}$ of open subsets of $\mathbb{R}^d$ such that $F = \cap_n U_n$. For this define U_n by

$$U_n = \left\{x \in \mathbb{R}^d : \|x - y\| < 1/n \text{ for some } y \text{ in } F\right\}$$

(Note that U_n is empty if F is empty.) It is clear that each U_n is open and that $F \subseteq \cap_n U_n$. The reverse inclusion follows from the fact that F is closed (note that each point in $\cap_n U_n$ is the limit of a sequence of points in F). Hence each closed subset of $\mathbb{R}^d$ is a G_δ

If U is open, then U^c is closed and so is a G_δ. Thus there is a sequence $\{U_n\}$ of open sets such that $U^c = \cap_n U_n$. The sets U_n^c are then closed, and $U = \cup_n U_n^c$; hence U is an F_σ ■

Proposition 9.0.8 Let X be a set, and let $\mathscr{A}$ be an algebra on X. Then $\mathscr{A}$ is a σ-algebra if either

1. $\mathscr{A}$ is closed under the formation of unions of increasing sequences of sets, or
2. $\mathscr{A}$ is closed under the formation of intersections of decreasing sequences of sets.

Proof. let $B_n = \cup_{i=1}^n A_i$. ■

Definition 9.0.7 Let X be a set, and let $\mathscr{A}$ be a σ -algebra on X. A function μ whose domain is the σ -algebra $\mathscr{A}$ and whose values belong to the extended half-line $[0,+\infty]$ is said to be **countably additive** if it satisfies

$$\mu\left(\cup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu\left(A_i\right)$$

for each infinite sequence $\{A_i\}$ of disjoint sets that belong to $\mathscr{A}$.

(Since $\mu(A_i)$ is nonnegative for each i, the sum $\sum_{i=1}^{\infty}\mu(A_i)$ always exists, either as a real number or as $+\infty$; see Appendix B.) A **measure** (or a countably additive measure) on $\mathscr{A}$ is a function $\mu : \mathscr{A} \to [0,+\infty]$ that satisfies $\mu(\varnothing) = 0$ and is countably additive.

Definition 9.0.8 Let $\mathscr{A}$ be an algebra (not necessarily a σ -algebra) on the set X. A function μ whose domain is $\mathscr{A}$ and whose values belong to $[0,+\infty]$ is finitely additive if it satisfies

$$\mu(\cup_{i=1}^{n}A_i) = \sum_{i=1}^{n}\mu(A_i)$$

for each finite sequence $A_1,\ldots,A_n$ of disjoint sets that belong to $\mathscr{A}$. A **finitely additive measure on the algebra** $\mathscr{A}$ is a function $\mu : \mathscr{A} \to [0,+\infty]$ that satisfies $\mu(\varnothing) = 0$ and is finitely additive.

Definition 9.0.9 If X is a set, if $\mathscr{A}$ is a σ -algebra on X, and if μ is a measure on $\mathscr{A}$, then the triplet $(X,\mathscr{A},\mu)$ is often called a measure space. Likewise, if X is a set and if $\mathscr{A}$ is a σ-algebra on X, then the pair $(X,\mathscr{A})$ is often called a measurable space. If $(X,\mathscr{A},\mu)$ is a measure space, then one often says that μ is a measure on $(X,\mathscr{A})$, or, if the σ -algebra $\mathscr{A}$ is clear from context, a measure on X

1. Let X be an arbitrary set, and let $\mathscr{A}$ be a σ -algebra on X. Define a function $\mu : \mathscr{A} \to [0,+\infty]$ by letting $\mu(A)$ be n if A is a finite set with n elements and letting $\mu(A)$ be $+\infty$ if A is an infinite set. Then μ is a measure; it is often called **counting measure** on $(X,\mathscr{A})$
2. Let X be a nonempty set, and let $\mathscr{A}$ be a σ -algebra on X. Let x be a member of X. Define a function $\delta_x : \mathscr{A} \to [0,+\infty]$ by letting $\delta_x(A)$ be 1 if $x \in A$ and letting $\delta_x(A)$ be 0 if $x \notin A$. Then δ_x is a measure; it is called a **point mass** concentrated at x
3. Consider the set $\mathbb{R}$ of all real numbers and the σ -algebra $\mathscr{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$. In Sect. 1.3 we will construct a measure on $\mathscr{B}(\mathbb{R})$ that assigns to each subinterval of $\mathbb{R}$ its length; this measure is known as **Lebesgue measure** and will be denoted by λ in this book.

Proposition 9.0.9 Let$(X,\mathscr{A},\mu)$ be a measure space, and let A and B be subsets of X that belong to $\mathscr{A}$ and satisfy $A \subseteq B$. Then $\mu(A) \leq \mu(B)$. If in addition A satisfies $\mu(A) < +\infty$, then $\mu(B-A) = \mu(B) - \mu(A)$

Definition 9.0.10 Let μ be a measure on a measurable space $(X,\mathscr{A})$.

1. Then μ is a **finite measure** if $\mu(X) < +\infty$ and is a σ -finite measure if X is the union of a sequence $A_1,A_2,\ldots$ of sets that belong to $\mathscr{A}$ and satisfy $\mu(A_i) < +\infty$

for each i.

2. More generally, a set in $\mathscr{A}$ is σ **-finite** under μ if it is the union of a sequence of sets that belong to $\mathscr{A}$ and have finite measure under μ. The measure space $(X, \mathscr{A}, \mu)$ is also called finite or σ -finite if μ is finite or σ -finite.

Theorem 9.0.10 If the measure space $(X, \mathscr{A}, \mu)$ is σ -finite, then X is the union of a sequence $\{B_i\}$ of disjoint sets that belong to $\mathscr{A}$ and have finite measure under μ; such a sequence $\{B_i\}$ can be formed by choosing a sequence $\{A_i\}$ as in the definition of σ -finiteness, and then letting $B_1 = A_1$ and $B_i = A_i - \left(\cup_{j=1}^{i-1} A_j\right)$ if $i > 1$

Theorem 9.0.11 Let$(X, \mathscr{A}, \mu)$ be a measure space. If $\{A_k\}$ is an arbitrary sequence of sets that belong to $\mathscr{A}$, then

$$\mu\left(\cup_{k=1}^{\infty} A_k\right) \le \sum_{k=1}^{\infty} \mu\left(A_k\right)$$

Proof. $B_k = A_k - \left(\cup_{i=1}^{k-1} A_i\right)$ if $k > 1$. $\mu\left(B_k\right) \le \mu\left(A_k\right)$. Since in addition the sets B_k are disjoint and satisfy $\cup_k B_k = \cup_k A_k$, it follows that

$$\mu\left(\cup_k A_k\right) = \mu\left(\cup_k B_k\right) = \sum_k \mu\left(B_k\right) \le \sum_k \mu\left(A_k\right)$$

In other words, the countable additivity of μ implies the countable subadditivity of μ ■

Theorem 9.0.12 Let$(X, \mathscr{A}, \mu)$ be a measure space.

1. If $\{A_k\}$ is an increasing sequence of sets that belong to $\mathscr{A}$, then
2.

$$\mu\left(\cup_k A_k\right) = \lim_k \mu\left(A_k\right)$$

3. If $\{A_k\}$ is a decreasing sequence of sets that belong to $\mathscr{A}$, and if $\mu\left(A_n\right) < +\infty$ holds for some n, then

$$\mu\left(\cap_k A_k\right) = \lim_k \mu\left(A_k\right)$$

Theorem 9.0.13 Let$(X, \mathscr{A})$ be a measurable space, and let μ be a finitely additive measure on $(X, \mathscr{A})$. Then μ is a measure if either

1. $\lim_k \mu\left(A_k\right) = \mu\left(\cup_k A_k\right)$ holds for each increasing sequence $\{A_k\}$ of sets that belong to $\mathscr{A}$, or
2. $\lim_k \mu\left(A_k\right) = 0$ holds for each decreasing sequence $\{A_k\}$ of sets that belong to $\mathscr{A}$ and satisfy $\cap_k A_k = \varnothing$

Definition 9.0.11 A measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ is often called a **Borel measure** on $\mathbb{R}^d$. More generally, if X is a Borel subset of $\mathbb{R}^d$ and if $\mathscr{A}$ is the σ -algebra consisting of those Borel subsets of $\mathbb{R}^d$ that are included in X, then a measure on $(X, \mathscr{A})$ is called a Borel measure on X

Now suppose that $(X, \mathscr{A})$ is a measurable space such that for each x in X the set$\{x\}$ belongs to $\mathscr{A}$. A finite or σ -finite measure μ on $(X, \mathscr{A})$ is **continuous** if $\mu(\{x\}) = 0$ holds for each x in X and is discrete if there is a countable subset D of X such that $\mu(D^c) = 0$. (More elaborate definitions are needed if $\mathscr{A}$ does not contain each $\{x\}$ or if μ is not σ -finite.)

Definition 9.0.12 — outer measure. Let X be a set, and let $\mathscr{P}(X)$ be the collection of all subsets of X. An outer measure on X is a function $\mu^* : \mathscr{P}(X) \to [0, +\infty]$ such tha

1. $\mu^*(\varnothing) = 0$
2. if $A \subseteq B \subseteq X$, then $\mu^*(A) \leq \mu^*(B)$, and
3. if $\{A_n\}$ is an infinite sequence of subsets of X, then $\mu^*(\cup_n A_n) \leq \sum_n \mu^*(A_n)$

Thus an outer measure on X is a monotone and countably subadditive function from $\mathscr{P}(X)$ to $[0, +\infty]$ whose value at $\varnothing$ is 0.

R Note that a measure can fail to be an outer measure; in fact, a measure on X is an outer measure if and only if its domain is $\mathscr{P}(X)$.

Theorem 9.0.14 Countable union of countable set is countable, proved by draw a table.

■ **Example 9.2** (d) Lebesgue outer measure on $\mathbb{R}$, which we will denote by λ^*, is defined as follows. For each subset A of $\mathbb{R}$, let $\mathscr{C}_A$ be the set of all infinite sequences $\{(a_i, b_i)\}$ of bounded open intervals such that $A \subseteq \cup_i (a_i, b_i)$. Then $\lambda^* : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ is defined by

$$\lambda^*(A) = \inf\left\{\sum_i (b_i - a_i) : \{(a_i, b_i)\} \in \mathscr{C}_A\right\}$$

(Note that the set of sums involved here is nonempty and that the infimum of the set consisting of $+\infty$ alone is $+\infty$. We check in the following proposition that λ^* is indeed an outer measure.)

Theorem 9.0.15 Lebesgue outer measure on $\mathbb{R}$ is an outer measure, and it assigns to each subinterval of $\mathbb{R}$ its length.

R The outer measure of an arbitrary bounded interval is its length.

■ **Example 9.3** Lebesgue outer measure on $\mathbb{R}^d$, which we will denote by λ^* (or, if necessary in order to avoid ambiguity, by λ_d^*) is defined as follows. A d-dimensional interval is a subset of $\mathbb{R}^d$ of the form $I_1 \times \cdots \times I_d$, where $I_1, \ldots, I_d$ are subintervals of $\mathbb{R}$ and $I_1 \times \cdots \times I_d$ is given by

$$I_1 \times \cdots \times I_d = \{(x_i, \ldots, x_d) : x_i \in I_i \text{ for } i = 1, \ldots, d\}$$

Note that the intervals $I_1, \ldots, I_d$, and hence the d -dimensional interval $I_1 \times \cdots \times I_d$ can be open, closed, or neither open nor closed. The volume of the d -dimensional interval $I_1 \times \cdots \times I_d$ is the product of the lengths of the intervals $I_1, \ldots, I_d$, and will be denoted by $\operatorname{vol}(I_1 \times \cdots \times I_d)$. For each subset A of $\mathbb{R}^d$ let $\mathscr{C}_A$ be the set of all sequences $\{R_i\}$ of bounded and open d -dimensional intervals for which $A \subseteq \cup_{i=1}^{\infty} R_i$ Then $\lambda^*(A)$, the outer measure of A, is the infimum of the set

$$\left\{\sum_{i=1}^{\infty} \operatorname{vol}(R_i) : \{R_i\} \in \mathscr{C}_A\right\}$$

Theorem 9.0.16 Lebesgue outer measure on $\mathbb{R}^d$ is an outer measure, and it assigns to each d -dimensional interval its volume.

Definition 9.0.13 Let X be a set, and let μ^* be an outer measure on X. A subset B of X is μ^* measurable (or measurable with respect to μ^*) if

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

holds for every subset A of X. Thus a μ^* -measurable subset of X is one that divides each subset of X in such a way that the sizes (as measured by μ^*) of the pieces add properly. A Lebesgue measurable subset of $\mathbb{R}$ or of $\mathbb{R}^d$ is of course one that is measurable with respect to Lebesgue outer measure. Note that the subadditivity of the outer measure μ^* implies that

$$\mu^*(A) \leq \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

holds for all subsets A and B of X. Thus to check that a subset B of X is μ^*- measurable, we need only check that

$$\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c) \tag{9.1}$$

holds for each subset A of X. Note also that inequality 9.1 certainly holds if $\mu^*(A) = +\infty$. Thus the μ^* -measurability of B can be verified by checking that 9.1 holds for each A that satisfies $\mu^*(A) < +\infty$.

Proposition 9.0.17 Let X be a set, and let μ^* be an outer measure on X. Then each subset B of X that satisfies $\mu^*(B) = 0$ or that satisfies $\mu^*(B^c) = 0$ is μ^* -measurable.

Proof.

$$\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

However our assumption about B and the monotonicity of μ^* imply that one of the terms on the right-hand side of this inequality vanishes and that the other is at most $\mu^*(A)$; thus the required inequality follows. ∎

It follows that the sets $\varnothing$ and X are measurable for every outer measure on X.

Theorem 9.0.18 Let X be a set, let μ^* be an outer measure on X, and let $\mathscr{M}_{\mu^*}$ be the collection of all μ^*-measurable subsets of X. Then

1. $\mathscr{M}_{\mu^*}$ is a σ -algebra, and
2. the restriction of μ^* to $\mathscr{M}_{\mu^*}$ is a measure on $\mathscr{M}_{\mu^*}$

Theorem 9.0.19 Every Borel subset of $\mathbb{R}$ is Lebesgue measurable.

R The collection $\mathscr{M}_{\lambda^*}$ of Lebesgue measurable sets is a σ -algebra on $\mathbb{R}$ that contains each interval of the form $(-\infty, b]$. However $\mathscr{B}(\mathbb{R})$ is the smallest σ -algebra on $\mathbb{R}$ that contains all these intervals , and so $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{M}_{\lambda^*}$

Definition 9.0.14 We will also use $\mathscr{M}_{\lambda^*}$ to denote the collection of Lebesgue measurable subsets of $\mathbb{R}^d$

Theorem 9.0.20 Every Borel subset of $\mathbb{R}^d$ is Lebesgue measurable.

Definition 9.0.15 The restriction of Lebesgue outer measure on $\mathbb{R}$ (or on $\mathbb{R}^d$) to the collection $\mathscr{M}_{\lambda^*}$ of Lebesgue measurable subsets of $\mathbb{R}$ (or of $\mathbb{R}^d$) is called Lebesgue measure and will be denoted by λ or by λ_d. The restriction of Lebesgue outer measure to $\mathscr{B}(\mathbb{R})$ or to $\mathscr{B}(\mathbb{R}^d)$ is also called **Lebesgue measure**, and it too will be denoted by λ or by λ_d. We can specify which version of Lebesgue measure we intend by referring, for example, to Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ or to Lebesgue measure on $(\mathbb{R}, \mathscr{M}_{\lambda^*})$ We will deal most often with Lebesgue measure on the Borel sets.

Theorem 9.0.21 Two questions arise immediately. Is every subset of $\mathbb{R}$ Lebesgue measurable? Is every Lebesgue measurable set a Borel set? The answer to each of these questions is no

Theorem 9.0.22 Let μ be a finite measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, and let $F_\mu : \mathbb{R} \to \mathbb{R}$ be defined by $F_\mu(x) = \mu((-\infty, x])$. Then F_μ is bounded, nondecreasing, and right continuous, and satisfies $\lim_{x\to-\infty} F_\mu(x) = 0$

Theorem 9.0.23 Let μ and F_μ be as in previous theorem. The interval $(a,b]$ is the difference of the intervals $(-\infty, b]$ and $(-\infty, a]$, and so

$$\mu((a,b]) = F_\mu(b) - F_\mu(a) \tag{9.2}$$

since F_μ is bounded and nondecreasing, the limit of $F_\mu(t)$ as t approaches x from the left exists for each x in $\mathbb{R}$; this limit is equal to $\sup\{F_\mu(t) : t < x\}$ and will be denoted by $F_\mu(x-)$. Now let $\{a_n\}$ be a sequence that increases to the real number b; if we apply Eq 9.2 to each interval $(a_n, b]$ and then we find that

$$\mu(\{b\}) = F_\mu(b) - F_\mu(b-) \tag{9.3}$$

Consequently F_μ is continuous at b if $\mu(\{b\}) = 0$, and is discontinuous there, with a jump of size $\mu(\{b\})$ in its graph, if $\mu(\{b\}) \neq 0$. Thus the measure μ is continuous if and only if the function F_μ is continuous.

Equations 9.2 and 9.3 allow one to use F_μ to recover the measure under μ of certain subsets of $\mathbb{R}$; however, the following proposition allows us to say more, namely that the measure under μ of every Borel subset of $\mathbb{R}$ is in fact determined by F_μ.

Theorem 9.0.24 For each bounded, nondecreasing, and right-continuous function $F : \mathbb{R} \to \mathbb{R}$ that satisfies $\lim_{x\to-\infty} F(x) = 0$, there is a unique finite measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ such that $F(x) = \mu((-\infty, x])$ holds at each x in $\mathbb{R}$.

Theorem 9.0.25 Let A be a Lebesgue measurable subset of $\mathbb{R}^d$. Then

1. $\lambda(A) = \inf\{\lambda(U) : U \text{ is open and } A \subseteq U\}$, and
2. $\lambda(A) = \sup\{\lambda(K) : K \text{ is compact and } K \subseteq A\}$

R Previous theorem can be put more briefly, namely as the assertion that Lebesgue measure is **regular**.

Theorem 9.0.26 Each open subset of $\mathbb{R}^d$ is the union of a countable disjoint collection of half-open cubes, each of which is of the form given in expression

$$\left\{(x_1, \ldots, x_d) : j_i 2^{-k} \leq x_i < (j_i + 1)2^{-k} \text{ for } i = 1, \ldots, d\right\} \tag{9.4}$$

for some integers $j_1, \ldots, j_d$ and some positive integer k

Theorem 9.0.27 Lebesgue measure is the only measure on $\left(\mathbb{R}^d, \mathscr{B}\left(\mathbb{R}^d\right)\right)$ that assigns to each d-dimensional interval, or even to each half-open cube of the form given in expression 9.5 , its volume.

Definition 9.0.16 For each element x and subset A of $\mathbb{R}^d$ we will denote by $A+x$ the subset of $\mathbb{R}^d$ defined by

$$A+x=\left\{y \in \mathbb{R}^d : y=a+x \text{ for some } a \text{ in } A\right\}$$

the set $A+x$ is called the **translate** of A by x. We turn to the invariance of Lebesgue measure under such translations.

Theorem 9.0.28 Lebesgue outer measure on $\mathbb{R}^d$ is translation invariant, in the sense that if $x \in \mathbb{R}^d$ and $A \subseteq \mathbb{R}^d$, then $\lambda^*(A)=\lambda^*(A+x)$. Furthermore, a subset B of $\mathbb{R}^d$ is Lebesgue measurable if and only if $B+x$ is Lebesgue measurable.

Theorem 9.0.29 Proposition 1.4.5. Let μ be a nonzero measure on $\left(\mathbb{R}^d, \mathscr{B}\left(\mathbb{R}^d\right)\right)$ that is finite on the bounded Borel subsets of $\mathbb{R}^d$ and is translation invariant, in the sense that $\mu(A)=\mu(A+x)$ holds for each A in $\mathscr{B}\left(\mathbb{R}^d\right)$ and each x in $\mathbb{R}^d$. Then there is a positive number c such that $\mu(A)=c\lambda(A)$ holds for each A in $\mathscr{B}\left(\mathbb{R}^d\right)$.

Note that for the concept of translation invariance for measures on $\left(\mathbb{R}^d, \mathscr{B}\left(\mathbb{R}^d\right)\right)$ to make sense, the Borel σ -algebra on $\mathbb{R}^d$ must be translation invariant, in the sense that if $A \in \mathscr{B}\left(\mathbb{R}^d\right)$ and $x \in \mathbb{R}^d$, then $A+x \in \mathscr{B}\left(\mathbb{R}^d\right)$. To check this translation invariance of $\mathscr{B}\left(\mathbb{R}^d\right)$, note that $\left\{A \subseteq \mathbb{R}^d : A+x \in \mathscr{B}\left(\mathbb{R}^d\right)\right\}$ is a σ -algebra that contains the open sets and hence includes $\mathscr{B}\left(\mathbb{R}^d\right)$

■ **Example 9.4 — (The Cantor Set).** Let K_0 be the interval $[0,1]$. Form K_1 by removing from K_0 the interval $(1/3,2/3)$. Thus $K_1=[0,1/3] \cup [2/3,1]$. Continue this procedure, forming K_n by removing from K_{n-1} the open middle third of each of the intervals making up K_{n-1}. Thus K_n is the union of 2^n disjoint closed intervals, each of length $(1/3)^n$ The Cantor set (which we will temporarily denote by K) is the set of points that remain; thus $K=\cap_n K_n$.

Of course K is closed and bounded. Furthermore, K has no interior points, since an open interval included in K would for each n be included in one of the intervals making up K_n and so would have length at most $(1/3)^n$. The cardinality of K is that of the continuum: it is easy to check that the map that assigns to a sequence $\{z_n\}$ of 0 's and 1 's the number $\sum_{n=1}^{\infty} 2z_n/3^n$ is a bijection of the set of all such sequences onto K; hence the cardinality of K is that of the set of all sequences of 0 's and 1 's and so that of the continuum .

Theorem 9.0.30 The Cantor set is a compact set that has the cardinality of the continuum but has Lebesgue measure zero.

Proof. We have already noted that the Cantor set (again call it K) is compact and has the cardinality of the continuum. To compute the measure of K, note that for each n it is included in the set K_n constructed above and that $\lambda(K_n) = (2/3)^n$. Thus $\lambda(K) \leq (2/3)^n$ holds for each n, and so $\lambda(K)$ must be zero. (For an alternative proof, check that the sum of the measures of the intervals removed from [0,1] during the construction of K is the sum of the geometric series

$$\frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} + \left(\frac{2}{3}\right)^2 \cdot \frac{1}{3} + \left(\frac{2}{3}\right)^3 \cdot \frac{1}{3} + \dots$$

and so is 1.) ■

Theorem 9.0.31 — (A Nonmeasurable Set). We now prove that there is a subset of $\mathbb{R}$ that is not Lebesgue measurable. There is a subset of $\mathbb{R}$, and in fact of the interval $(0,1)$, that is not Lebesgue measurable.

Proof. Define a relation $\sim$ on $\mathbb{R}$ by letting $x \sim y$ hold if and only if $x - y$ is rational. It is easy to check that $\sim$ is an equivalence relation: it is reflexive ($x \sim x$ holds for each x), symmetric $(x \sim y$ implies $y \sim x)$, and transitive $(x \sim y$ and $y \sim z$ imply $x \sim z)$ Note that each equivalence class under $\sim$ has the form $\mathbb{Q} + x$ for some x and so is dense in $\mathbb{R}$. since these equivalence classes are disjoint, and since each intersects the interval $(0,1)$, we can use the axiom of choice to form a subset E of (0,1) that contains exactly one element from each equivalence class. We will prove that the set E is not Lebesgue measurable.

Let $\{r_n\}$ be an enumeration of the rational numbers in the interval $(-1,1)$, and for each n let $E_n = E + r_n$. We will check that

1. the sets E_n are disjoint,
2. $\cup_n E_n$ is included in the interval $(-1,2)$, and
3. the interval (0,1) is included in $\cup_n E_n$

To check (a), note that if $E_m \cap E_n \neq \varnothing$, then there are elements e and e' of E such that $e + r_m = e' + r_n$; it follows that $e \sim e'$ and hence that $e = e'$ and $m = n$. Thus (a) is proved. Assertion (b) follows from the inclusion $E \subseteq (0,1)$ and the fact that each term of the sequence $\{r_n\}$ belongs to $(-1,1)$. Now consider assertion (c). Let x bean arbitrary member of $(0,1)$, and let e be the member of E that satisfies $x \sim e$. Then $x - e$ is rational and belongs to $(-1,1)$ (recall that both x and e belong to $(0,1)$) and so has the form r_n for some n. Hence $x \in E_n$, and assertion (c) is proved.

Suppose that the set E is Lebesgue measurable. Then for each n the set E_n is measurable (Proposition 1.4 .4), and so property (a) above implies that

$$\lambda\left(\cup_n E_n\right)=\sum_n \lambda\left(E_n\right)$$

furthermore, the translation invariance of λ implies that $\lambda\left(E_n\right)=\lambda(E)$ holds for each n. Hence if $\lambda(E)=0$, then $\lambda\left(\cup_n E_n\right)=0$, contradicting assertion (c) above, while if $\lambda(E)\neq 0$, then $\lambda\left(\cup_n E_n\right)=+\infty$, contradicting assertion (b). Thus the assumption that E is measurable leads to a contradiction, and the proof is complete. ■

Definition 9.0.17 Let A be a subset of $\mathbb{R}$. Then diff (A) is the subset of $\mathbb{R}$ defined by

$$\operatorname{diff}(A)=\{x-y: x\in A \text{ and } y\in A\}$$

Theorem 9.0.32 Let A be a Lebesgue measurable subset of $\mathbb{R}$ such that $\lambda(A)>0$. Then diff(A) includes an open interval that contains 0

Theorem 9.0.33 There is a subset A of $\mathbb{R}$ such that each Lebesgue measurable set that is included in A or in A^c has Lebesgue measure zero.

Definition 9.0.18 — complet. Let $(X,\mathscr{A},\mu)$ be a measure space. The measure $\mu($ or the measure space $(X,\mathscr{A}$ is **complete** if the relations $A\in\mathscr{A},\mu(A)=0$, and $B\subseteq A$ together imply that $B\in\mathscr{A}$. It is sometimes convenient to call a subset B of X μ -negligible (or μ -null) if there is a subset A of X such that $A\in\mathscr{A},B\subseteq A$, and $\mu(A)=0$. Thus the measure μ is complete if and only if every μ -negligible subset of X belongs to $\mathscr{A}$

Definition 9.0.19 If μ^* is an outer measure on the set X and if $\mathscr{M}_{\mu^*}$ is the σ -algebra of all μ^* -measurable subsets of X, then the restriction of μ^* to $\mathscr{M}_{\mu^*}$ is **complete**. In particular, Lebesgue measure on the σ -algebra ofLebesgue measurable subsets of $\mathbb{R}^d$ is complete. On the other hand, as we will soon see, the restriction of Lebesgue measure to the σ -algebra of Borel subsets of $\mathbb{R}$ is not complete.

Definition 9.0.20 Let $(X,\mathscr{A})$ be a measurable space, and let μ be a measure on $\mathscr{A}$. The completion of $\mathscr{A}$ under μ is the collection $\mathscr{A}_\mu$ of subsets A of X for which there are sets E and F in $\mathscr{A}$ such that

$$E\subseteq A\subseteq F$$

and

$$\mu(F-E)=0$$

A set that belongs to $\mathscr{A}_\mu$ is sometimes said to be μ -measurable.

Definition 9.0.21 Suppose that A, E, and F are as in the preceding paragraph. It follows immediately that $\mu(E) = \mu(F)$. Furthermore, if B is a subset of A that belongs to $\mathscr{A}$ then

$$\mu(B) \leq \mu(F) = \mu(E)$$

Hence

$$\mu(E) = \sup\{\mu(B) : B \in \mathscr{A} \text{ and } B \subseteq A\}$$

and so the common value of $\mu(E)$ and $\mu(F)$ depends only on the set A (and the measure μ), and not on the choice of sets E and F satisfying (1) and (2). Thus we can define a function $\bar{\mu} : \mathscr{A}_\mu \rightarrow [0, +\infty]$ by letting $\bar{\mu}(A)$ be the common value of $\mu(E)$ and $\mu(F)$, where E and F belong to $\mathscr{A}$ and satisfy (1) and (2). This function $\bar{\mu}$ is called the completion of μ

Definition 9.0.22 Let$(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $\mathscr{A}$. Then $\mathscr{A}_\mu$ is a σ -algebra on X that includes $\mathscr{A}$, and $\bar{\mu}$ is a measure on $\mathscr{A}_\mu$ that is complete and whose restriction to $\mathscr{A}$ is μ

Proposition 1.5.2. Lebesgue measure on ($\boxtimes$ $^{d}, \mathscr{M}_{\lambda^*}$) is the completion of Lebesgue measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ We begin with the following lemma. Lemma 1.5.3. Let A be a Lebesgue measurable subset of $\mathbb{R}^d$. Then there exist Borel subsets E and F of $\mathbb{R}^d$ such that $E \subseteq A \subseteq F$ and $\lambda(F - E) = 0$

there are Lebesgue measurable subsets of $\mathbb{R}$ that are not Borel sets, and the restriction of Lebesgue measure to $\mathscr{B}(\mathbb{R})$ is not complete.

Let $(X, \mathscr{A})$ be a measurable space, let μ be a measure on $\mathscr{A}$, and let A be an arbitrary subset of X. Then $\mu^*(A)$, the outer measure of A, is defined by

$$\mu^*(A) = \inf\{\mu(B) : A \subseteq B \text{ and } B \in \mathscr{A}\} \tag{9.5}$$

and $\mu_*(A)$, the inner measure of A, is defined by

$$\mu_*(A) = \sup\{\mu(B) : B \subseteq A \text{ and } B \in \mathscr{A}\}$$

It is easy to check that $\mu_*(A) \leq \mu^*(A)$ holds for each subset A of X

Theorem 9.0.34 Let$(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $(X, \mathscr{A})$. Then the function $\mu^* : \mathscr{P}(X) \rightarrow [0, +\infty]$ defined by Eq.9.5 is an outer measure (as defined in Sect.1.3) on X.

Theorem 9.0.35 Let$(X,\mathscr{A})$ be a measurable space, let μ be a measure on $\mathscr{A}$ and let A be a subset of X such that $\mu^*(A)<+\infty$. Then A belongs to $\mathscr{A}_\mu$ if and only if $\mu_*(A)=\mu^*(A)$

Definition 9.0.23 Let $\mathscr{A}$ be a σ -algebra on $\mathbb{R}^d$ that includes the σ -algebra $\mathscr{B}\left(\mathbb{R}^d\right)$ of Borel sets. A measure μ on $\left(\mathbb{R}^d,\mathscr{A}\right)$ is regular if (a) each compact subset K of $\mathbb{R}^d$ satisfies $\mu(K)<+\infty$ (b) each set A in $\mathscr{A}$ satisfies $\mu(A)=\inf\{\mu(U): U \text{ is open and } A\subseteq U\}$, and (c) each open subset U of $\mathbb{R}^d$ satisfies $\mu(U)=\sup\{\mu(K): K \text{ is compact and } K\subseteq U\}$

Proposition 1.5.6. Let μ be a finite measure on $\left(\mathbb{R}^d,\mathscr{B}\left(\mathbb{R}^d\right)\right)$. Then μ is regular. Moreover, each Borel subset A of $\mathbb{R}^d$ satisfies

$$\mu(A)=\sup\{\mu(K): K\subseteq A \text{ and } K \text{ is compact }\}$$

Let us first prove the following weakened form of Proposition 1.5 .6 Lemma 1.5.7. Let μ be a finite measure on $\left(\mathbb{R}^d,\mathscr{B}\left(\mathbb{R}^d\right)\right)$. Then each Borel subset A of $\mathbb{R}^d$ satisfies $\mu(A)=\inf\{\mu(U): A\subseteq U \text{ and } U \text{ is open }\}$ and

$$\mu(A)=\sup\{\mu(C): C\subseteq A \text{ and } C \text{ is closed }\}$$

Proposition 2.1.1.

Theorem 9.0.36 Let $(X,\mathscr{A})$ be a measurable space, and let A be a subset of X that belongs to $\mathscr{A}$. For a function $f: A\to[-\infty,+\infty]$ the conditions (a) for each real numbert the set $\{x\in A: f(x)\leq t\}$ belongs to $\mathscr{A}$ (b) for each real number t the set $\{x\in A: f(x)<t\}$ belongs to $\mathscr{A}$ (c) for each real number the set $\{x\in A: f(x)\geq t\}$ belongs to $\mathscr{A}$, and (d) for each real number the set $\{x\in A: f(x)>t\}$ belongs to $\mathscr{A}$ are equivalent.

Definition 9.0.24 Let $(X,\mathscr{A})$ be a measurable space, and let A be a subset of X that belongs to $\mathscr{A}$ A function $f: A\to[-\infty,+\infty]$ is measurable with respect to $\mathscr{A}$ if it satisfies one, and hence all, of the conditions of 9.0.36 . A function that is measurable with respect to $\mathscr{A}$ is sometimes called $\mathscr{A}$ -measurable or, if the σ -algebra $\mathscr{A}$ is clear from context, simply measurable. In case $X=\mathbb{R}^d$, a function that is measurable with respect to $\mathscr{B}\left(\mathbb{R}^d\right)$ is called Borel measurable or a Borel function, and a function that is measurable with respect to $\mathscr{M}_{\lambda^*}$ is called Lebesgue measurable (recall that $\mathscr{M}_{\lambda^*}$ is the σ -algebra of Lebesgue measurable subsets of $\mathbb{R}^d$). Of course every Borel measurable function on $\mathbb{R}^d$ is Lebesgue measurable.

■ **Example 9.5** 1. (c) Let $(X,\mathscr{A})$ be a measurable space, and let B be a subset of X. Then χ_B, the characteristic function of B, is $\mathscr{A}$ -measurable if and only if $B\in\mathscr{A}$. (d) A function is called simple if it has only finitely many values. Let $(X,\mathscr{A})$ be a measurable space, let $f: X\to[-\infty,+\infty]$ be simple, and let $\alpha_1,\ldots,\alpha_n$ be the values of f. Then f is $\mathscr{A}$ -measurable if and only if $\{x\in X: f(x)=\alpha_i\}\in\mathscr{A}$ for $i=1,\ldots,n$

Theorem 9.0.37 Proposition 2.1.3. Let$(X,\mathscr{A})$ be a measurable space, let A be a subset of X that belongs to $\mathscr{A}$, and let f and g be $[-\infty,+\infty]$ -valued measurable functions on A. Then the sets $\{x \in A : f(x) < g(x)\}, \{x \in A : f(x) \leq g(x)\}$, and $\{x \in A : f(x) = g(x)\}$ belong to $\mathscr{A}$

Definition 9.0.25 Let f and g be $[-\infty,+\infty]$ -valued functions having a common domain A. The maximum and minimum of f and g, written $f \vee g$ and $f \wedge g$, are the functions from A to $[-\infty,+\infty]$ defined by

$$(f \vee g)(x) = \max(f(x), g(x))$$

and

$$(f \wedge g)(x) = \min(f(x), g(x))$$

Equivalently, we can define $f \vee g$ by

$$(f \vee g)(x) = \begin{cases} f(x) & \text{if } f(x) > g(x) \text{ and} \\ g(x) & \text{otherwise} \end{cases}$$

with $f \wedge g$ getting a corresponding definition. If $\{f_n\}$ is a sequence of $[-\infty,+\infty]$ -valued functions on A, then $\sup_n f_n : A \to [-\infty,+\infty]$ is defined by

$$\left(\sup_n f_n\right)(x) = \sup\{f_n(x) : n = 1,2,\ldots\}$$

and $\inf_n f_n, \limsup_n f_n, \liminf_n f_n$, and $\lim_n f_n$ are defined in analogous ways. The domain of $\lim_n f_n$ consists of those points in A at which $\limsup_n f_n$ and $\liminf_n f_n$ agree; the domain of each of the other four functions is A. Each of these functions can have infinite values, even if all the f_n 's have only finite values; in particular, $\lim_n f_n(x)$ can be $+\infty$ or $-\infty$

Theorem 9.0.38 Proposition 2.1.4. Let$(X,\mathscr{A})$ be a measurable space, let A be a subset of X that belongs to $\mathscr{A}$, and let f and g be $[-\infty,+\infty]$ -valued measurable functions on A. Then $f \vee g$ and $f \wedge g$ are measurable.

Theorem 9.0.39 Proposition 2.1.5. Let$(X,\mathscr{A})$ be a measurable space, let A be a subset of X that belongs to $\mathscr{A}$, and let $\{f_n\}$ be a sequence of $[-\infty,+\infty]$ -valued measurable functions on A. Then (a) the functions sup $_n f_n$ and inf $_n f_n$ are measurable, (b) the functions limsup f_n and $\liminf_n f_n$ are measurable, and (c) the function $\lim_n f_n$ (whose domain is $\{x \in A : \limsup_n f_n$ is measurable.

Theorem 9.0.40 Proposition 2.1.6. Let$(X,\mathscr{A})$ be a measurable space, let A be a subset of X that belongs to $\mathscr{A}$, let f and g be $[0,+\infty]$ -valued measurable functions on A, and let α be a nonnegative real number. Then α fand $f+g$ are measurable.

Theorem 9.0.41 Proposition 2.1.7. Let$(X,\mathscr{A})$ be a measurable space, let A be a subset of X that belongs to $\mathscr{A}$, let f and g be measurable real-valued functions on A, and let α be a real number. Then $\alpha f, f+g, f-g$, fg, and f/g (where the domain of f/g is $\{x \in A : g(x) \neq 0\}$) are measurable.

Theorem 9.0.42 Let A be a set, and let f be an extended real-valued function [2] on A. The positive part f^+ and the negative part f^- of f are the extended real-valued functions defined by

$$f^+(x) = \max(f(x), 0)$$

and

$$f^-(x) = -\min(f(x), 0)$$

Thus $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$. It is easy to check that if $(X,\mathscr{A})$ is a measurable space and if f is a $[-\infty,+\infty]$ -valued function defined on a subset of X, then f is measurable if and only if f^+ and f^- are both measurable. It follows from this remark, together with Proposition 2.1.6, that the absolute value $|f|$ of a measurable function f is measurable (note that $|f| = f^+ + f^-$)

Theorem 9.0.43 (a) If f is $\mathscr{A}$ -measurable and if B is a subset of A that belongs to $\mathscr{A}$, then the restriction f_B of f to B is $\mathscr{A}$ -measurable; this follows from the identity

$$\{x \in B : f_B(x) < t\} = B \cap \{x \in A : f(x) < t\}$$

(b) If $\{B_n\}$ is a sequence of sets that belong to $\mathscr{A}$, if $A = \cup_n B_n$, and if for each n the restriction f_{B_n} of f to B_n is $\mathscr{A}$ -measurable, then f is $\mathscr{A}$ -measurable; this follows from the identity

$$\{x \in A : f(x) < t\} = \bigcup_n \{x \in B_n : f_{B_n}(x) < t\}$$

Theorem 9.0.44 Proposition 2.1.8. Let$(X,\mathscr{A})$ be a measurable space, let A be a subset of X that belongs to $\mathscr{A}$, and let f be $a[0,+\infty]$ -valued measurable function on A. Then there is a sequence $\{f_n\}$ of simple $[0,+\infty)$ -valued measurable functions on A that satisfy

$$f_1(x) \leq f_2(x) \leq \dots$$

and

$$f(x) = \lim_n f_n(x)$$

at each x in A

Theorem 9.0.45 Proposition 2.1.9. Let$(X, \mathscr{A})$ be a measurable space, and let A be a subset of X that belongs to $\mathscr{A}$. For a function $f : A \to \mathbb{R}$, the conditions (a) f is measurable with respect to $\mathscr{A}$ (b) for each open subset U of $\mathbb{R}$ the set $f^{-1}(U)$ belongs to $\mathscr{A}$ (c) for each closed subset C of $\mathbb{R}$ the set $f^{-1}(C)$ belongs to $\mathscr{A}$, and (d) for each Borel subset B of $\mathbb{R}$ the set $f^{-1}(B)$ belongs to $\mathscr{A}$ are equivalent.

■ **Example 9.6** Recall the construction of the Cantor set given in Sect. 1.4. There we let K_0 be the interval $[0,1]$, and for each positive integer n we constructed a compact set K_n by removing from K_{n-1} the open middle third of each of the intervals making up K_{n-1}. The Cantor set K is given by $K = \cap_n K_n$

The Cantor function (also known as the Cantor singular function) is the function $f : [0,1] \to [0,1]$ defined as follows (the concept of singularity will be defined and studied in Chap. 4). For each x in the interval $(1/3,2/3)$ let $f(x) = 1/2$. Thus f is now defined at each point removed from [0,1] in the construction of K_1. Next define f at each point removed from K_1 in the construction of K_2 by letting $f(x) = 1/4$ if $x \in (1/9,2/9)$ and letting $f(x) = 3/4$ if $x \in (7/9,8/9)$. Continue in this way, letting $f(x)$ be $1/2^n, 3/2^n, 5/2^n, \dots$ on the various intervals removed from K_{n-1} in the construction of K_n. After all these steps, f is defined on the open set $[0,1] - K$, is nondecreasing, and has values in $[0,1]$. Extend it to all of [0,1] by letting $f(0) = 0$ and letting

$$f(x) = \sup\{f(t) : t \in [0,1] - K \text{ and } t < x\}$$

if $x \in K$ and $x \neq 0$. This completes the definition of the Cantor function. It is easy to check that f is nondecreasing and continuous, and it is clear that $f(0) = 0$ and $f(1) = 1$. The intermediate value theorem (Theorem C.13) thus implies that for each y in [0,1] there is at least one x in [0,1] such that $f(x) = y$ and so we can define a function $g : [0,1] \to [0,1]$ by

$$g(y) = \inf\{x \in [0,1] : f(x) = y\}$$

The continuity of f implies that $f(g(y)) = y$ holds for each y in [0,1]; hence g is injective. It is easy to check that all the values of g lie in the Cantor set. The fact that f is nondecreasing implies that g is nondecreasing and hence that g is Borel measurable (see Example 2.1.2(b))

Theorem 9.0.46 There is a Lebesgue measurable subset of $\mathbb{R}$ that is not Borel set

Proof. Let g be the function constructed above, let A be a subset of [0,1] that is not Lebesgue measurable (see Theorem 1.4 .9), and let $B = g(A)$. Then B is a subset of the Cantor set and so is Lebesgue measurable (recall that $\lambda(K) = 0$ and that Lebesgue measure on the σ -algebra of Lebesgue measurable sets is complete). If B were a Borel set, then $g^{-1}(B)$ would also be a Borel set (recall that g is Borel measurable, and see Proposition 2.1 .9). However the injectivity of g implies that $g^{-1}(B)$ is the set A, which is not Lebesgue measurable and hence is not a Borel set. Consequently the Lebesgue measurable set B is not a Borel set. ■

■ **Example 9.7** The proof of 9.0.46 gives a Borel set of Lebesgue measure 0 (the Cantor set) that has a subset that is not a Borel set. It follows that Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is not complete.

Definition 9.0.26 — μ -almost everywhere. Let $(X, \mathscr{A}, \mu)$ be a measure space. A property of points of X is said to hold μ -almost everywhere if the set of points in X at which it fails to hold is μ -negligible. In other words, a property holds μ -almost everywhere if there is a set N that belongs to $\mathscr{A}$ satisfies $\mu(N) = 0$, and contains every point at which the property fails to hold. More generally, if E is a subset of X, then a property is said to hold μ -almost everywhere on E if the set of points in E at which it fails to hold is μ -negligible. The expression μ -almost everywhere is often abbreviated to μ -a.e. or to a.e. $[\mu]$.

Theorem 9.0.47 Let$(X, \mathscr{A}, \mu)$ be a measure space, and let f and g be extended real-valued functions on X that are equal almost everywhere. If μ is complete and if f is $\mathscr{A}$ -measurable, then g is $\mathscr{A}$ -measurable.

Theorem 9.0.48 — 2.2.3. Let$(X, \mathscr{A}, \mu)$ be a measure space, let $\{f_n\}$ be a sequence of extended real-valued functions on X, and let f be an extended real-valued function on X such that $\{f_n\}$ converges to f almost everywhere. If μ is complete and if each f_n is $\mathscr{A}$ -measurable, then f is $\mathscr{A}$ -measurable.

Theorem 9.0.49 Proposition 2.2.5. Let$(X, \mathscr{A}, \mu)$ be a measure space, and let $\mathscr{A}_\mu$ be the completion of $\mathscr{A}$ under μ. Then a function $f : X \to [-\infty, +\infty]$ is $\mathscr{A}_\mu$ -measurable if and only if there are $\mathscr{A}$ -measurable functions $f_0, f_1 : X \to [-\infty, +\infty]$ such that

$$f_0 \leq f \leq f_1 \text{ holds everywhere on } X$$

and $f_0 = f_1$ holds μ -almost everywhere on X In the context of Proposition 2.2.5, it is natural to ask whether it is always possible, given an $\mathscr{A}_\mu$ -measurable function f with values in $\mathbb{R}$, rather than in $[-\infty, +\infty]$, to find real -valued functions f_0 and f_1 that satisfy (2) and (3). It turns out that the answer is no; see Exercise 8.3 .3

Definition 9.0.27 We begin with the simple functions. Let $(X, \mathscr{A})$ be a measurable space. We will denote by $\mathscr{S}$ the collection of all simple real-valued $\mathscr{A}$ -measurable functions on X and by $\mathscr{S}_+$ the collection of nonnegative functions in $\mathscr{S}$

Definition 9.0.28 Let μ be a measure on $(X, \mathscr{A})$. If f belongs to $\mathscr{S}_+$ and is given by $f = \sum_{i=1}^m a_i \chi_{A_i}$ where $a_1, \ldots, a_m$ are nonnegative real numbers and $A_1, \ldots, A_m$ are disjoint subsets of X that belong to $\mathscr{A}$, then $\int f d\mu$, the integral of f with respect to μ, is defined to be $\sum_{i=1}^m a_i \mu(A_i)$ (note that this sum is either a nonnegative real number or $+\infty$).

Theorem 9.0.50 Let $(X, \mathscr{A}, \mu)$ be a measure space, let fand g belong to $\mathscr{S}_+$ and let α be a nonnegative real number. Then

1. $\int \alpha f d\mu = \alpha \int f d\mu$
2. $\int (f+g) d\mu = \int f d\mu + \int g d\mu$, and
3. if $f(x) \leq g(x)$ holds at each x in X, then $\int f d\mu \leq \int g d\mu$

Theorem 9.0.51 Let$(X, \mathscr{A}, \mu)$ be a measure space, let f belong to $\mathscr{S}_+$, and let $\{f_n\}$ be a nondecreasing sequence of functions in $\mathscr{S}_+$ such that $f(x) = \lim_n f_n(x)$ holds at each x in X. Then $\int f d\mu = \lim_n \int f_n d\mu$

We define the integral of an arbitrary $[0, +\infty]$ -valued $\mathscr{A}$ measurable function on X. For such a function f, let

$$\int f d\mu = \sup \left\{ \int g d\mu : g \in \mathscr{S}_+ \text{and } g \leq f \right\}$$

It is easy to see that for functions f in $\mathscr{S}_+$, this agrees with the previous definition.

Theorem 9.0.52 Let$(X, \mathscr{A}, \mu)$ be a measure space, let f be a $[0, +\infty]$ -valued $\mathscr{A}$ measurable function on X, and let $\{f_n\}$ be a nondecreasing sequence of functions in $\mathscr{S}_+$such that $f(x) = \lim_n f_n(x)$ holds at each x in X. Then $\int f d\mu = \lim_n \int f_n d\mu$

Theorem 9.0.53 Proposition 2.3.4. Let$(X, \mathscr{A}, \mu)$ be a measure space, let fand g be $[0, +\infty]$ -valued $\mathscr{A}$ -measurable functions on X, and let α be a nonnegative real number. Then

1. $\int \alpha f d\mu = \alpha \int f d\mu$
2. $\int (f+g) d\mu = \int f d\mu + \int g d\mu$, and
3. if $f(x) \leq g(x)$ holds at each x in X, then $\int f d\mu \leq \int g d\mu$

Definition 9.0.29 Finally, let f be an arbitrary $[-\infty, +\infty]$ -valued $\mathscr{A}$ -measurable function on X. If $\int f^+ d\mu$ and $\int f^- d\mu$ are both finite, then f is called integrable (or μ -integrable or summable), and its integral $\int f d\mu$ is defined by

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu$$

The integral of f is said to exist if at least one of $\int f^+ d\mu$ and $\int f^- d\mu$ is finite, and again in this case, $\int f d\mu$ is defined to be $\int f^+ d\mu - \int f^- d\mu$. In either case one sometimes writes $\int f(x) \mu(dx)$ or $\int f(x) d\mu(x)$ in place of $\int f d\mu$

Definition 9.0.30 In case $X = \mathbb{R}^d$ and $\mu = \lambda$, one often refers to Lebesgue integrability and the Lebesgue integral. The Lebesgue integral of a function f on $\mathbb{R}$ is often written $\int f(x) dx$. In case we are integrating over the interval $[a, b]$, we may write $\int_a^b f$ or $\int_a^b f(x) dx$ or, if we need to emphasize that we mean the Lebesgue integral, $(L) \int_a^b f$ or$(L) \int_a^b f(x) dx$

Definition 9.0.31 We define $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$ (or sometimes simply $\mathscr{L}^1$) to be the set of all realvalued (rather than $[-\infty, +\infty]$ -valued) integrable functions on X. According to Proposition 2.3 .6 below, $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$ is a vector space and the integral is a linear functional on $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$

Theorem 9.0.54 Let$(X, \mathscr{A}, \mu)$ be a measure space, and let f_1, f_2, g_1, and g_2 be nonnegative real-valued integrable functions on X such that $f_1 - f_2 = g_1 - g_2$. Then $\int f_1 d\mu - \int f_2 d\mu = \int g_1 d\mu - \int g_2 d\mu$

Theorem 9.0.55 Proposition 2.3.6. Let$(X, \mathscr{A}, \mu)$ be a measure space, let f and g be real-valued integrable functions on X, and let α be a real number. Then (a) αf and $f+g$ are integrable, (b) $\int \alpha f d\mu = \alpha \int f d\mu$ (c) $\int (f+g) d\mu = \int f d\mu + \int g d\mu$, and (d) if $f(x) \leq g(x)$ holds at each x in X, then $\int f d\mu \leq \int g d\mu$

■ **Example 9.8** 1. (a) If μ is a finite measure, then every bounded measurable function on $(X, \mathscr{A}, \mu)$ is integrable. (b) In particular, every bounded Borel function, and hence every continuous function, on $[a, b]$ is Lebesgue integrable. (We'll see in Sect.2.5 that the Lebesgue integral of a continuous function on $[a, b]$ can be found by calculating its Riemann integral.) (c) Suppose that $\mathscr{A}$ is the σ -algebra on $\mathbb{N}$ containing all subsets of $\mathbb{N}$ and that μ is counting measure on $\mathscr{A}$. It follows from

Proposition 2.3 .3 that a nonnegative function f on $\mathbb{N}$ is μ -integrable if and only if the infinite series $\sum_n f(n)$ is convergent, and that in that case the integral and the sum of the series agree. since a not necessarily nonnegative function f is integrable if and only if f^+ and f^- are integrable, it follows that f is integrable if and only if the infinite series $\sum_n f(n)$ is absolutely convergent. Once again, the integral and the sum of the series have the same value.

Theorem 9.0.56 Proposition 2.3.8. Let$(X, \mathscr{A}, \mu)$ be a measure space, and let f be a $[-\infty, +\infty]$ valued $\mathscr{A}$ -measurable function on X. Then f is integrable if and only if $|f|$ is integrable. If these functions are integrable, then $|\int f d\mu| \leq \int |f| d\mu$

Theorem 9.0.57 Proposition 2.3.9. Let$(X, \mathscr{A}, \mu)$ be a measure space, and let f and g be $[-\infty, +\infty]$ valued $\mathscr{A}$ -measurable functions on X that agree almost everywhere. If either $\int f d\mu$ or $\int g d\mu$ exists, then both exist, and $\int f d\mu = \int g d\mu$

Theorem 9.0.58 Proposition 2.3.10. Let$(X, \mathscr{A}, \mu)$ be a measure space, and let f be a $[0, +\infty]$ valued $\mathscr{A}$ -measurable function on X. If t is a positive real number and if A_t is defined by $A_t = \{x \in X : f(x) \geq t\}$, then

$$\mu(A_t) \leq \frac{1}{t} \int_{A_t} f d\mu \leq \frac{1}{t} \int f d\mu$$

Theorem 9.0.59 Corollary 2.3.11. Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be $a[-\infty, +\infty]$ valued integrable function on X. Then $\{x \in X : f(x) \neq 0\}$ is σ-finite under μ

Theorem 9.0.60 Corollary 2.3.12. Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be a $[-\infty, +\infty]$ valued $\mathscr{A}$ -measurable function on X that satisfies $\int |f| d\mu = 0$. Then f vanishes μ -almost everywhere.

Theorem 9.0.61 Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be $a[-\infty, +\infty]$ valued integrable function on X such that $\int_A f d\mu \geq 0$ holds for all A in $\mathscr{A}$ (or even just for all A in the smallest σ-algebra on X that makes f measurable). Then $f \geq 0$ holds μ -almost everywhere.

Theorem 9.0.62 Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be $a[-\infty, +\infty]$ valued integrable function on X. Then $|f(x)| < +\infty$ holds at μ -almost every x in X

Theorem 9.0.63 Corollary 2.3.15. Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be $a[-\infty, +\infty]$ valued $\mathscr{A}$ -measurable function on X. Then f is integrable if and only if there is a function in $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$ that is equal to f almost everywhere.

Theorem 9.0.64 Theorem 2.4.1 (The Monotone Convergence Theorem). Let $(X, \mathscr{A}, \mu)$ be a measure space, and let fand $f_1, f_2, \ldots$ be $[0, +\infty]$ -valued $\mathscr{A}$ -measurable functions on X. Suppose that

$$f_1(x) \leq f_2(x) \leq \ldots$$

and

$$f(x) = \lim_n f_n(x)$$

hold at μ -almost every x in X. Then $\int f d\mu = \lim_n \int f_n d\mu$

In this theorem the functions f and $f_1, f_2, \ldots$ are only assumed to be nonnegative and measurable; there are no assumptions about whether they are integrable.

Theorem 9.0.65 Corollary 2.4 .2 (Beppo Levi's Theorem). Let$(X, \mathscr{A}, \mu)$ be a measure space, and let $\sum_{k=1}^{\infty} f_k$ be an infinite series whose terms are $[0, +\infty]$ -valued $\mathscr{A}$ -measurable functions on X. Then

$$\int \sum_{k=1}^{\infty} f_k d\mu = \sum_{k=1}^{\infty} \int f_k d\mu$$

Theorem 9.0.66 Theorem 2.4.4 (Fatou's Lemma). Let $(X, \mathscr{A}, \mu)$ be a measure space, and let $\{f_n\}$ be a sequence of $[0, +\infty]$ -valued $\mathscr{A}$ -measurable functions on X. Then

$$\int \underline{\lim}_n f_n d\mu \leq \underline{\lim}_n \int f_n d\mu$$

Theorem 9.0.67 — Lebesgue's Dominated Convergence Theorem. Let $(X, \mathscr{A}, \mu)$ be a measure space, let g be a $[0, +\infty]$ -valued integrable function on X, and let f and $f_1, f_2, \ldots$ be $[-\infty, +\infty]$ -valued $\mathscr{A}$ -measurable functions on X such that

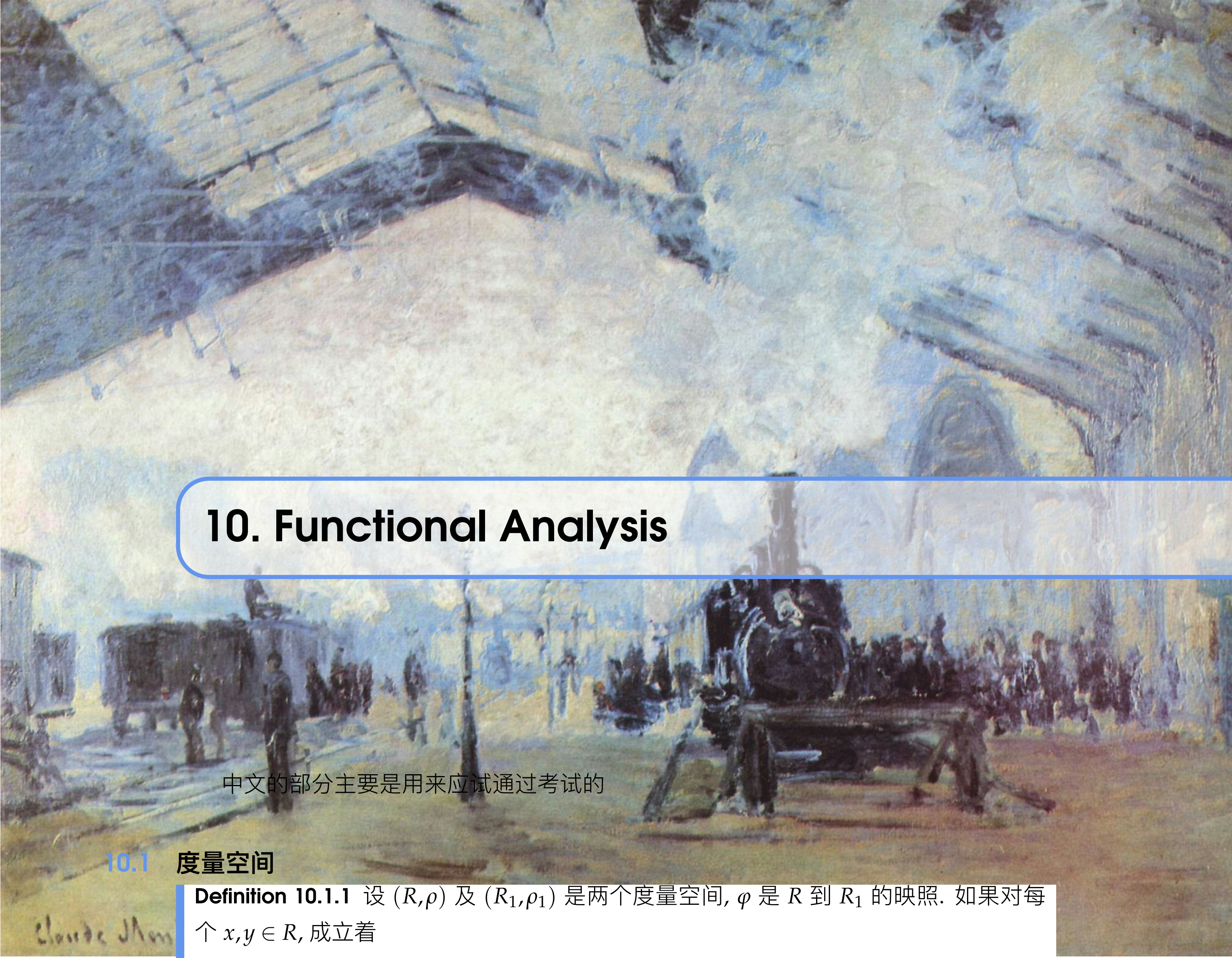
$$f(x) = \lim_n f_n(x)$$

and

$$|f_n(x)| \leq g(x), n = 1, 2, \ldots$$

hold at μ -almost every x in X. Then f and $f_1, f_2, \ldots$ are integrable, and $\int f d\mu =$

$\lim_n \int f_n d\mu$



10. Functional Analysis

中文的部分主要是用来应试通过考试的

10.1 度量空间

Definition 10.1.1 设 (R,ρ) 及 (R_1,ρ_1) 是两个度量空间, φ 是 R 到 R_1 的映照. 如果对每个 $x,y \in R$, 成立着

$$\rho(x,y) = \rho_1(\varphi x, \varphi y)$$

那么称 φ 是 (R,ρ) 到 (R_1,ρ_1) 上的等距映照. 进一步, 如果还有 $\varphi(R) = R_1$, 那称这两个度量空间是等距同构的。

Definition 10.1.2 — 映射连续. 设 f 是度量空间 X 到度量空间 Y 中的映照，那末下面三件事情等价。

1. 设 $x_0 \in X$, 假如对于 $f(x_0)$ 的任何邻域 $O(f(x_0))$, 存在 x_0 的一个邻域 $O(x_0)$, 使得 $f(O(x_0)) \subset O(f(x_0))$, 即当 $x \in O(x_0)$ 时, $f(x) \in O(f(x_0))$, 就称映照 f 在, 点 x_0 是连续的。
2. 映照 f 在, 点 x_0 是连续的.
3. 对于 $f(x_0)$ 的任 $-\varepsilon$ -邻域 $O(f(x_0),\varepsilon)$, 存在 x_0 在 X 中的 δ -邻域 $O(x_0,\delta)$, 使得:

$$f(O(x_0,\delta)) \subset O(f(x_0),\varepsilon)$$

4. 对于 X 中任意一列收敛 x_0 的点列 $\{x_n\}_{n=1}^{\infty}$, 成立着

$$\lim_{n\to\infty} f(x_n) = f(x_0)$$

5. 设 X 和 Y 是度量空间，f 是 X 到 Y 中的映照。那末 f 在 X 上连续的充要条件是象空间 Y 中任一闭集 F 的原象 $f^{-1}(F)$ 是 X 中的闭集.
6. 开集合的原像是开集

Theorem 10.1.1 — 距离函数是连续的. 如果 $x_n \to x_0, y_n \to y_0$, 那么 $\rho(x_n, y_n) \to \rho(x_0, y_0)$ (也就是说, 距离 $\rho(x,y)$ 是两个变元 x,y 的“连续函数"，证明利用三角不等式

Definition 10.1.3 — 一致离散. 对任何非空集 R，可引入如下的距离：

$$\rho_0(x,y) = \begin{cases} 0, & x = y \in R \\ 1, & x \neq y, x, y \in R \end{cases}$$

在一个度量空间 R 中, 如果存在正的常数 α, 使得任何 $x,y \in R, x \neq y$, 都 $y \in R, x \neq y$ 有 $\rho(x,y) \geqslant \alpha$ 时, 称 R 是一致离散的距离空间. 例如对任何非空集 $R, (R, \rho_0)$ 是一致离散的距离空间

■ **Example 10.1** 开球可能只含一点。例如前面提到的一致离散的度量空间，对于不同的两个点，他们的距离为 1，对于任何的小于 1 的开球，其中只含有一个点。$O(x_0, r)$ 中只能含有一点.

■ **Example 10.2** E^n 中按距离收敛就是按每个坐标收敛.

R 证明距离的三角不等式的时候，想到单调性

Theorem 10.1.2 常用不等式 $|a-b|^2 \leqslant 2\left(|a|^2 + |b|^2\right)$

$$|a+b|^p \leqslant (2\max\{|a|,|b|\})^p \leqslant 2^p\left(|a|^p + |b|^p\right)$$

■ **Example 10.3** 空间 $C^{(k)}[a,b]$：设 k 是一个非负整数, $x(t)$ 是区间 $[a,b]$ 上的连续函数, 而且具有连续的 k 阶导函数 (当 $k=0$ 时就表示只要求 $x(t)$ 本身连续)，这种函数 $x(t)$ 的全体记为 $C^{(k)}[a,b]$, 特别简记 $C^0[a,b]$ 为 $C[a,b]$. 对于 $x(t), y(t) \in C^{(k)}[a,b]$, 令

$$\rho_k(x,y) = \max_{0 \leqslant j \leqslant k} \max_{a \leqslant t \leqslant b} \left| x^{(j)}(t) - y^{(j)}(t) \right|$$

容易证明 $\rho_k(x,y)$ 是距离. 在 $C^{(k)}[a,b]$ 中函数列 $\{x_n(t)\}$ 依距离收敛于 $x(t)$ 的充要条件是, $\{x_n(t)\}$ 以及它们的前 k 阶导函数列在 $[a,b]$ 上都分别均匀收敛于 $x(t)$ 及其前 k 阶导函数.

Definition 10.1.4 — 线性基. 设 A 是线性空间 R 中的一个线性无关向量组. 如果对于每一个非寶向量 $x \in R$, 都是 A 中的向量的线性组合，即有不全为奥的 n 个实 (或复)$\alpha_1, \cdots, \alpha_n$, 使得

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n, x_1, x_2, \cdots, x_n \in A$$

就称 A 是线性空间 R 的一组线性基. 线性基又称为 Hamel 基. 用 Zorn 引理可以证明: 任何线性空间总存在 Hamel 基.

Definition 10.1.5 — 线性同构. 设 R', R'' 同是实或复的两个线性空间, 如果存在 R' 到 R'' 上的一一对应 φ, 使得对任何一对 $x, y \in R'$ 及任何数 α, 成立着

$$\varphi(x+y) = \varphi(x) + \varphi(y), \varphi(\alpha x) = \alpha\varphi(x)$$

那么称 R' 和 R'' 是线性同构的, 而映照 φ 称为 R' 到 R'' 的线性同构映照

10.2 赋范空间

Definition 10.2.1 线性空间的定义参考 Algebra 中的定义一致。线性空间中的元素称之为向量。由于线性空间中一定有零元，故定义向量的长度为到零元的距离，引出了距离的定义。

Definition 10.2.2 — 范数. 设 X 是域 $\mathbb{K}$(实数域或复数域) 上的线性空间, 函数 $||\cdot||: X \to R$ 满足条件:

1. 对任意 $x \in X, \|x\| \geqslant 0$; 且 $\|x\| = 0$, 当且仅当 $x = 0$
2. 对任意 $x \in X$ 及 $\alpha \in K, \|\alpha x\| = |\alpha|\|x\|$(齐次性)
3. 对任意 $x, y \in X, \|x+y\| \leqslant \|x\| + \|y\|$ (三角形不等式).

称 $||\cdot||$ 是 X 上的一个范数, X 上定义了范数 $||\cdot||$ 称为赋范 (线性) 空间. 在一个赋范线性空间中, 通过范数可以自然地定义一个距离,

$$d(x,y) = \|x-y\|, \quad x, y \in X \tag{10.1}$$

称赋范空间中这个距离是由范数诱导的距离. 赋范线性空间可以看作是一个集合定义了代数结构加上拓扑结构，代数结构的加法数乘，拓扑结构的距离。由于这两个结构是不同的，所以不能随便定义一个距离函数，需要照顾到加法数乘，故引入范数的概念。特别地, 设 $\{x_n\}$ 是赋范空间 X 中的点列, $x \in X$, 如果

$$\|x_n - x\| \to 0 \quad (n \to \infty)$$

称 $\{x_n\}$ 强 (或按范) 收敛于 x, 记为

$$x_n \to x \quad (n \to \infty)$$

由于存在范数，故有诱导距离，故有可以引入极限

Definition 10.2.3 — 半范数. 如果 R 上的实值函数 $p(\cdot)$ 满足下列条件:

1. $p(x) \geqslant 0, x \in R$
2. $p(\alpha x) = |\alpha| p(x), x \in R, \alpha \in F$
3. $p(x+y) \leqslant p(x) + p(y), x, y \in R$

我们称 $p(x)$ 是 x 的半范数或称为拟范数. 如果半范数 $p(x)$ 又满足如下条件:
如果 $p(x)=0$, 那么 $x=0$ 便称 $p(x)$ 是 x 的范数

Example 10.4 — C a,b

(空间 $C[a,b]$) 设 $C[a,b]$ 是闭区间 $[a,b]$ 上的连续函数全体所成的线性空间. 当 $f\in C[a,b]$ 时, 规定

$$\|f\|=\max_{x\in[a,b]}|f(x)|$$

$C[a,b]$ 按范数 $\|\cdot\|$ 成为赋范线性空间.

Example 10.5 — L a,b

(空间 $L[a,b]$) 设 $L[a,b]$ 是区间 $[a,b]$ 上的 Lebesgue 可积函数全体所成的线性空间. 对于 $f\in L[a,b]$, 令

$$p(f)=\int_a^b|f(t)|dt$$

那么 $p(f)$ 是 $L[a,b]$ 上的半范数, 但不是范数, 因为 $p(f)=0$ 时并不能推出 $f=0$ 而只能得出 $f(t)\doteq 0$. 但是 $p(f)$ 限制在 $L[a,b]$ 的线性子空间 $C[a,b]$ 上时, 它成为范数. 这是因为在 $C[a,b]$ 中当 $f(t)\doteq 0$ 时, $f=0$

Definition 10.2.4 — 按距离收敛. 设 R 是一个度量空间, $x_n(n=1,2,3,\cdots),x\in R$, 假如当 $n\to\infty$ 时数列 $\rho(x_n,x)\to 0$, 就说点列 $\{x_n\}$ 按照距离 $\rho(x,y)$ 收敛于记作

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

或 $x_n\to x$. 这时称 $\{x_n\}$ 为收敛点列, x 为 $\{x_n\}$ 的极限.

Definition 10.2.5 — 依范数收敛. 设 R 是赋范线性空间, $x_n\in R,n=1,2,3,\cdots$ 如果存在 $x\in R$, 使得 x_n 按距离收敛于 x, 即

$$\lim_{n\to\infty}\|x_n-x\|=0$$

那么称 $\{x_n\}$ 依范数收敛于 x, 记作 $\lim_{n\to\infty}x_n=x$ 或 $x_n\to x(n\to\infty)$

Theorem 10.2.1 在依范数收敛意义之下， 只要 $x_n\to x_0$, 就有 $\|x_n\|\to\|x_0\|$, 就是说范数 $||x\|$ 是 x 的“连续函数"。事实上, 取 $y_n=y_0=0$, 那么

$$\|x_n\|=\rho(x_n,0)\to\rho(x_0,0)=\|x_0\|$$

因此, 如果 $\{x_n\}$ 是赋范线性空间中的收敛点列, 那么它们的范数 $\{\|x_n\|\}$ 是有界的.

Theorem 10.2.2 距离是由范数决定的充要条件就是 $\rho(x,y)$ 适合

$$\rho(x,y)=\rho(x-y,0),\rho(\alpha x,0)=|\alpha|\rho(x,0)$$

当距离适合上述条件时, 定义 $\|x\| = \rho(x,0)$, 就成范数. 所以该条件也是线性的度量空间成为赋范线性空间的充要条件.

Definition 10.2.6 — 有界变差. 设 $f(x)$ 是 $[a,b]$ 上的有限函数, 在 $[a,b]$ 上任取一组分点

$$a = x_0 < x_1 < \cdots < x_n = b$$

作和式

$$V_f(x_0,\cdots,x_n) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

称它为 f 对分点组 $x_0,x_1,\cdots,x_n$ 的变差. 如果对一切可能的分点组, 变差所形成的数集 $\{V_f(x_0,\cdots,x_n)\}$ 有界, 即

$$\sup_{x_0,\cdots,x_n} V_f(x_0,\cdots,x_n) < \infty$$

就称 f 是 $[a,b]$ 上的有界变差函数. 记

$$\mathbf{V}_a(f) = \sup_{x_0,\cdots,x_n} V_f(x_0,\cdots,x_n)$$

称 $\mathbf{V}_a^b(f)$ 是 f 在 $[a,b]$ 上的全变差. 当 x 在 $[a,b]$ 上变化时, 称 f 在 $[a,x]$ 上的全变差 $\underset{a}{\mathbf{V}}(f)$ 为 f 在 $[a,b]$ 上的全变差函数.

区间 $[a,b]$ 上有界变差函数全体所成的函数类记为 $V[a,b]$

■ Example 10.6 — V a,b

(空间 $V[a,b]$) 设 $V[a,b]$ 是区间 $[a,b]$ 上的实 (或复) 有界变差函数的全体, 依照通常的线性运算, 它是一个线性空间. 对于 $f \in V[a,b]$, 规定

$$\|f\| = |f(a)| + V_a^b(f)$$

那么 $V[a,b]$ 按范数 $\|f\|$ 成为赋范线性空间. 我们令

$$V_0[a,b] = \left\{ f | f \in V[a,b], \quad \begin{array}{l} f \text{ 在}(a,b)\text{ 中每点是右连续的} \\ \text{而且} f(a) = 0 \end{array} \right\}$$

它是 $V[a,b]$ 的线性子空间. 在 $V_0[a,b]$ 上, 范数 $\|f\|$ 等于全变差 $\mathbf{V}_a^b(f)$.

Definition 10.2.7 — 凸集. 设 X 是线性空间, A 是 X 的子集, 如果对任意 $x,y \in A$, 及满足 $0 < \alpha < 1$ 的数 α

$$\alpha x + (1-\alpha)y \in A$$

称 A 是 X 中的凸集.

R 从定义不难看出, 任意个凸集的交集是凸集. 设 A 是空间 X 中任意子集, 所有包含集 A 的凸集交集是凸集, 称这个凸集是集 A 生成的凸集或集 A 的凸包, 记为 $Co(A)$，显然集 A 的凸包 $Co(A)$ 是 X 中包含集 A 的最小口集.

R 全空间是凸集，空集合是凸集，空集合的验证：写出不是凸集的定义，发现不发验证。

Exercise 10.1 给出凸包的构造

Definition 10.2.8 — 商空间. 设 R 是线性空间, E 是 R 的一个线性子空间, 我们在 R 中规定: 当 $x-y\in E$ 时为 $x\sim y$, 容易证明 $\sim$ 是 R 中的等价关系. 我们把商集 $R/\sim$ 记为 R/E, 并记 x 所在的等价类为 $\tilde{x}$. 在 R/E 中规定线性运算如下:

$$\begin{aligned}&\tilde{x}+\tilde{y}=(\widetilde{x+y})\\&\alpha\tilde{x}=(\widetilde{\alpha x}),\alpha \text{ 是数.}\end{aligned}$$

R/E 按这样规定的线性运算成为线性空间, 称 R/E 为 R 关于 E 的商空间. 也容易看出 R/E 中的零向量就是 E, 即 $\tilde{0}=E$. 事实上, 对任何 $x\in E, x-0=x\in E$. 直观地说, 在商空间 R/E 中, E 被“缩成”为零向量.

■ **Example 10.7** 去掉一个 f 处处等于 0 的集合后，剩下的元素都是等价的

Definition 10.2.9 — L^p. 设 $(X,\boldsymbol{B},\boldsymbol{\mu})$ 是一个测度空间, $E\in\boldsymbol{B}, f(t)$ 是 E 上的实值或复值) 函数, 取定正数 p. 设 f 是 E 上的可测函数, 而且 $|f|^p$ 在 E 上是可积的。这种函数 f 的全体记作 $L^p(E,\boldsymbol{B},\mu)$, 简记为 $L^p(E,\mu)$, 简称 $L^p(E,\mu)$ 中的函数是 p 方可积函数. 有时也用 $L^p(E)$ 表示 E 上关于 Lebesgue 测度的 p 方可积函数空间.

Theorem 10.2.3 $L^p(E,\mu)(p>0)$ 按通常的线性运算成一线性空间.

事实上, 对于 $f,g\in L^p(E,\mu), f+g$ 在 E 上可测, 所以 $|f+g|^p$ 是 E 上的可测函数. 对于任意的数 a,b, 成立着不等式

$$(|a|+|b|)^p\leqslant[2\max(|a|,|b|)]^p\leqslant 2^p\left(|a|^p+|b|^p\right)$$

在这个空间上定义范数:

$$\|f\|_p=\left(\int_E|f(t)|^p\mathrm{d}\mu\right)^{\frac{1}{p}}(p\geqslant 1)$$

Definition 10.2.10 — 平均收敛. 在 $L^p(E,\mu)$ 中, 设函数列 f_n 依范数 $\|\cdot\|_p$ 收敛于 f, 即

$$\int_E|f_n(x)-f(x)|^p\mathrm{d}\mu\to 0, n\to\infty$$

这种收敛在经典分析中称为 $f_n(x)$ 在 E 上 p 方平均收敛于 $f(x)$ (有时也省略地说

$f_n(x)$ 平均收敛于$f(x)$).

Theorem 10.2.4 设 $f_n(x)(n=1,2,3,\cdots)$ 及 $f(x)$ 是 $L^p(E,\boldsymbol{B},\mu)$ 中的函数. 如果函数列 $\{f_n(x)\}$ 是 p 方平均收敛于 $f(x)$, 那么函数列 $\{f_n(x)\}$ 必然在 E 上依测度收敛于 $f(x)$

Proof. 对于任何正数 σ, 有

$$\begin{aligned}\int_E |f_n(x)-f(x)|^p \mathrm{d}\mu &\geqslant \int_{E(|f_n-f|\geqslant\sigma)} |f_n(x)-f(x)|^p \mathrm{d}\mu \\ &\geqslant \sigma^p \mu\left(E\left(|f_n(x)-f(x)|\geqslant\sigma\right)\right)\end{aligned}$$

令 $n\to\infty$, 就有 $\mu\left(E\left(|f_n(x)-f(x)|\right)\geqslant\sigma\right)\to 0$ ■

R

1. 设函数列 $\{f_n(x)\}$ 是 p 方平均收敛于函数 $f(x)$, 那么必有子函数列 $\{f_{n_k}(x)\}$ 概收敛于 $f(x)$
2. 然而定理的逆命题不正确. 即使函数列 $\{f_n(x)\}$ 在有限可测集 E 上处处收敛于 $f(x)$, 也不能保证 $\{f_n(x)\}$ 平均收敛于 $f(x)$

■ **Example 10.8** 我们作 [0,1] 区间上的函数列 $\{f_n(x)\}$ 如下:

$$f_n(x)=\begin{cases}0, & \text{当} x=0 \text{ 或} \frac{1}{n}\leqslant x\leqslant 1 \\ e^n, & \text{当} 0<x<\frac{1}{n}\end{cases}$$

显然, $\{f_n(x)\}$ 在 [0,1] 上处处收敛于零, 但是当 $n\to\infty$ 时, 对于任何的正数 p

$$\int_0^1 |f_n(x)|^p \mathrm{d}x=\int_0^{\frac{1}{n}} e^{pn}\mathrm{d}x=\frac{1}{n}e^{pn}\to\infty, (n\to\infty)$$

所以 $\{f_n(x)\}$ 并不 p 方平均收敛于零.

Definition 10.2.11 — $L^\infty(E,\mu)$. 空间 $L^\infty(E,\mu)$ 设 E 是测度空间 $(\Omega,\boldsymbol{B},\mu)$ 上一个可测集, $f(x)$ 是 E 上的可测函数。如果 $f(x)$ 和 E 上的一个有界函数几乎处处相等—换句话说, 如果有 E 中 (关于 μ) 的零集 E_0, 使得 $f(x)$ 在 $E-E_0$ 上是有界的 —— 那么我们称 $f(x)$ 是 E 上 (关于μ) 的本性有界可测函数. E 上的本性有界可测函数全体记作 $L^\infty(E,\mu)$. 显然, 由于有限个零集的和集也是零集, 所以任意有限个本性有界可测函数的线性组合是本性有界的, 因此, $L^\infty(E,\mu)$ 按通常的线性运算是一线性空间。设 $f(x)$ 是 E 上的本性有界可测函数, 令

$$\|f\|_\infty=\inf_{\substack{\mu(E_0)=0\\E_0\subset E}}\left(\sup_{E-E_0}|f(x)|\right)$$

这里下确界是对于 E 中所有使得 $f(x)$ 在 $E-E_0$ 上成为有界函数的零集 E_0 而取的, 称

为 f 的本性最大模, 有时也记作

$$\|f\|_\infty=\inf_{\substack{\mu(E_0)=0\\E_0\subset E}}\left(\sup_{E-E_0}|f(x)|\right)=\operatorname*{ess\,sup}_{x\in E}|f(x)|$$

定义中的下确界 inf 是可达的, 就是说必有含于 E 的零集 E_0 使得 $\|f\|_\infty$ 等于 $|f(x)|$ 在 $E-E_0$ 上的上确界. 这是因为, 由 inf 的意义, 对每个 n, 有 $E_n\subset E$ 使得 $\mu(E_n)=0$, 并日

$$\sup_{x\in E-E_n}|f(x)|<\|f\|_\infty+\frac{1}{n}$$

令 $n\to\infty$, 就得到 $\|f\|_\infty=\sup_{E-E_0}|f(x)|$. 还可以证明 $\|f\|_\infty$ 是与 $f(x)$ 几乎处处相等的各个有界函数的绝对值的上界的最小值。

Theorem 10.2.5 $L^\infty(E,\mu)$ 关于 $\|\cdot\|_\infty$ 成为赋范线性空间. 现在来考察空间 $L^\infty(E,\mu)$ 中点列 $\{f_n\}$ 收敛的情况. 设 $f_n,f\in L^\infty(E,\mu),n=1,2,3,\cdots$, 而且 $\|f_n-f\|_\infty\to0$, 那么有 $F_n\subset E,\mu(F_n)=0$, 使得

$$\|f_n-f\|_\infty=\sup_{E-F_n}|f_n(x)-f(x)|\to0$$

取 $F_0=\bigcup_{n=1}^\infty F_n$, 那么, F_0 是一零集, 并且

$$\sup_{E-F_0}|f_n(x)-f(x)|\to0 \tag{10.2}$$

由于 F $_0$ 是 E 中的零集, 说明了 $\{f_n(x)\}$ 在 E 上除去一个零集 F_0 后是均匀收敌于 $f(x)$ 的. 这时我们就说 $\{f_n(x)\}$ 在 E 上几乎匀敛于 $f(x)$. 显然, 10.2也是使 $\|f_n-f\|_\infty\to0$ 的充分条件因此, 度量空间 $L^\infty(E,\mu)$ 中依距离收敛就是几乎匀敛.

R 度量空间 $L^\infty(E,\mu)$ 中依距离收敛就是几乎匀敛.

Theorem 10.2.6 如果 $\mu(E)<\infty$, 显然对一切正数 $p,L^\infty(E,\mu)\subset L^p(E,\mu)$. 现在来证明: 当 $\mu(E)<\infty$ 时

$$\|f\|_\infty=\lim_{p\to\infty}\|f\|_p$$

Proof. 事实上，只要考察 $\mu(E)>0,\|f\|_\infty\neq0$ 的情况好了. 取 E 中的零集 E_0 使得 $\|f\|_\infty=\sup_{E-E_0}|f(x)|$, 于是

$$\int_E|f(x)|^p\mathrm{d}\mu=\int_{E-E_0}|f(x)|^p\mathrm{d}\mu\leqslant\|f\|_\infty^p\mu(E)$$

由于 $\mu(E)^{\frac{1}{p}}\to1(p\to\infty)$, 从得到

$$\varlimsup_{p\to\infty}\|f\|_p\leqslant\|f\|_\infty$$

另一方面, 任取一个正数 $\varepsilon < \|f\|_\infty$, 集 $E_\varepsilon = E(|f(x)| \geqslant \|f\|_\infty - \varepsilon)$ 不会是零集. 因为如果这是零集的话, 在 E 中去掉这个集后, $|f(x)|$ 在剩下来的 $E - E_\varepsilon$ 中的上确界不超过 $\|f\|_\infty - \varepsilon$, 这显然和 $\|f\|_\infty$ 的定义冲突. 因此

$$\begin{aligned}\|f\|_p &\geqslant \left(\int_{E_\varepsilon} |f(x)|^p \mathrm{d}\mu\right)^{\frac{1}{p}} \\ &\geqslant (\|f\|_\infty - \varepsilon)\left[\mu(E_\varepsilon)\right]^{\frac{1}{p}}\end{aligned}$$

令 $p \to \infty$ 就得到

$$\varliminf_{p\to\infty} \|f\|_p \geqslant \|f\|_\infty - \varepsilon$$

再令 $\varepsilon \to 0$, 即可得到结论 ■

Proposition 10.2.7 设 A 是度量空间 R 中的点集, $x_0 \in R$. 那么下面四件事是彼此等价的。

1. x_0 是集 A 的极限点.
2. x_0 的任何一个环境 $O(x_0)$ 中必含有 A 中异于 x_0 的点, 即 $(O(x_0) - \{x_0\}) \cap A \neq \varnothing$
3. 在集 A 中存在一列点 $\{x_n\}$, 适合 $x_n \neq x_0$ 而且 $x_n \to x_0$
4. 在集 A 中必有一列互不相同的点 $\{x_n\}$, 而目 $x_n \neq x_0$, 使得 $x_n \to x_0$

Theorem 10.2.8 在离散的度量空间中, 任一点集都没有极限点。因而每个点集都是闭集, 同时还是开集. 由此可见, 在有的度量空间中, 既开又闭的集可能很多。

Definition 10.2.12 如果度量空间 R 不能被分解为两个都不空的互不相交的闭集 R_1 及 R_2 的和, 那么称 R 是联络的空间.

1. 每个区间都是直线上的联络点集.
2. 赋范线性空间是联络空间。

Definition 10.2.13

1. 我们仍然用 $\boldsymbol{R}_0$ 表示 E^n 中形如

$$C(\{a_i\},\{b_i\}) = \{(x_1,x_2,\cdots,x_n)\,|a_i < x_i \leqslant b_i, i = 1,2,\cdots,n\}$$

(其中 $-\infty < a_i \leqslant b_i < +\infty$) 的半开半闭室全体所张成的环.

2. 称由 R_0 张成的 E^n 中的最小 σ -环 $\boldsymbol{S}(\boldsymbol{R}_0)$ (其实, 还是$\sigma-$ 代数) 中的集是 E^n 中的 Borel 集, 这种集的全体记作 $\boldsymbol{B}$
3. E^n 中的开集、闭集都是 Borel 集, 而且 E^n 中开集全体 (或者闭集全体) 所张成的 σ-环就是 B.

■ **Example 10.9**

1. 在一般的度量空间中, 可能 $O(x_0,r) = B(x_0,r)$, 例如: 在平凡的度量空间中, 取 $r = \frac{1}{2}$

$$\text{那么} O(x_0,r) = B(x_0,r) = \{x_0\}$$

2. n 维 Euclid 空间中包含在闭球中的点列都存在收敛子列,一般的度量空间没有这个性质,例如 l^2 中的点列 $\{e_n : n = 1,2,\cdots\}$,每个 e_n 都落在闭球 $B(0,1)$ 中,和 $\|e_n - e_m\| = \sqrt{2}(n \neq m)$
3. 一般度量空间中半径小的球可以包含半径大的球。例如，考虑度量空间 $X = [0,1]$, 作为直线的子空间。对手闭球 $B\left(0,\frac{1}{2}\right) = \left[0,\frac{1}{2}\right]$,取其中的一个点 $y_0 = \frac{1}{4}$, 则以 y_0 为中心
$$\frac{1}{3}\text{为半径的闭球}B\left(y_0,\frac{1}{3}\right) = \left[0,\frac{7}{12}\right] \supset B\left(0,\frac{1}{2}\right)$$
4. $C[0,1]$ 作为度量空间 $L^1[0,1]$ 的子集，其中的任何点都不是内点.

Definition 10.2.14 — 连续的定义. 设 R 和 K 是度量空间,f 是 $A \subset R$ 到 K 中的映照 (映射). 设 $x_0 \in A$. 假如对于 $f(x_0)$ 的任何环境 $O(f(x_0))(\subset K)$, 必有 x_0 在 R 中的一个环境 $O(x_0)$,使得 $f(O(x_0) \cap A) \subset O(f(x_0))$ (就是说当 $x \in O(x_0) \cap A$ 时, $f(x) \in O(f(x_0))$), 就称映照 f 在点 x_0 是连续的.

Definition 10.2.15 — 稠密. 设 R 是度量空间, A 及 E 是 R 中的点集. 如果 E 中任何一点 x 的任何环境中都含有集 A 中的点, 就称 A 在 E 中稠密.
1. A 在 E 中稠密的充要条件是 $\bar{A} \supset E$
2. A 在 E 中稠密的充要条件是对任一 $x \in E$, 有 A 中的点列 $\{x_n\}, x_n \to x(n \to \infty)$
3. P 在 $C[a,b]$ 中是稠密的.(Weierstrass 的逼近定理)

Theorem 10.2.9 设 E 是 μ 可测集, $L^p(E,\mu)$ 中的有界可测函数全体 $B(E)$ 是 $L^p(E,\mu)(\infty > p \geqslant 1)$ 的稠密子集.

Proof. 设 $f(x) \in L^p(E,\mu)$, 对每个自然数 n, 造函数
$$f_n(x) = \begin{cases} f(x) & |f(x)| \leqslant n \\ 0, & |f(x)| > n \end{cases}$$
那么 $f_n(x)$ 是 $L^p(E,\mu)$ 中的有界可测函数, 而且
$$\int_E |f_n(x) - f(x)|^p \mathrm{d}\mu = \int_{E(|f|>n)} |f(x)|^p \mathrm{d}\mu$$
由于 $|f|^p \in L(E,\mu)$, 由积分的全连续性, 对任一 $\varepsilon > 0$, 必有 $\delta > 0$, 使得当 $e \subset E, \mu(e) < \delta$ 时成立着
$$\int_e |f|^p \mathrm{d}\mu < \varepsilon^p$$
因为
$$n^p \mu(E(|f| > n)) \leqslant \int_{E(|f|>n)} |f(x)|^p \mathrm{d}\mu \leqslant \int_E |f|^p \mathrm{d}\mu$$
所以有正数 N, 使当 $n > N$ 时 $\mu(E(|f| > n)) < \delta$, 因而
$$\|f_n - f\|_p = \left(\int_{E(|f|>n)} |f|^p \mathrm{d}\mu\right)^{\frac{1}{p}} < \varepsilon$$

所以 E 上的有界可测函数全体 $B(E)$ 在 $L^p(E,\mu)$ 中稠密. ■

Theorem 10.2.10 对于直线上任一 Lebesgue 可测集 E, 当 $1 \leqslant p < \infty$ 时, $L^p(E)$ 中的有界连续函数全体在 $L^p(E)$ 中是稠密的.

Theorem 10.2.11 设 $[a,b]$ 是有限区间, $p \geqslant 1$, 那么 P 和 $C[a,b]$ 在 $L^p([a,b])$ 中稠密.

Definition 10.2.16 设 R 是度量空间, A 是 R 中的子集. 如果存在有限集或可列集 $\{x_k\} \subset R$ 在 A 中稠密, 就称 A 是可析点集. 可析点集.

例 1 n 维欧几里得空间 E^n 按通常的距离是可析空间. 因为坐标为有理数的点全体是可列集合, 并且在 E^n 中稠密. 例 2 当 $1 \leqslant p < \infty$ 时, 空间 l^p 是可析的. 因为形如 $y = \{y_1,y_2,\cdots y_m\ 0,\cdots\}$, 而 $y_1,y_2,\cdots,y_m$ 是有理数的点的全体 A 是可列集, 它在 l^p 中稠密. 事实上, 任取 $x = \{x_v\}$, 设 $x \in l^p$, 今证 $x \in \bar{A}$. 由于 $\sum_v |x_v|^p < \infty$, 对任意的正数 ε 必有 m 使得 $\sum_{v=m+1}^{\infty} |x_v|^p < \left(\frac{\varepsilon}{2}\right)^p$, 再取有理数 $y_1,\cdots,y_m$ 使得 $\sum_{v=1}^{m} |x_v - y_v|^p < \left(\frac{\varepsilon}{2}\right)^p$, 因此 A 中的点 $y = \{y_1,\cdots,y_m,0,\cdots\}$ 与 x 的距离 $\|y-x\|_p < \varepsilon$, 即

$$y \in O(x,\varepsilon) \cap A, \text{所以} x \in \bar{A}$$

例 3 $C[a,b]$ 和 $L^p[a,b](\infty > p \geqslant 1)$ 是可析空间. 因为对任意的多项式 $P(x)$, 总有以有理数为系数的多项式 $p(x)$ 适合

$$\|P-p\| = \max_x |P(x) - p(x)| < \varepsilon$$

例 4 有界数列全体组成的空间 l^∞ 是不可析的. 证 l^∞ 中形如 $\{x_i\}, x_i = 0$ 或 1 的点, 其全体记为 K, 则 K 是不可列集. 对于 K 中任意的相异两点 x,y, 必有 $\rho(x,y) = 1$, 即 l^∞ 中有一个不可列的集 K, 其中每两点之间的距离都是 1. 如果 l^∞ 是可析的, 那么有可列集 $\{y_k\}$ 在 l^∞ 中稠密, 空间 l^∞ 中以 K 中点为中心, $\frac{1}{3}$ 为半径的每一球内至少有一个 y_k, 因为这种球有不可列个, 但是 $\{y_k\}$ 中只有可列个点, 所以至少有一个 y_k 同时属于两个不同的球, 例如属于 $O\left(x^{(1)},\frac{1}{3}\right), O\left(x^{(2)},\frac{1}{3}\right)$, 其中 $x^{(1)},x^{(2)} \in K$ 这样一来

$$1 = \rho\left(x^{(1)},x^{(2)}\right) \leqslant \rho\left(x^{(1)},y_{k_0}\right) + \rho\left(x^{(2)},y_{k_0}\right) \leqslant \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

这是矛盾. 所以空间 l^∞ 是不可析的.

Definition 10.2.17 设 R 是度量空间, A 是 R 的子集. 如果 A 不在 R 的任何一个非空的开集中稠密, 那么 A 称作疏朗集.

Theorem 10.2.12 显然, 这个定义中的非空开集可以换成半径不为零的开球。我们用 $S(a,\rho)$ 表示闭球 $\{x|\rho(x,a) \leqslant \rho\}$, 那么 A 在度量空间 R 中疏朗的充要条件是任何闭球 $S(a,\rho)(\rho > 0)$ 中必有闭球 $S(b,r)(r > 0)$ 与 A 不交. 事实上, 如果 A 是疏朗的, 那么 A 不在开球 $O(a,\rho) \subset S(a,\rho)$ 中稠密, 所以必有 $b \in O(a,\rho)$ 以及b 的ε -环境$O(b,\varepsilon)$ (不妨设$O(b,\varepsilon) \subset S(a,\rho)$) 使得 $O(b,\varepsilon)$ 和 A 不交. 取 $0 < r < \varepsilon$, 那么 $S(b,r) \subset O(b,\varepsilon) \subset S(a,\rho)$, 而且 $S(b,r)$

与 A 不交. 条件的充分性是显然的.

Definition 10.2.18 在度量空间中的点集 A 如果能表示成为最多可列个疏朗集 $M_\nu(\nu = 1,2,3,\cdots)$ 的和, 就称 A 是第一类型的集. 度量空间中的不是第一类型的集称作第二类型的集. 第一、二类型集又分别称为第一、二纲集.

■ **Example 10.10** 由于单元素集显然在欧氏空间中是疏朗集。所以欧几里得空间中任一可列集都是第一类型的集.

10.3 完备性

Definition 10.3.1 设 (R,ρ) 是度量空间, $\{x_n\}$ 是 R 中的点列. 如果对于任一正数 ε, 存在正数 $N(\varepsilon)$, 使得当自然数 $n,m \geqslant N(\varepsilon)$ 时

$$\rho(x_n, x_m) < \varepsilon$$

就称 $\{x_n\}$ 是 R 中基本点列, 或称为 Cauchy 点列.

Theorem 10.3.1
1. 度量空间 R 中收敛点列必是基本点列.
2. 设 $\{x_n\}$ 是度量空间 R 中基本点列, 如果 $\{x_n\}$ 有子点列 $\{x_{n_k}\}$ 收敛于 R 中的点 x, 那么 $\{x_n\}$ 也收敛于 x

证明都是利用三角不等式

Definition 10.3.2 如果度量空间 R 中每个基本点列都收敛, 称 R 是完备 (度量) 空间. 完备赋范线性空间又称为 Banach (巴拿赫) 空间. 如果 R 是度量空间, 是 R 的子空间, 当 A 作为度量空间是完备的, 那么称 A 是 R 的完备子空间.

Definition 10.3.3
1. 完备的赋范线性空间称为 Banach 空间。
2. 完备的内积空间称为 Hilbert 空间
3. n 维 Euclid 空间 R^n 是 Hilbert 空间.
4. $l^p(1 \leqslant p \leqslant \infty)$ 是 Banach 空间. 特别的, l^2 是 Hilbert 空间.
5. $C[a,b]$ 是 Banach 空间.
6. 空间 $L^p[a,b](1 \leqslant p < \infty)$ 是 Banach 空间.　　特别的, $L^2[a,b]$ 是 Hilbert 空间.
7. $L^\infty[a,b]$ 是 Banach 空间.
8. 作为直线的子空间, 整数集合 Z 是完备的度量空间。
9. 作为直线的子空间, 集合 $X = \{\frac{1}{n} : n = 1,2,\cdots\} \cup \{0\}$ 是完备的度量空间.
10. 集合 $X = \{\frac{1}{n} : n = 1,2,\cdots\}$ 作为直线的子空间, 不是完备的度量空间.
11. 有理数集作为直线的子空间不是完备的。例如 $\left\{\left(1+\frac{1}{n}\right)^n\right\}$ 是基本列, 但是在有理数空间中它不收敛。

Proposition 10.3.2
1. 完备度量空间的闭子集是完备子空间.
2. 任何度量空间的完备子空间是闭子集.

3. 设 X 是度量空间。如果在 X 上闭球套定理成立，那末 X 是完备的。

■ **Example 10.11** 一致离散的度量空间是完备的。事实上，如果 $\{x_n\}$ 是基本点列, 那么由一致离散性可知必存在 N, 当 $n \geqslant N$ 时, $x_N = x_{N+1} = \cdots = x_{N+k} = \cdots$ 因而 $\{x_n\}$ 收敛于 x_N, 从而空间是完备的.

■ **Example 10.12** n 维欧几里得空间 E^n 是完备的.

Proof. 设 $\left\{x_m | x_m = \left(x_1^{(m)}, \cdots, x_n^{(m)}\right), m = 1,2,3,\cdots\right\}$ 是 E^n 中的一个点列. 由于

$$\begin{aligned}\left|x_i^{(m)} - x_i^{(k)}\right| &\leqslant \|x_m - x_k\| = \left(\sum_{j=1}^{n}\left|x_j^{(m)} - x_j^{(k)}\right|^2\right)^{\frac{1}{2}} \\ &\leqslant \sqrt{n}\max_{1\leqslant j\leqslant n}\left|x_j^{(m)} - x_j^{(k)}\right|, i = 1,2,\cdots,n\end{aligned}$$

■

Theorem 10.3.3 设 (X) 是赋范空间, 则:
1. 范数是一个连续函数, 即当 $x_n \to x(n \to \infty)$ 时, $\|x_n\| \to \|x\|(n \to \infty)$.
2. 线性运算是连续的，即当 $x_n \to x(n \to \infty)$ 及 $y_n \to y(n \to \infty)$ 时 ,$x_n + y_n \to x + y(n \to \infty)$; 当 $\alpha_n \to \alpha(n \to \infty)$ 及 $x_n \to x(n \to \infty)$ 时 ,$\alpha_n x_n \to \alpha x(n \to \infty)$

Proof. 证设 $x_n \to x(n \to \infty)$, 由三角形不等式

$$\|x\| \leqslant \|x - x_n\| + \|x_n\|$$

及

$$\|x_n\| \leqslant \|x_n - x\| + \|x\|$$

所以

$$|\|x_n\| - \|x\|| \leqslant \|x_n - x\|$$

由此立刻得 $\|x_n\| \to \|x\|(n \to \infty)$, 即1) 成立. 其次, 由于

$$\|(x_n + y_n) - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\|$$

及

$$\begin{aligned}\|\alpha_n x_n - \alpha x\| &\leqslant \|\alpha_n x_n - \alpha x_n\| + \|\alpha x_n - \alpha x\| \\ &= |\alpha_n - \alpha|\,\|x_n\| + |\alpha|\,\|x - x_n\|\end{aligned}$$

注意由本定理的结论 1)

1. $\|x_n\|$ 是有界的, 则当 $x_n \to x$ 及 $y_n \to y$ 时 ,$x_n + y_n \to x + y_i$ 当 $\alpha_n \to \alpha$ 及 $x_n \to x$ 时 ,$\alpha_n x_n \to \alpha x$

2. 可见，在一个赋范空间中，作为线性空间的代数结构与作为距离空间的拓扑结构以线性运算的连续性把两种结构联系起来.

■

Theorem 10.3.4 若度量函数想与加法数乘融合，则一定要满足范数条件

Exercise 10.2 是否任何一个距离函数都是范数函数诱导出的，若成立的话，则范数的定义则没有意义。找到例子。

设 s 是所有的实 (或复) 数列所成的线性空间. 在数列空间 s 中, 如果令

$$\|x\| = \rho(x,0) = \sum_{\nu=1}^{\infty} \frac{1}{2^\nu} \frac{|x_\nu|}{1+|x_\nu|}$$

那么, 对于 $\alpha \neq 0$, 并不满足齐次性条件 $\|\alpha x\| = |\alpha|\|x\|$ ■

Theorem 10.3.5 如果 $\sum_{\nu=1}^{\infty}\|a_\nu\| < \infty$, 那么 Banach 空间中的级数 $\sum_{\nu=1}^{\infty} a_\nu$ 收敛

$$\left(\text{因为} \left\| \sum_{\nu=m+1}^{n} a_\nu \right\| \leqslant \sum_{\nu=m+1}^{n} \|a_\nu\| \right)$$

Theorem 10.3.6 1. $C[a,b]$ 是一个 Banach 空间.: 证明思路:

$$\|f_n - f_m\| = \sup_{a \leqslant x \leqslant b} |f_n(x) - f_m(x)| < \varepsilon$$

令 $m \to \infty$

2. 如果在连续函数族 $C[a,b](-\infty < a < b < +\infty)$ 上, 定义范数为

$$\|f\|_1 = \int_a^b |f(t)| \mathrm{d}t$$

就是把 $C[a,b]$ 看成 $L[a,b]$ 的子空间, 那么 $(C[a,b], \|\cdot\|_1)$ 是不完备的空间. $C[a,b]$ 中按 $\|\cdot\|_1$ 基本点列 $\{f_n\}$, 它不收敛于 $C[a,b]$ 中任何一点: 任取 $c, a < c < b$, 函数列 $f_n(x)(n$ 充分大) 如下:

$$f_n(x) = \begin{cases} 1, & c + \frac{1}{n} \leqslant x \leqslant b \\ \text{线性}, & c - \frac{1}{n} \leqslant x \leqslant c + \frac{1}{n} \\ -1, & a \leqslant x \leqslant c - \frac{1}{n} \end{cases}$$

那么 $f_n(x) \in C[a,b]$, 不难证明, 在 $[a,b]$ 上每一点 $x, f_n(x)$ 收敛于函数

$$f(x) = \begin{cases} 1, & c < x \leqslant b \\ 0, & x = c \\ -1, & a \leqslant x < c \end{cases}$$

显然，$f(x) \in L[a,b]$, 通过直接计算就知道

$$\|f_n - f_m\|_1 \leqslant \frac{2}{n} + \frac{2}{m} \to 0(n,m \to \infty)$$

所以 $\{f_n\}$ 是空间 $(C[a,b], \|\cdot\|_1)$ 中的基本点列. 但 $\{f_n\}$ 在 $(C[a,b], \|\cdot\|_1)$ 中并不收敛. 因为如果有 $g \in C[a,b]$ 使得 $\|f_n - g\|_1 \to 0$, 则由 Lebesgue 积分的控制收敛定理得到 $\|f-g\|_1 = \lim_{n\to\infty} \|f_n - g\|_1 = 0$, 所以 $f(x) \doteq g(x)$. 但是容易看出, $f(x)$ 不可能在 $[a,b]$ 上几乎处处等于 $[a,b]$ 上的一个连续函数, 所以 $(C[a,b], \|\cdot\|_1)$ 是不完备的赋范线性空间。

■ **Example 10.13** 空间 $C[a,b]$ 在区间 [a,b] 上的连续函数空间 C[a,b](参看 $1.1) 中按通常方式规定线性运算是一个线性空间, 定义

$$\|x\| = \max_{a \le t < \le b} |x(t)|$$

不难验证,$C[a,b]$ 是一个赋范空间. 显然以前我们在 $C[a,b]$ 中定义的距离正是由范数诱导的距离, 作为距离空间它是完备的、可分的。因此,$C[a,b]$ 是一个可分的 *Banach* 空间. $C[a,b]$ 是一个十分重要的 *Banach* 空间, 它在分析中有着广泛的应用. 一般地, 设 K 是紧 *Hausdorff* 空间, 可以类似地定义 *Banach* 空间 $C(K)$ ■

Theorem 10.3.7 X 是赋范空间, 如果 X 是完备的且级数

$$\sum_{k=1}^{\infty} \|x_k\| = \|x_1\| + \|x_2\| + \cdots + \|x_n\| + \cdots \tag{10.3}$$

收敛, 则级数

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots + x_n + \cdots \tag{10.4}$$

收敛且 $\|\sum_{k=1}^{\infty} x_k\| \leqslant \sum_{k=1}^{\infty} \|x_k\|$ 反之, 如果在一个赋范空间中, 任意无穷级数10.3 收敛必有级数10.4 收敛, 则空间是 *Banach* 空间.（这是一个充分必要的命题）

R 无穷级数可以看作是部分和数列的极限，判断部分和的 *Cauchy* 列，这就提到是否空间完备。注意10.3这个式子是数项技术，而10.4是赋范空间中的级数。类似于之前的数项级数绝对收敛则原级数收敛。但是此时需要空间的完备性。当时所有的欧氏空间都是完备的。

Analysis 10.1 复杂的一部分主要证明思路是：根据级数的形式构造出一串 *Cauchy* 序列的子列是有极限的，在利用 *Cauchy* 的整体的性质得到整体会收敛 ■

Proof. 证设 $s_n = x_1 + \cdots + x_n$ 是级数10.4的部分和，对任意自然数 p

$$\begin{aligned}\left\|s_{n+p} - s_n\right\| &= \left\|x_{n+1} + \cdots + x_{n+p}\right\| \\ &\leqslant \left\|x_{n+1}\right\| + \cdots + \left\|x_{n+p}\right\|\end{aligned}$$

范数的三角不等式（有限项）由于级数10.3收敛，可见 $\{s_n\}$ 是 Cauchy 列，而 X 是完备的，所以级数 10.4收敛. 在不等式

$$\left\|\sum_{k=1}^{n} x_k\right\| \leqslant \sum_{k=1}^{n} \|x_k\|$$

两边令 $n \to \infty$, 则有

$$\left\|\sum_{k=1}^{\infty} x_k\right\| \leqslant \sum_{k=1}^{\infty} \|x_k\|$$

反之, 设空间 X 中, 任意级数10.3收敛必有级数10.4收敛, 且 $\{x_n\}$ 是 X 中任一 Cauchy 列. 从 $\{x_n\}$ 中选取子列 $\{x_{n_k}\}$, 使得

$$\|x_{n_{k+1}} - x_{n_k}\| < \frac{1}{2^k} \quad (k = 1, 2, \cdots)$$

于是级数 $\sum_{k=1}^{\infty} \|x_{n_{k+1}} - x_{n_k}\|$ 收敛, 因此级数

$$x_{n_1} + \sum_{k=1}^{\infty} (x_{n_{k+1}} - x_{n_k}) = x_{n_1} + (x_{n_2} - x_{n_1}) + \cdots + (x_{n_k} - x_{n_{k-1}}) + \cdots$$

必收敛, 其前 k 项的部分和是 x_{n_k}, 设 $x_{n_k} \to s(k \to \infty)$, 这样, 存在 $\{x_n\}$ 的一个子列 $\{x_{n_k}\}$ 收敛. 由于 $\{x_n\}$ 是 *Cauchy* 列, 对任意 $\varepsilon > 0$, 存在 N, 当 $m, n > N$ 时

$$\|x_n - x_m\| < \varepsilon$$

因此, 对于充分大的 k,

$$\|x_n - x_{n_k}\| < \varepsilon$$

命 $k \to \infty$, 则当 $n > N$ 时 $\|x_n - s\| \leqslant \varepsilon$ 所以 $\{x_n\}$ 收敛, 即 X 是完备的. 注意, 在以上定理的证明中, 我们得到一个有用的事实, 当一个 *Cauchy* 列有一个子列收敛时, 则点列本身必收敛并且收敛于同一极限.

■

三大不等式

Theorem 10.3.8 设 p, q 是正数, 且满足

$$\frac{1}{p} + \frac{1}{q} = 1$$

则对任意数 a, b

$$|ab| \leqslant \frac{|a|^p}{p} + \frac{|b|^q}{q}$$

Proof. 证不妨设 a,b 都是正数, 记 $s=\frac{1}{p}$. 考虑 $(0,\infty)$ 上的函数 $\varphi(t)=t^s-s$ 由于 $\varphi'(t)=s\left(t^{-1}-1\right)$, 所以当 $t=1$ 时 $\varphi(t)$ 取最大值. 因此当 $t>0$ 时 $\varphi(t)\leqslant\varphi(1)$. 由此 $t'-1\leqslant s(t-1)$. 用 $t=\frac{a^p}{b^q}$ 代人这个不等式则得

$$ab^{-\frac{q}{p}}-1\leqslant\frac{1}{p}\left(a^pb^{-q}-1\right)$$

在上式两边乘 b^q, 并注意 $q-\frac{q}{p}=1$, 即得不等式 (2. 2. 2). ■

Proof. 证明二：设 $y=\phi(x)$ 为单调增函数，并且满足 $\phi(0)=0$，则一定存在反函数 ψ。考虑 $\phi(a)$ 的函数值有三种可能：$\left[\begin{array}{l}\varphi(a)>b\\ \varphi(a)=b\\ \varphi(a)<b\end{array}\right.$ 很显然根据函数图像可以得到下面的不等式：

$$ab\leq\int_0^a\varphi(x)dx+\int_0^b\psi(y)dy$$

取 $\phi(x)=y=x^{p-1}=x^{\frac{p}{q}}$，则反函数为 $\psi(y)=x=y^{\frac{q}{p}}=y^{q-1}$，这里用到了共轭数的性质：

$$\begin{aligned}\frac{1}{p}+\frac{1}{q}=1\Rightarrow pq=p+q\Rightarrow p-1&=\frac{p}{q}\\ \Rightarrow q-1&=\frac{q}{p}\end{aligned}$$

带入得到 *Yong* 不等式

$$ab\leqslant\int_0^a x^{p-1}dx+\int_0^b y^{q-1}dy=\frac{a^p}{p}+\frac{b^q}{q}$$

■

Theorem 10.3.9 — Hölder 不等式. 设 E 是 *Lebesgue* 可测集, $x(t),y(t)$ 是 E 上的可测函数. 则有不等式

$$\int_E|x(t)y(t)|\mathrm{d}t\leqslant\left(\int_E|x(t)|^p\mathrm{d}t\right)^{\frac{1}{p}}\left(\int_E|y(t)|^q\mathrm{d}t\right)^{\frac{1}{q}},$$

其中 p,q 满足式共轭数.

Analysis 10.2 aaa

Proof. 证 记 $A^p=\int_E|x(t)|^p\mathrm{d}t,B^q=\int_E|y(t)|^q\mathrm{d}t$, 则不妨设 $0<A^p<\infty$ 且 $0<B^4<\infty$, 因为如果 A^p,B^q 中有一个为 0 或无穷, 不等式 (2. 2.3) 显然成立. 对于任意的 $t\in E$, 使用 Yong 不等式，则

$$\frac{|x(t)y(t)|}{AB}\leqslant\frac{1}{p}\left|\frac{x(t)}{A}\right|^p+\frac{1}{q}\left|\frac{y(t)}{B}\right|^q \tag{10.5}$$

对上式两边积分, 则得

$$
\begin{aligned}
\frac{1}{AB}\int_E |x(t)y(t)|\mathrm{d}t &\leqslant \frac{A^{-p}}{p}\int_E |x(t)|^p \mathrm{d}t + \frac{B^{-q}}{q}\int_E |y(t)|^q \mathrm{d}t \\
&= \frac{1}{p} + \frac{1}{q} = 1
\end{aligned}
$$

所以

$$
\int_E |x(t)y(t)|\mathrm{d}t \leqslant AB = \left(\int_E |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \left(\int_E |y(t)|^q \mathrm{d}t\right)^{\frac{1}{q}}
$$

■

Theorem 10.3.10 — (Minkowski 不等式). 设 E 是 Lebesgue 可测集, $x(t),y(t)$ 是 E 上可测函数, $p \geqslant 1$, 则有不等式

$$
\left(\int_E |x(t)+y(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \leqslant \left(\int_E |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} + \left(\int_E |y(t)|^p \mathrm{d}t\right)^{\frac{1}{p}}
$$

Proof. 证只需证明 $p>1$ 的情形. 如果式 (2. 2. 4) 右边有一个积分为无穷, 则不等式 (2.2.4) 显然成立. 其次, 由于对任意数 a,b

$$
(|a|+|b|)^p \leqslant (2\max(|a|,|b|))^p \leqslant 2^p\left(|a|^p+|b|^p\right)
$$

则有

$$
\int_E |x(t)+y(t)|^p \mathrm{d}t \leqslant 2^p \left(\int_E |x(t)|^p \mathrm{d}t + \int_E |y(t)|^p \mathrm{d}t\right)
$$

由此, 如果式 (2.2.4) 左边为无穷则右边的积分至少有一个为无穷. 因此, 可以认为所有积分是有穷的. 应用 $Hlder$ 不等式, 并注意 $\frac{1}{p}+\frac{1}{q}=1$, 则有（利用 Holder 不等式）

$$
\begin{aligned}
\int_E |x(t)+y(t)|^p \mathrm{d}t &\leqslant \int_E |x(t)||x(t)+y(t)|^{p-1}\mathrm{d}t + \int_E |y(t)||x(t)+y(t)|^{p-1}\mathrm{d}t \\
&\leqslant \left(\int_E |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \left(\int_E |x(t)+y(t)|^{q(p-1)} \mathrm{d}t\right)^{\frac{1}{q}} \\
&\quad + \left(\int_E |y(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \left(\int_E |x(t)+y(t)|^{q(p-1)} \mathrm{d}t\right)^{\frac{1}{q}}
\end{aligned}
$$

所以

$$
\left(\int_E |x(t)+y(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \leqslant \left(\int_E |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} + \left(\int_E |y(t)|^p \mathrm{d}t\right)^{\frac{1}{p}}
$$

■

10.3.1 $L^p(E)(p \geqslant 1)$

现在我们研究空间 $L^p(E)$. 设 E 是 R 上的 *Lebesgue* 可测集, $x(t)$ 是 E 上的可测函数, $p \geqslant 1$. 如界 $|x(t)|^p$ 在 E 上可积. 称 $x(t)$ 是 E 上的 p 次幂可积函数. 用 $L^p(E)$ 表示所有 E 上 p 次列可积函数的全体, 其中两个几乎处处相等的函数可看作是同一元（这一步解决一个几乎处处等于零的函数就直接看成是零，这里引用了一个等价类的概念，也就是范数等于零则函数等于零）, 在 $L^p(E)$ 中按通常方式定义线性运算, $L^p(E)$ 是线性空间. 对于每一 $x \in L^p(E)$ 定义

$$\|x\| = \left(\int_{\varepsilon} |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}}$$

由 *Minkowski* 不等式, 它满足三角形不等式, 至于范数的另两条公理显然成立，所以 $L^p(E)$ 建一个赋范空间。

下面我们研究 $L^p(E)$ 的离散形式 $l^p(p \geqslant 1)$. 考些满足条件

$$\sum_{k=1}^{\infty} |\xi_k|^P < \infty$$

的数列 $x = \langle \varepsilon_k \rangle$ 的全体 $l^p(p \geqslant 1)$, 在其中按照坐标定义线性运算, ,l^p 是一个线性空间. 对于 $x \in l^p, x = \{\xi_k\}$ 定义 $\|x\| = (\sum_{k=1}^{\infty} |\xi_k|^p)^{\frac{1}{p}}$ 要证明这是一个范数，显然只需证明三角形不等式成立，为此可以与引理 2.2.3 类似地证明离散形的 *Minkowski* 不第式. 对任意 $x, y \in l^p, x = \{\xi_k\}, y \in \{\eta_k\}$, 有不等式

$$\left(\sum_{k=1}^{\infty} |\varepsilon_k + \eta_k|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |\eta_k|^p\right)^{\frac{1}{p}}$$

据此, 三角形不等式成立. 所以 $l^p(p \geqslant 1)$ 是一个赋范空间, 它是一个可分的 Banach 空间 (参香本率的习览).

R l^2 就是欧氏空间的无穷维推广。

Theorem 10.3.11 — 离散形式. $\sum_{n=1}^{\infty} |x_n y_n| \leq \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} \left(\sum_{n=1}^{\infty} |y_n|^q\right)^{1/q}$

Theorem 10.3.12
1. l^{∞} : 有界数列的全体.
2. c : 收敛数列 $x = (x_1, x_2, \cdots, x_n, \cdots)$ (即 $\lim_{n \to \infty} x_n$ 存在) 的全体.
3. c_0 : 收敛于 0 的数列全体.
4. $l^p : l^p = \left\{x : x = (x_1, x_2, \cdots), \sum_{n=1}^{\infty} |x_n|^p < \infty\right\} (1 \leqslant p < \infty)$

显然, $l^p \subset c_0 \subset c \subset l^{\infty} \subset \mathbb{R}^{\infty}$

Exercise 10.3 直线上有理点顶上画个球，无理点顶上画个球，两个分别作为集合，证明两个集合交集非空。 ■

度量空间	度量	点列按照度量收块 $\lim_{k\to\infty} x_k = x_0$
R^1	$\rho(x,y)=\|x-y\|$	$\Leftrightarrow$ 通常的数列收剑
R^n	$\rho(x,y)=\sqrt{\sum_{i=1}^n (x_i-y_i)^2}$	$\Leftrightarrow \lim_{k\to\infty} x_k^{(i)} = x_0^{(i)} (i=1,2,\cdots,n)$
R_1^n	$\rho_1(x,y)=\sum_{i=1}^n \|x_i-y_i\|$	$\Leftrightarrow \lim_{k\to\infty} x_k^{(i)} = x_0^{(i)} (i=1,2,\cdots,n)$
R_∞^n	$\rho_\infty(x,y)=\max_{1\le i\le n}\|x_i-y_i\|$	$\Leftrightarrow \lim_{k\to\infty} x_k^{(i)} = x_0^{(i)} (i=1,2,\cdots,n)$
R_p^n	$\rho_p(x,y)=\left(\sum_{i=1}^n \|x_i-y_i\|^p\right)^{\frac{1}{p}}$	$\Leftrightarrow \lim_{k\to\infty} x_k^{(i)} = x_0^{(i)} (i=1,2,\cdots,n)$
l^p	$\rho_p(x,y)=\left(\sum_{n=1}^\infty \|x_n-y_n\|^p\right)^{\frac{1}{p}}$	$\Rightarrow \lim_{k\to\infty} x_k^{(n)} = x_0^{(n)} (n=1,2,\cdots)$
l^∞	$\rho_\infty(x,y)=\sup_{1\le n<\infty}\|x_n-y_n\|$	$\Rightarrow \lim_{k\to\infty} x_k^{(n)} = x_0^{(n)} (n=1,2,\cdots)$
$C[a,b]$	$\rho(x,y)=\max_{a\le t\le b}\|x(t)-y(t)\|$	$\Leftrightarrow \lim_{k\to\infty} x_k(t)=x_0(t)$ 一致收敘
$C_p[a,b]$	$\rho(x,y)=\left(\int_a^b \|x(t)-y(t)\|^p dt\right)^{\frac{1}{p}}$	$\Leftrightarrow$ 函数列 $\{x_k\}$ $p-$ 方收敘手 x_0
平凡空间 X	$\rho(x,y)=\delta_{xy}$	$\Leftrightarrow$ 存在 $N, x_n = x_0($ 当$n \geqslant N$ 时$)$

Proof. 有理点构成的球记为 A，无理点构成的球记为 B。回忆疏朗集：取了闭包之后还是没有内点。不是疏朗集：取了闭包之后非空，也就是存在一个点为内部，根据内点的定义，存在一个小球落在里面。定义两个集合之间的稠密：取了闭包之后包含，例子有理数在无理数中稠密，取了闭包之后是直线，当然包含了无理数。不是疏朗集的定义：存在一个小球使得该集合在小球的内部是稠密的。闭包：包含的闭集合的交。是疏朗集：随便画点，构成的小球，疏朗集取了闭包也不能盖住小球。疏朗闭集合：没有内点，不能盖住任何一个球。如果一个闭集合只要能够盖住一个小球，则不是疏朗集。

回到本题：假设 $A\cap B=\varnothing$, 以球的半径大小来分类，对于无理数分类：将半径大于等于 1 的点记为集合 $\{x^{(1)}\}$, 第二步半径大于等于 $\frac{1}{2}$，记为 $\{x^{(2)}\}$，分析第一个集合，无理点上做球后，可以知道这部分球不是稠密的在任何开区间上，这是因为做完球后，限制了有理点做球的空间，不能无限接近做球。那么这接下去的继续做，只要限制了球的半径就行。回忆刚刚的结论：这批点都是疏朗集，无理点 = 至多可数的疏朗集的并。有理点 = 本来就是至多可数的，单点集合也是疏朗集。直线 = 至多疏朗集的并，但是纲领的结论是：完备的度量空间一定是第二纲纪，不是第一岗集，即可以写成至多可数个疏朗集的并，与直线上完备的矛盾。 ■

■ **Example 10.14** 第一章第十五题，可能是不可数的。 ■

Theorem 10.3.13 $L^p(E)(p\geqslant 1)$ 是 *Banach* 空间.

Proof. (证明的时候注意这里极限函数与 $\lim x_n$ 等价，证明式子中注意是 n 还是 m 取极限) 设 $\{x_n\}$ 是 $L^p(E)$ 中任意 *Cauchy* 列，从 $\{x_n\}$ 中可选出子列 $\langle x_{n_k}\}$, 使得

$$\left\|x_{n_{k+1}}-x_{n_k}\right\|<\frac{1}{2^k} \quad (k=1,2,\cdots)$$

由 *Hlder* 不等式, 对于每个具有有穷测度的可测集 $E_1\subset E$

$$\int_{E_1}\left|x_{n_{k+1}}(t)-x_{n_k}(t)\right|\mathrm{d}t\leqslant\left(m\left(E_1\right)\right)^{\frac{1}{q}}\left\|x_{n_{k+1}}-x_{n_k}\right\|$$

应用 *Fatou* 引理, 并注意式 (2.2.5)，(2.2.6)，则有

$$\begin{aligned}
&\int_{E_1}\lim_{n\to\infty}\sum_{k=1}^{n}\left|x_{n_{k+1}}(t)-x_{n_k}(t)\right|\mathrm{d}t\\
\leqslant&\lim_{n\to\infty}\int_{E_1}\sum_{k=1}^{n}\left|x_{n_{k+1}}(t)-x_{n_k}(t)\right|\mathrm{d}t\\
=&\sum_{k=1}^{\infty}\int_{E_1}\left|x_{n_{k+1}}(t)-x_{n_k}(t)\right|\mathrm{d}t<\infty
\end{aligned}$$

因此级数

$$|x_{n_1}(t)|+|x_{n_2}(t)-x_{n_1}(t)|+\cdots+\left|x_{n_{k+1}}(t)-x_{n_k}(t)\right|+\cdots$$

在 E_1 上几乎处处收敛. 但是 $E_1\subset E$ 是任意有穷测度可测子集. 所以实际上，它在 E 上几乎处处收敛，从而级数

$$x_{n_1}(t)+(x_{n_2}(t)-x_{n_1}(t))+\cdots+\left(x_{n_{k+1}}(t)-x_{n_k}(t)\right)+\cdots$$

在 E 上几乎处处收敛, 即 $\{x_{n_k}(t)\}$ 在 E 上几乎处处收敛, 设 $x_{n_k}(t)\to x(t)\quad(k\to\infty)\quad$ a. e. (实分析有结论：依测度收敛的 Cauchy 列存在收敛极限，这里由于依测度收敛，则存在处处收敛的子列) 我们证明 $x\in L^p(E)$ 并且 $\|x_n-x\|\to 0(n\to\infty)$. 由于 $\{x_n\}$ 是 *Cauchy* 列, 对任意 $\varepsilon>0$, 存在 N, 当 $m,n>N$ 时

$$\|x_n-x_m\|<\varepsilon$$

再次应用 *Fatou* 引理, 则有

$$\int_E\lim_{m\to\infty}|x_n(t)-x_m(t)|^p\,\mathrm{d}t\leqslant\lim_{m\to\infty}\int_E|x_n(t)-x_m(t)|^p\,\mathrm{d}t\leqslant\varepsilon^p$$

所以当 $n>N$ 时（这里用到了 Cauchy 的定义，上面的式子里有 m 趋于无穷，这里再加上 n 趋于无穷）

$$\begin{aligned}
\|x_n-x\|&\leqslant\varepsilon\\
x=(x-x_n)+x_n&\in L^p(E)
\end{aligned}$$

并且

$$\|x_n-x\|\to 0\quad(n\to\infty)$$

■

Exercise 10.4 闭球套来证明完备，注意每次的球之间的包含关系：第一次取半径为 1/2 时，但实际上半径取 1，是为了保证在边界上取 1/4 的球时能够保证套住。

Theorem 10.3.14 $L^p[a,b](p \geqslant 1)$ 是可分的

Proof. 证我们采取逐步逼近方式证明, 有理系数多项式全体是 $L^p[a,b]$ 中的可数稠密子集首先, 对于每一 $x \in L^p[a,b]$, 设 x_n 是 x 的截断函数, 即

$$x_n(t) = \begin{cases} x(t), & \text{当} |x(t)| \leqslant n \text{ 时} \\ 0, & \text{当} |x(t)| > n \text{ 时}, n = 1,2,\cdots \end{cases} \tag{10.6}$$

则 $x_n \in L^P[a,b]$ 且 $|x_n(t)| \leqslant n$, 由于

$$n^p m\{t : |x(t)| > n\} \leqslant \int_{\{t:|x(t)|>n)} |x(t)|^p \mathrm{d}t \leqslant \int_a^b |x(t)|^p \mathrm{d}t$$

$$\lim_{n\to\infty} m\{t : |x(t)| > n\} = 0$$

因此，由积分的绝对连续性, 对任意 $\varepsilon > 0$, 存在 N, 当 $n > N$ 时

$$\|x - x_n\| = \left(\int_{(t:|x(t)|>n)} |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} < \varepsilon$$

其次, 任取一满足上式的 n, 由 *Luzin* 定理, 存在连续函数 $y(t)$, 使得除去一个可测子集 A 之外, $x_n(t) = y(t)$, 并且可使 $mA \leqslant (\frac{\varepsilon}{2n})^p$, 且 $|y(t)| \leqslant n$. 这样我们有

$$\begin{aligned} \|x_n - y\| &= \left(\int_A |x_n(t) - y(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \\ &\leqslant \left(\int_A (2n)^p \mathrm{d}t\right)^{\frac{1}{p}} = 2n(mA)^{\frac{1}{p}} < \varepsilon \end{aligned}$$

最后，由 *Weierstrass* 定理, $y(t)$ 可由多项式一致逼近, 因此可选取有理系数多项式 $p(t)$, 使得在 $[a,b]$ 上

$$|y(t) - p(t)| < \frac{\varepsilon}{(b-a)^{\frac{1}{p}}}$$

从而 $\|y - p\| < \varepsilon$. 所以 $\|x - p\| < 3\varepsilon$

■

R 其中有一个小结论：依范数收敛可以推出以测度收敛。

R 之前几种收敛之间的反例

Theorem 10.3.15 我们研究 $p=\infty$ 的情形, 设 E 是 *Lebesgue* 可测集 ,$x(t)$ 是 E 上的可测函数, 如果存在可测子集 $E_0\subset E$, 使得 $mE_0=0$ 且 $x(t)$ 在 $E\backslash E_0$ 上是有界的, 称 $x(t)$ 在 E 上是本质有界的 (几乎处处有界). 用 $L^\infty(E)$ 表示 E 上本质有界可测函数全体按通常方式定义线性运算构成的线性空间. 同样地, 在 $L^\infty(E)$ 中两个几乎处处相等的欧数香作是同一元. 在 $L^\infty(E)$ 上定义

$$\|x\|_\infty=\inf_{\substack{mE_0=0\\ E_0CE}}\sup_{E/E_0}|x(t)|$$

零测集的选取可能有很多种，找那些上确界中最小的。我们首先证明式 (2.2.7) 中的下确界是可达到的, 即存在 E 的零测度子集 E_0 使得 $\|x\|=\sup_{E/E_0}|x(t)|$. 这是因为由下确界的定义, 存在一串零测集合，对于每一个 n, 存在 $E_n\subset E$ 使得 $mE_n=0$ 且

$$\sup_{E/E_n}|x(t)|<\|x\|+\frac{1}{n}$$

记 $E_0=\bigcup_{n=1}^\infty E_n$, 至多可数个零测集的并还是零测集，则 $E_0\subset E,mE_0=0$ 且对每一个 n

$$\|x\|\leqslant\sup_{E/E_0}|x(t)|\leqslant\sup_{E/E_n}|x(t)|<\|x\|+\frac{1}{n}$$

所以 $\|x\|=\sup_{E/E_0}|x(t)|$ 对于每一 $x\in L^\infty(E)$, 称 $\|x\|$ 为 $x(t)$ 在 E 上的本质上界, 记为 $\|x\|=\text{ess}_E\sup|x(t)|$ 不难验证, $||x||$ 是 $L^\infty(E)$ 上的范数, 在 $L^\infty(E)$ 中点列 $\langle x_n\}$ 按范收敛于 x 等价于函数列 $\{x_n(t)\}$ 在 E 上除去一个零测度集之外一致收敛于 $x(t)$. $L^\infty(E)$ 是一个不可分的 Banach 空间. 我们可以把 $L^\infty(E)$ 看作 $L^p(E)$ 的极限情形 (参看本章后面的习题)。

R 上面的证明太经典了，就是实分析里的常用方法。

R 不可分的例子 l^∞ 不可分。不可数多个点,任意两点还是 1,则不可分。$L^\infty[0,1]$, $x_\lambda'(t)=\begin{cases}1 & t\in[0,x]\\ 0 & t\in(\lambda,1]\end{cases}$

Theorem 10.3.16 — 可分. 1. n 维 Euclid 空间 R^n 按通常的距离是可分空间.

2. 空间 $l^p(1\leqslant p<\infty)$ 是可分的.
3. $C[a,b]$ 和 $L^p[a,b](1\leqslant p<\infty)$ 是可分空间.
4. 有界数列全体组成的空间 l^∞ 是不可分的。: 证明: 集合 $A=\{(x_1,x_2,\cdots):x_i=0,1\}$ 是 l^∞ 的子集, 它是不可数的, 并且对于 A 中的任意两个不同元素 x,y 有 : $\|x-y\|_\infty=1$, 所以 l^∞ 是不可分的.

Definition 10.3.4 设 X 是度量空间，A 是 X 的子集。如果 A 中的任何点列都有在 X 中收敛的子点列, 就称 A 是 (X 中的) 相对列紧集。如果 X 自身是相对列紧集，就称 X 是列紧空间。

Proposition 10.3.17
1. 有限点集是相对列紧集。
2. 有限个相对列紧集的并集是相对列紧集。
3. 相对列紧集的子集是相对列紧集，因此，任意一族相对列紧集的交集是相对列紧集。
4. 相对列紧集的闭包是相对列紧集。
5. 相对列紧集中的基本点列都收敛。因此, 列紧的度量空间是完备的
6. 相对列紧集是有界集
7. n 维 Euclid 空间 R^n 中的有界集是相对列紧集.

Theorem 10.3.18 $l^p(p \geqslant 1)$ 是 Banach 空间.

Proof. 令 N 是自然数全体，B 是 N 的子集全体，μ 是 B 上如下的测度：当 $M \in \boldsymbol{B}$ 时, $\mu(M) = M$ 中元素的个数. 当 $\{x_n\} \in l^p$ 时, 把它看成函数 $x(n) = x_n$ 那么 l^p 就可以看成 $l^p(N, \boldsymbol{B}, \boldsymbol{\mu})$. 由 L^p 定理就知道 l^p 是完备的. ■

Theorem 10.3.19 — (闭球套定理). 设 R 是完备的度量空间, 又设 $S_\nu = \{x|\rho(x, x_\nu) \leqslant \varepsilon_\nu\}$ 是 R 中的一套闭球:

$$S_1 \supset S_2 \supset \cdots \supset S_n \supset \cdots$$

如果球的半径 $\varepsilon_n \to 0$, 则必有唯一的点 $x \in \bigcap_{\nu=1}^{\infty} S_\nu$

Theorem 10.3.20 设 R 是度量空间, 如果在 R 上闭球套定理成立, 那么 R 必是完备的。

Definition 10.3.5 设 A 是度量空间 X 的一个子集,
1. 如果 A 的闭包 $\overline{A}$ 不含有 X 的非空开集，即 $\overline{A}$ 没有内点, 称 A 在 X 中是疏朗的。
2. 如果 $\bar{A} = X$, 称 A 在 X 中稠密.
3. 如果 X 可以表示成可数个疏朗集的并，就称 X 是第一纲的度量空间,
4. 否则，就称 X 为第二纲的度量空间。

。

■ **Example 10.15**
1. 如果 A 是疏朗集，那么 $\bar{A}$ 是疏朗的
2. 如果 A 是疏朗的，那么 A 的余集 $A^c = X \backslash A$ 是稠密的
3. 整数集 Z 作为直线的子集是疏朗集。
4. 集合 $A = \left\{\frac{1}{n} : n = 1, 2, \cdots\right\}$ 作为直线的子集是疏朗集.
5. 单点集作为直线的子集是疏朗的。
6. Cantor 集是疏朗集。
7. 赋范线性空间真的闭线性子空间是疏朗集。

8. 有理数集在直线中是稠密的。
9. 无理数集在直线中也是稠密的。
10. 有理数集作为度量空间是第一纲的，因为单点集在有理数集中也是疏朗集，有理数是可数集。
11. 对于平凡的度量空间，由于每个点都是开集，所以它的任何子集都不是疏朗的。

Theorem 10.3.21 (Baire) 完备度量空间必是第二类型的集。

Proof. 我们用反证法。设 R 是完备的度量空间, 而且是第一类型的. 下面我们要从这里推出矛盾。设 $R=\bigcup_{\nu=1}^{\infty} M_\nu$, 其中每个子集 M_ν 都是疏朗集. 任取一个球 $S(a,1)$, 由于 M_1 是疏朗的, 必有 R 中的非空闭球 $S(a_1,\rho_1)\subset S(a,1)$, 使得 $S(a_1,\rho_1)$ 中不含有 M_1 的点; 由于 M_2 是疏朗的, 必有 $S(a_2,\rho_2)\subset S(a_1,\rho_1)$ 一不妨取 $0<\rho_2<\frac{1}{2}$— 使得 $S(a_2,\rho_2)\cap M_2=\varnothing$, 如此可以选得一套非空闭球

$$S(a_1,\rho_1)\supset S(a_2,\rho_2)\supset\cdots,\quad S(a_\nu,\rho_\nu)\cap M_\nu=\varnothing$$

而且 $0<\rho_\nu<\frac{1}{\nu}$. 由引理 2, 通集 $\bigcap_{\nu=1}^{\infty} S(a_\nu,\rho_\nu)$ 中存在一点 x_0. 因为 $S(a_\nu,\rho_\nu)$ 和 M_ν 不相交, 所以 x_0 不在每个 M_ν 之中, $\nu=1,2,3,\cdots$, 因此 $x_0\bar{\epsilon}\bigcup_{\nu=1}^{\infty} M_\nu$, 但是 $\bigcup_{\nu=1}^{\infty} M_\nu=R$. 这是矛盾. 所以 R 不是第一类型的. ■

■ **Example 10.16** 把闭区间 [0,1] 看作完备空间 E^1 的闭子集是一完备子空间. 利用 Baire 定理可以给出区间 [0,1] 是不可列集的另一个证明。因为每个单元素集 $\{x\}$ 显然是 [0,1] 中的疏朗集. 由于 $[0,1]=\bigcup_{0\leqslant x\leqslant 1}\{x\}$. 因为 [0,1] 是完备空间, 所以. $\bigcup_{0\leqslant x\leqslant 1}\{x\}$ 不可能是可列和, 所以 [0,1] 是不可列的。

10.3.2 压缩映射

Definition 10.3.6 设 R 是度量空间, A 是 R 到它自身的一个映照。如果存在数 $\alpha, 0\leqslant\alpha<1$ 使得对一切 $x,y\in R$ 成立着

$$\rho(Ax,Ay)\leqslant\alpha\rho(x,y)$$

那么就称 A 是 R 上的一个压缩映照 (对于线性空间, 往往又称之为压缩算子)

R 一个点集经压缩映照后, 集中任意两点的距离经映照后被缩短了，至多等于原像距离的 $\alpha(\alpha<1)$ 倍. 压缩映照是连续的, 即对任何收敛点列 $x_n\to x_0$, 必有 $Ax_n\to Ax_0$. 事实上

$$\rho(Ax_n,Ax_0)\leqslant\alpha\rho(x_n,x_0)$$

当 $n\to\infty$ 时, 由 $\rho(x_n,x_0)\to 0$ 就得到 $\rho(Ax_n,Ax_0)\to 0$

Definition 10.3.7 设 R 为一集, A 是 R 到自身的映照. 如果 $x^* \in R$, 使得 $Ax^* = x^*$, 那么称 x^* 为映照 A 的一个不动点.

Theorem 10.3.22 — (Banach). 在完备的度量空间中的压缩映照必然有唯一的不动点。

Definition 10.3.8 — 列紧集. 设 R 是度量空间, A 是 R 中的集. 如果 A 中的任何点列必有在 R 中收敛的子点列, 就称 A 是 (R 中的) 致密集 (或称作列紧集). 如果 R 自身是致密集, 就称 R 是致密空间 (或列紧空间)

Proposition 10.3.23
1. 有限点集是致密集.
2. 有限个致密集的和集是致密集.
3. 致密集的任何子集是致密集, 因此, 任意一族致密集的交集是致密集.
4. 致密集的闭包是致密集.
5. 致密集中的基本点列必然收銅, 致密的度量空间是完备的.

Theorem 10.3.24 n 维欧几里得空间 E^n 中的有界集必是致密集.

Definition 10.3.9 设 A 是度量空间 R 中点集, B 是 A 的子集. 如果正数 ε, 使得以 B 中各点为心, 以 ε 为半径的开球全体覆盖 A, 即

$$\bigcup_{x \in B} O(x,\varepsilon) \supset A$$

那么称 B 是 A 的 ε -网. 如果对任何 $\varepsilon > 0$, 集 A 总有有限的 $\varepsilon -$ 网 $\{x_1, \cdots, x_n\} \subset A$ (点的个数 n 可以随 ε 而变), 那么称 A 是完全有界的集.

R 显然, 度量空间中集 A 的有限 ε -网概念正是有限个半径不超过 ε_n 的正方体覆盖 A 的抽象.

Proposition 10.3.25
1. 有限点集是完全有界集。
2. 有限个完全有界集的并集是完全有界集。
3. 完全有界集的子集是完全有界集，因此，任意一族完全有界集的交集是完全有界集.
4. 完全有界集的闭包是完全有界集.
5. 基本点列构成的集合是完全有界的
6. 完全有界集是有界集, 但是反过来不对.
7. n 维 Euclid 空间 R^n 中的有界集是完全有界的。事实上, 设 A 是 Euclid 空间 R^n 中的有界集, 则它的闭包 $\bar{A}$ 也是有界集，由上面的说明它是完全有界的。A 是 $\bar{A}$ 的子集也是完全有界的。

Theorem 10.3.26 集 A 是度量空间 R 中完全有界集的充要条件是 A 中任何一个点列 $\{x_n\}$ 必有一个基本的子点列.

相对列紧集	完全有界集
有限点集是相对列紧的	有限点集是完全有界集
有限个相对列紧集的并是相对列紧的	有限个完全有界集的并是完全有界集
相对列紧集的子集是相对列紧的	完全有界集的子集是完全有界集
任意个相对列紧集的交是相对列紧的	任意个完全有界集的交是完全有界集
相对列紧集的闭包是相对列紧的	完全有界集的闭包是完全有界集
相对列紧集是有界集	完全有界集是有界集
$\mathbb{R}^n$ 的有界集是相对列紧的	$\mathbb{R}^n$ 的有界集是完全有界集
列紧的度量空间是完备的	基本点列是完全有界集

Theorem 10.3.27 — (Hausdorff). 1. 度量空间中致密集必是完全有界集;

2. 在完备度量空间中, 完全有界集必是致密集.

Theorem 10.3.28 完全有界集必是有界集, 因而致密集必是有界集.

Theorem 10.3.29 设 R 是度量空间, 如果 R 中每个完全有界集都是致密集, 那么 R 必是完备的.

Theorem 10.3.30 完全有界集是可析的, 即其中含有有限的或可列的稠密子集。

Corollary 10.3.31 致密集是可析的.

Corollary 10.3.32 相对列紧集是可分的。

Theorem 10.3.33 设 B_n 是 n 维的赋范线性空间, $e_1, e_2, \cdots, e_n$ 是 B_n 的一个基, 那么必有正数 c_1, c_2, 使得对于 B_n 中的每个 $x = \sum_{\nu=1}^n x_\nu e_\nu$, 成立着

$$c_2\sqrt{\sum_{\nu=1}^n |x_\nu|^2} \leqslant \|x\| \leqslant c_1\sqrt{\sum_{\nu=1}^n |x_\nu|^2}$$

而且映照 $A : (x_1, x_2, \cdots, x_n) \mapsto \sum_{\nu=1}^n x_\nu e_\nu$ 是 n 维欧几里得空间 E^n 到 B_n 的拓扑映照。

Theorem 10.3.34 设在有限维线性空间 B_n 上定义了两个范数 $\|.\|$ 和 $\|.\|_1$, 那么必有常数 $M > 0$ 及 $K > 0$, 使得对于任何一点 $\psi \in B_n$ 成立着

$$K\|\psi\| \leqslant \|\psi\|_1 \leqslant M\|\psi\|$$

Definition 10.3.10 设 R 是一线性空间, $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 是在 R 上定义的两个范数. 如果存在正数 c_1 和 c_2, 使得对于每一点 $x \in R$, 有

$$c_1\|x\|_2 \leqslant \|x\|_1 \leqslant c_2\|x\|_2$$

就称范数 $||||_1$ 和 $||.||_2$ 是等价的。

Theorem 10.3.35 有限维赋范线性空间是完备的。

Theorem 10.3.36 任意赋范线性空间的有限维线性子空间是闭子空间

Theorem 10.3.37 有限维赋范线性空间中任何有界集是致密的。

Theorem 10.3.38 如果赋范线性空间 R 是无限维的, 那么 R 中必有不致密的有界集.

Theorem 10.3.39 设 E 是赋范线性空间 R 的闭子空间, 并且 $E \neq R$. 那么对于任一 $\varepsilon, 0 < \varepsilon < 1$, 必存在 R 中的单位向量 $x_0, \|x_0\| = 1$, 使得

$$\rho(x_0, E) > \varepsilon$$

Definition 10.3.11 — 赋范空间的子空间. 设 $(X, ||\cdot||)$ 是赋范线性空间, X_1 是 X 的线性子空间, 如果我们在 X_1 中取原来 X 上的范数, 那么 $(X_1, ||\cdot||)$ 是赋范空间, 称它为 $(X, ||\cdot||)$ 的(赋范)子空间. 不难证明, 赋范空间的任一完备子空间是闭子空间, *Banach* 空间的任一闭子空间是 *Banach* 空间.

■ **Example 10.17** 线性子空间例子：过原点，0 集合，大空间 ■

Theorem 10.3.40 完备空间的闭子空间是完备的，第一章的习题。完备子空间是闭子空间。这是度量空间的结论。第一章的结论在第二章都成立。

Theorem 10.3.41 Banach 空间 c

Proof. 设 c 是所有收敛数列全体,按坐标规定线性运算构成的线性空间,对于 $x \in c, x = \{\xi_k\}$, 定义

$$\|x\| = \sup_n |\xi_n|$$

易见,c 是赋范空间并且是 l^∞(参看例2.1.3 及本章习题) 的子空间, 为证明c 是 Banach 空间只需证明, c 是 Banach 空间 l^∞ 的闭子空间.

设 $\{x_n\}$ 是收敛于 $x_0 \in l^\infty$ 的 c 中点列 (要证明 x_0 也属于 c), 其中 $x_n = \left\{\xi_k^{(n)}\right\}, x_0 =$

$\left\{\xi_k^{(0)}\right\}$, 则对任意 $\varepsilon>0$, 存在 N, 当 $n\geqslant N$ 时

$$\|x_n-x_0\|=\sup_k\left|\xi_k^{(n)}-\xi_k^{(0)}\right|<\frac{\varepsilon}{3}$$

于是，当 $n\geqslant N$ 时, 对于每一个 k

$$\left|\xi_k^{(m)}-\xi_k^{(0)}\right|<\frac{\varepsilon}{3}$$

(坐标列，列于列之间的）取定 $n\geqslant N$, 由于 $x_n\in\left\{c,\xi_k^{(n)}\right\}$ 当 $k\to\infty$ 时收敛，因此对充分大的 k,k_1

$$\left|\xi_k^{(n)}-\xi_{k_1}^{(n)}\right|<\frac{\varepsilon}{3}$$

(这一步是判断一列中的尾部是 *Cauchy*) 从而

$$\left|\xi_k^{(0)}-\xi_{k_1}^{(0)}\right|\leqslant\left|\xi_k^{(0)}-\xi_k^m\right|+\left|\xi_k^{(n)}-\xi_{k_1^{(n)}}\right|+\left|\xi_k^{(n)}-\xi_{k_1}^{(0)}\right|$$

■

Analysis 10.3 要考虑两部分，列于列之间的，一个列的尾部的。

Theorem 10.3.42 *Banach* 空间 c_0 设 c_0 是所有收敛于 0 的数列的全体, 线性运算与范数定义与空间 c 相同. c_0 是 c 的子空间, c_0 是一个 *Banach* 空间 (参看本章习题).

R 只有收敛于零的才是线性子空间，收敛于 1 的不是。

Definition 10.3.12 商空间：从已知赋范空间构造新赋范空间的方法之一是构造商空间. 设 M 是线性空间 X 的线性子空间. 对于 $x_1,x_2\in X$, 如果 $x_1-x_2\in M$, 我们认为 $x_1\sim x_2$, 不难验证 $\sim$ 是等价关系, 对于 $x\in X$, 用 $\tilde{x}$ 表示以 x 为代表的等价类, $\widetilde{X}$ 表示所有 X 中元的等价类全体. 我们在 $\widetilde{X}$ 中定义线性运算:

$$\tilde{x}+\tilde{y}=\widetilde{x+y} \tag{10.7}$$

$$\alpha\tilde{x}=\widetilde{\alpha x} \tag{10.8}$$

这样定义的运算不依赖代表的选取. 事实上, 如果 $x,x_1\in\widetilde{X},y,y_1\in\widetilde{Y}$, 则 $x-x_1\in M,y-y_1\in M$, 因此 $(x+y)-(x_1+y_1)=(x-x_1)+(y-y_1)\in M\quad \alpha x-\alpha x_1=\alpha(x-x_1)\in M$ 即 $\widetilde{x+y}=\widetilde{x_1+y_1},\widetilde{\alpha x}=\alpha\tilde{x}_1$. 范数的定义需要是闭子空间，照顾定义。闭子空间的验证：按照定义，或者完备子空间一定是闭的。

Definition 10.3.13 — 赋范空间的乘积. 从已知赋范空间构造新赋范空间的另一个方法是构造乘积赋范空间. 设 $(X_1,||\cdot||_1),(X_2,||\cdot||_2)$ 是赋范空间. 在积集 $X_1\times X_2$ 中按坐标定义线性运第, 显然这时 $X_1\times X_2$ 是一个线性空间. 如果 $z\in X_1\times X_2,z=(x,y),x\in X_1,y\in$

X_2, 定义

$$\|z\| = \|x\|_1 + \|y\|_2$$

不难验证, 这时 $X_1 \times X_2$ 是赋范空间, 并且如果 X_1, 及 X_2 都是 *Banach* 空间, 则 $X_1 \times X_2$ 也是 *Banach* 空间.

Exercise 10.5 如果乘积是完备的，则单个是否是完备的 ■

Definition 10.3.14 — 赋范线性空间的基. 如果 X 是一个有穷维赋范空间, 例如 n 维的, 则在 X 中存在 n 个线性无关元 $\{e_1 \cdots, e_n\}$ 使得 X 中每一元 x 可惟一地表示为

$$x = \xi_1 e_1 + \cdots + \xi_n e_n$$

的形式. $\{e_1, \cdots, e_n\}$ 是线性空间 X 的一个基. 基的重要性是显而易见的.

对于一般无穷维线性空间 X, 我们有另外一种基的概念. 设 $\{x_a\}$ 是 X 的子集, $\{x_a\}$ 中元的所有有穷线性组合的全体显然是一个线性空间, 它是包含 $\{x_\alpha\}$ 的最小线性子空间, 称这个子空间为由 $\{x_\alpha\}$ 生成的子空间, 或者称它是 $\{x_\alpha\}$ 的线性包 (无穷集合做有限次)。

如果 $\{x_a\}$ 是 X 的线性无关子集且 $\{x_a\}$ 的线性包为 X. 则称 $\{x_\alpha\}$ 是 X 的 *Hamel* 基. 可以证明, 每一个线性空间都有 *Hamel* 基 (参考本章后面的习题)。如果 $\{ x_a\}$ 是 X 的 Hamel 基, 则每一 $x \in X$, 可惟一地表示为 $\{x_\alpha\}$ 中元的有穷线性组合的形式. *Hamel* 基可看作是有穷维线性空间中的基的一种推广。

Definition 10.3.15 设 X 是赋范空间, 以上基的概念都没有涉及空间的收敛性, 以下基的概念似乎更自然些. 设 $\{e_n\}$ 是 X 中的点列, 如果每一 $x \in X$ 可惟一地表示为

$$x = \sum_{k=1}^{\infty} \xi_k e_k, \quad \xi_k \in \mathrm{K}$$

的形式，则称 $\{e_n\}$ 是 X 的 *Schauder* 基。显然有穷维空间中通常的基是 *Schauder* 基.

■ **Example 10.18** 在空间 $l^p(p \geqslant 1)$ 中, 设 $e_1 = \{1, 0, \cdots\}, e_2 = \{0, 1, 0, \cdots\rangle, \cdots, e_n = \{0, \cdots, 0, 1, 0, \cdots\rangle, \cdots$ 则 $\{e_n\}$ 是一个 Schauder 基 ■

Proof. 因为对任 $x \in l^p, x = \{\xi_k\}, \sum_{k=1}^{\infty} |\xi_k|^p < \infty$ 于是

$$\left\| x - \sum_{k=1}^{n} \xi_k e_k \right\| = \left(\sum_{k=n+1}^{\infty} |\xi_k|^p \right)^{\frac{1}{p}} \to 0 \quad (n \to \infty)$$

即 $x = \sum_{k=1}^{\infty} \xi_k e_k$. 此外如果

$$x = \sum_{k=1}^{\infty} \xi_k e_k = \sum_{k=1}^{\infty} \xi'_k e_k \sum_{k=1}^{\infty} \left(\xi_k - \xi'_k\right) e_k = 0, \xi_k = \xi'_k, k = 1, 2, \cdots,$$

不难证实, 如果 *Banach* 空间 X 具有 *Schauder* 基, X 必须是可分的. 但是, 是否每一个可分 *Banach* 空间都具有 *Schauder* 基? 答案是否定的. ■

Exercise 10.6 找 c 中的 Schauder 基，或者是 l^∞ ■

Definition 10.3.16 设 $||\cdot||_1$ 与 $||\cdot||_2$ 是线性空间 X 上的两个范数，如果存在常数 a,b 使得对于每一 $x \in X$

$$a\|x\|_1 \leqslant \|x\|_2 \leqslant b\|x\|_1$$

称这两个范数等价。注意这里的范数控制是相互的

Definition 10.3.17 如果线性空间 X 上两个范数 $||\cdot||_1$ 与 $||\cdot||_2$ 等价, 那么这两个赋范空间代数同构拓扑同胚, 在这两个空间中关于收敛性是同样的.

Theorem 10.3.43 可以证明：两个等价范数,如果一个序列在一个范数的意义下是 Cauchy 列，则在另外一个下也是。

■ **Example 10.19** 欧式范数与积空间意义下的范数是等价的 $\frac{1}{\sqrt{2}}(|x|+|y|) \leq \sqrt{x^2+y^2} \leq |x|+|y|$ ■

在这一节中我们讨论有穷维赋范空间. 我们首先证明, 在代数同构与拓扑同胚意义下有穷维赋范空间只有一个, 即 R^n，其次通过紧性给出有穷维赋范空间一个特征性质.

R 线性同构：存在双射，原像的运算与像的运算一致。同胚：范数等价。

Theorem 10.3.44 任意 n 维赋范空间必与 R^n 代数同构拓扑同胚.

Proof. 证设 X 是任意 n 维赋范空间且 $\{e_1,\cdots,e_n\}$ 是这个空间的一个基. 于是对任意 $x \in X$, 可惟一地表示为

$$x = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$$

对于元 $x \in X$, 令

$$\bar{x} = (\xi_1, \xi_2, \cdots, \xi_n) \in \mathbb{R}^n$$

与之对应. 显然, 这样在 X 与 R^n 之间建立的映射是映上的并且是一对一的，它是 X 到 R^n 上的一个同构映射. 现在我们证明这个映射是同胚映射. 对于 $x \in X$, 我们有

$$\begin{aligned} \|x\| &= \left\| \sum_{k=1}^{n} \xi_k e_k \right\| \leqslant \sum_{k=1}^{n} |\xi_k| \, \|e_k\| \\ &\leqslant \left(\sum_{k=1}^{n} \|e_k\|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} |\xi_k|^2 \right)^{\frac{1}{2}} = \beta \|\bar{x}\| \end{aligned}$$

第二步是范数的三角不等式，然后是两个数相乘，Holder 不等式共轭数取 2. 其中常数 β 不依赖 x. 找到了一边的整数。可以得到 $T: R^n \mapsto X$ 是连续的。

另一方面, 在空间 R^n 的单位球面 $S=\left\{(\xi_1,\cdots,\xi_n)\in\mathbb{R}^n:\sum_{k=1}^n|\xi_k|^2=1\right\}$ 上考虑函数, 单位球面是有界闭集，所以是紧致集 + 映射是连续的，故一定可以取到最大最小值。

$$f(\bar{x})=f(\xi_1,\xi_2,\cdots,\xi_n)=\|x\|=\|\xi_1e_1+\xi_2e_2+\cdots+\xi_ne_n\|$$

因为在 S 上 ξ_k 不能同时为 0 且 $\{e_1,\cdots,e_n\}$ 线性无关, 所以

$$f(\xi_1,\xi_2,\cdots,\xi_n)>0$$

由于

$$\begin{aligned}|f(\xi_1,\xi_2,\cdots,\xi_n)-f(\eta_1,\eta_2,\cdots,\eta_n)|&=|\|x\|-\|y\||\\&\leqslant\|x-y\|\leqslant\beta\|\bar{x}-\bar{y}\|\end{aligned}$$

$f(\xi_1,\xi_2,\cdots,\xi_n)$ 是连续函数, 而 S 是 R^n 中紧集, 因此 f 在 S 上有最小值 $\alpha>0$ 由此对于每一个 $\bar{x}\in S$(考虑球面和球的内部)

$$f(\bar{x})=\|x\|\geqslant\alpha$$

所以对于每一 $x\in X$, 且 $x\neq 0$

$$f(\bar{x})=\|x\|=\|\bar{x}\|\left\|\frac{\sum_{k=1}^n\xi_ke_k}{\left(\sum_{k=1}^n|\xi_k|^2\right)^{\frac{1}{2}}}\right\|\geqslant\alpha\|\bar{x}\|$$

■

R 在任意有穷维赋范空间中点列收敛等价于按坐标收敛; 任意有穷维赋范空间是 Banach 空间; 在任意有穷维赋范空间中有界集是列紧集

■ **Example 10.20** 一个度量空间的子集合 E 是紧致的，则 E 外部一点到 E 的距离是可以取到的，但是如果当 E 为有界闭集时，结果不成立。 ■

Proof. 这是由于度量空间中的距离函数是连续函数，连续函数 + 紧致集和的最大最小可以取到。反例子：l^∞:$E=\{(0,1,1,\cdots),(1,0,1,1,\cdots),(1,1,\cdots,0,1,\ldots)\}$ 在单位球面上，有界闭集，两两个距离是 1，所以没有极限点，故是闭集。取 $x=\left(1+1,1+\frac{1}{2},\cdots,1+\frac{1}{n},\cdots\right)$, 这个点是有界数列，到 E 中的第一个距离是 2，第二个是 1.5，所以距离是 1，但是去不到 1。 ■

引理 2.4.2(F. Riesz)　设 X_0 是赋范空间 X 的真闭子空间, 则对任意 $\varepsilon>0$, 存在 $x_0\in X$, 使得 $\|x_0\|=1$ 且对于每一 $x\in X_0$

$$\|x-x_0\|>1-\varepsilon$$

证任取 $x_1 \in X\backslash X_0$, 记

$$d = \inf_{x \in X_0} \|x_1 - x\|$$

因为 X_0 是 X 的闭子空间, 所以 $d > 0$. 因为如不然则存在 $x_n \in X_0$, 使得 $x_n \to x_1 (n \to \infty)$, 从而 $x_1 \in X_0$, 矛盾. 不妨设 $\epsilon < 1$, 于是 $\frac{d}{1-\epsilon} > d$, 由 d 的定义, 存在 $x_2 \in X_0$, 使得

$$\|x_1 - x_2\| < \tfrac{d}{1-\varepsilon}$$

令

$$x_0 = \frac{x_1 - x_2}{\|x_1 - x_2\|}$$

则 $\|x_0\| = 1$, 并且对任意 $x \in X_0$

$$\begin{aligned} \|x - x_0\| &= \left\| x - \frac{x_1 - x_2}{\|x_1 - x_2\|} \right\| \\ &= \frac{1}{\|x_1 - x_2\|} \|(\|x_1 - x_2\| x + x_2) - x_1\| \\ &\geqslant \frac{d}{\|x_1 - x_2\|} > 1 - \varepsilon \end{aligned}$$

Exercise 10.7 — 期末考. 两个度量空间双射已定义，如果 cauchy 一致（一个是另外一个也是），则两个度量空间同构。在赋范空间中，取两个范数。若这两个范数下收敛一致，则范数等价

有限维的空间一定是完备的，闭的。

注意有界集合的刻画：直径，半径为 m 的球盖住，或者在赋范空间用范数来度量，注意零向量在度量空间没有

区分有界算子的定义和线性算子有界的定义前者是把有界集合映射成有界集合后者是像集合的范数小于原像的几倍

赋范空间才有零向量，零向量的像一定是零向量

注意：根据线性算子有界定义推导出

有界算子的定义只用到了线性算子的数乘的性质，

故对于不满足线性算子的第一条性质的算子依旧可以使用这种有界的定义，只用到了非负数数乘可以提出。

线性算子只需要讨论单位圆上的点即可，再利用线性性 + 单位化

求范数上确界的方法：找到一个元素使其达到上界 M，方法二：找到一串 x 的范数为 1，的像的范数，使其的极限可以趋近于一个 M。

10.4 有界线性算子

Definition 10.4.1 设 Λ 是实数或复数域, X 及 Y 是域 A 上的两个线性空间, D 是 X 的线性子空间, T 是 D 到 Y 中的一个映照, 对 $x \in D$, 记 x 经 T 映照后的像为 Tx 或者 $T(x)$. 如果对任何 $x,y \in D$ 及数 $\alpha,\beta \in \Lambda$, 成立着

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty$$

就称 T 是**线性算子**, 称 D 是 T 的定义域, 也记为 $\mathscr{D}(T)$. 而称集 $TD = \{Tx | x \in D\}$ 是 T 的值域 (或像域), 记为 $\mathscr{R}(T)$. 取值为实数或复数的线性算子 T (即 $\mathscr{R}(T) \subset \Lambda$) 分别称作实的或复的线性泛函, 通称为线性泛函.

Theorem 10.4.1 — 一点连续，则点点连续. 设 T 是赋范线性空间 X 到赋范线性空间 Y 的线性算子. 假如 T 在某一点 $x_0 \in \mathscr{D}(T)$ 连续, 那么在 $\mathscr{D}(T)$ 上处处连续.

Theorem 10.4.2 有限维赋范空间中一切线性算子是连续的。

Theorem 10.4.3 — 有界算子--有界集合映成有界集合. 如果算子 T 将其定义域 $\mathscr{D}(T)$ 中的每个有界集映照成一个有界集, 就称 T 是有界算子. 不是有界的算子就称为无界算子.

Definition 10.4.2 — 有界算子的定义.

1. 有界算子--有界集合映成有界集合：如果算子 T 将其定义域 $\mathscr{D}(T)$ 中的每个有界集映照成一个有界集, 就称 T 是有界算子. 不是有界的算子就称为无界算子.
2. 设 T 是赋范线性空间 X 到赋范线性空间 Y 的线性算子. 那么 T 是有界算子的充要条件是存在常数 $M \geqslant 0$, 使得对一切 $x \in X$

$$\|Tx\| \leqslant M\|x\|$$

Definition 10.4.3 — 保范算子. 设 X,Y 是两个赋范线性空间. U 是 X 到 Y 的映照, 而且对一切 $x \in X$, 有 $\|Ux\| = \|x\|$, 那么称 U 是 X 到 Y 的一个**保范算子**. 如果 U 不但是保范的, 又是线性的, 而且还实现 X 到 Y 上的一一对应, 那么我们就称 U 是 X 到 Y 上的 (保范) 同构映照。如果空间 X,Y 之间存在一个从 X 到 Y 上的 (保范) 同构映照, 我们就称 X 和 Y 同构。

Definition 10.4.4 — 共轭空间. 设 X 是赋范线性空间. X 上的连续线性泛函全体记作 X^*, 它按通常的线性运算及泛函的范数作范数构成一个赋范线性空间, 称为 X 的共轭空间.

Definition 10.4.5 有界算子的范数：

1. $\|T\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}$
2. $\|T\| = \inf\{M : \|Tx\| \leq M\|x\|, \quad \text{for all } x \in X\}$
3. $\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} = \sup\{\|Tx\| : \|x\| = 1\} = \sup\{\|Tx\| : \|x\| \leq 1\}$

Theorem 10.4.4 1. l^1 的共轭空间 $\left(l^1\right)^*$ 是 l^∞
2. $l^p(1<p<+\infty)$ 的共轭空间 $\left(l^p\right)^*$ 是 $l^q\left(\frac{1}{q}+\frac{1}{p}=1\right)$
3. $\left(l^2\right)^*=l^2$
4. (l^∞) * 并不就是 l^1, 而 c_0 的共轭空间才是 l^1
5. $L^p[a,b](1\leqslant p<\infty)$ 的共轭空间是 $L^q[a,b]\left(\frac{1}{p}+\frac{1}{q}=1\right)$
6. $\left(L^2[a,b]\right)^*=L^2[a,b]$

R 求证算子是有界的，主要是分离出 x 的范数。

■ **Example 10.21 — 无界算子的例子.** $C[0,1]$ 中考虑

$$Tx(t)=x'(t)$$

事实上, 取$x_n(t)=\sin nt\quad(n=2,3,\cdots)$, 则 $\|x_n\|=1$, 但是 $\|Tx_n\|=n$

$\|\cos nt\|=n\rightarrow\infty\quad(n\rightarrow\infty)$. 可见 T 把定义域中单位球上的元映为 $C[0,1]$ 中的无界集, 所以 T 是无界的.

■ **Example 10.22** 设 $k(t,s)$ 是 $a\leqslant t\leqslant b,a\leqslant s\leqslant b$ 上的连续函数.

$$Tx(t)=\int_a^b k(t,s)x(s)\mathrm{d}s\quad(x\in C[a,b])$$

显然 T 是 $C[a,b]$ 上到 $C[a,b]$ 中的线性算子. 由于

$$\|Tx\|=\max_{a\leqslant t\leqslant b}\left|\int_a^b k(t,s)x(s)\mathrm{d}s\right| \tag{10.9}$$

$$\leqslant\left(\max_{a\leqslant t\leqslant b}\int_a^b|k(t,s)|\mathrm{d}s\right)\|x\|=\beta\|x\| \tag{10.10}$$

其中, $\beta=\max_{a\leqslant t\leqslant b}\int_a^b|k(t,s)|\mathrm{d}s$. 由此可知 ,$T$ 是有界算子. 我们证明

$$\|T\|=\max_{a\leqslant t\leqslant b}\int_a^b|k(t,s)|\mathrm{d}s$$

由式10.10，只需证明 $\|T\|\geqslant\beta$. 由于 $\int_a^b k(t,s)\mathrm{d}s$ 是 t 的连续函数，所以存在 $t_0\in[a,b]$, 使得

$$\beta=\int_a^b\left|k\left(t_0,s\right)\right|\mathrm{d}s$$

取 $Z_0(s)=\operatorname{sgn}k\left(t_0,s\right)$, 则 $Z_0(s)$ 可测且 $|Z_0(s)|\leqslant 1$. 由 Luzin 定理, 对于每一个自然数 n, 存在 $[a,b]$ 上的连续函数 $x_n(t)$, 使得 $|x_n(t)|\leqslant 1$, 并且除去一个测度小于 $\frac{1}{2Mn}$ 的可测集 E_n

之外, 在 $[a,b]/E_n$ 上 $x_n(s)=Z_0(s)$, 其中 $M=\max_{a\leqslant t,s\leqslant b}|k(t,s)|$. 于是

$$\beta=\int_a^b|k(t_0,s)|\mathrm{d}s=\left|\int_a^b k(t_0,s)Z_0(s)\mathrm{d}s\right|$$
$$\leqslant\left|\int_a^b k(t_0,s)x_n(s)\mathrm{d}s\right|+\int_a^b|k(t_0,s)||Z_0(s)-x_n(s)|\mathrm{d}s$$
$$\leqslant\|T\|\|x_n\|+2M\cdot mE_n<\|T\|+\frac{1}{n}$$

令 $n\to\infty$, 则有$\beta\leqslant\|T\|$. 再由式 (10.10), $\|T\|=\beta$

Definition 10.4.6 — 范子列按范收敛等价于一致收敛. 在空间 $\mathscr{B}(X,X_1)$ 中按范收敛等价于算子列在 X 中的单位球面上一致收敛. 事实上, 设 $A,A_n\in\mathscr{B}(X,X_1)\quad(n=1,2,\cdots)$ 及 $S=\{x\in X:\|x\|=1\}$. 如果 $A_n\to A\quad(n\to\infty)$. 则对任意 $\varepsilon>0$, 存在 N, 当 $n>N$ 时, 对于每一个 $x\in S$

$$\|A_nx-Ax\|\leqslant\sup_{\|x\|=1}\|A_nx-Ax\|=\|A_n-A\|<\varepsilon$$

即 $\{A_n\}$ 在S 上一致收敛于A.

Theorem 10.4.5 设 X 是赋范空间, X_1 是 Banach 空间, 则 $\mathscr{B}(X,X_1)$ 是 Banach 空间.

R 所有的共轭空间都是完备的

Proof. 设 $\{T_n\}$ 是 $\mathscr{B}(X,X_1)$ 中任意 Cauchy 列, 则对任意 $\varepsilon>0$, 存在 N, 当 $m,n>N$ 时, $\|T_n-T_m\|<\varepsilon$ 于是对任意 $x\in X$

$$\|T_nx-T_mx\|\leqslant\|T_n-T_m\|\|x\|<\varepsilon\|x\|$$

由此可知, 对于每一 $x\in X$, $\{T_nx\}$ 是 X_1 中的 Cauchy 列, 由于 X_1 是完备的, 所以存在 $y\in X_1$, 使得

$$T_nx\to y\quad(n\to\infty)$$

这样, 对于每一 $x\in X$, 有 $y\in X_1$ 与之对应. 令 $Tx=y$ 表示这个对应关系, 显然，T 是 X 上取值于 X$_1$ 中的线性算子. 此外, 由于

$$|\|T_n\|-\|T_m\||\leqslant\|T_n-T_m\|\to 0\quad(n,m\to\infty)$$

$\{\|T_n\|\}$ 是 Cauchy 数列, 所以存在 $M>0$, 使得 $\|T_n\|\leqslant M\quad(n=1,2)$. 因此对于每一 $x\in X$

$$\|Tx\|=\left\|\lim_{n\to\infty}T_nx\right\|=\lim_{n\to\infty}\left\|T_nx\right\|\leqslant M\|x\|$$

即 T 是有界算子, 从而 $T\in\mathscr{B}(X,X_1)$. 最后，由于对于每一 $x\in X$, 当 $n>N$ 时

$$\|T_n-T\|\leqslant\varepsilon$$

即 $T_n\to T\quad(n\to\infty)$, $\mathscr{B}(X,X_1)$ 是 Banach 空间. ■

Definition 10.4.7 设 $T, T_n \in \mathscr{B}(X, X_1) \quad (n = 1, 2,)$. 如果对于每一 $x \in X$

$$T_n x \to Tx \quad (n \to \infty)$$

称 $\{T_n\}$ 逐点收敛于 T 或 $\{T_n\}$ 强收敛于 T

R 一致收敛一定强收敛

Theorem 10.4.6 如果 X, X_1 都是 Banach 空间, 那么空间 $\mathscr{B}(X, X_1)$ 在强收敛意义下也是完备的.

Theorem 10.4.7 — (Banach-Steinhaus)--每一点有界则一致有界. 设 $\{T_\alpha\}(\alpha \in I)$ 是 Banach 空间 X 上到赋范空间 X_1 中的有界线性算子族, 如果对于每一 $x \in X$, $\sup_{\alpha \in I} \|T_\alpha x\| < \infty$, 则 $\{\|T_\alpha\|\}(\alpha \in I)$ 是有界集.

R 对于每一 $x \in X$, $\{T_\alpha x\}(\alpha \in I)$ 是算子族 $\{T_\alpha\}(\alpha \in I)$ 在 x 点的"轨道", 因此 Banach-Steinhaus 定理说, 如果 Banach 空间 X 上的有界线性算子族 $\{T_\alpha\}$ $(\alpha \in I)$ 在每一点 $x \in X$ 轨道有界, 则算子族一致有界, 即存在常数 M, 使得

$$\|T_\alpha\| \leqslant M \quad (\alpha \in I)$$

Proof. 设

$$p(x) = \sup_{\alpha \in I} \|T_\alpha x\| \quad (x \in X)$$

及对每一个自然数 k

$$M_k = \{x \in X : p(x) \leqslant k\} = \bigcap_{\alpha \in I} \{x \in X : \|T_\alpha x\| \leqslant k\}$$

因为每一个 T_α 是有界线性算子, $\|T_\alpha x\|$ 是 x 的连续函数, 因此对于每一个 $\alpha \in I$, $\{x \in X : \|T_\alpha x\| \leqslant k\}$ 是 X 中的闭集, 从而每一个 M_k 是闭集. 由给定条件可知

$$X = \bigcup_{k=1}^{\infty} M_k$$

因为 X 是 Banach 空间, 由 Baire 纲定理, X 是第二纲集, 必存在 k_0, 使得 M_{k_0} 在某个闭球 $\bar{S} = \{x \in X : \|x - x_0\| \leqslant r_0\}$ 中稠密, 所以

$$\bar{S} \subset \bar{M}_{k_0} = M_{k_0}$$

任取 $x \in X, x \neq 0$, 则 $x_0 \pm \frac{x}{\|x\|} r_0 \in \bar{S}$, 于是

$$\begin{aligned} p\left(\frac{2r_0 x}{\|x\|}\right) &= p\left(x_0 + \frac{x}{\|x\|} r_0 - x_0 + \frac{x}{\|x\|} r_0\right) \\ &\leqslant p\left(x_0 + \frac{x}{\|x\|} r_0\right) + p\left(\frac{x}{\|x\|} r_0 - x_0\right) \leqslant 2k_0 \end{aligned}$$

因此

$$p(x) \leqslant \frac{k_0}{r_0}\|x\| \quad (x \in X)$$

从而对于每一个 $\alpha \in I, \|T_\alpha\| \leqslant \frac{k_0}{r_0}$ ■

Theorem 10.4.8 设 $\{T_n\}$ 是赋范空间 X 上到 Banach 空间 X_1 中的有界线性算子列，如果：

1. $\{\|T_n\|\}$ 有界
2. 对于一个稠密子集 G 中的元 $x, \{T_n x\}$ 收敛

则 $\{T_n\}$ 强收敛于一个有界线性算子 T, 并且

$$\|T\| \leqslant \varliminf_{n\to\infty} \|T_n\|$$

Proof. 由条件 1)，存在常数 M, 使得 $\|T_n\| \leqslant M \quad (n = 1,2,\cdots)$. 任取 $x \in X$, 由于 G 在 X 中稠密, 对任意 $\varepsilon > 0$, 存在 $y \in G$, 使得

$$\|x - y\| < \frac{\varepsilon}{3M}$$

由条件 2)，$\{T_n y\}$ 收敛, 故存在 N, 当 $n > N$ 时对任意的 k

$$\|T_{n+k}y - T_n y\| < \frac{\varepsilon}{3}$$

于是，$\|T_{n+k}x - T_n x\| \leqslant \|T_{n+k}x - T_{n+k}y\| + \|T_{n+k}y - T_n y\|$

$$+ \|T_n y - T_n x\| < \frac{M\varepsilon}{3M} + \frac{\varepsilon}{3} + \frac{M\varepsilon}{3M} = \varepsilon$$

所以 $\{T_n x\}$ 是 X_1 中的 Cauchy 列, 由于 X_1 完备 ,$\{T_n x\}$ 收敛, 记

$$Tx = \lim_{n\to\infty} T_n x \quad (x \in X)$$

不难看出，T 是 X 上到 X $_1$ 中的线性算子, 并且由于

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leqslant \varliminf_{n\to\infty} \|T_n\| \, \|x\|$$

可知 T 有界且 $\|T\| \leqslant \varliminf_{n\to\infty} \|T_n\|$ ■

R 如果我们进一步假定 X 也是 Banach 空间，由 BanachSteinhaus 定理可知，条件 1),2) 是有界线性算子列 $\{T_n\}$ 强收敛于一个有界线性算子的充要条件.（Banach 保证了极限可以取到，之前不是完备空间不一定极限存在故只能取下极限）

Theorem 10.4.9 — (Fourier 级数的发散性). 用 $C_{2\pi}$ 表示数直线上以 2π 为周期的实值连续函数全体构成的线性空间,(这是 $c[-\pi,\pi]$ 的子空间，保证了起点终点的函数值相等，可以连起来，可以验证这个空间是 Banach 空间，希望可以在傅立叶展开中写成等号) 在 $C_{2\pi}$ 中定义

$$\|x\| = \max_{-\infty<t<\infty} |x(t)| \quad (x \in C_{2\pi})$$

则 $C_{2\pi}$ 是一个 Banach 空间. 对于 $x \in C_{2\pi}$, 设 x 的 Fourier 级数为

$$x(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt)$$

上面级数前 $n+1$ 项的部分和为

$$\begin{aligned}
&\frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt) \\
=&\frac{1}{\pi}\int_{-\pi}^{\pi} x(s)\left(\frac{1}{2} + \sum_{k=1}^{n} \cos k(s-t)\right) \mathrm{d}s \\
=&\int_{-\pi}^{\pi} x(s) K_n(s,t) \mathrm{d}s
\end{aligned}$$

其中 $K_n(s,t) = \frac{\sin\left(n+\frac{1}{2}\right)(s-t)}{2\pi \sin\frac{1}{2}(s-t)}$ 为 Dirichlet 核.

我们证明, 对任一点 $t_0 \in [-\pi,\pi]$, 存在 $x \in C_{2\pi}$, 使得 x 的 Fourier 级数在 t_0 点发散. 因为 $C_{2\pi}$ 中函数以2π 为周期, 不失一般性可设$t_0 = 0$. 此时算子变成了泛函，并且线性有界。

对于每个 n, 作 $C_{2\pi}$ 上的线性泛函

$$f_n(x) = \int_{-\pi}^{\pi} x(s) K_n(s,0) \mathrm{d}s$$

其中, $K_n(s,0) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{n} \cos ks$, 显然 $K_n(s,0)$ 连续. 因此 f_n 是有界的，利用与例 3.1.5 类似的方法可以证明

$$\|f_n\| = \int_{-\pi}^{\pi} |K_n(s,0)| \, \mathrm{d}s \quad (n = 1,2,\cdots)$$

由于

$$
\begin{aligned}
\int_{-\pi}^{\pi}|K_n(s,0)|\,\mathrm{d}s &= \int_0^{2\pi}|K_n(s,0)|\,\mathrm{d}s \\
&= \frac{1}{2\pi}\int_0^{2\pi}\frac{\left|\sin\left(n+\frac{1}{2}\right)s\right|}{\left|\sin\frac{1}{2}s\right|}\mathrm{d}s \\
&= \frac{1}{\pi}\int_0^{\pi}\frac{|\sin(2n+1)t|}{\sin t}\mathrm{d}t \\
&\geqslant \frac{1}{\pi}\int_0^{\pi}\frac{|\sin(2n+1)t|}{t}\mathrm{d}t \\
&= \frac{1}{\pi}\int_0^{(2n+1)\pi}\frac{|\sin u|}{u}\mathrm{d}u \to \infty \quad (n\to\infty)
\end{aligned}
$$

由 Banach-Steinhaus 定理可知，存在 $x_0 \in C_{2\pi}$, 使得 $\{f_n(x_0)\}$ 发散, 即 $x_0(t)$ 的 Fourier 级数在 $t=0$ 点发散。

设 T 是赋范空间 X 上到赋范空间 X_1 上的线性算子且存在常数 $m>0$, 使得

$$\|Tx\| > m\|x\| \quad (x \in X)$$

则 T 有有界逆算子 T^{-1} 证首先,T 一对一地把 X 映到 X_1 上. 因为如果 $Tx_1 = Tx_2$, 即 $T(x_1 - x_2) = 0,)$

$$m\|x_1 - x_2\| \leqslant \|T(x_1 - x_2)\| = 0$$

由此 $x_1 = x_2$, 这样 , 由前面的讨论, 存在逆线性算子 T^{-1}. 其次, 由式 (3.3. 4) 对任意 $y \in X_1$

$$\left\|T^{-1}y\right\| \leqslant \frac{1}{m}\left\|TT^{-1}y\right\| = \frac{1}{m}\|y\|$$

即 T^{-1} 是有界的.

R $\|T^n\| \leqslant \|T\|^n$

Theorem 10.4.10 设 X 是 Banach 空间, $T \in \mathscr{B}(X)$. 如果 $\|T\| < 1$, 则算子 $I-T$ 有有界逆算子, 并且

$$\left\|(I-T)^{-1}\right\| \leqslant \frac{1}{1-\|T\|}$$

Proof. 证考虑级数

$$\sum_{k=0}^{\infty} T^k = I + T + T^2 + \cdots + T^{n-1} + \cdots \tag{10.11}$$

记

$$S_n = \sum_{k=0}^{n-1} T^k$$

则对任意自然数 $m,n(m>n)$

$$\|S_m - S_n\| = \left\|\sum_{k=n}^{m-1} T^k\right\| \leqslant \sum_{k=n}^{m-1} \|T\|^k$$

由条件 $\|T\|<1$ 可知，$\{S_n\}$ 是 $\mathscr{B}(X)$ 中的 Cauchy 列, 因为 X 是 Banach 空间从而 $\mathscr{B}(X)$ 是 Banach 空间, 所以 $\{S_n\}$ 按算子范数收敛于一个有界线性算子, 即级数10.11 按算子范数收敛. 由于

$$(I-T)\left(I+T+\cdots+T^{n-1}\right) \tag{10.12}$$

$$=\left(I+T+\cdots+T^{n-1}\right)(I-T)=I-T^n \tag{10.13}$$

及

$$\lim_{n\to\infty}\|T^n\| \leqslant \lim_{n\to\infty}\|T\|^n = 0$$

在式10.12 两边令 $n\to\infty$, 则有

$$(I-T)\left(\sum_{k=0}^{\infty} T^k\right) = \left(\sum_{k=0}^{\infty} T^k\right)(I-T) = I$$

这说明, 算子 $I-T$ 有逆算子, 并且

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

由此我们还得到

$$\left\|(I-T)^{-1}\right\| = \left\|\sum_{k=0}^{\infty} T^k\right\| \leqslant \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1-\|T\|}$$

■

Theorem 10.4.11 — Banach 逆算子定理. 设 T 是 Banach 空间 X 上到 Banach 空间 X_1 上的一对一的有界线性算子, 则 T 的逆算子 T^{-1} 是有界算子.

Proof. 根据定理的条件, 逆算子 T^{-1} 存在并且是线性算子. 由定理 3.3. 4 的证明中的式 (3.3. 7), 存在 $\delta>0$, 使得 $T\bar{S}(0,1)\supset S_1\left(0,\frac{1}{2}\delta\right)$, 因此对任意 $y\in S_1\left(0,\frac{1}{2}\delta\right)$, $T^{-1}y\in\bar{S}(0,1)$. 对任意 $z\in X_1$, $\frac{\delta z}{4\|z\|}\in S_1\left(0,\frac{1}{2}\delta\right)$, 所以

$$\left\|T^{-1}z\right\| \leqslant \frac{4}{\delta}\|z\|$$

即 T^{-1} 是有界算子. ■

设线性空间 X 上的两个范数 $\|\|_1$ 及 $\|\|_2$ 都使 X 成为 Banach 空间, 并且存在常数 C, 使得

$$\|x\|_2 \leqslant C\|x\|_1 \quad (x\in X)$$

则 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价.

Proof. 设 I 是 X 上的恒等算子, 由给定条件，I 是 Banach 空间 $(X, ||||1)$ 上到 Banach 空间 $(X, \|\cdot\|_2)$ 上的一对一的有界线性算子, 由 Banach 逆算子定理, 存在常数 C_1 使得

$$\|x\|_1 \leqslant C_1\|x\|_2 \quad (x \in X)$$

$$||I^{-1}x|| = \|x\|_1 \leq \|I^{-1}\|\|x\|_2$$

所以 $||\cdot||_1$ 与 $||\cdot||_2$ 等价. ■

R 要验证 I 是连续的，即连续有界的。有条件可以知道由 1 到 2 的恒等映射是连续的，根据连续的定义

Theorem 10.4.12 — 开映射定理. 设 T 是 Banach 空间 X 上到 Banach 空间 X_1 上的有界线性算子, 则 T 是一个开映射。

Theorem 10.4.13 — (Banach 逆算子定理). 设 T 是 Banach 空间 X 上到 Banach 空间 X_1 上的一对一的有界线性算子, 则 T 的逆算子 T^{-1} 是有界算子.

Theorem 10.4.14 设线性空间 X 上的两个范数 $\|\cdot||_1$ 及 $\|\cdot\|_2$ 都使 X 成为 Banach 空间, 并且存在常数 C, 使得

$$\|x\|_2 \leqslant C\|x\|_1 \quad (x \in X)$$

则 $\|\cdot||_1$ 与 $\|\cdot||_2$ 等价

Proof. 设 I 是 X 上的恒等算子，由给定条件，I 是 Banach 空间 $(X, \|\cdot\|_1)$ 上到 Banach 空间 $(X, \|\cdot\|_2)$ 上的一对一的有界线性算子, 由 Banach 逆算子定理，存在常数 C_1 使得

$$\|x\|_1 \leqslant C_1\|x\|_2 \quad (x \in X)$$

所以 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价. ■

Definition 10.4.8 设 XX_1 是赋范空间，T 是 X 中到 X_1 中的线性算子. 考虑乘积赋范空间 $X \times X_1$. 记

$$G(T) = \{(x, Tx) \in X \times X_1 : x \in D(T)\}$$

称 $G(T)$ 为算子 T 的图像. 如果 $G(T)$ 是乘积赋范空间 $X \times X_1$ 中的闭集, 则称 T 是闭算子.

Theorem 10.4.15 — 闭算子的充要条件. 设 X,X_1 是赋范空间,T 是 X 中到 X_1 中的线性算子, 则 T 是闭算子, 当且仅当, 对任意 $\{x_n\}\subset D(T), x_n\to x$ 及 $Tx_n\to y(n\to\infty)$, 这里 $x\in X, y\in X_1$, 此时必有 $x\in D(T)$ 并且 $Tx=y$

Proof. 设 T 是闭算子，即 $(x,y)\in\overline{G(T)}$, 则存在 $\{x_n\}\subset D(T)$, 使得

$$(x_n,Tx_n)\to(x,y)\quad(n\to\infty)$$

于是

$$\|(x_n-x,Tx_n-y)\|=\|x_n-x\|+\|Tx_n-y\|\to 0\quad(n\to\infty)$$

从而 $x_n\to x, Tx_n\to y(n\to\infty)$. 如果定理中条件满足, 则 $(x,y)\in G(T)$, 即 T 是闭算子。

反之, 设 $\{\ x_n\}\subset D(T)$, 且$x_n\to x$ 及$Tx_n\to y(n\to\infty)$, 于是 $(x_n,Tx_n)\to(x,y)$. 如果 $G(T)$ 是闭集 , 则 $(x,y)\in G(T)$, 即 $x\in D(T)$ 且 $Tx=y$ ■

Corollary 10.4.16 在全空间上的有界 (连续) 线性算子一定是闭线性算子

R

1. 对于闭线性算子来说，在上述条件下, 极限运算可以和算子交换顺序。
2. 闭的线性算子与连续线性算子有很“类似”的性质
3. 在开映射定理中，T 连续的条件可改为 T 是闭算子. 即: $\boldsymbol{X},\boldsymbol{X}_1$ 是 Banach 空间, $\boldsymbol{T}$ 是在上的 $(\boldsymbol{TX}=\boldsymbol{X}_1)$ $\boldsymbol{T}$ 是闭算子, 则 $\boldsymbol{T}$ 是开映射.

■ **Example 10.23 — 十分重要的无界线性算子—微分算子是闭算子.** $X=C[0,1],\mathscr{D}(T)=C^1[0,1]\neq X$, 定义

$$\boldsymbol{T}:\mathscr{D}(\boldsymbol{T})\to\boldsymbol{C}[0,1],\quad \boldsymbol{T}=\frac{\mathrm{d}}{\mathrm{d}t}$$

则 $\boldsymbol{T}$ 是闭算子.

Analysis 10.4 要证 T 是闭算子, 即要证明: 由 $x_n\in\mathscr{D}(T), x_n\to x,\quad Tx_n=\frac{\mathrm{d}}{\mathrm{d}t}x_n\to y(n\to\infty)$ 可推出 $x\in\mathscr{D}(T)$ 且 $Tx=y$ ■

Proof. (1) 由于

$$\int_0^t x_n'(s)\mathrm{d}s=\int_0^t \mathrm{d}x_n(s)=x_n(t)-x_n(0)$$
$$\therefore \lim_{n\to\infty}\int_0^t x_n'(s)\mathrm{d}s=\lim_{n\to\infty}[x_n(t)-x_n(0)]=x(t)-x(0)$$

(因为 $x_n\to x$, 一致收敛可推出点点收敛). (2) 因为 $x_n'(s)\to y(n\to\infty)$ 是一致收敛 (按范数收敛)，所以积分和极限可以交换顺序，结合 (4.5.6) 式, 有

$$\begin{aligned}x(t)-x(0)&=\lim_{n\to\infty}\int_0^t x_n'(s)\mathrm{d}s\\&=\int_0^t\lim_{n\to\infty}x_n'(s)\mathrm{d}s=\int_0^t y(s)\mathrm{d}s\end{aligned}$$

即

$$x(t) = x(0) + \int_0^t y(s)\mathrm{d}s$$

于是 $x'(t) = y(t) \in C[0,1]$. 所以 $x(t) \in C^1[0,1]$, 且

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = y(t), \text{即} Tx = y$$

因而 T 是闭算子。但 T 是无界线性算子。 ■

R 从数学分析中函数项级数逐项求导的例子可以进一步地体会闭算子的性质。

■ **Example 10.24** 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 满足:

1. $u_n(x)(n = 1, 2, \cdots)$ 在 $[a,b]$ 上连续可导;
2. $\sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上点点收敛到 $S(x)$
3. $\sum_{n=1}^{\infty} u_n'(x)$ 在 $[a,b]$ 上一致收敛到 $\sigma(x)$ (由此可推出 $\sum_{n=1}^{\infty} u_n(x)$ 在$[a,b]$ 上一致收敛到$S(x)$)

则 $S(x) = \sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 可导, 且

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}u_n(x)$$

即微分运算可以与无限求和运算交换顺序。

从泛函分析的角度看，上述条件相当于: 在 $C[a,b]$ 空间中考虑,

1. 函数项级数前 n 项和 $S_n(x) = \sum_{k=1}^{n} u_k(x)$ 在 $\boldsymbol{C}[\boldsymbol{a},\boldsymbol{b}]$ 中按范数收敛到 $S(x)$
2. $\frac{\mathrm{d}}{\mathrm{d}x}(S_n(x)) = \sum_{k=1}^{n} \frac{\mathrm{d}}{\mathrm{d}x}(u_k(x))$ 在 $C[a,b]$ 中按范数收敛到 $\sigma(x)$

由于 $\frac{d}{\mathrm{d}x}$ 是闭算子, 于是有 $\sigma(x) = \frac{\mathrm{d}}{\mathrm{d}x}S(x)$, 即:

$$\sigma(x) = \sum_{n=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}u_n(x) = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=1}^{\infty} u_n(x)$$

成立, 微分可以和级数运算交换顺序。

1. 可与数学分析 (广义积分、含参变量积分) 中有关求导数和极限交换顺序的有关定理相对照,
2. 由于微分运算是闭算子 (不是有界线性算子), 这些定理中都有类似例 4.5.5 中条件 (3) 的要求,
3. 而对于积分和极限交换顺序则没有这样的要求。

■ **Example 10.25** 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 满足:

1. $u_n(x)(n = 1, 2, \cdots)$ 在 $[a,b]$ 连续;
2. $\sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上一致收敛到 $S(x)$

则 $S(x) = \sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 可积, 且 $\int_a^b S(x)\mathrm{d}x = \int_a^b \sum_{n=1}^{\infty} u_n(x)\mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b u_n(x)\mathrm{d}x$ 上述例子表明:

1. 在一致收敛 (即在 $C[a,b]$ 中按范数收敛) 的条件下, 积分运算可以与无限求和运算交换顺序。
2. 积分算子是有界 (连续) 线性算子,
3. 微分算子是无界线性算子，但是它是闭算子。
4. 微分与级数运算交换顺序比积分与级数运算交换顺序多了一个条件 (3)

Theorem 10.4.17 如何证明界
1. 一种是正好有一个点取到了上确界
2. 有一串逼近

$$(||K^n||)^* = \mathbb{K}^n$$

Theorem 10.4.18 — (闭图像定理). 设 T 是 Banach 空间 X 上到 Banach 空间 X_1 中的闭线性算子, 则 T 是有界算子.

Proof. 因为 X, X_1 都是 Banach 空间, 所以乘积赋范空间 $X \times X_1$ 是 Banach 空间. 由于 $G(T)$ 是 $X \times X_1$ 中的闭集及 $G(T)$ 是 $X \times X_1$ 的线性子空间, 从而 $G(T)$ 也是 Banach 空间. 定义从 $G(T)$ 上到 X 中的算子 $\widetilde{T}$:

$$\widetilde{T}(x, Tx) = x \quad (x \in X)$$

显然, $\widetilde{T}$ 是 $G(T)$ 上到 X 上的一对一的有界线性算子. 由 Banach 逆算子定理 $\widetilde{T}^{-1}$ 有界, 即

$$\|(x, Tx)\| = \left\|\widetilde{T}^{-1}x\right\| \leqslant \left\|\widetilde{T}^{-1}\right\| \|x\| \quad (x \in X)$$
$$\text{所以} \|Tx\| \leqslant \left\|\widetilde{T}^{-1}\right\| \|x\| (x \in X)$$

■

10.4.1 Hahn-Banach 定理

Definition 10.4.9 设 f 是线性空间 L 上非零的线性泛函, c 是给定常数，集合:

$$M_c = \{x : f(x) = c, x \in L\}$$

称为 L 的超平面。

Theorem 10.4.19 — (Zorn 引理). 设 $(\mathscr{F}, <)$ 是一个半序集, 如果 $(\mathscr{F}, <)$ 中的任意全序子集皆有上界, 则 $(\mathscr{F}, <)$ 中必有极大元.

定理 3. 4. 2 (实空间的 Hahn-Banach 定理) 设 M 是实线性空间 X 的线性子空间, $p : X \to \mathbb{R}$, 对任意 $x, y \in X$ 及 $\alpha \geqslant 0$ 满足

$$p(x+y) \leqslant p(x) + p(y); \quad p(ax) = \alpha p(x)$$

f 是 M 上的线性泛函且满足

$$f(x) \leqslant p(x) \quad (x \in M)$$

则存在 X 上的线性泛函 F，使得

$$F(x) = f(x) \quad (x \in M)$$

并且

$$-p(-x) \leqslant F(x) \leqslant p(x) \quad (x \in X)$$

证设 $M \neq X$, 任取 $x_1 \in X \backslash M$, 用 M_1 表示由 x_1 与 M 张成的线性子空间, 即

$$M_1 = \{x + \alpha x_1 : x \in M, \alpha \in \mathbb{R}\}$$

由于对任意 $x, y \in M$

$$f(x) + f(y) = f(x+y) \leqslant p(x+y) \leqslant p(x - x_1) + p(x_1 + y)$$

我们有

$$f(x) - p(x - x_1) \leqslant p(y + x_1) - f(y)$$

设 β 是上式左边当 x 取遍 M 中元的上确界, 则有

$$f(x) - \beta \leqslant p(x - x_1) \quad (x \in M) \tag{10.14}$$

及

$$f(y) + \beta \leqslant p(y + x_1) \quad (y \in M) \tag{10.15}$$

现在在 M_1 上定义

$$f_1(x + \alpha x_1) = f(x) + \alpha\beta \quad (x \in M, \alpha \in \mathbb{R})$$

不难看出, f_1 是 M_1 上的线性泛函并且在 M 上 $f_1 = f$. 取 $\alpha < 0$, 用 $-\frac{x}{\alpha}$ 代替式10.14中的 x 再用 $(-\alpha)$ 乘所得不等式的两边; 再取 $\alpha > 0$, 用 $\frac{y}{\alpha}$ 代替10.15中的 y, 然后用 α 乘所得不等式的两边，则有式

$$f_1(x + \alpha x_1) \leqslant p(x + a x_1) \quad (x \in M, \alpha \in \mathbb{R})$$

这样, 我们把 f 保持关系 $f \leqslant p$ 延拓到 M_1 上. 为了能够把 f 保持这种关系延拓到全空间上, 我们需要用 Zorn 引理: 用 $\mathcal{F}$ 表示 f 的保持 $f \leqslant p$ 延拓的全体, 在 $\mathcal{F}$ 中引进关系 <: 设 $F_1, F_2 \in \mathscr{F}$ $D(F_1)$ 与 $D(F_2)$ 分别是它们的定义域, 如果 $D(F_1) \subset D(F_2)$, 并且当 $x \in D(F_1)$ 时 ,$F_1(x) = F_2(x)$, 即 F_2 是 F_1 的延拓时, 定义 $F_1 < F_2$. 易见 < 是中的半序, 从而 ($\mathcal{F}$,<) 是一个半序集. 设 $\mathscr{G}$ 是 $\mathcal{F}$ 的任一全序子集，令

$$D = \bigcup_{F \in \mathscr{G}} D(F)$$

并在 D 上定义泛函 Φ : 任取 $x \in D$, 存在 $F \in \mathscr{G}, x \in D(F)$, 此时令 $\Phi(x) = F(x)$. 由于 $\mathscr{G}$ 是 $\mathscr{F}$ 的全序子集 ,D 是线性子空间, 且 Φ 在 D 上是惟一确定的线性泛函, 并且对 $x \in D$ 是满足 $\Phi(x) \leqslant p(x)$ 的 f 的延拓, 即 $\Phi \in \mathscr{F}$. 显然 Φ 是 $\mathscr{G}$ 一个上界. 由 Zorn 引理,$\mathscr{F}$ 中存在极大元 F_0, 这时必有 $D(F_0) = X$. 因为如果不然, 则存在 $x_0 \in X \backslash D(F_0)$. 由证明的第一步 , 可把 F_0 保持 $F_0 \leqslant p$ 延拓到由 x_0 与 $D(F_0)$ 张成的线性子空间上, 这显然与 F_0 的极大性矛盾.

Theorem 10.4.20 — (复) 赋范空间上的 Hahn-Banach 定理. 设 G 是复赋范空间 X 的子空间, f 是 G 上的有界线性泛函, 则 f 可保持范数不变延拓到全空间 X 上, 即存在 X 上的有界线性泛函 F, 使得

1. 对于 $x \in G, F(x) = f(x)$
2. $\|F\| = \|f\|_G$

这里 |f(x)$\|_G$ 表示 f 作为 G 上的有界线性泛函的范数.

Corollary 10.4.21 设 X 是赋范空间, 则对任意 $x_0 \in X, x_0 \neq 0$ 必存在 X 上的有界线性泛函 f, 使得

$$\|f\| = 1, \quad f(x_0) = \|x_0\|$$

Proof. 设 G 是由 $\{x_0\}$ 张成的线性子空间, 即

$$G = \{ax_0 : \alpha \in \mathbb{K}\}$$

在 G 上定义

$$f_0(\alpha x_0) = \alpha \|x_0\| \quad (\alpha \in \mathbb{K})$$

则 f_0 是 G 上的线性泛函, 并且当 $x = \alpha x_0$ 时

$$|f_0(x)| = |f_0(\alpha x_0)| = |\alpha| \|x_0\| = \|\alpha x_0\| = \|x\|$$

所以 $\|f_0\|_G = 1$. 于是由 H-B 定理, 存在 X 上的有界线性泛函 f, 使得

$$\|f\| = 1 \text{ 及 } f(x_0) = \|x_0\|$$

■

R 由这个推论可以看出, 如果赋范空间 $X \neq \{0\}$, 那么 X 上必存在非零有界线性泛函

Corollary 10.4.22 设 G 是赋范空间 X 的子空间, $x_0 \in X$, 如果

$$d = d(x_0, G) = \inf_{x \in G} \|x - x_0\| > 0$$

则存在 X 上的有界线性泛函 f, 使得

$$\|f\| = \frac{1}{d}; f(x_0) = 1; \quad f(x) = 0 \quad (x \in G)$$

Proof. 设 G_1 是由 x_0 及 G 张成的线性子空间, 即

$$G_1 = \{\alpha x_0 + x : \alpha \in \mathbb{K}, x \in G\}$$

在 G_1 上定义

$$f_1(\alpha x_0+x)=\alpha \quad (\alpha\in\mathbb{K},x\in G)$$

显然 f_1 是 G_1 上的线性泛函, 并且

$$f_1(x_0)=1;\quad f_1(x)=0\quad (x\in G)$$

由于

$$\|\alpha x_0+x\|=|\alpha|\left\|x_0+\frac{x}{\alpha}\right\|\geqslant|\alpha|d$$

所以

$$|f_1(\alpha x_0+x)|\leqslant\frac{1}{d}\|\alpha x_0+x\|$$

即 $\|f_1\|_{G_1}\leqslant\frac{1}{d}$ 另一方面，取 $x_n\in G\quad(n=1,2,\cdots)$, 使得 $\|x_n-x_0\|\to d(n\to\infty)$, 于是有

$$\begin{aligned}\|f_1\|_{G_1}\|x_n-x_0\|&\geqslant|f_1(x_n-x_0)|=|f_1(x_0)|=1\\ \|f_1\|_{G_1}&\geqslant\tfrac{1}{\|x_n-x_0\|}\quad(n=1,2,\cdots)\end{aligned}$$

在后面不等式的右边，令 $n\to\infty$, 则得 $\|f_1\|_{G_1}\geqslant\frac{1}{d}$, 所以 $\|f_1\|_{G_1}=\frac{1}{d}$. 由 H-B 定理，$f_1$ 可保持范数不变延拓成全空间 X 上的泛函 f 就是所要求的泛函。 ■

R 当 G 是 X 的闭子空间时，任取 $x_0\in X\backslash G$, 则 $d(x_0,G)>0$, 因此该推论的结论成立.

Corollary 10.4.23 满足 H-B 定理中条件 1）,2) 的延拓一般不是惟一的

Theorem 10.4.24 常见的凸集合

1. 在三维 Euclid 空间 R^3 中，线段、平面、三角形、立方体、球、四面体以及半空间等都是凸集。
2. $C[a,b]$ 的子集$A=\{f:|f|\leqslant 1\}$ 是凸集.
3. l^2 中的集合$A=\left\{x:x=(x_1,x_2,\cdots)\in l^2,\sum_{n=1}^{\infty}n^2x_n^2\leqslant 1\right\}$ 是乃集.
4. 线性空间 L 中的任意多个凸集的交仍是凸集。

共轭空间，共轭算子

Definition 10.4.10 设 X 是赋范空间, 记

$$X^*=\mathscr{B}(X,\mathbb{K})$$

称 X^* 是 X 的共轭空间 (或对偶空间). 即 X 的共轭空间是 X 上所有有界线性泛函构成的赋范空间.

R 任意赋范空间的共轭空间是 Banach 空间.

Definition 10.4.11 设 XX_1 是赋范空间, $T\in\mathscr{B}(X,X_1)$, 对于每一个 $f\in X_1^*$, 令

$$(T^*f)(x)=f(Tx)\quad(x\in X)$$

称 T^* 是 T 的共轭算子. 由以上定义算子 T 的共轭算子 T^* 是 X_1^* 上到 X^* 的线性算子。

Theorem 10.4.25 有界线性算子 T 的共轭算子具有以下性质:

1. T^* 是有界线性算子并且 $\|T^*\|=\|T\|$
2. 对于每一个 $\alpha\in\mathrm{K},(\alpha T)^*=\alpha T^*$
3. $(T_1+T_2)^*=T_1^*+T_2^*$
4. $(T_1T_2)^*=T_2^*T_1^*$
5. 设 T 有有界逆算子, 则 T^* 也有有界逆算子并且 $(T^*)^{-1}=(T^{-1})^*$

Definition 10.4.12

$$\|f\|_\infty=\inf_{\substack{\mu(E_0)=0\\E_0\subset E}}\left(\sup_{E-E_0}|f(x)|\right)=\operatorname*{ess\,sup}_{x\in E}|f(x)|$$

■ **Example 10.26** 空间 $L^p[a,b](1<p<\infty)$ 上的有界线性泛函设 f 是空间 $L^p[a,b]$ 上的有界线性泛函，则存在惟一的 $y\in L^q[a,b]$, 其中$\frac{1}{p}+\frac{1}{q}=1$, 使得

$$f(x)=\int_a^b x(t)y(t)\mathrm{d}t\tag{10.16}$$

并且

$$\|f\|=\|y\|=\left(\int_a^b|y(t)|^q\mathrm{d}t\right)^{\frac{1}{q}}$$

反之, 任意 $y\in L^q[a,b]$, 上式定义了 $L^P[a,b]$ 上的一个有界线性泛函.

在 $p=1$ 的情形, $L^1[a,b]$ 上的每一个有界线性泛函 f, 存在惟一的 $y\in L^\infty\,[a,b]$, 使得表达式10.16 成立并且 $\|f\|=\text{esssup}_{a\leqslant k\leqslant b}|y(t)|$

实际上, 在上述证明中对于引进的函数 x_s 及 $g(s)=f(x_s)$, 我们证明了 $g(s)$ 是绝对连续的. 此外对任意 $s_1,s_2\in[a,b]$, 由于 $x_s\in L^\infty[a,b]$, 同时

$$\begin{aligned}|g(s_2)-g(s_1)|&=|f(x_{s_2})-f(x_{s_1})|=|f(x_{s_2}-x_{s_1})|\\&\leqslant\|f\|\,\|x_{s_2}-x_{s_1}\|=\|f\|\,|s_2-s_1|\end{aligned}$$

$y(t)=g'(t)$, 可知 $|y(t)|\leqslant\|f\|$ a. e, 即 $y\in L^\infty[a,b]$ 并且 $\text{esssup}_{a\leqslant r\leqslant b}|y(t)|\leqslant\|f\|$. 反之，由表达式10.16得 $\|f\|\leqslant\text{esssup}_{a\leqslant k\leqslant b}|y(t)|$. 所以 $\|f\|=$

$$\text{esssup}_{a\leq t\leq b}|y(t)|.$$

因此

$$\left(L^1[a,b]\right)^*=L^\infty[a,b]$$

从以上的讨论,我们得到 $(L^p[a,b])^* = L^q[a,b]$, 其中 $\frac{1}{p}+\frac{1}{q}=1$. 如果我们约定 $p=1$ 时 $q=\infty$, 那么这个结论对任意 $p(1\leqslant p<\infty)$ 成立。

R 对于离散情形,类似地我们有 $(l^p)^* = l^q(1\leqslant p<\infty)$

■ **Example 10.27 — 空间 c 上的有界线性泛函.** 对于每一个 $f\in c^*$, 存在数 α 及 $\{\alpha_n\}\in l^1$, 使得对于每一个 $x\in c, x=\{\xi_n\}$

$$f(x)=\alpha\lim_{n\to\infty}\xi_n+\sum_{n=1}^{\infty}\alpha_n\xi_n$$

且

$$\|f\|=|\alpha|+\sum_{n=1}^{\infty}|\alpha_n|$$

反之,如果给定 α 及 $\{\alpha_n\}\in l^1$, 则上式决定了 c 上的一个有界线性泛函.

R $c^*=l^1$

■ **Example 10.28** c_0 的共轭空间 $(c_0)^*\cong l^1$ c_0 是收敛于零的数列 $x=\{x_n\}$ 的全体, 按照通常的线性运算和范数 $\|x\|=\sup_n|x_n|$ 所成的 Banach 空间在 c_0 中取 $\mathfrak{F}$ 为一列“单位”向量 $e_n=(0,0,\cdots,0,1,0,\cdots)(n=1,2,\cdots)$(第$n$ 个位置为 1, 其他为 0), 则对任何 $x=(x_1,x_2,\cdots)\in c_0$, 显然

$$x=\lim_{n\to\infty}\textstyle\sum_{i=1}^n x_ie_i$$

设$f\in(c_0)^*$, 记$\alpha_i=f(e_i)(i=1,2,\cdots)$, 则$f(x)=\lim_{n\to\infty}\sum_{i=1}^n x_i\alpha_i$. 作$c_0$ 中的, 点列 $\left\{x^{(m)}\right\}, x^{(m)}=\left(x_1^{(m)},x_2^{(m)},\cdots,x_m^{(m)},0,0,\cdots\right)$, 其中

$$x_i^{(m)}=\begin{cases} e^{-\sqrt{-1}\theta_i}, & \theta_i \text{ 是}\alpha_i \text{ 的幅角}(i\leqslant m)\\ 0, & i>m\end{cases}$$

则

$$\sum_{i=1}^{m}|\alpha_i|=f\left(x^{(m)}\right)\leqslant\|f\|$$

所以 $\alpha\in l^1$, 并且 $\|\alpha\|_1\leqslant\|f\|$ 定义从 $(c_0)^*$ 到 l^1 的映照 $U:f\mapsto\alpha=(\alpha_1,\alpha_2,\cdots)$, 其中 $\alpha_i=f(e_i)$. 显然, U 是线性映照, 易知 U 将非零元 f 映照成非零元 $\alpha=Uf$, 并且 $\|Uf\|_1\leqslant\|f\|$ 由 $f(x)=\lim_{n\to\infty}\sum_{i=1}^n x_i\alpha_i$, 得到

$$|f(x)|=\lim_{n\to\infty}\left|\sum_{i=1}^{n}x_i\alpha_i\right|\leqslant\|x\|\|\alpha\|_1$$

手是, $||\ f\|\leqslant\|Uf\|_1$, 所以 U 是保范的. 另一方面, 对手任意的 $\alpha=(\alpha_1,\alpha_2,\cdots)\in l^1$, 定义 c_0 上的泛函:

$$f(x)=\sum_{n=1}^{\infty}x_n\alpha_n$$

由手, $|f(x)| \leqslant \sum_{n=1}^{\infty} |x_n\alpha_n| \leqslant \|x\|\|\alpha\|_1$, 所以 f 是 c_0 上的有界线性泛函, 并且 $Uf = \alpha$, 即 U 是到上的。

Definition 10.4.13 设 X 是赋范空间, $x_0, x_n \in X \quad (n = 1, 2, \cdots)$, 如果对于每一个 $f \in X^*, f(x_n) \to f(x_0) \quad (n \to \infty)$, 称 $\{x_n\}$ 弱收敛于 x_0. 记为 $x_n \xrightarrow{W} x_0 (n \to \infty)$, 称 x_0 为 $\{x_n\}$ 的弱极限.

Proposition 10.4.26 由弱收敛的定义可得下列性质成立:
1. 弱收敛的极限是惟一的.
2. 如果 $\{x_n\}$ 弱收敛, 则 $\{\|x_n\|$ 有界.
3. 如果 $\{x_n\}$ 强收敛于 x_0, 则 $\{x_n\}$ 必弱收敛于 x_0. 反之则不然.

Theorem 10.4.27 在空间 R^n 中弱收敛与强收敛等价

Theorem 10.4.28 空间 $C[a,b]$ 中点列 $\{x_n\}$ 强收敛于 $x_0 \in C[a,b]$, 当且仅当
1. $\{\|x_n\|\}$ 有界
2. $\{x_n(t)\}$ 在 $[a,b]$ 上逐点收敛于 $x_0(t)$

R 有穷维赋范空间中强弱收敛等价

Theorem 10.4.29 — 无穷维的情况. (Schur) 在空间 l^1 中, 点列强收敛与弱收敛等价。

Theorem 10.4.30 — 期末要考的. 巴纳盒烟拖定理证明要点, 推论, 烟拖定理的几何意义, 推论的几何意义, 大定理的几何意义

$c^* = l' \quad c_0^* = l'$ 共轭空间一样是否原空间一样? 错

Theorem 10.4.31 强调所有的共轭空间, 看讲义

Theorem 10.4.32 弱收敛, 弱有界--函数列转化成数列

Theorem 10.4.33 弱有界一定有界

Theorem 10.4.34 共轭空间一定是完备的, 原空间不一定

Theorem 10.4.35 范数收敛可以推出弱收敛

Theorem 10.4.36 H-B 定理的推论的逆否命题

10.5 Hilbert 空间

Definition 10.5.1 — 内积. 设 Λ 是实数域或复数域, H 是 Λ 上的线性空间, 如果对于 H 中任何两个向量 x,y, 都对应着一个数 $(x,y) \in \Lambda$, 满足条件:

1. 共轭对称性: 对任何 $x,y \in H, (x,y) = \overline{(y,x)}$
2. 对第一变元的线性: 对任何 $x,y,z \in H$ 及任何两数 $\alpha,\beta \in \Lambda$, 成立着
$$(\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z)$$
3. 正定性: 对于一切 $x \in H, (x,x) \geqslant 0$, 而且 $(x,x) = 0$ 的充要条件是 $x = 0$
4. 对于第二个变元来说, 是共轭线性的: 即对于任何 $x,y,z \in H$ 及任何两个数 α,β, 成立着
$$(z, \alpha x + \beta y) = \bar{\alpha}(z,x) + \bar{\beta}(z,y)$$

那么 $(\cdot,\cdot)$ 称为 H 中的内积. 如果 H 上定义了内积, 当 Λ 是实数 (或复数) 域时, 称 H 为实 (或复) 内积空间

Theorem 10.5.1 — 施瓦茨不等式. 如果 H 是内积空间, 那么对于任何 $x,y \in H$
$$|(x,y)|^2 \leqslant (x,x)(y,y)$$

Proof. $(x + \lambda y, x + \lambda y)$ 展开 $+\lambda = -\frac{(x,y)}{(y,y)}$ ■

R 复配方：$\lambda\bar{\lambda} + \alpha\bar{\lambda} + \bar{\alpha}\lambda + \beta = (\lambda + \alpha)(\bar{\lambda} + \bar{\alpha}) + \beta - \alpha\bar{\alpha}$

Theorem 10.5.2 假设 H 是内积空间,$(\cdot,\cdot)$ 是 H 上的内积, 记 $\|x\| = \sqrt{(x,x)}$ 那么 $\|\cdot\|$ 是一个范数。

R 内积可以诱导范数，距离满足条件可以诱导范数；范数可以诱导度量；范数需要满足平行四边形法则可以诱导内积

R 完备的赋范空间 Banach 完备的内积空间 Hilbert 空间

■ **Example 10.29** L^2 是完备的，但是区间上的连续函数不是完备的，例子：一半 0，一半 1，当中是随着 n 增加越来愈陡，是柯西列，但是不收敛

Theorem 10.5.3 设 H 是内积空间, 那么内积关于两个变元是连续的，也就是当
$$x_n \to x_0, y_n \to y_0 \text{ 时}, (x_n, y_n) \to (x_0, y_0)$$

R 范数是连续的，也就是范数和内积与极限可交换

Definition 10.5.2 完备的内积空间称为 Hilbert (希尔伯特) 空间.

■ **Example 10.30** 设 l^2 是满足条件 $\sum_{v=1}^{\infty}|x_v|^2<\infty$ 的数列 $\{x_n\}$ 全体按通常的线性运算所成的线性空间, 当

$$x=(x_1,x_2,\cdots,x_n,\cdots)\in l^2, y=(y_1,y_2,\cdots,y_n,\cdots)\in l^2$$

时规定

$$(x,y)=\sum_{i=1}^{\infty}x_i,\bar{y}_i$$

可以证明这样定义的 (· , ·) 确实满足内积的三个条件, 今后在 l^2 中都是这样取内积. l^2 是一个 Hilbert 空间

■ **Example 10.31** 设 $L^2(\Omega,\boldsymbol{R},\mu)$ (这里 $\boldsymbol{R}$ 是 σ -代数) 是定义在 Ω 上的关于 μ 平方可积的函数全体 (几乎处处相等的两个函数看作是同一个函数) 按通常的线性运算所成线性空间 (见 §4.3), 对于 $f,g\in L^2(\Omega,\boldsymbol{R},\mu)$, 规定

$$(f,g)=\int_{\Omega}f(x)\overline{g(x)}\mathrm{d}\mu$$

容易验证这确实是内积, 所以 $L^2(\Omega,\boldsymbol{R},\mu)$ 是一个 Hilbert 空间.

Theorem 10.5.4 — 极化恒等式--范数诱导内积. 我们注意, 当 H 是内积空间, $\|x\|$ 是由内积所导出的范数时, 内积也可以用范数来表达. 当 H 是实内积空间时

$$(x,y)=\frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2\right)$$

当 H 是复的内积空间时

$$(x,y)=\frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2+\mathrm{i}\|x+\mathrm{i}y\|^2-\mathrm{i}\|x-\mathrm{i}y\|^2\right)$$

这些可以直接从内积的定义导出. 等式被称为极化恒等式, 这是一个重要的等式.

Theorem 10.5.5 — 四边形法则. 如果 H 是内积空间, $||\cdot||$ 是由内积导出的范数, 则对任何 $x,y\in H$ 成立

$$\|x+y\|^2+\|x-y\|^2=2\left(\|x\|^2+\|y\|^2\right)$$

平行四边形的对角线长度的平方和等于四边的长度平方和--由内积决定的范数必须适合平行四边形公式

■ **Example 10.32** 并非每个赋范线性空间都是内积空间. 例如 $L^p[a,b]$ 当 $p\geqslant 1,p\neq 2$ 时, 它就不是内积空间. 因为容易验证向量 c (常数函数) 和另一个适当的非常数函数 x 就是 $L^p[a,b]$ 中不满足平行四边形公式.

Theorem 10.5.6 L^p 只有在 $p=2$ 的时候才是内积空间：取 $x=(1,0,\cdots),y=(0,1\cdots)$，可以验证不满足平行四边形法则，除了 2 的情况

Definition 10.5.3 设 H 是内积空间，$(\cdot,\cdot)$ 是其中的内积．如果 H 中两个向量 x,y，使 $(x,y)=0$，就说 x 与 y 直交，记作 $x\perp y$.

R 正交补一定是子空间

Definition 10.5.4 闭的线性子空间也简称为闭线性子空间，或线性闭子空间.

设A 是赋范线性空间R 中的子集，记$L(A)$ (或span$\{A\}$) 为 A 中向量的线性组合全体所成的线性子空间是由 A 张成的线性闭子空间.

Proposition 10.5.7
1. 直交是相互的，即 $x\perp y$ 时，$y\perp x$
2. 如果 x 与空间中的一个稠密子集合正交，则该向量为 0 向量。(证明思路：利用内积的连续性 $+x$ 与自身内积为零，则为零向量。
3. $x\perp H$ 的充要条件是 $x=0$
4. 当 $M\subset N$ 时，$M^{\perp}\supset N^{\perp}$
5. 对任何 $M\subset H, M\cap M^{\perp}=\{0\}$
6. 勾股弦定理：当 $x\perp y$ 时，

$$\|x+y\|^2=\|x\|^2+\|y\|^2$$

Theorem 10.5.8 设 H 是内积空间，$M\subset H$，那么 $M^{\perp}$ 是 H 的闭线性子空间.

Proof. M 不需要是空间，只需要是集合，加法数乘很好验证；证明它的闭集只需要闭集的定义就可以了，依旧是利用内积的连续性。 ∎

Theorem 10.5.9 设 $M\subset H$，span $\{M\}$ 是 M 张成的闭线性子空间. 那么 $(\overline{\text{span}}\{M\})^{\perp}=M^{\perp}$

Theorem 10.5.10 设 M 是内积空间 H 中的完备凸集，则对任意 $x\in H$，存在 $x_0\in M$，使得

$$\|x-x_0\|=d(x,M)=\inf_{y\in M}\|x-y\|$$

凸组合表示的是两点连线上的所有点。只要是完备凸集，那么距离就能取到

Proof. 根据下确界的定义，可以知道不一定可以达到，但是是可以趋于的。不失一般性，我们可以设 M 是 H 的真子集并且 $x\notin M$. 记 $\alpha=\inf_{y\in M}\|x-y\|$，于是存在 $\{x_n\}\subset M$，使得

$$\|x-x_n\|\to\alpha\quad(n\to\infty)$$

由于 M 是凸集，对任意自然数 m,n

$$\frac{x_m+x_n}{2}\in M$$

因此

$$\left\| x - \frac{x_m + x_n}{2} \right\| \geqslant \alpha$$

由平行四边形法则

$$\begin{aligned} \|x_m - x_n\|^2 &= \|x_m - x + x - x_n\|^2 \\ &= 2\|x_m - x\|^2 + 2\|x - x_n\|^2 - 4\left\| x - \frac{x_m + x_n}{2} \right\|^2 \\ &\leqslant 2\|x_m - x\|^2 + 2\|x - x_n\|^2 - 4\alpha^2 \end{aligned}$$

因此当 $m, n \to \infty$ 时 $\|x_m - x_n\| \to 0$, 即 $\{x_n\}$ 是 M 中的 Cauchy 列，由 M 的完备性，必存在 $x_0 \in M$, 使得 $x_n \to x_0 (n \to \infty)$, 因为范数是连续的

$$\|x - x_0\| = \alpha$$

■

R 凸集合的作用：可以从两条路径上跳跃的选点列可以满足范数趋于 d，但是这不就是柯西列；凸集和可以使这种情况（两条路径趋于 d，两条路径的中间的更加近）不存在。上面的 4 是提出 2^2

R 任何一个子空间都是凸集和，这是由于子空间任意做线性组合封闭。完备性是在大空间的完备性下，闭子空间一定是完备的。

Definition 10.5.5 设 H 是内积空间, M_1 及 M_2 是 H 的两个线性子空间, 如果 $M_1 \perp M_2$, 那么称 $M = \{x_1 + x_2 | x_1 \in M_1, x_2 \in M_2\}$ 为 M_1 与 M_2 的直交和, 记作 $M_1 \oplus M_2$

Definition 10.5.6 — 投影. 设 M 是内积空间 H 的线性子空间, $x \in H$, 如果有 $x_0 \in M, x_1 \perp M$, 使得

$$x = x_0 + x_1$$

那么称 x_0 是 x 在 M 上的 (直交) 投影或 x 在 M 上的 (投影) 分量.

Theorem 10.5.11 设 M 是内积空间 H 的线性子空间, $x \in H$, 如果 x_0 是 x 在 M 上的投影, 那么

$$\|x - x_0\| = \inf_{y \in M} \|x - y\|$$

而且 x_0 是 M 中使上式成立的唯一的向量.

下面我们要证明当 M 为 H 的完备子空间时, H 中任何元 x 在 M 上的投影必定存在. 因此 H 可以分解成 M 与 $M^\perp$ 的直交和.

Theorem 10.5.12 — (变分引理). 设 M 是内积空间 H 中完备的凸集, $x \in H$. 记 d 为 x 到 M 的距离

$$d = d(x,M) = \inf_{y \in M} \|x-y\|$$

那么必有唯一的 $x_0 \in M$ 使得 $\|x-x_0\| = d$

Analysis 10.5 要证明距离是可以取到的，由于下确界是一定存在的，于是等价于证明逼近下确界这一串序列的极限存在。证明极限存在 + 条件中的完备想到 Cauchy 条件。凸集的条件是用来控制距离展开式中的交叉项。

Proof. 由距离的定义，必定有 M 中点列 $\{x_n\}$ 使得 $\lim_{n\to\infty}\|x_n - x\| = d$. 这样的点列称为“极小化”序列。下面证明 $\{x_n\}$ 是基本点列。由平行四边形公式得到

$$2\left\|\frac{x_m - x_n}{2}\right\|^2 = \|x_m - x\|^2 + \|x_n - x\|^2 - 2\left\|\frac{x_m + x_n}{2} - x\right\|^2$$

因为 M 是凸集, $\frac{x_m + x_n}{2} \in M$, 所以

$$\left\|\frac{x_m + x_n}{2} - x\right\| \geqslant d$$

由上式得到

$$0 \leqslant 2\left\|\frac{x_m - x_n}{2}\right\|^2 \leqslant \|x_m - x\|^2 + \|x_n - x\|^2 - 2d^2$$

令 $m,n \to \infty$, 就有 $\lim_{m,n\to\infty}\|x_m - x_n\|^2 = 0$, 所以 $\{x_n\}$ 是基本点列.

因为 M 是个完备的度量空间, 所以有 $x_0 \in M$, 使 $x_n \to x_0$. 这时

$$\|x - x_0\| = \lim_{n\to\infty}\|x - x_n\| = d$$

如果 M 中还有元 y_0 使 $\|x - y_0\| = d$, 那么点列 $\{x_0, y_0, x_0, y_0, \cdots\}$ 显然是“极小化", 序列, 因此是基本点列，这就说明 $x_0 = y_0$. 也就是说在 M 中使 $\|x - x_0\| = d$ 的元 x_0 是唯一的. ■

R 说明用 M 中的元 y 来逼近 x 时，当且仅当 y 等手 x 在 M 上的投影 x_0 时，逼近的程度最好。投影的这个性质常常被用来研究最佳逼近问题。

Theorem 10.5.13 设 H 是内积空间, M 是 H 的线性子空间, $x \in H, x_0 \in M$, 如果

$$\|x - x_0\| = d(x,M), \text{那么} x - x_0 \perp M$$

Theorem 10.5.14 — (投影定理). 设 M 是内积空间 H 的完备线性子空间, 那么对任何 $x \in H, x$ 在 M 上的投影唯一地存在. 也就是说有 $x_0 \in M, x_1 \perp M$ 使 $x = x_0 + x_1$, 而且这种分解是唯一的. 特别地, 当 $x \in M$ 时, $x_0 = x$

Analysis 10.6 子空间一定是凸集合；完备空间的闭子空间是完备的；注意 λ 的巧用 ■

Proof. 由于 M 是 Hilbert 空间 H 的闭子空间, 因此 M 是完备的, 于是对任意 $x \in H$, 存在 $x_0 \in M$, 使得

$$\|x - x_0\| = d(x, M) = \alpha$$

任取 $z \in M, z \neq 0$, 则对任意 $\lambda \in \mathbb{C}, x_0 + \lambda z \in M$（凸集和）, 因此

$$\alpha^2 \leqslant \|x - x_0 - \lambda z\|^2 = \|x - x_0\|^2 - \bar{\lambda}(x - x_0, z) - \lambda(z, x - x_0) + |\lambda|^2 \|z\|^2$$

取

$$\lambda = \frac{(x - x_0, z)}{\|z\|^2}$$

并把它代入上式则得

$$|(x - x_0, z)|^2 \leq 0$$

因此 $(x - x_0, z) = 0$, 即 $(x - x_0) \in M^{\perp}$. 记$y = x - x_0$, 则有

$$x = x_0 + y, \quad x_0 \in M, \quad y \in M^{\perp}$$

下面证明惟一性, 设还有 $x = x_0' + y'$, 其中 $x_0' \in M, y' \in M^{\perp}$. 则可得

$$x_0' - x_0 = y - y', \text{ 因此}, y - y' \in M \bigcap M^{\perp}, \text{ 所以} y' = y \text{ 且} x_0' = x_0$$

■

1. 由上面两个定理有 $x_0 \in M$, 使 $\|x - x_0\| = \inf_{y \in M} \|x - y\|$, 又由 ,$x - x_0 \perp M$ 因此记 $x_1 = x - x_0$ 时, x_0, x_1 就满足定理的要求, x_0 就是 x 在 M 上的投影. 唯一性前面已经证明过了. 特别地, 当 $x \in M$ 时, $\inf_{y \in M} \|x - y\| = 0$, 所以 $x_0 = x$ 证毕

R 投影定理是 Hilbert 空间理论中极其重要的一个基本定理. 这个定理在一般的 Banach 空间中并不成立. 因为在一般情况下并没有直交概念, 即使是性质较好的空间如 $L^p(\Omega, \boldsymbol{B}, \mu)$ 和 l^p 等, 当 $p \neq 2$ 时, 这个定理也不成立. 有了这个定理可以把 Hilbert 空间中的闭线性子空间与投影算子一一对应起来, 而且空间之间的关系与投影算子之间的关系是完全对应的, 简直可以把闭线性子空间和投影算子看成一回事。

R 有限维赋范空间一定是完备的。完备的一定是闭的。

R $H = M \oplus M^{\perp}$

R 容易看出, 当 H 是 Hilbert 空间，M 是 H 的闭线性子空间时, 上述四个定理都成立.

R H 空间如果是可数无穷维的话，可以一直找出正交系：找一个向量，由它构成生成空间和它的正交补是全空间的直和，接着找第二个向量不属于第一个向量的长成空间里，这两个向量长成的空间依旧是闭子空间，从而依旧构成直和，一致做下去

R

1. Banach 可以找到两个闭子空间是直和吗？成为 M 在 X 中是可补的
2. 如果 Banach 中，每个闭子空间 M 都是可补的，则 B 与 H 同构，也就是可以构造内积。也即是投影定理可以来决定 H，类似于 H 的等价定义
3. 所以多了一个判断 H 是否是完备的条件，也即是如果能找到一个闭子空间没有正交补则大空间 H 一定不是完备的

Theorem 10.5.15 设 M 是内积空间 H 中的完备线性子空间, 而且 $M \neq H$, 那么 $M^{\perp}$ 中有非零元素.

Corollary 10.5.16 设 H 是 Hilbert 空间, M 是 H 的线性子空间, 那么 $\bar{M} = \left(M^{\perp}\right)^{\perp}$. 特别地，如果 $M^{\perp} = \{0\}$, 那么 M 在 H 中稠密.

■ **Example 10.33** 设 $(\Omega, \boldsymbol{R}, P)$ 是一个测度空间，$\boldsymbol{R}$ 是 σ -代数，而且 $P(\Omega) = 1$，这时称 $(\Omega, \boldsymbol{R}, P)$ 为概率测度空间，P 称为概率测度.（$\Omega, \boldsymbol{R}$）上的可测函数称为一个随机变量. 设 X 和 $X_1, X_2, \cdots, X_n$ 是 $L^2(\Omega, \boldsymbol{R}, P)$ 中给定的 $n+1$ 个随机变量, 在概率论中, 常要求出 n 个数 $\alpha_1, \cdots, \alpha_n$ 使二阶矩

$$\left\|X - \sum_{\nu=1}^{n} \alpha_\nu X_\nu\right\|^2 = \left(\int_\Omega \left|X(\omega) - \sum_{\nu=1}^{n} \alpha_\nu X_\nu(\omega)\right|^2 \mathrm{d}P(\omega)\right)$$

达到最小, 这是用 $X_1, \cdots, X_n$ 的线性组合来估计 X 的最佳估值问题

Theorem 10.5.17 可以抽象成如下的问题: 设 H 是内积空间, $x, x_1, x_2, \cdots, x_n$ 是 H 中 $n+1$ 个向量, 要求出 n 个数 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 使得

$$\left\|x - \sum_{\nu=1}^{n} \alpha_\nu x_\nu\right\| = \min_{\lambda_1, \cdots, \lambda_n} \left\|x - \sum_{\nu=1}^{n} \lambda_\nu x_\nu\right\|$$

这个问题的解法如下：我们不妨设 $x_1, \cdots, x_n$ 是线性无关的, 不然的话, 只要从 $x_1, x_2, \cdots, x_n$ 中取出 $k(k \leqslant n)$ 个线性无关向量, 暨如 $\{x_1, \cdots, x_k\}$, 使得其余的都是这 k 个向量的线性组合，这时我们只要考虑 $x_1, \cdots, x_k$ 的线性组合好了. 令 M 表示 $x_1, \cdots, x_n$ 的线性组合全

体所成的 n 维线性空间, 由于 M 是完备的, 因此由引理 2 必有 $x_0=\sum_{\nu=1}^{n}\alpha_\nu x_\nu\in M$ 使得 $\|x-x_0\|$ 达到最小, 由引理 3

$$(x-x_0,y)=0,\quad y\in M$$

这显然等价于

$$\left(x-x_0,x_\mu\right)=0,\quad \mu=1,2,\cdots,n$$

而这就是代数方程组

$$\sum_{\nu=1}^{n}\alpha_\nu\left(x_\nu,x_\mu\right)=\left(x,x_\mu\right),\quad \mu=1,2,\cdots,n$$

由于引理 2, 达到最小的 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是唯一的, 所以上述代数方程组的解是唯一的，因此方程组的系数行列式不等于 0，其解就是

$$\alpha_\nu=\frac{\begin{vmatrix}(x_1,x_1)\cdots & (x,x_1)\cdots & (x_n,x_1)\\(x_1,x_2)\cdots & (x,x_2)\cdots & (x_n,x_2)\\\vdots & \vdots & \vdots\\(x_1,x_n)\cdots & (x,x_n)\cdots & (x_n,x_n)\end{vmatrix}}{\begin{vmatrix}(x_1,x_1)\cdots & (x_\nu,x_1)\cdots & (x_n,x_1)\\(x_1,x_2)\cdots & (x_\nu,x_2)\cdots & (x_n,x_2)\\\vdots & \vdots & \vdots\\(x_1,x_n)\cdots & (x_\nu,x_n)\cdots & (x_n,x_n)\end{vmatrix}},\nu=1,2,\cdots,n$$

正交系

在内积空间 H 上, 一个向量 x 在完备线性子空间 M 上的投影 x_0 是达到极值 $\inf_{y\in M}\|x-y\|$ 的向量. 如何求出 x_0 ?

Definition 10.5.7 — 就范直交系. 设 $\mathcal{F}$ 是内积空间 H 中的一族非零向量, 如果 $\mathcal{F}$ 中任何两个不同向量都直交, 就称 $\mathcal{F}$ 是 H 中的一个直交系. 如果直交系 $\mathcal{F}$ 中每个向量的范数都是 1, 就称 $\mathcal{F}$ 是就范直交系（标准正交系）.

R 只要两两正交就是正交系，但是正交基要求长成的空间的闭包是全空间

■ **Example 10.34** 在实空间 $L^2[0,2\pi]$ 中, 规定内积为

$$(f,g)=\frac{1}{\pi}\int_0^{2\pi}f(x)g(x)\mathrm{d}x\quad\left(f(x),g(x)\in L^2[0,2\pi]\right)$$

这时 $\frac{1}{\sqrt{2}},\cos x,\sin x,\cos 2x,\sin 2x,\cdots,\cos nx,\sin nx,\cdots$ 组成 $L^2[0,2\pi]$ 中的就范直交系

R 由定义立即可知, 如果 $\mathcal{F}$ 是内积空间 H 中的直交系 (就范直交系), 那么 F 的任何子集也是 H 中的直交系 (就范直交系).

Definition 10.5.8 设 $\mathcal{F}$ 是内积空间 H 中的就范直交系， $x \in H$, 数集

$$\{(x,e)|e \in \mathcal{F}\}$$

称为向量 x 关于就范直交系 $\mathcal{F}$ 的 Fourier 系数集, 而数 (x,e) 称为 x 关于 $e(\in \mathcal{F})$ 的 Fourier 系数。

Theorem 10.5.18 设 $\{e_1,e_2,\cdots,e_n\}$ 是内积空间 H 中的就范直交系 , $M=\operatorname{span}\{e_1,\cdots,e_n\}$, $x \in H$, 那么 $x_0=\sum_{i=1}^n (x,e_i)\,e_i$ 是 x 在 M 上的投影, 而且

$$\|x_0\|^2=\sum_{i=1}^n |(x,e_i)|^2, \|x-x_0\|^2=\|x\|^2-\|x_0\|^2$$

R 表明: 如果 M 是内积空间中有限维线性子空间 (它必是完备的), 这时只要在 M 中选个数为 dimM 的就范直交向量 $\{e_1,\cdots,e_n\}\,(n=\dim M)$. 那么, 任何 $x \in H$, 在 M 上投影 x_0 就是 $\sum_{i=1}^n (x,e_i)\,e_i$. 特别地, 当 M 是一维子空间 $\{\alpha e\}\,(\alpha$ 是数) 时, 如取 $\|e\|=1$, 那么x 在M 上投影是 $x_0=(x,e)e$

Corollary 10.5.19 设 $\{e_1,\cdots,e_n\}$ 是内积空间 H 中就范直交系, 那么对 $x \in H$

$$\sum_{i=1}^n |(x,e_i)|^2 \leqslant \|x\|^2$$

利用范数大于零

Corollary 10.5.20 设 $\{e_1,\cdots,e_n\}$ 是内积空间 H 中就范直交系, $x \in H$, 那么对任何 n 个数 $\alpha_1,\alpha_2,\cdots,\alpha_n$

$$\left\|x-\sum_{i=1}^n \alpha_i e_i\right\| \geqslant \left\|x-\sum_{i=1}^n (x,e_i)\,e_i\right\|$$

而且当式等号成立时必定 $\alpha_i=(x,e_i)$

Proof. 由于

$$\left(x-\sum_{k=1}^n (x,e_k)\,e_k,e_i\right)=(x,e_i)-(x,e_i)=0,(i=1,2,\cdots,n)$$

其与每个 e_i 都正交，故与他的线性组合也是正交的，由勾股定理，则有

$$\begin{aligned}\left\|x-\sum_{k=1}^{n}\alpha_k e_k\right\|^2&=\left\|x-\sum_{k=1}^{n}(x,e_k)e_k+\sum_{k=1}^{n}((x,e_k)-\alpha_k)e_k\right\|^2\\&=\left\|x-\sum_{k=1}^{n}(x,e_k)e_k\right\|^2+\left\|\sum_{k=1}^{n}((x,e_k)-\alpha_k)e_k\right\|^2\\&=\left\|x-\sum_{k=1}^{n}(x,e_k)e_k\right\|^2+\sum_{k=1}^{n}|(x,e_k)-\alpha_k|^2\end{aligned}$$

由此可知, 当且仅当 $\alpha_k=(x,e_k)\,(k=1,2,\cdots,n)$ 时 , $\|x-\sum_{k=1}^{n}\alpha_k e_k\|$ 取最小值.

这个定理的 是，设 M 是由 $\{e_1,\cdots,e_n\}$ 张成的 n 维子空间, 则 $x\in H$ 在 M 上的投影为 $x_0=\sum_{k=1}^{n}(x,e_k)e_k$, 而 x 到 M 上的最短距离为 $\|x-x_0\|$。

几何意义：投影的平方和不会超过 X 的平方和 ■

Theorem 10.5.21 — 投影定理无限维时的推广–(Bessel (贝塞尔) 不等式). 设 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中的就范直交系, 那么对每个 $x\in\boldsymbol{H},x$ 的 Fourier 系数 $\{(x,e_\lambda)|\lambda\in\Lambda\}$ 中最多只有可列个不为零而且适合 Bessel 不等式:

$$\sum_{\lambda\in\Lambda}|(x,e_\lambda)|^2\leqslant\|x\|^2$$

Proof. 由于对任意 n，取上个式子中的 $\alpha_k=0$

$$\left\|x-\sum_{k=1}^{n}(x,e_k)e_k\right\|^2=\left(x-\sum_{k=1}^{n}(x,e_k)e_k,x-\sum_{k=1}^{n}(x,e_k)e_k\right)\tag{10.17}$$

$$=\|x\|^2-\sum_{k=1}^{n}|(x,e_k)|^2\geqslant 0\tag{10.18}$$

即 $\sum_{k=1}^{n}|(x,e_k)|^2\leqslant\|x\|^2$. 令 $n\to\infty$ 即得要证的不等式 ■

R 之前不能直接在范数里面取极限，因为没法保证级数是收敛的。几何意义：

R 证明中的结论非常重要

Corollary 10.5.22 — R-L 引理. 设 $\{e_n\}$ 是内积空间 H 中的就范直交系, 那么对任何 $x\in\boldsymbol{H}$, 必有

$$\lim_{n\to\infty}(x,e_n)=0$$

Proof. 由级数 $\sum_{n=1}^{\infty}|(x,e_n)|^2$ 收敛, 即得 $\lim_{n\to\infty}(x,e_n)=0$ ■

R 在 $L^2[0,2\pi]$ 中用于就范直交三角函数系 (见例1) 的情况下, 就是 Riemann-Lebesgue 引理。

R Bessel 不等式表示向量在 $\mathcal{F}$ 中每个就范向量 e_λ 上投影 $(x,e_\lambda)e_\lambda$ 的“长度" 平方的和不超过 x 的“长度”平方。

Corollary 10.5.23 设 $\{e_\alpha\}\alpha\in I$ 是内积空间 H 中的标准正交系. 证明对于每一个 $x\in H,x$ 关于这个标准正交系的 Fourier 系数 $\{\ (x,e_a):\alpha\in I\}$ 中最多有可数个不为零.

Theorem 10.5.24 — Parseval--勾股定理推广--完备正交系的定理. 设 $\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中就范直交系，如果对任何 $x\in H$, 成立下列 Parseval (帕塞瓦尔) 等式

$$\|x\|^2=\sum_{\lambda\in\Lambda}|(x,e_\lambda)|^2$$

称直交系 $\{e_\lambda|\lambda\in\Lambda\}$ 是 H 中完备直交系. 相等意味着每个维数都有，每个轴上都有。

Definition 10.5.9 — 傅立叶级数. 设 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中的就范直交系, 对 $x\in H$ 形式级数 $\sum_{\lambda\in\Lambda}(x,e_\lambda)e_\lambda$(不管它是否收敛) 称为向量$x$ 关于$\mathscr{F}$ 的 Fourier 级数, 或 Fourier 展开式. 当 $x=\sum_{\lambda\in\Lambda}(x,e_\lambda)e_\lambda$ 成立时, 就称 x 关于 $\mathscr{F}$ 可以展开成 Fourier 级数。

Theorem 10.5.25 当 x 关于 $\mathscr{F}$ 可以展开成 Fourier 级数时, 展开式 $x=\sum_{\lambda\in\Lambda}(x,e_\lambda)e_\lambda$ 的几何意义就是向量 x 等于它在 $\mathscr{F}$ 的每个 e_λ 方向的分量 $(x,e_\lambda)e_\lambda$ (根据 Bessel 不等式, 最多只有可列个分量不是零) 的和。

Theorem 10.5.26 设 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中的就范直交系, E 为 $\mathscr{F}$ 张成的线性闭子空间, 那么对于 $x\in H$, 下面三个结论是等价的:

1. $x\in E$
2. $\|x\|^2=\sum_{\lambda\in\Lambda}|(x,e_\lambda)|^2$.
3. $x=\sum_{\lambda\in\Lambda}(x,e_\lambda)e_\lambda$

Theorem 10.5.27 内积空间 H 中就范直交系 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是完备的充要条件是 $\mathscr{F}$ 张成的闭线性子空间 $E=H$, 或者充要条件是对任何 $x\in H$, 成立

$$x=\sum_{\lambda\in\Lambda}(x,e_\lambda)e_\lambda$$

Theorem 10.5.28 (CrekappaoB (斯切克洛夫)) 设 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中就范直交系，如果有 H 中的稠密子集 D，使得对于 $x\in D$ 都成立 Parseval 等式 $\|x\|^2=$

$\sum_{\lambda\in\Lambda}|(x,e_\lambda)|^2$, 那么 $\mathscr{F}$ 是完备的.

Proof. 设 $\{e_n\}$ 是完备的, 则对任意 $x\in H$

$$\left\|x-\sum_{k=1}^{n}(x,e_k)e_k\right\|\to 0\quad(n\to\infty)$$

因此 $\{e_n\}$ 张成的子空间 L 在 H 中稠密. 反之, 对任意 $x\in H$ 及 $\varepsilon>0$, 存在 $x_n=\sum_{k=1}^{n_0}\alpha_k^{(n_0)}e_k$, 使得

$$\|x-x_{n_0}\|<\varepsilon$$

于是

$$\left\|x-\sum_{k=1}^{n_0}(x,e_k)e_k\right\|\leqslant\left\|x-\sum_{k=1}^{n_0}\alpha_k^{(n_0)}e_k\right\|=\|x-x_{n_0}\|<\varepsilon$$

因此由式11.4.14得, 当 $n>n_0$ 时

$$\left\|x-\sum_{k=1}^{n}(x,e_k)e_k\right\|\leqslant\left\|x-\sum_{k=1}^{n_0}(x,e_k)e_k\right\|<\varepsilon$$

所以 $\{e_n\}$ 是完备的. ■

Proof. 记 E 为 $\mathcal{F}$ 张成的闭线性子空间, 由于 D 中向量都成立 Parseval 等式所以 $D\subset E$. 但因为 E 是闭的, 所以 $\bar{D}\subset E$. 由假设 $\bar{D}=H$, 所以 $E=H$. 则, $\mathcal{F}$ 是完备的。 ■

R 有限维的线性空间没有闭包的必要，因为已经是完备的

Definition 10.5.10 如果 $\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中完备就范直交系, 那么对一切 $x,y\in H, x=\sum_{\lambda\in\Lambda}(x,e_\lambda)e_\lambda, y=\sum_{\lambda\in\Lambda}(y,e_\lambda)e_\lambda$. 由此易知 $(x,y)=\sum_{\lambda\in\Lambda}(x,e_\lambda)\overline{(y,e_\lambda)}$. 这个等式也称为 Parseval 等式。

Theorem 10.5.29 $L^2[0,2\pi]$ 中的就范直交系 $\left\{\frac{1}{\sqrt{2}}\cos t,\sin t,\cdots,\cos nt,\sin nt,\cdots\right\}$ 是完备的.

Theorem 10.5.30 设 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中的就范直交系, $\mathcal{F}$ 张成的闭线性子空间为 E。对任何 $x\in H$, 如果 x 在 E 中有投影 x_0, 那么 x_0 就是 x 的 Fourier 级数 $\sum_\lambda(x,e_\lambda)e_\lambda$. 如果 H 是 Hilbert 空间, 那么对任何 $x\in H, \sum_\lambda(x,e_\lambda)e_\lambda$ 就是 x 在 E 上的投影

Theorem 10.5.31 — (Riesz-Fischer (里斯 - 费味)). 设 $\mathcal{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是 Hilbert 空间 H 中的就范直交系, $\mathcal{F}$ 张成的闭线性子空间为 E. 又设 $\{c_\lambda|\lambda\in\Lambda\}$ 是一族数, 并且 $\sum_{\lambda\in\Lambda}|c_\lambda|^2<\infty$. 那么必有唯一的 $x\in E, x$ 是以 $\{c_\lambda\}$ 为关于 $\{e_\lambda\}$ 的 Fourier 系数, 且 $x=\sum_{\lambda\in\Lambda}c_\lambda e_\lambda$

(Riesz-Fischer) 设 H 是 Hilbert 空间, $\{e_n\}$ 是 H 中的标准正交系, $\{\xi_n\} \in l^2$, 则存在 $x \in H$, 使得

$$\xi_k = (x, e_k) \quad (k = 1, 2, \cdots)$$

并且

$$\sum_{k=1}^{\infty} |\xi_k|^2 = \|x\|^2$$

Proof. 由于 $\sum_{\lambda \in \Lambda} |c_\lambda|^2 < \infty$, 所以最多只有可列个 $\{c_\nu\}$ 不为零. 作序列

$$x_n = \sum_{\nu=1}^{n} c_\nu e_\nu, n = 1, 2, 3, \cdots$$

易知当$n < m$ 时,

$$\|x_n - x_m\|^2 = \left\| \sum_{n+1}^{m} c_\nu e_\nu \right\|^2 = \sum_{n+1}^{m} |c_\nu|^2$$

由于 $\sum_{\nu=1}^{\infty} |c_\nu|^2 < \infty$, 故 $\lim_{m,n\to\infty} \|x_n - x_m\|^2 = 0$, 因此 $\{x_n\}$ 是基本点列. 因为 H 是完备的, 故有 $x = \lim_{n\to\infty} x_n . x \in E$ 是显然的. 而对于任何自然数 ν

$$(x, e_\nu) = \lim_{n\to\infty} (x_n, e_\nu) = c_\nu$$

而当 $e_\lambda \neq e_\nu (\nu = 1, 2, 3, \cdots)$ 时, 因为 $(x_n, e_\lambda) = 0$, 所以 $(x, e_\lambda) = \lim_{n\to\infty} (x_n, e_\lambda) = 0 = c_\lambda$, 从而对一切 $\lambda \in \Lambda, (x, e_\lambda) = c_\lambda$. 唯一性由定理 6.3.3 即知.

法二：令部分和

$$x_n = \sum_{k=1}^{n} \xi_k e_k \quad (n = 1, 2, \cdots)$$

则对任意自然数（正交基）

$$\|x_{n+p} - x_n\|^2 = \left\| \sum_{k=n+1}^{n+p} \xi_k e_k \right\|^2 = \sum_{k=n+1}^{n+p} |\xi_k|^2$$

由于 $\{\xi_k\} \in l^2, \{x_n\}$ 是 H 中的 Cauchy 列, 由 H 的完备性, 存在 $x \in H$, 使得

$$\|x - x_n\| \to 0 \quad (n \to \infty)$$

其次，由于

$$(x_n, e_i) = \left(\sum_{k=1}^{n} \xi_k e_k, e_i \right) = \xi_i \quad (n \geqslant i)$$

$$|(x - x_n, e_i)| \leqslant \|x - x_n\| \|e_i\| = \|x - x_n\|$$

$$(x, e_i) = (x_n, e_i) + (x - x_n, e_i)$$

注意上式左边与 n 无关. 在上式右边令 $n\to\infty$ 则有

$$(x,e_i)=\xi_i \quad (i=1,2,\cdots)$$

并且

$$\begin{aligned}\lim_{n\to\infty}\|x-x_n\|^2&=\lim_{n\to\infty}\left(x-\sum_{k=1}^{n}\xi_k e_k, x-\sum_{k=1}^{n}\xi_k e_k\right)\\&=\|x\|^2-\sum_{k=1}^{\infty}|\xi_k|^2\end{aligned}$$

■

R 利用内积连续性，直接计算

R l_2 平方可和的和标准正交基是等价的；给一串平方可和的，在子空间内可以找到一个向量，以这一串为坐标；反过来，根据 Bessel 不等式可以得到；由可数的坐标系长成的坐标系里的向量都可以用坐标来表达，那么与 l2 没有差别，可以构成双射；

Definition 10.5.11 — 直交系的完全性. 设 $\mathscr{F}$ 是内积空间 H 中的就范直交系, 如果 $\mathscr{F}^{\perp}=\{0\}$, 那么就称 $\mathscr{F}$ 是完全的

R 由定义可知, $\mathscr{F}$ 是完全的, 就是说在 H 中不存在与 $\mathscr{F}$ 直交的非零向量. 因此, 它的意思就是直交系 $\mathscr{F}$ 已经不能再扩大了, 即 $\mathscr{F}$ 是 H 中极大的 (就范) 直交系.

Theorem 10.5.32 设 $\mathscr{F}=\{e_\lambda|\lambda\in\Lambda\}$ 是内积空间 H 中的就范直交系, 如果 $\mathscr{F}$ 是完备的, 那么 $\mathscr{F}$ 是完全的。如果 H 是 Hilbert 空间, 那么完全的就范直交系必定是完备的。

Proof. 证设 $\{e_n\}$ 是完备的, 如果 $x\in H$, 使得

$$(x,e_n)=0 \quad (n=1,2,\cdots)$$

由 Parseval 等式

$$\|x\|^2=\sum_{n=1}^{\infty}|(x,e_n)|^2=0$$

因此 $x=0$, 所以 $\{e_n\}$ 是完全的. 反之, 设 $\{e_n\}$ 是完全的, 反证：假设 $\{e_n\}$ 张成的子空间 M 在 H 中不稠密, 于是存在 $x\in H/\bar{M}$, 由正交分解定理, 存在 $x_0\in\bar{M}$ 及 $y\in\bar{M}^{\perp}$, 使得

$$x=x_0+y$$

显然 $y\neq 0$ 且 y 与所有 e_n 正交, 这与 $\{e_n\}$ 的完全性矛盾, 所以 $\bar{M}=H$, 从而 $\{e_n\}$ 是完备的。 ■

R H 空间可能是不可数维的

Theorem 10.5.33 — (Gram-Schmidt (格拉姆 - 施密特)). 设 $G=\{g_1,g_2,g_3,\cdots\}$ 是内积空间 H 中有限个或可列个线性无关的向量, 那么必定有 H 中的就范直交系 $\mathcal{F}=\{h_1,h_2,h_3,\cdots\}$ 使得对于每个自然数 n,g_n 是 $h_1,h_2,\cdots,h_n$ 的线性组合, h_n 也是 $g_1,g_2,\cdots,g_n$ 的线性组合. 这种 h_n 除去一个绝对值为 1 的常数因子外，由 $g_1,g_2,\cdots,g_n$ 完全确定。也就是两组的张成子空间等价，且每一个对应分量的张成子空间也等价

Proof. $e_1=\frac{x_{n_1}}{\|x_{n_1}\|}$，设 x_{n_k} 是第一个不属于 M_{k-1} 的元，记

$$h_k=x_{n_k}-\sum_{i=1}^{k-1}(x_{n_k},e_i)e_i$$

则 $h_k\neq 0$ 并且 $h_k\perp \dot{e}_i(i=-1,\cdots,k-1)$ 令

$$e_k=\frac{h_k}{\|h_k\|}$$

即我们可以作出 e_k ■

Corollary 10.5.34 可分的内积空间，存在完备的标准正交系

Theorem 10.5.35 任何可分的内积空间中，存在完备的标准正交关系。（可数个就可以了--之多可数维）可分的一定是至多可数维的

Theorem 10.5.36 设 H 是任一可分的无穷维的 Hilbert 空间, 则存在 H 上到 l^2 上同构映射 φ, 且 φ 保持内积.

Proof. 由于 H 可分且是无穷维的，H 中存在由可数个元构成的完备的标准正交系 $\{e_n\}$. 对于任意 $x\in H$, 令

$$\varphi(x)=\{(x,e_n)\}$$

由 Bessel 不等式，$\varphi(x)\in l^2$. 此外对任意 $x,y\in H$ 及 $\alpha,\beta\in \mathrm{K}$:

$$\begin{aligned}\varphi(\alpha x+\beta y)&=\{(\alpha x+\beta y,e_n)\}=\{\alpha(x,e_n)+\beta(y,e_n)\}\\&=a\varphi(x)+\beta\varphi(y)\end{aligned}$$

及由 Riesz-Fisher 定理, φ 是映上的（满射）, 因此 φ 是 H 上到 l^2 上的一个同构映射. 最后, 对任意 $x,y\in H$

$$\begin{aligned}(x,y)&=\left(\sum_{k=1}^{\infty}(x,e_k)e_k,\sum_{k=1}^{\infty}(y,e_k)e_k\right)\\&=\sum_{k=1}^{\infty}(x,e_k)\overline{(y,e_k)}\\&=(\varphi(x),\varphi(y))_l^2\end{aligned}$$

即 φ 保持内积，从而是保持范数的，从而是单射。定理 4.2.10 表明, 任何一个无穷维可分 Hilbert 空间都可以表示为“坐标形式" l^2 ■

R 有限维的赋范空间和有限维的 C^n 是同构的；有限维赋范一定是完备的，由于欧式空间是由内积的，所以有限维赋范空间也是有角度的

Definition 10.5.12 — 保范线性同构. 设 H_1 和 H_2 是两个内积空间, 如果有 H_1 到 H_2 上的一一对应 φ 保持线性运算及内积, 即对任何 $x_1,y_1 \in H_1$ 及两个数 α,β, 都成立

$$\begin{aligned}\varphi(\alpha x_1+\beta y_1)&=\alpha\varphi(x_1)+\beta\varphi(y_1)\\(\varphi(x_1),\varphi(y_1))&=(x_1,y_1)\end{aligned}$$

就说内积空间 H_1 和 H_2 是保范线性同构, 简称同构.

Theorem 10.5.37 任何 n 维内积空间 H 必和 n 维欧几里得空间 E^n 同构。

Theorem 10.5.38 任何可析的 Hilbert 空间 H 必和某个 E^n 或 l^2 同构.

Theorem 10.5.39 设 H 为 Hilbert 空间, M 是 H 的闭子空间. 证明 M 为 H 上某个非零连续线性泛函的零空间，当且仅当 $M^{\perp}$ 是一维子空间.

Proof. 如果 M 为非零连续线性泛函 f 的零空间。于是

$$M=\{x\in H: f(x)=0, x\in H\}=\{y\}^{\perp}$$

假设不是一维的, 则假设两个线性无关向量, $f(x_1)=a\neq 0, f(x_2)=b\neq 0$, $ax_2-bx_1\in M$。于是 $ax_2-bx_1=0$ 与线性无关矛盾

如果正交补是 1 维的，设 $M^{\perp}=\{y\}$, 其中 $\|y\|=1$

$$\begin{aligned}d^2(y,M)&=\inf_{x\in M}\{\|y-x\|^2\}=\inf_{x\in M}\{\|y\|^2+\|x\|^2\}\\&=\|y\|^2=1\end{aligned}$$

正交用勾股定理，x 取 0 时其他都是大于零的，所以 $d(y,M)=||y||=1$，根据 HB 延拖定理可知存在全空间的泛函 $f|_M=0, f(y)=||y||=1$。任取一个元素 $\forall z\in H$, 由投影定理, $z=z_1+z_2, z_1\in M, z_2=\lambda y\in M^{\perp}$, 所以 $f(z)=f(z_1)+\lambda f(z_2)=\lambda$, 所以 $f(z)=0$ 时,$\lambda=0$, 所以 $z=z_1\in M$，所以 M 是 f 的零空间

（有界线性泛函的零空间一定是闭子空间，有界一定连续，0 单点集一定是闭集--则正交补是一条线--所以整个空间把每一层算成一个数，相当于在直线上当作 Z 轴，每一个片看成一个元素）

■

Theorem 10.5.40 对于内积函数来说，固定第二个位置的数，则构成一个有界线性泛函。但是如果固定第一个数，则不是线性泛函，不满足数乘，故叫做共轭线性泛函

Theorem 10.5.41 (F. Riesz) 设 H 是 Hilbert 空间, f 是 H 上任意有界线性泛函, 则存在惟一的 $y_f \in H$, 使得对于每一个 $x \in H$

$$f(x) = (x, y_f)$$

并且

$$\|f\| = \|y_f\|$$

找到一个固定的 y，换 x 做内积

Proof. 存在 $y_0 \in \mathcal{N}(f)^{\perp}, y_0 \neq 0$, 取

$$y_f = \frac{\overline{f(y_0)}}{\|y_0\|^2} y_0$$

我们证明 y_f 满足定理的要求. 首先，如果 $x \in \mathcal{N}(f)$, 则 $f(x) = (x, y_f)$, 因为这时等式两边都是零, 如果 $x = \alpha y_0 (\alpha \in \mathbb{C})$, 则

$$(x, y_f) = (\alpha y_0, y_f) = \left(\alpha y_0, \frac{\overline{f(y_0)}}{\|y_0\|^2} y_0\right) = f(x)$$

由于 $y_0 \in \mathcal{M} f)^{\perp}$ 且 $y_0 \neq 0$, 所以 $f(y_f) \neq 0$, 于是对任意 $x \in H$

$$f\left(x - \frac{f(x)}{f(y_f)} y_f\right) = f(x) - \frac{f(x)}{f(y_f)} f(y_f) = 0$$

即 $x - \frac{f(x)}{f(y_f)} y_f \in \mathcal{N}(f)$.日

$$x = \left(x - \frac{f(x)}{f(y_f)} y_f\right) + \frac{f(x)}{f(y_f)} y_f$$

所以

$$\begin{aligned}
f(x) &= f\left(x - \frac{f(x)}{f(y_f)} y_f\right) + f\left(\frac{f(x)}{f(y_f)} y_f\right) \\
&= \left(x - \frac{f(x)}{f(y_f)} y_f, y_f\right) + \left(\frac{f(x)}{f(y_f)} y_f, y_f\right) \\
&= (x, y_f)
\end{aligned}$$

其次，由于

$$\|f\| = \sup_{\|x\| \leqslant 1} |f(x)| = \sup_{\|x\| \leqslant 1} |(x, y_f)| \leqslant \|y_f\|$$

及

$$\|f\| = \sup_{\|x\|\leqslant 1} |f(x)| \geqslant \left| f\left(\frac{y_f}{\|y_f\|}\right) \right| = \left(\frac{y_f}{\|y_f\|}, y_f\right) = \|y_f\|$$

所以 $\|f\| = \|y_f\|$

最后，如果还有 $y'_f \in H$, 使得对于每一个 $x \in H, f(x) = \left(x, y'_f\right)$. 则有

$$(x, y_f) = \left(x, y'_f\right) \quad (x \in H)$$

所以 $y'_f = y_f$ ■

Proof. 法二：因为 f 是连续线性泛函, 所以 f 的零空间

$$\mathcal{N}(f) = \{x \in H : f(x) = 0\}$$

(零空间看成是一片一片的，其他的空间可以表示成加法的陪集 + 可以确定垂直的空间)

是 H 的闭子空间. 如果 $\mathcal{N}(f) = H$, 定理的结论显然成立 , 这时只需取 $y_f = 0$ 即可; 如果 $\mathcal{N}(f) \neq H$, 由投影定理,$H = \mathcal{N}(f) \oplus \mathcal{N}(f)^{\perp}$, 于是

$$\exists z \in \mathcal{N}(f)^{\perp}, \|z\| = 1, f(z) \neq 0$$

若表现 y_f 存在，则 $y_f = y_1 + y_2, y_1 \in \mathcal{N}(f), y_2 \in \mathcal{N}(f)^{\perp}$. 假设 $y_1 \neq 0$, 则 $f(y_1) = (y_1, y_1 + y_2) = \|y_1\| \neq 0$，这与 $y_1 \in \mathcal{N}(f)$ 矛盾，故 $y_f = y_2 \in \mathcal{N}(f)^{\perp}$. (也就是：表现如果存在的话只有第二部分，第一部分为零)。根据习题的结论可知：$\mathcal{N}(f)^{\perp}$ 为 1 维子空间，即 $\mathcal{N}(f)^{\perp} = \{z\}$，故问题转化成寻找 λ 使得，$\forall x \in H, f(x) = (x, \lambda z)$.

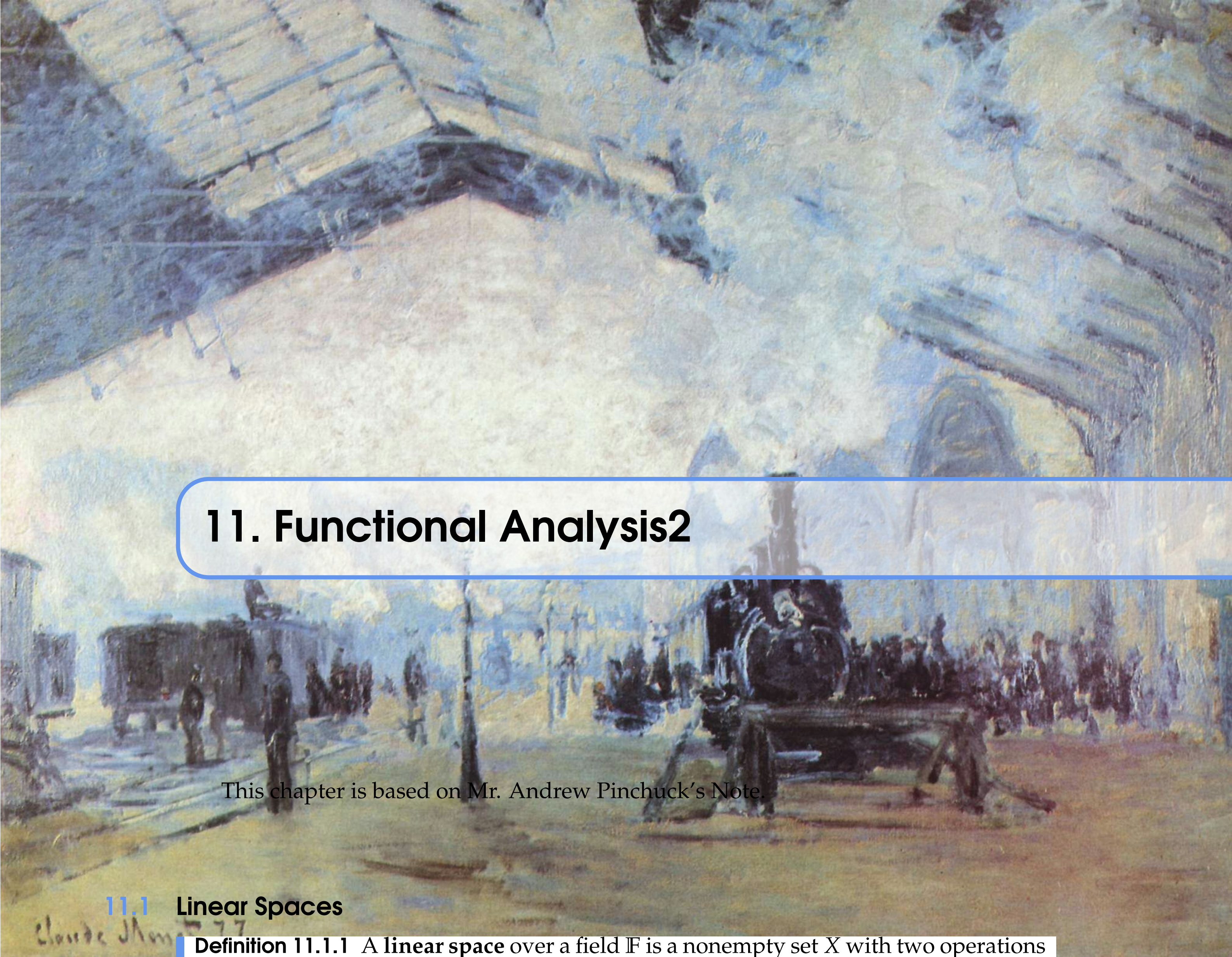
$\forall x \in H$, 根据投影定理可知，$x = x_1 + x_2 = x_1 + \alpha\lambda, x_1 \in \mathcal{N}(f), x_2 \in \mathcal{N}(f)^{\perp}$, 于是

$$\begin{aligned} f(x) = f(x_1 + \alpha z) = f(x_1) + \alpha f(z) = \alpha f(z) \\ (x, \lambda z) = (x_1 + \alpha z, \lambda z) = \alpha\bar{\alpha}(z, z) = \alpha\lambda \end{aligned}$$

所以，$\forall x \in H, f(x) = (x, \overline{f(z)} z)$ ■

R H 空间一定是 B 空间，R 表示定理说明了：对于任何一个 H 空间上的有界线性泛函，都可以表示成固定第二个位置的内积

R 也即是说 H 到 H^* 可以存在双射 + 保范 ==



11. Functional Analysis2

This chapter is based on Mr. Andrew Pinchuck's Note.

11.1 Linear Spaces

Definition 11.1.1 A **linear space** over a field $\mathbb{F}$ is a nonempty set X with two operations

$$
\begin{aligned}
+: \quad & X \times X \to X \quad \text{(called addition), and} \\
\cdot: \quad & \mathbb{F} \times X \to X \quad \text{(called multiplication)}
\end{aligned}
$$

satisfying the following properties:

1. $x + y \in X$ whenever $x, y \in X$
2. $x + y = y + x$ for all $x, y \in X$
3. There exists a unique element in X, denoted by 0, such that $x + 0 = 0 + x = x$ for all $x \in X$
4. Associated with each $x \in X$ is a unique element in X, denoted by $-x$, such that $x + (-x) = -x + x = 0$
5. $(x + y) + z = x + (y + z)$ for all $x, y, z \in X$
6. $\alpha \cdot x \in X$ for all $x \in X$ and for all $\alpha \in \mathbb{F}$
7. $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$ for all $x, y \in X$ and all $\alpha \in \mathbb{F}$
8. $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$ for all $x \in X$ and all $\alpha, \beta \in \mathbb{F}$
9. $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$ for all $x \in X$ and all $\alpha, \beta \in \mathbb{F}$
10. $1 \cdot x = x$ for all $x \in X$

Definition 11.1.2 Sequence Spaces: Informally, a sequence in X is a list of numbers indexed by $\mathbb{N}$. Equivalently, a sequence in X is a function $x : \mathbb{N} \to X$ given by $n \mapsto x(n) = x_n$. We shall denote a sequence $x_1, x_2, \ldots$ by

$$x = (x_1, x_2, \ldots) = (x_n)_1^\infty$$

Definition 11.1.3 The sequence space s. Let s denote the set of all sequences $x = (x_n)_1^\infty$ of real or complex numbers. Define the operations of addition and scalar multiplication pointwise: For all $x = (x_1, x_2, \ldots), y = (y_1, y_2, \ldots) \in \mathbf{s}$ and all $\alpha \in \mathbb{F}$, define

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots)$$
$$\alpha \cdot x = (\alpha x_1, \alpha x_2, \ldots)$$

Then s is a linear space over $\mathbb{F}$.

Definition 11.1.4 The sequence space ℓ_∞. Let $\ell_\infty = \ell_\infty(\mathbb{N})$ denote the set of all bounded sequences of real or complex numbers. That is, all sequences $x = (x_n)_1^\infty$ such that

$$\sup_{i \in \mathbb{N}} |x_i| < \infty$$

Define the operations of addition and scalar multiplication pointwise as in example (3). Thenℓ_∞ is a linear space over $\mathbb{F}$.

Definition 11.1.5 The sequence space $\ell_p = \ell_p(\mathrm{N}), \quad 1 \le p < \infty$. Let ℓ_p denote the set of all sequences $x = (x_n)_1^\infty$ of real or complex numbers satisfying the condition

$$\sum_{i=1}^\infty |x_i|^p < \infty$$

Define the operations of addition and scalar multiplication pointwise: For all $x = (x_n), y = (y_n)$ in ℓ_p and all $\alpha \in \mathbb{F}$, define

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots)$$
$$\alpha \cdot x = (\alpha x_1, \alpha x_2, \ldots)$$

Then ℓ_p is a linear space over $\mathbb{F}$.

Proof. Let $x = (x_1, x_2, \ldots), y = (y_1, y_2, \ldots) \in \ell_p$. We must show that $x + y \in \ell_p$. since, for each $i \in \mathbb{N}$

$$|x_i + y_i|^p \le [2\max\{|x_i|, |y_i|\}]^p \le 2^p \max\{|x_i|^p, |y_i|^p\} \le 2^p \left(|x_i|^p + |y_i|^p\right)$$

it follows that

$$\sum_{i=1}^\infty |x_i + y_i|^p \le 2^p \left(\sum_{i=1}^\infty |x_i|^p + \sum_{i=1}^\infty |y_i|^p\right) < \infty$$

Thus, $x+y \in \ell_p$. Also, if $x=(x_n) \in \ell_p$ and $\alpha \in \mathbb{F}$, then

$$\sum_{i=1}^{\infty} |\alpha x_i|^p = |\alpha|^p \sum_{i=1}^{\infty} |x_i|^p < \infty$$

That is, $\alpha \cdot x \in \ell_p$ ■

Definition 11.1.6 The sequence space $c = c(\mathbb{N})$. Let c denote the set of all convergent sequences $x = (x_n)_1^{\infty}$ of real or complex numbers. That is, c is the set of all sequences $x = (x_n)_1^{\infty}$ such that $\lim_{n\to\infty} x_n$ exists. Define the operations of addition and scalar multiplication pointwise as in example(3). Then c is a linear space over $\mathbb{F}$.

Definition 11.1.7 The sequence space $c_0 = c_0(\mathbb{N})$. Let c_0 denote the set of all sequences $x = (x_n)_1^{\infty}$ of real or complex numbers which converge to zero. That is, $\mathbf{c}_0$ is the space of all sequences $x = (x_n)_1^{\infty}$ such that $\lim_{n\to\infty} x_n = 0$. Define the operations of addition and scalar multiplication pointwise as in example (3). Then $\mathbf{c}_0$ is a linear space over $\mathbb{F}$.

Definition 11.1.8 The sequence space $\ell_0 = \ell_0(\mathbb{N})$. Let ℓ_0 denote the set of all sequences $x = (x_n)_1^{\infty}$ of real or complex numbers such that $x_i = 0$ for all but finitely many indices i. Define the operations of addition and scalar multiplication pointwise as in example (3). Then ℓ_0 is a linear space over $\mathbb{F}$.

Definition 11.1.9 A subset M of a linear space X is called a linear subspace of X if

1. $x+y \in M$ for all $x, y \in M$, and
2. $\lambda x \in M$ for all $x \in M$ and for all $\lambda \in \mathbb{F}$.

Clearly, a subset M of a linear space X is a linear subspace if and only if $M + M \subset M$ and $\lambda M \subset M$ for all $\lambda \in \mathbb{F}$

Definition 11.1.10 Let K be a subset of a linear space X. The linear hull of K, denoted by $\lim(K)$ or $\operatorname{span}(K)$, is the intersection of all linear subspaces of X that contain K

The linear hull of K is also called the linear subspace of X spanned (or generated) by K.

It is easy to check that the intersection of a collection of linear subspaces of X is a linear subspace of X. It therefore follows that the linear hull of a subset K of a linear space X is again a linear subspace of X

In fact, the linear hull of a subset K of a linear space X is the smallest linear subspace of X which contains$K$

Definition 11.1.11 Let K be a subset of a linear space X. Then the linear hull of K is the set of all finite linear combinations of elements of K. That is,

$$\operatorname{lin}(K) = \left\{ \sum_{j=1}^{n} \lambda_j x_j | x_1, x_2, \ldots, x_n \in K, \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}, n \in \mathbb{N} \right\}$$

Definition 11.1.12 If $\{x_1, x_2, \ldots, x_n\}$ is a linearly independent subset of X and In $X = \lim\{x_1, x_2, \ldots, x_n\}$, then X is said to have dimension n. In this case we say that $\{x_1, x_2, \ldots, x_n\}$ is a basis for the linear space X. If a linear space X does not have a finite basis, we say that it is infinite dimensional.

Definition 11.1.13 Let K be a subset of a linear space X. We say that

1. K is convex if $\lambda x + (1-\lambda)y \in K$ whenever $x, y \in K$ and $\lambda \in [0,1]$
2. K is balanced if $\lambda x \in K$ whenever $x \in K$ and $|\lambda| \leq 1$
3. K is absolutely convex if K is convex and balanced.

Proposition 11.1.1 1. It is easy to verify that K is absolutely convex if and only if $\lambda x + \mu y \in K$ whenever $x, y \in K$ and $|\lambda| + |\mu| \leq 1$

2. Every linear subspace is absolutely convex.

Definition 11.1.14 — Convex hull. Let S be a subset of the linear space X. The convex hull of S, denoted $\mathrm{co}(S)$, is the intersection of all convex sets in X which contain S. Since the intersection of convex sets is convex, it follows that $\mathrm{co}(S)$ is the smallest convex set which contains S.

Theorem 11.1.2 Let S be a nonempty subset of a linear space X. Then $\mathrm{co}(S)$ is the set of all convex combinations of elements of S. That is,

$$\mathrm{co}(S) = \left\{ \sum_{j=1}^{n} \lambda_j x_j | x_1, x_2, \ldots, x_n \in S, \lambda_j \geq 0 \forall j = 1, 2, \ldots, n, \sum_{j=1}^{n} \lambda_j = 1, n \in \mathbb{N} \right\}$$

Proof. Let C denote the set of all convex combinations of elements of S. That is,

$$C = \left\{ \sum_{j=1}^{n} \lambda_j x_j | x_1, x_2, \ldots, x_n \in S, \lambda_j \geq 0 \forall j = 1, 2, \ldots, n, \sum_{j=1}^{n} \lambda_j = 1, n \in \mathbb{N} \right\}$$

Let $x, y \in C$ and $0 \leq \lambda \leq 1$. Then $x = \sum_1^n \lambda_i x_i, y = \sum_1^m \mu_i y_i$, where $\lambda_i, \mu_i \geq 0, \sum_1^n \lambda_i = 1$ $\sum_1^m \mu_i = 1$, and $x_i, y_i \in S$. Thus

$$\lambda x + (1-\lambda)y = \sum_1^n \lambda \lambda_i x_i + \sum_1^m (1-\lambda)\mu_i y_i$$

is a linear combination of elements of S, with nonnegative coefficients, such that

$$\sum_1^n \lambda \lambda_i + \sum_1^m (1-\lambda)\mu_i = \lambda \sum_1^n \lambda_i + (1-\lambda)\sum_1^m \mu_i = \lambda + (1-\lambda) = 1$$

That is, $\lambda x + (1-\lambda)y \in C$ and C is convex. Clearly $S \subset C$. Hence $\mathrm{co}(S) \subset C$

We now prove the inclusion $C \subset \mathrm{co}(S)$. Note that, by definition, $S \subset \mathrm{co}(S)$. Let $x_1, x_2 \in S$ $\lambda_1 \geq 0, \lambda_2 \geq 0$ and $\lambda_1 + \lambda_2 = 1$. Then, by convexity of $\mathrm{co}(S), \lambda_1 x_1 + \lambda_2 x_2 \in \mathrm{co}(S)$. Assume that $\sum_1^{n-1} \lambda_i x_i \in \mathrm{co}(S)$ whenever $x_1, x_2, \ldots, x_{n-1} \in S, \lambda_j \geq 0, j = 1, 2, \ldots, n-1$ and $\sum_{j=1}^{n-1} \lambda_j = 1$. Let $x_1, x_2, \ldots, x_n \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_n$ be such that $\lambda_j \geq 0, j = 1, 2, \ldots, n$

and $\sum_{j=1}^{n}\lambda_j = 1$. If $\sum_{j=1}^{n-1}\lambda_j = 0$, then $\lambda_n = 1$. Hence $\sum_1^n \lambda_j x_j = \lambda_n x_n \in \mathrm{co}(S)$. Assume that $\beta = \sum_{j=1}^{n-1}\lambda_j > 0$. Then $\frac{\lambda_j}{\beta} \geq 0$ for all $j = 1,2,\dots,n-1$ and $\sum_{j=1}^{n-1}\frac{\lambda_j}{\beta} = 1$. By the induction assumption, $\sum_{j=1}^{n-1}\frac{\lambda_j}{\beta}x_j \in \mathrm{co}(S)$. Hence

$$\sum_{j=1}^{n}\lambda_j x_j = \beta\left(\sum_{j=1}^{n-1}\frac{\lambda_j}{\beta}x_j\right) + \lambda_n x_n \in \mathrm{co}(S)$$

Thus $C \subset \mathrm{co}(S)$. ■

11.1.1 Quotient Space

Definition 11.1.15 Let M be a linear subspace of a linear space X over $\mathbb{F}$. For all $x,y \in X$, define

$$x \equiv y(\operatorname{mod} M) \iff x - y \in M$$

It is easy to verify that $\equiv$ defines an equivalence relation on X For $x \in X$, denote by

$$[x] := \{y \in X : x \equiv y(\operatorname{mod} M)\} = \{y \in X : x - y \in M\} = x + M$$

the coset of x with respect to M. The quotient space X/M consists of all the equivalence classes $[x]$ $x \in X$. The quotient space is also called a factor space.

Theorem 11.1.3 Let M be a linear subspace of a linear space X over $\mathbb{F}$. For $x,y \in X$ and $\lambda \in \mathbb{F}$, define the operations

$$[x] + [y] = [x+y] \text{ and } \lambda \cdot [x] = [\lambda \cdot x]$$

Then X/M is a linear space with respect to these operations.

Note that the linear operations on X/M are equivalently given by: For all $x,y \in X$ and $\lambda \in \mathbb{F}$

$$(x+M) + (y+M) = x+y+M \text{ and } \lambda(x+M) = \lambda x + M$$

Definition 11.1.16 Let M be a linear subspace of a linear space X over $\mathbb{F}$. The codimension of M in X is defined as the dimension of the quotient space X/M. It is denoted by $\mathrm{codim}(M) = \dim(X/M)$

■ **Example 11.1** Clearly, if $X = M$, then $X/M = \{0\}$ and so $\mathrm{codim}(X) = 0$ ■

11.1.2 Direct Sums and Projections

Definition 11.1.17 Let M and N be linear subspaces of a linear space X over $\mathbb{F}$. We say that X is a direct sum of M and N if

$$X = M + N \text{ and } M \cap N = \{0\}$$

If X is a direct sum of M and N, we write $X = M \oplus N$. In this case, we say that M (resp. N) is an algebraic complement of N (resp. M).

Proposition 11.1.4 Let M and N be linear subspaces of a linear space X over $\mathbb{F}$. If $X = M \oplus N$, then each $x \in X$ has unique representation of the form $x = m + n$ for some $m \in M$ and $n \in N$.

Theorem 11.1.5 Let M and N be linear subspaces of a linear space X over $\mathbb{F}$ such that $X = M \oplus N$. Then $\operatorname{codim}(M) = \dim(N)$. Also, since $X = M \oplus N, \dim(X) = \dim(M) + \dim(N)$. Hence

$$\dim(X) = \dim(M) + \operatorname{codim}(M)$$

It follows that if $\dim(X) < \infty$, then $\operatorname{codim}(M) = \dim(X) - \dim(M)$

Definition 11.1.18 The operator $P : X \to X$ is called an algebraic projection if P is linear (i.e., $P(\alpha x + y) = \alpha Px + Py$ for all $x, y \in X$ and $\alpha \in \mathbb{F}$) and $P^2 = P$, i.e., P is idempotent.

Theorem 11.1.6 Let M and N be linear subspaces of a linear space X over $\mathbb{F}$ such that $X = M \oplus N$. Define $P : X \to X$ by $P(x) = m$, where $x = m + n$, with $m \in M$ and $n \in N$. Then P is an algebraic projection of X onto$M$ along N. Moreover $M = P(X)$ and $N = (I - P)(X) = \ker(P)$ Conversely, if $P : X \to X$ is an algebraic projection, then $X = M \oplus N$, where $M = P(X)$ and $N = (I - P)(X) = \ker(P)$

Proof. Linearity of $\mathbf{P}$: Let $x = m_1 + n_1$ and $y = m_2 + n_2$, where $m_1, m_2 \in M$ and $n_1, n_2 \in N$. For $\alpha \in \mathbb{F}$

$$P(\alpha x + y) = P((\alpha m_1 + m_2) + (\alpha n_1 + n_2)) = \alpha m_1 + m_2 = \alpha Px + Py$$

Idempotency of P : since $m = m + 0$, with $m \in M$ and $0 \in N$, we have that $Pm = m$ and hence $P^2 x = Pm = m = Px$. That is, $P^2 = P$ Finally, $n = x - m = (I - P)x$. Hence $N = (I - P)(X)$. Also, $Px = 0$ if and only if $x \in N$, i.e., $\ker(P) = N$ Conversely, let $x \in X$ and set $m = Px$ and $n = (I - P)x$. Then $x = m + n$, where $m \in M$ and $n \in N$. We show that this representation is unique. Indeed, if $x = m_1 + n_1$ where $m_1 \in M$ and $n_1 \in N$ then $m_1 = Pu$ and $n_1 = (I - P)v$ for some $u, v \in X$. since $P^2 = P$, it follows that $Pm_1 = m_1$ and $Pn_1 = 0$. Hence $m = Px = Pm_1 + Pn_1 = Pm_1 = m_1$. Similarly $n = n_1$

■

11.1.3 The Hölder and Minkowski Inequalities

Definition 11.1.19 — Conjugate exponents. Let p and q be positive real numbers. If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, or if $p = 1$ and $q = \infty$, or if $p = \infty$ and $q = 1$, then we say that p and q are conjugate exponents.

Theorem 11.1.7 — Young's Inequality. Let p and q be conjugate exponents, with $1 < p, q < \infty$ and $\alpha, \beta \geq 0$. Then

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$$

Proof. Proof. If $p = 2 = q$, then the inequality follows from the fact that $(\alpha - \beta)^2 \geq 0$. Notice also, that if $\alpha = 0$ or $\beta = 0$, then the inequality follows trivially. If $p \neq 2$, then consider the function $f : [0, \infty) \to \mathbb{R}$ given by

$$f(\alpha) = \frac{\alpha^p}{p} + \frac{\beta^q}{q} - \alpha\beta, \text{ for fixed } \beta > 0$$

Then, $f'(\alpha) = \alpha^{p-1} - \beta = 0$ when $\alpha^{p-1} = \beta$. That is, when $\alpha = \beta^{\frac{1}{p-1}} = \beta^{\frac{q}{p}} > 0$. We now apply the second derivative test to the critical point $\alpha = \beta^{\frac{q}{p}}$

$$f''(\alpha) = (p-1)\alpha^{p-2} > 0, \quad \text{for all } \alpha \in (0, \infty)$$

Thus, we have a global minimum at $\alpha = \beta^{\frac{q}{p}}$. It is easily verified that

$$0 = f\left(\beta^{\frac{q}{p}}\right) \leq f(\alpha) = \frac{\alpha^p}{p} + \frac{\beta^q}{q} - \alpha\beta \Leftrightarrow \alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$$

for each $\alpha \in [0, \infty)$

∎

Theorem 11.1.8 — Hölder's Inequality for sequences. Let $(x_n) \in \ell_p$ and $(y_n) \in \ell_q$, where $p > 1$ and $1/p + 1/q = 1$ Then

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}$$

Proof. If $\sum_{k=1}^{\infty} |x_k|^p = 0$ or $\sum_{k=1}^{\infty} |y_k|^q = 0$, then the inequality holds. Assume that $\sum_{k=1}^{\infty} |x_k|^p \neq 0$ and $\sum_{k=1}^{\infty} |y_k|^q \neq 0$. Then for $k = 1, 2, \ldots$, we have, by 11.1.7 , that

$$\frac{|x_k|}{\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}}} \cdot \frac{|y_k|}{\left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} \frac{|x_k|^p}{\sum_{k=1}^{\infty} |x_k|^p} + \frac{1}{q} \frac{|y_k|^q}{\sum_{k=1}^{\infty} |y_k|^q}$$

Hence,

$$\frac{\sum_{k=1}^{\infty} |x_k y_k|}{\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

That is,

$$\sum_{k=1}^{\infty}|x_k y_k| \leq \left(\sum_{k=1}^{\infty}|x_k|^p\right)^{\frac{1}{p}}\left(\sum_{k=1}^{\infty}|y_k|^q\right)^{\frac{1}{q}}$$

■

Theorem 11.1.9 — Minkowski's Inequality for sequences. Let $p > 1$ and (x_n) and (y_n) sequences in ℓ_p. Then

$$\left(\sum_{k=1}^{\infty}|x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{\infty}|x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty}|y_k|^p\right)^{\frac{1}{p}}$$

Proof. Let $q = \frac{p}{p-1}$. If $\sum_{k=1}^{\infty}|x_k + y_k|^p = 0$, then the inequality holds. We therefore assume that $\sum_{k=1}^{\infty}|x_k + y_k|^p \neq 0$. Then

$$\begin{aligned}
\sum_{k=1}^{\infty}|x_k + y_k|^p &= \sum_{k=1}^{\infty}|x_k + y_k|^{p-1}|x_k + y_k| \\
&\leq \sum_{k=1}^{\infty}|x_k + y_k|^{p-1}|x_k| + \sum_{k=1}^{\infty}|x_k + y_k|^{p-1}|y_k| \\
&\leq \left(\sum_{k=1}^{\infty}|x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}}\left[\left(\sum_{k=1}^{\infty}|x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty}|y_k|^p\right)^{\frac{1}{p}}\right] \\
&= \left(\sum_{k=1}^{\infty}|x_k + y_k|^p\right)^{\frac{1}{q}}\left[\left(\sum_{k=1}^{\infty}|x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty}|y_k|^p\right)^{\frac{1}{p}}\right]
\end{aligned}$$

Dividing both sides by $\left(\sum_{k=1}^{\infty}|x_k + y_k|^p\right)^{\frac{1}{q}}$, we have

$$\left(\sum_{k=1}^{\infty}|x_k + y_k|^p\right)^{\frac{1}{p}} = \left(\sum_{k=1}^{\infty}|x_k + y_k|^p\right)^{1-\frac{1}{q}} \leq \left(\sum_{k=1}^{\infty}|x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty}|y_k|^p\right)^{\frac{1}{p}}$$

■

11.2 Normed Linear Spaces

Definition 11.2.1 A norm on a linear space X is a real-valued function $\|\cdot\| : X \to \mathbb{R}$ which satisfies the following properties: For all $x, y \in X$ and $\lambda \in \mathbb{F}$

1. $\|x\| \geq 0$
2. $\|x\| = 0 \iff x = 0$
3. $\|\lambda x\| = |\lambda|\|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$ (Triangle Inequality).

A normed linear space is a pair $(X, \|\cdot\|)$, where X is a linear space and $\|\cdot\|$ a norm on X. The number $\|x\|$ is called the norm or length of x

■ **Example 11.2** Let $X = \mathbb{F}$. For each $x \in X$, define $\|x\| = |x|$. Then $(X, \|\cdot\|)$ is a normed linear space. We give the proof for $X = \mathbb{C}$. Properties N1 - N3 are easy to verify. We only verify N4. Let $x, y \in \mathbb{C}$. Then

$$\begin{aligned}
\|x+y\|^2 = |x+y|^2 = (x+y)\overline{(x+y)} = (x+y)(\bar{x}+\bar{y}) &= x\bar{x} + y\bar{x} + x\bar{y} + y\bar{y} \\
&= |x|^2 + \bar{x}\bar{y} + x\bar{y} + |y|^2 = |x|^2 + 2\Re e(x\bar{y}) + |y|^2 \\
&\leq |x|^2 + 2|x\bar{y}| + |y|^2 = |x|^2 + 2|x||\bar{y}| + |y|^2 \\
&= |x|^2 + 2|x||y| + |y|^2 \\
&= (|x| + |y|)^2 = (\|x\| + \|y\|)^2
\end{aligned}$$

Taking the positive square root both sides yields N4. ■

■ **Example 11.3** Let n be a natural number and $X = \mathbb{F}^n$. For each $x = (x_1, x_2, \ldots, x_n) \in X$, define

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \text{ for } 1 \leq p < \infty, \text{ and}$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Then $(X, \|\cdot\|_p)$ and $(X, \|\cdot\|_\infty)$ are normed linear spaces. We give a detailed proof that $(X, \|\cdot\|_p)$ is a normed linear space for $1 \leq p < \infty$ ■

Proof. N4. For any $x, y \in X$

$$\begin{aligned}
\|x+y\|_p &= \left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \\
&\leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}} \qquad \text{(by Minkowski's Inequality)} \\
&= \|x\|_p + \|y\|_p
\end{aligned}$$

■

Definition 11.2.2 Let a be an element of a normed linear space $(X, \|\cdot\|)$ and $r > 0$

$B(a, r) = \{x \in X | \|x - a\| < r\}$ (Open ball with centre a and radius r);

$B[a, r] = \{x \in X | \|x - a\| \leq r\}$ (Closed ball with centre a and radius r);

$S(a, r) = \{x \in X | \|x - a\| = r\}$ (Sphere with centre a and radius r)

Definition 11.2.3 — Norm equivalent. Let $\|\cdot\|$ and $\|\cdot\|_0$ be two different norms defined on the same linear space X. We say that $\|\cdot\|$ is equivalent to $\|\cdot\|_0$ if there are positive numbers α and β such that

$$\alpha\|x\| \leq \|x\|_0 \leq \beta\|x\|, \text{ for all } x \in X$$

Theorem 11.2.1 Let $X = \mathbb{F}^n$. For each $x = (x_1, x_2, \ldots, x_n) \in X$, let

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}, \text{ and } \|x\|_\infty = \max_{1\leq i\leq n} |x_i|$$

We show they are equivalent.

Proof. Equivalence of $\|\cdot\|_1$ and $\|\cdot\|_\infty$: Let $x = (x_1, x_2, \ldots, x_n) \in X$. For each $k = 1, 2, \ldots, n$

$$|x_k| \leq \sum_{i=1}^{n} |x_i| \Rightarrow \max_{1\leq k\leq n} |x_k| \leq \sum_{i=1}^{n} |x_i| \iff \|x\|_\infty \leq \|x\|_1$$

Also, for $k = 1, 2, \ldots, n$

$$|x_k| \leq \max_{1\leq k\leq n} |x_k| = \|x\|_\infty \Rightarrow \sum_{i=1}^{n} |x_i| \leq \sum_{i=1}^{n} \|x\|_\infty = n\|x\|_\infty \iff \|x\|_1 \leq n\|x\|_\infty$$

Hence, $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$ We now show that $\|\cdot\|_2$ is equivalent to $\|\cdot\|_\infty$. Let $x = (x_1, x_2, \ldots, x_n) \in X$. For each The the stres of the solution of the solution the form $k = 1$

$$|x_k| \leq \|x\|_\infty \Rightarrow |x_k|^2 \leq (\|x\|_\infty)^2 \Rightarrow \sum_{i=1}^{n} |x_i|^2 \leq \sum_{i=1}^{n} (\|x\|_\infty)^2 = n\,(\|x\|_\infty)^2$$
$$\iff \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

Also, for each $k = 1, 2, \ldots, n$

$$|x_k| \leq \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} = \|x\|_2 \Rightarrow \max_{1\leq k\leq n} |x_k| \leq \|x\|_2 \iff \|x\|_\infty \leq \|x\|_2$$

Consequently, $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$, which proves equivalence of the norms $\|\cdot\|_2$ and $\|\cdot\|_\infty$

■

R We shall see later that all norms on a finite-dimensional normed linear space are equivalent.

11.2.1 Open

Definition 11.2.4 A subset S of a normed linear space $(X, \|\cdot\|)$ is open if for each $s \in S$ there is an $\epsilon > 0$ such that $B(s, \epsilon) \subset S$ A subset F of a normed linear space $(X, \|\cdot\|)$ is closed if its complement $X \backslash F$ is open.

Definition 11.2.5 Let S be a subset of a normed linear space $(X, \|\cdot\|)$. We define the closure of S, denoted by S, to be the intersection of all closed sets containing S.

It is easy to show that S is closed if and only if $S = \bar{S}$

Definition 11.2.6 Recall that a metric on a set X is a real-valued function $d : X \times X \to \mathbb{R}$ which satisfies the following properties: For all $x, y, z \in X$

1. $d(x,y) \geq 0$
2. $d(x,y) = 0 \iff x = y$
3. $d(x,y) = d(y,x)$
4. $d(x,z) \leq d(x,y) + d(y,z)$

Theorem 11.2.2

1. If $(X, \|\cdot\|)$ is a normed linear space, then

$$d(x,y) = \|x - y\|$$

defines a metric on X. Such a metric d is said to be induced or generated by the norm $\|\cdot\|$. Thus, every normed linear space is a metric space, and unless otherwise specified, we shall henceforth regard any normed linear space as a metric space with respect to the metric induced by its norm.

2. If d is a metric on a linear space X satisfying the properties: For all $x, y, z \in X$ and for all $\lambda \in \mathbb{F}$
   (a) $d(x,y) = d(x+z, y+z)$ (Translation Invariance)
   (b) $d(\lambda x, \lambda y) = |\lambda| d(x,y)$ (Absolute Homogeneity)

   then

$$\|x\| = d(x,0)$$

defines a norm on X

R A metric d on a linear space X is induced by a norm on X if and only if d is translation-invariant and positive homogeneous.

R Every normed linear space is a metric space

Definition 11.2.7 We now want to introduce a norm on a quotient space. Let M be a closed linear subspace of a normed linear space X over $\mathbb{F}$. For $x \in X$, define

$$\|[x]\| := \inf_{y \in [x]} \|y\|$$

If $y \in [x]$, then $y - x \in M$ and hence $y = x + m$ for some $m \in M$. Hence

$$\|[x]\| = \inf_{y \in [x]} \|y\| = \inf_{m \in M} \|x + m\| = \inf_{m \in M} \|x - m\| = d(x, M)$$

Theorem 11.2.3 Let M be a closed linear subspace of a normed linear space X over $\mathbb{F}$. The quotient space X/M is a normed linear space with respect to the norm

$$\|[x]\| := \inf_{y\in[x]} \|y\|, \text{ where } [x] \in X/M$$

The norm on X/M as defined in theorem 11.2.3 is called the quotient norm on X/M.

Proof. N1. It is clear that for any $x \in X, \|[x]\| = d(x,M) \geq 0$

N2. For any $x \in X$

$$\|[x]\| = 0 \iff d(x,M) = 0 \iff x \in \overline{M} = M \iff x + M = M = [0]$$

M is close set. N3. For any $x,y \in X$ and $\lambda \in \mathbb{F}\backslash\{0\}$

$$\|\lambda[x]\| = \|[\lambda x]\| = d(\lambda x, M) = \inf_{y\in M} \|\lambda x - y\| = \inf_{y\in M} \left\|\lambda\left(x - \frac{y}{\lambda}\right)\right\|$$
$$= |\lambda| \inf_{z\in M} \|x - z\| = |\lambda| d(x,M) = |\lambda|\|[x]\|$$

N4. Let $x,y \in X$. Then

$$\begin{aligned}
\|[x] + [y]\| = \|[x+y]\| = d(x+y,M) &= \inf_{z\in M} \|x+y-z\| \\
&= \inf_{z_1,z_2\in M} \|x+y-(z_1+z_2)\| \\
&= \inf_{z_1,z_2\in M} \|(x-z_1)+(y-z_2)\| \\
&\leq \inf_{z_1,z_2\in M} \|x-z_1\| + \|y-z_2\| \\
&= \inf_{z_1\in M} \|x-z_1\| + \inf_{z_2\in M} \|y-z_2\| \\
&= d(x,M) + d(y,M) = \|[x]\| + \|[y]\|
\end{aligned}$$

■

Definition 11.2.8 — Quotient map. Let M be a closed subspace of the normed linear space X. The mapping Q_M from $X \to X/M$ defined by

$$Q_M(x) = x + M, \quad x \in X$$

is called the **quotient map** (or natural embedding) of X onto X/M

11.2.2 Completeness of Normed Linear Spaces

Definition 11.2.9 Let $(x_n)_{n=1}^{\infty}$ be a sequence in a normed linear space $(X, \|\cdot\|)$

1. $(x_n)_{n=1}^{\infty}$ is said to converge to x if given $\epsilon > 0$ there exists a natural number $N = N(\epsilon)$ such that

$$\|x_n - x\| < \epsilon \text{ for all } n \geq N$$

Equivalently, $(x_n)_{n=1}^{\infty}$ converges to x if

$$\lim_{n\to\infty} \|x_n - x\| = 0$$

If this is the case, we shall write

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

Convergence in the norm is called **norm convergence or strong convergence**.

2. $(x_n)_{n=1}^{\infty}$ is called a Cauchy sequence if given $\epsilon > 0$ there exists a natural number $N = N(\epsilon)$ such that

$$\|x_n - x_m\| < \epsilon \text{ for all } n, m \geq N$$

Equivalently, (x_n) is Cauchy if

$$\lim_{n,m\to\infty} \|x_n - x_m\| = 0$$

Theorem 11.2.4 Let C be a closed set in a normed linear space $(X, \|\cdot\|)$ over $\mathbb{F}$, and let (x_n) be a sequence contained in C such that $\lim_{n\to\infty} x_n = x \in X$. Then $x \in C$

Theorem 11.2.5 Let X be a normed linear space and A a nonempty subset of X

1. $|d(x,A) - d(y,A)| \leq \|x - y\|$ for all $x, y \in X$
2. $|\|x\| - \|y\|| \leq \|x - y\|$ for all $x, y \in X$
3. If $x_n \to x$, then $\|x_n\| \to \|x\|$
4. If $x_n \to x$ and $y_n \to y$, then $x_n + y_n \to x + y$
5. If $x_n \to x$ and $\alpha_n \to \alpha$, then $\alpha_n x_n \to \alpha x$
6. The closure of a linear subspace in X is again a linear subspace;
7. Every Cauchy sequence is bounded;
8. Every convergent sequence is a Cauchy sequence.

Proof. (1). For any $a \in A$

$$d(x,A) \leq \|x - a\| \leq \|x - y\| + \|y - a\|$$

so $d(x,A) \leq \|x - y\| + d(y,A)$ or $d(x,A) - d(y,A) \leq \|x - y\|$. Interchanging the roles of x and y gives the desired result.

(2) follows from (1) by taking $A = \{0\}$.

(3) is an obvious consequence of (2).

(4), (5) and (8) follow from the triangle inequality and, in the case of (5), the absolute homogeneity.

(6) follows from (4) and (5).

(7). Let (x_n) be a Cauchy sequence in X. Choose n_1 so that $\|x_n - x_{n_1}\| \leq 1$ for all $n \geq n_1$. By (2) , $\|x_n\| \leq 1 + \|x_{n_1}\|$ for all $n \geq n_1$. Thus

$$\|x_n\| \leq \max\{\|x_1\|, \|x_2\|, \|x_3\|, \ldots, \|x_{n_1-1}\|, 1 + \|x_{n_1}\|\}$$

for all n.

(8) Let (x_n) be a sequence in X which converges to $x \in X$ and let $\epsilon > 0$. Then there is a natural number N such that $\|x_n - x\| < \frac{\epsilon}{2}$ for all $n \geq N$. For all $n, m \geq N$

$$\|x_n - x_m\| \leq \|x_n - x\| + \|x - x_m\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, (x_n) is a Cauchy sequence in X ■

Theorem 11.2.6 Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$. A Cauchy sequence in X which has a convergent subsequence is convergent.

Proof. Proof. Let (x_n) be a Cauchy sequence in X and (x_{n_k}) its subsequence which converges to $x \in X$. Then, for any $\epsilon > 0$, there are positive integers N_1 and N_2 such that

$$\|x_n - x_m\| < \frac{\epsilon}{2} \text{ for all } n, m \geq N_1$$

and

$$\|x_{n_k} - x\| < \frac{\epsilon}{2} \text{ for all } k \geq N_2$$

Let $N = \max\{N_1, N_2\}$. If $k \geq N$, then since $n_k \geq k$

$$\|x_k - x\| \leq \|x_k - x_{n_k}\| + \|x_{n_k} - x\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Hence $x_n \to x$ as $n \to \infty$ ■

Definition 11.2.10 — Complete. A metric space (X, d) is said to be complete if every Cauchy sequence in X converges in X.

Definition 11.2.11 — Banach space. A normed linear space that is complete with respect to the metric induced by the norm is called a Banach space.

Theorem 11.2.7 — Subspace complete equal to closed. Let $(X, \|\cdot\|)$ be a Banach space and let M be a linear subspace of X. Then M is complete if and only if the M is closed in X

Proof. Assume that M is complete. We show that M is closed. To that end, let $x \in \overline{M}$. Then there is a sequence (x_n) in M such that $\|x_n - x\| \to 0$ as $n \to \infty$. since (x_n) converges, it is Cauchy. Completeness of M guarantees the existence of an element $y \in M$ such that

$\|x_n - y\| \to 0$ as $n \to \infty$. By uniqueness of limits, $x = y$. Hence $x \in M$ and, consequently, M is closed.

Assume that M is closed. We show that M is complete. Let (x_n) be a Cauchy sequence in M. Then (x_n) is a Cauchy sequence in X. Since X is complete, there is an element $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$. But then $x \in M$ since M is closed. Hence M is complete. ■

■ **Example 11.4** Let $1 \le p < \infty$. The sequence space ℓ_p is a Banach space.

The classical sequence space ℓ_p is complete. ■

Proof. Let $(x_n)_1^\infty$ be a Cauchy sequence in ℓ_p. We shall denote each member of this sequence by

$$x_n = (x_n(1), x_n(2), \ldots)$$

Then, given $\epsilon > 0$, there exists an $N(\epsilon) = N \in \mathbb{N}$ such that

$$\|x_n - x_m\|_p = \left(\sum_{i=1}^{\infty} |x_n(i) - x_m(i)|^p\right)^{\frac{1}{p}} < \epsilon \text{ for all } n, m \ge N$$

For each fixed index i, we have

$$|x_n(i) - x_m(i)| < \epsilon \quad \text{for all} \quad n, m \ge N$$

That is, for each fixed index i, $(x_n(i))_1^\infty$ is a Cauchy sequence in $\mathbb{F}$. since $\mathbb{F}$ is complete, there exists $x(i) \in \mathbb{F}$ such that

$$x_n(i) \to x(i) \text{ as } n \to \infty$$

Define $x = (x(1), x(2), \ldots)$. We show that $x \in \ell_p$, and $x_n \to x$. To that end, for each $k \in \mathbb{N}$

$$\left(\sum_{i=1}^{k} |x_n(i) - x_m(i)|^p\right)^{\frac{1}{p}} \le \|x_n - x_m\|_p = \left(\sum_{i=1}^{\infty} |x_n(i) - x_m(i)|^p\right)^{\frac{1}{p}} < \epsilon$$

That is,

$$\sum_{i=1}^{k} |x_n(i) - x_m(i)|^p < \epsilon^p, \text{ for all } k = 1, 2, 3, \ldots$$

Keep k and $n \ge N$ fixed and let $m \to \infty$. since we are dealing with a finite sum, What is the solution of the solution

$$\sum_{i=1}^{k} |x_n(i) - x(i)|^p \le \epsilon^p$$

Now letting $k \to \infty$, then for all $n \ge N$

$$\sum_{i=1}^{\infty} |x_n(i) - x(i)|^p \le \epsilon^p \tag{11.1}$$

which means that $x_n - x \in \ell_p$. since $x_n \in \ell_p$, we have that $x = (x - x_n) + x_n \in \ell_p$. It also follows from eq. (11.1) that $x_n \to x$ as $n \to \infty$ ■

■ **Example 11.5** The space ℓ_0 of all sequences $(x_i)_1^\infty$ with only a finite number of nonzero terms is an incomplete normed linear space. It suffices to show that ℓ_0 is not closed in ℓ_2 (and hence not complete). To that end, consider the sequence $(x_i)_1^\infty$ with terms

$$
\begin{aligned}
x_1 &= (1,0,0,0,\ldots)\\
x_2 &= \left(1,\frac{1}{2},0,0,0,\ldots\right)\\
x_3 &= \left(1,\frac{1}{2},\frac{1}{2^2},0,0,0,\ldots\right)\\
&\vdots\\
x_n &= \left(1,\frac{1}{2},\frac{1}{2^2},\ldots,\frac{1}{2^{n-1}},0,0,0,\ldots\right)
\end{aligned}
$$

This sequence $(x_i)_1^\infty$ converges to

$$x = \left(1,\frac{1}{2},\frac{1}{2^2},\ldots,\frac{1}{2^{n-1}},\frac{1}{2^n},\frac{1}{2^{n+1}},\ldots\right)$$

Indeed, since $x - x_n = \left(0,0,0,\ldots,0,\frac{1}{2^n},\frac{1}{2^{n+1}},\ldots\right)$, it follows that

$$\|x_n - x\|^2 = \sum_{k=n}^{\infty} \frac{1}{2^{2k}} \to 0 \text{ as } n \to \infty$$

That is, $x_n \to x$ as $n \to \infty$, but $x \notin \ell_0$ ■

■ **Example 11.6** The space $\mathcal{C}_2[-1,1]$ of continuous real-valued functions on [-1,1] with the norm

$$\|x\|_2 = \left(\int_{-1}^{1} x^2(t)dt\right)^{1/2}$$

is an incomplete normed linear space. To see this, it suffices to show that there is a Cauchy sequence in $\mathcal{C}_2[-1,1]$ which converges to an element which does not belong to $\mathcal{C}_2[-1,1]$. Consider the sequence $(x_n)_1^\infty \in \mathcal{C}_2[-1,1]$ defined by

$$x_n(t) = \begin{cases} 0 & \text{if} \quad -1 \le t \le 0 \\ nt & \text{if} \quad 0 \le t \le \frac{1}{n} \\ 1 & \text{if} \quad \frac{1}{n} \le t \le 1 \end{cases}$$

We show that $(x_n)_1^\infty$ is a Cauchy sequence in $\mathcal{C}_2[-1,1]$. To that end, for positive integers

m and n such that $m > n$

$$\begin{aligned}
\|x_n - x_m\|_2^2 &= \int_{-1}^{1} [x_n(t) - x_m(t)]^2 \, dt \\
&= \int_0^{1/m} [nt - mt]^2 dt + \int_{1/m}^{1/n} [1 - nt]^2 dt \\
&= \int_0^{1/m} \left[m^2t^2 - 2mnt^2 + n^2t^2\right] dt + \int_{1/m}^{1/n} \left[1 - 2nt + n^2t^2\right] dt \\
&= \left(m^2 - 2mn + n^2\right) \frac{t^3}{3}\bigg|_0^{1/m} + \left(t - nt^2 + n^2\frac{t^3}{3}\right)\bigg|_{1jm}^{1/n} \\
&= \frac{m^2 - 2mn + n^2}{3m^2n} = \frac{(m-n)^2}{3m^2n} \to 0 \quad \text{as} \quad n, m \to \infty
\end{aligned}$$

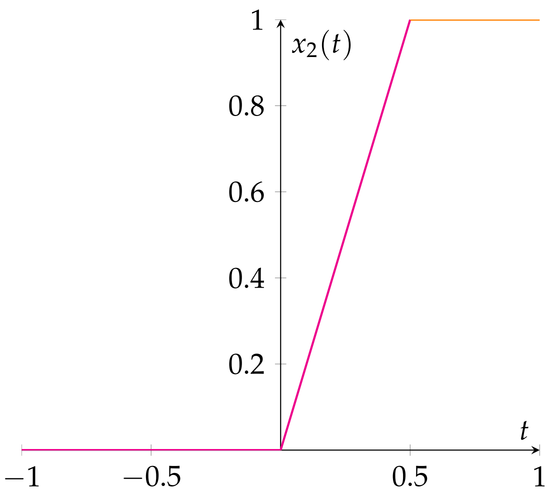
Define

$$x(t) = \begin{cases} 0 & \text{if } -1 \le t \le 0 \\ 1 & \text{if } \ \ 0 < t \le 1 \end{cases}$$

Then $x \notin \mathcal{C}_2[-1,1]$, and

$$\|x_n - x\|_2^2 = \int_{-1}^{1} [x_n(t) - x(t)]^2 \, dt = \int_0^{\frac{1}{n}} [nt - 1]^2 dt = \frac{1}{3n} \to 0 \quad \text{as} \quad n \to \infty$$

That is, $x_n \to x$ as $n \to \infty$ ■



11.2.3 Series in Normed Linear Spaces

Definition 11.2.12 Let (x_n) be a sequence in a normed linear space $(X, \|\cdot\|)$. To this sequence we associate another sequence (s_n) of partial sums, where $s_n = \sum_{k=1}^{n} x_k$.

Let (x_n) be a sequence in a normed linear space $(X, \|\cdot\|)$. If the sequence (s_n) of partial sums converges to s, then we say that the series $\sum_{k=1}^{\infty} x_k$ converges and that its sum is s. In this case we write $\sum_{k=1}^{\infty} x_k = s$ The series $\sum_{k=1}^{\infty} x_k$ is said to be absolutely convergent if $\sum_{k=1}^{\infty} \|x_k\| < \infty$

Theorem 11.2.8 A normed linear space $(X, \|\cdot\|)$ is a Banach space if and only if every absolutely convergent series in X is convergent.

Proof. Let X be a Banach space and suppose that $\sum_{j=1}^{\infty} \|x_j\| < \infty$. We show that the series $\sum_{j=1}^{\infty} x_j$ converges. To that end, let $\epsilon > 0$ and for each $n \in \mathbb{N}$, let $s_n = \sum_{j=1}^{n} x_j$. Let K be a positive integer such that $\sum_{j=K+1}^{\infty} \|x_j\| < \epsilon$. Then, for all $m > n > K$, we have

$$\|s_m - s_n\| = \left\| \sum_{1}^{m} x_j - \sum_{1}^{n} x_j \right\| = \left\| \sum_{n+1}^{m} x_j \right\| \le \sum_{n+1}^{m} \|x_j\| \le \sum_{n+1}^{\infty} \|x_j\| \le \sum_{K+1}^{\infty} \|x_j\| < \epsilon$$

Hence the sequence (s_n) of partial sums forms a Cauchy sequence in X. since X is complete, the sequence (s_n) converges to some element $s \in X$. That is, the series $\sum_{j=1}^{\infty} x_j$ converges.

Conversely, assume that $(X, \|\cdot\|)$ is a normed linear space in which every absolutely convergent series converges. We show that X is complete. Let (x_n) be a Cauchy sequence in X. Then there is an $n_1 \in \mathbb{N}$ such that $\|x_{n_1} - x_m\| < \frac{1}{2}$ whenever $m > n_1$. Similarly, there is an $n_2 \in \mathbb{N}$ with $n_2 > n_1$ such that $\|x_{n_2} - x_m\| < \frac{1}{2^2}$ whenever $m > n_2$. Continuing in this way, we get natural numbers $n_1 < n_2 < \cdots$ such that $\|x_{n_k} - x_m\| < \frac{1}{2^k}$ whenever $m > n_k$. In particular, we have that for each $k \in \mathbb{N}$, $\|x_{n_{k+1}} - x_{n_k}\| < 2^{-k}$. For each $k \in \mathbb{N}$, let $y_k = x_{n_{k+1}} - x_{n_k}$. Then

$$\sum_{k=1}^{n} \|y_k\| = \sum_{k=1}^{n} \|x_{n_{k+1}} - x_{n_k}\| < \sum_{k=1}^{n} \frac{1}{2^k}$$

Hence, $\sum_{k=1}^{\infty} \|y_k\| < \infty$. That is, the series $\sum_{k=1}^{\infty} y_k$ is absolutely convergent, and hence, by our assumption, the series $\sum_{k=1}^{\infty} y_k$ is convergent in X. That is, there is an $s \in X$ such that $s_j = \sum_{k=1}^{j} y_k \to s$ as $j \to \infty$. It follows that

$$s_j = \sum_{k=1}^{j} y_k = \sum_{k=1}^{j} \left[x_{n_{k+1}} - x_{n_k}\right] = x_{n_{j+1}} - x_{n_1} \xrightarrow{j \to \infty} s$$

Hence $x_{n_{j+1}} \xrightarrow{j \to \infty} s + x_{n_1}$. Thus, the subsequence (x_{n_k}) of (x_n) converges in X. But if a Cauchy sequence has a convergent subsequence, then the sequence itself also converges (to the same limit as the subsequence). It thus follows that the sequence (x_n) also converges in X. Hence X is complete. ■

Theorem 11.2.9 Let M be a closed linear subspace of a Banach space X. Then the quotient space X/M is a Banach space when equipped with the quotient norm.

Proof. Let $([x_n])$ be a sequence in X/M such that $\sum_{j=1}^{\infty} \|[x_j]\| < \infty$. For each $j \in \mathbb{N}$, choose an element $y_j \in M$ such that

$$\|x_j - y_j\| \le \|[x_j]\| + 2^{-j}$$

It now follows that $\sum_{j=1}^{\infty}\|x_j - y_j\| < \infty$, i.e., the series $\sum_{j=1}^{\infty}(x_j - y_j)$ is absolutely convergent in X. since X is complete, the series $\sum_{i=1}^{\infty}(x_j - y_j)$ converges to some element $z \in X$. We show that the series $\sum_{j=1}^{\infty}[x_j]$ converges to [z]. Indeed, for each $n \in \mathbb{N}$

$$\begin{aligned}
\left\|\sum_{j=1}^{n}[x_j] - [z]\right\| &= \left\|\left[\sum_{j=1}^{n}x_j\right] - [z]\right\| = \left\|\left[\sum_{j=1}^{n}x_j - z\right]\right\| \\
&= \inf_{m\in M}\left\|\sum_{j=1}^{n}x_j - z - m\right\| \\
&\leq \left\|\sum_{j=1}^{n}x_j - z - \sum_{j=1}^{n}y_j\right\| \\
&= \left\|\sum_{j=1}^{n}(x_j - y_j) - z\right\| \to 0 \text{ as } n \to \infty
\end{aligned}$$

Hence, every absolutely convergent series in X/M is convergent, and so X/M is complete. ■

11.2.4 Bounded, Totally Bounded, and Compact Subsets of a Normed Linear Space

Definition 11.2.13 A subset A of a normed linear space $(X, \|\cdot\|)$ is bounded if $A \subset B[x,r]$ for some $x \in X$ and $r > 0$.

It is clear that A is bounded if and only if there is a $C > 0$ such that $\|a\| \leq C$ for all $a \in A$

Definition 11.2.14 — ϵ -net. Let A be a subset of a normed linear space $(X, \|\cdot\|)$ and $\epsilon > 0$. A subset $A_\epsilon \subset X$ is called an ϵ **-net** for A if for each $x \in A$ there is an element $y \in A_\epsilon$ such that $\|x - y\| < \epsilon$. Simply put, $A_\epsilon \subset X$ is an ϵ -net for A if each element of A is within an ϵ distance to some element of A_ϵ .

Definition 11.2.15 — Totally bounded. A subset A of a normed linear space $(X, \|\cdot\|)$ is totally bounded (or precompact) if for any $\epsilon > 0$ there is a finite ϵ -net $F_\epsilon \subset X$ for A. That is, there is a finite set $F_\epsilon \subset X$ such that

$$A \subset \bigcup_{x\in F_\epsilon} B(x,\epsilon)$$

Definition 11.2.16 Every totally bounded subset of a normed linear space $(X, \|\cdot\|)$ is bounded.

Proof. This follows from the fact that a finite union of bounded sets is also bounded. ■

■ **Example 11.7** Let $X = \ell_2$ and consider $B = B(X) = \{x \in X | \|x\| \leq 1\}$, the closed unit ball in X. Clearly, B is bounded. We show that B is not totally bounded.

Consider the elements of B of the form: for $j \in \mathbb{N}, e_j = (0,0,\ldots,0,1,0,\ldots)$, where 1 occurs in the j -th position. Note that $\|e_i - e_j\|_2 = \sqrt{2}$ for all $i \neq j$. Assume that an ϵ -

net $B_\epsilon \subset X$ existed for $0 < \epsilon < \frac{\sqrt{2}}{2}$. Then for each $j \in \mathbb{N}$ there is an element $y_j \in B_\epsilon$ such that $\|e_j - y_j\| < \epsilon$. This says that for each $j \in \mathbb{N}$, there is an element $y_j \in B_\epsilon$ such that $y_j \in B(e_j, \epsilon)$. But the balls $B(e_j, \epsilon)$ are disjoint. Indeed, if $i \neq j$, and $z \in B(e_i, \epsilon) \cap B(e_j, \epsilon)$, then by the triangle inequality

$$\sqrt{2} = \|e_i - e_j\|_2 \leq \|e_i - z\| + \|z - e_j\| < 2\epsilon < \sqrt{2}$$

which is absurd. since the balls $B(e_j, \epsilon)$ are (at least) countably infinite, there can be no finite ϵ -net for B. ∎

Theorem 11.2.10 A subset A of a normed linear space $(X, \|\cdot\|)$ is totally bounded if and only if for any $\epsilon > 0$ there is a finite set $F_\epsilon \subset A$ such that

$$A \subset \bigcup_{x \in F_\epsilon} B(x, \epsilon)$$

R In our definition of total boundedness of a subset $A \subset X$, we required that the finite ϵ -net be a subset of X. The following proposition suggests that the finite ϵ -net may actually be assumed to be a subset of A itself.

Theorem 11.2.11 A subset K of a normed linear space $(X, \|\cdot\|)$ is totally bounded if and only if every sequence in K has a Cauchy subsequence.

Proof. Assume that K is totally bounded and let (x_n) be an infinite sequence in K. There is a finite set of points $\{y_{11}, y_{12}, \ldots, y_{1r}\}$ in K such that

$$K \subset \bigcup_{j=1}^{r} B\left(y_{1j}, \frac{1}{2}\right)$$

At least one of the balls $B\left(y_{1j}, \frac{1}{2}\right), j = 1, 2, \ldots, r$, contains an infinite subsequence (x_{n1}) of (x_n). Again, there is a finite set $\{y_{21}, y_{22}, \ldots, y_{2s}\}$ in K such that

$$K \subset \bigcup_{j=1}^{s} B\left(y_{2j}, \frac{1}{2^2}\right)$$

At least one of the balls $B\left(y_{2j}, \frac{1}{2^2}\right), j = 1, 2, \ldots, s$, contains an infinite subsequence (x_{n2}) of (x_{n1}) Continuing in this way, at the m -th step, we obtain a subsequence (x_{nm}) of $\left(x_{n(m-1)}\right)$ which is contained in a ball of the form $B\left(y_{mj}, \frac{1}{2^m}\right)$.

Claim: The diagonal subsequence (x_{nn}) of (x_n) is Cauchy. Indeed, if $m > n$, then both x_{nn} and x_{mm} are in the ball of radius 2^{-n}. Hence, by the triangle inequality,

$$\|x_{nn} - x_{mm}\| < 2^{1-n} \to 0 \text{ as } n \to \infty$$

Conversely, assume that every sequence in K has a Cauchy subsequence and that K is not totally bounded. Then, for some $\epsilon > 0$, no finite ϵ -net exists for K. Hence, if $x_1 \in K$, then there is an $x_2 \in K$ such that $\|x_1 - x_2\| \geq \epsilon$. (Otherwise, $\|x_1 - y\| < \epsilon$ for all $y \in K$ and consequently $\{x_1\}$ is a finite ϵ -net for K, a contradiction. Similarly, there is an $x_3 \in K$ such that

$$\|x_1 - x_3\| \geq \epsilon \text{ and } \|x_2 - x_3\| \geq \epsilon$$

Continuing in this way, we obtain a sequence (x_n) in K such that $\|x_n - x_m\| \geq \epsilon$ for all $m \neq n$. Therefore (x_n) cannot have a Cauchy subsequence, a contradiction.

■

Definition 11.2.17 — Sequentially compact. A normed linear space $(X, \|\cdot\|)$ is sequentially compact if every sequence in X has a convergent subsequence.

R It can be shown that on a metric space, compactness and sequential compactness are equivalent. Thus, it follows, that on a normed linear space, we can use these terms interchangeably.

Theorem 11.2.12 A subset of a normed linear space is sequentially compact if and only if it is totally bounded and complete.

Proof. Let K be a sequentially compact subset of a normed linear space $(X, \|\cdot\|)$. We show that K is totally bounded. To that end, let (x_n) be a sequence in K. By sequential compactness of $K, (x_n)$ has a subsequence (x_{n_k}) which converges in K. since every convergent sequence is Cauchy, the subsequence (x_{n_k}) of (x_n) is Cauchy. Therefore, by Theorem 2.5.1, K is totally bounded.

Next, we show that K is complete. Let (x_n) be a Cauchy sequence in K. By sequential compactness of $K, (x_n)$ has a subsequence (x_{n_k}) which converges in K. But if a subsequence of a Cauchy sequence converges, so does the full sequence. Hence (x_n) converges in K and so K is complete.

Conversely, assume that K is a totally bounded and complete subset of a normed linear space $(X, \|\cdot\|)$ We show that K is sequentially compact. Let (x_n) be a sequence in K. By Theorem 2.5.1, (x_n) has a Cauchy subsequence (x_{n_k}). since K is complete, (x_{n_k}) converges in K. Hence K is sequentially compact. ■

Corollary 11.2.13 A subset of a Banach space is sequentially compact if and only if it is totally bounded and closed.

Corollary 11.2.14 A sequentially compact subset of a normed linear space is closed and bounded.

Corollary 11.2.15 A closed subset F of a sequentially compact normed linear space $(X, \|\cdot\|)$ is sequentially compact.

11.2.5 Finite Dimensional Normed Linear Spaces

Theorem 11.2.16 Let $(X, \|\cdot\|)$ be a finite-dimensional normed linear space with basis $\{x_1, x_2, \ldots, x_n\}$. Then there is a constant $m > 0$ such that for every choice of scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, we have

$$m \sum_{j=1}^{n} |\alpha_j| \leq \left\| \sum_{j=1}^{n} \alpha_j x_j \right\|$$

Proof. Proof. If $\sum_{j=1}^{n} |\alpha_j| = 0$, then $\alpha_j = 0$ for all $j = 1, 2, \ldots, n$ and the inequality holds for any $m > 0$ Assume that $\sum_{j=1}^{n} |\alpha_j| \neq 0$. We shall prove the result for a set of scalars $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ that satisfy the condition $\sum_{j=1}^{n} |\alpha_j| = 1$. Let

$$A = \left\{ (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n \,\middle|\, \sum_{j=1}^{n} |\alpha_j| = 1 \right\}$$

since A is a closed and bounded subset of $\mathbb{F}^n$, it is compact. Define $f : A \to \mathbb{R}$ by

$$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = \left\| \sum_{j=1}^{n} \alpha_j x_j \right\|$$

since for any $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $(\beta_1, \beta_2, \ldots, \beta_n)$ in A

$$\begin{aligned}
|f(\alpha_1, \alpha_2, \ldots, \alpha_n) - f(\beta_1, \beta_2, \ldots, \beta_n)| &= \left\| \sum_{j=1}^{n} \alpha_j x_j \right\| - \left\| \sum_{j=1}^{n} \beta_j x_j \right\| \\
&\leq \left| \sum_{j=1}^{n} \alpha_j x_j - \sum_{j=1}^{n} \beta_j x_j \right| \\
&= \left\| \sum_{j=1}^{n} (\alpha_j - \beta_j) x_j \right\| \leq \sum_{j=1}^{n} |\alpha_j - \beta_j| \|x_j\| \\
&\leq \max_{1 \leq j \leq n} \|x_j\| \sum_{j=1}^{n} |\alpha_j - \beta_j|
\end{aligned}$$

f is continuous on A. since f is a continuous function on a compact set A, it attains its minimum on A, i.e., there is an element $(\mu_1, \mu_2, \ldots, \mu_n) \in A$ such that

$$f(\mu_1, \mu_2, \ldots, \mu_n) = \inf \{ f(\alpha_1, \alpha_2, \ldots, \alpha_n) \,|\, (\alpha_1, \alpha_2, \ldots, \alpha_n) \in A \}$$

Let $m = f(\mu_1, \mu_2, \ldots, \mu_n)$. since $f \geq 0$, it follows that $m \geq 0$. If $m = 0$, then

$$\left\| \sum_{j=1}^{n} \mu_j x_j \right\| = 0 \Rightarrow \sum_{j=1}^{n} \mu_j x_j = 0$$

since the set $\{x_1, x_2, \ldots, x_n\}$ is linearly independent, $\mu_j = 0$ for all $j = 1, 2, \ldots, n$. This is a contradiction since $(\mu_1, \mu_2, \ldots, \mu_n) \in A$. Hence $m > 0$ and consequently for all $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in A$

$$0 < m \leq f(\alpha_1, \alpha_2, \ldots, \alpha_n) \iff m \sum_{j=1}^{n} |\alpha_j| \leq \| \sum_{j=1}^{n} \alpha_j x_j \|$$

Now, let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be any collection of scalars and set $\beta = \sum_{j=1}^{n} |\alpha_j|$. If $\beta = 0$, then the inequality holds vacuously. If $\beta > 0$, then $\left(\frac{\alpha_1}{\beta}, \frac{\alpha_2}{\beta}, \ldots, \frac{\alpha_n}{\beta} \right) \in A$ and consequently

$$\left\| \sum_{j=1}^{n} \alpha_j x_j \right\| = \left\| \sum_{j=1}^{n} \frac{\alpha_j}{\beta} x_j \right\| \beta = f\left(\frac{\alpha_1}{\beta}, \frac{\alpha_2}{\beta}, \ldots, \frac{\alpha_n}{\beta} \right) \beta \geq m\beta = m \sum_{j=1}^{n} |\alpha_j|$$

That is, $m \sum_{j=1}^{n} |\alpha_j| \leq \left\| \sum_{j=1}^{n} \alpha_j x_j \right\|$ ■

Theorem 11.2.17 Let X be a finite-dimensional normed linear space over $\mathbb{F}$. Then all norms on X are equivalent.

Proof. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis for X and $\|\cdot\|_0$ and $\|\cdot\|$ be any two norms on X. For any $x \in X$ there is a set of scalars $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ such that $x = \sum_{i=1}^{n} \alpha_j x_j$. By Lemma 2.6 .1 , there is an $m > 0$ such that

$$m \sum_{j=1}^{n} |\alpha_j| \leq \left\| \sum_{j=1}^{n} \alpha_j x_j \right\| = \|x\|$$

By the triangle inequality

$$\|x\|_0 \leq \sum_{j=1}^{n} |\alpha_j| \left\| x_j \right\|_0 \leq M \sum_{j=1}^{n} |\alpha_j|$$

where $M = \max_{1 \leq j \leq n} \left\| x_j \right\|_0$. Hence

$$\|x\|_0 \leq M \left(\frac{1}{m} \|x\| \right) \Rightarrow \frac{m}{M} \|x\|_0 \leq \|x\| \iff \alpha \|x\|_0 \leq \|x\| \text{ where } \alpha = \frac{m}{M}$$

Interchanging the roles of the norms $\|\cdot\|_0$ and $\|\cdot\|$, we similarly get a constant β such that $\|x\| \leq \beta \|x\|_0$ Hence, $\alpha \|x\|_0 \leq \|x\| \leq \beta \|x\|_0$ for some constants α and β

■

Theorem 11.2.18 Every finite-dimensional normed linear space $(X, \|\cdot\|)$ is complete.

Proof. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis for X and let (z_k) be a Cauchy sequence in X. Then, given any $\epsilon > 0$, there is a natural number N such that

$$\|z_k - z_\ell\| < \epsilon \text{ for all } k, \ell > N$$

Also, for each $k \in \mathbb{N}, z_k = \sum_{j=1}^{n} \alpha_{kj} x_j$. By 11.2.16 , there is an $m > 0$ such that

$$m \sum_{j=1}^{n} |\alpha_{kj} - \alpha_{\ell j}| \leq \|z_k - z_\ell\|$$

Hence, for all $k, \ell > N$ and all $j = 1, 2, \ldots, n$

$$|\alpha_{kj} - \alpha_{\ell j}| \leq \frac{1}{m} \|z_k - z_\ell\| < \frac{\epsilon}{m}$$

That is, for each $j = 1, 2, \ldots, n$, $(\alpha_{kj})_k$ is a Cauchy sequence of numbers. since $\mathbb{F}$ is complete, $\alpha_{kj} \to \alpha_j$ as $k \to \infty$ for each $j = 1, 2, \ldots, n$. Define $z = \sum_{j=1}^{n} \alpha_j x_j$. Then $z \in X$ and

$$\|z_k - z\| = \left\| \sum_{j=1}^{n} \alpha_{kj} x_j - \sum_{j=1}^{n} \alpha_j x_j \right\| = \left\| \sum_{j=1}^{n} (\alpha_{kj} - \alpha_j) x_j \right\| \leq \sum_{j=1}^{n} |\alpha_{kj} - \alpha_j| \|x_j\| \to 0$$

as $k \to \infty$. That is, the sequence (z_k) converges to $z \in X$. hence X is complete. ■

Corollary 11.2.19 Every finite-dimensional normed linear space X is closed.

Theorem 11.2.20 In a finite-dimensional normed linear space $(X, \|\cdot\|)$, a subset $K \subset X$ is sequentially compact if and only if it is closed and bounded.

Proof. We have seen (Corollary 11.2.14), that a compact subset of a normed linear space is closed and bounded.

Conversely, assume that a subset $K \subset X$ is closed and bounded. We show that K is compact. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis for X and let (z_k) be any sequence in K. Then for each $k \in \mathbb{N}, z_k = \sum_{j=1}^{n} \alpha_{kj} x_j$ since K is bounded, there is a positive constant M such that $\|z_k\| \leq M$ for all $k \in \mathbb{N}$. By Lemma 11.2.16 there is an $m > 0$ such that

$$m \sum_{j=1}^{n} |\alpha_{kj}| \leq \left\| \sum_{j=1}^{n} \alpha_{kj} x_j \right\| = \|z_k\| \leq M$$

It now follows that $|\alpha_{kj}| \leq \frac{M}{m}$ for each $j = 1, 2, \ldots, n$, and for all $k \in \mathbb{N}$. That is, for each fixed $j = 1, 2, \ldots, n$, the sequence $(\alpha_{kj})_k$ of numbers is bounded. Hence the sequence $(\alpha_{kj})_k$ has a subsequence $(\alpha_{k_r j})$ which converges to α_j for $j = 1, 2, \ldots, n$. Setting $z = \sum_{j=1}^{n} \alpha_j x_j$, we have that

$$\|z_{k_r} - z\| = \left\| \sum_{j=1}^{n} \alpha_{k_r j} x_j - \sum_{j=1}^{n} \alpha_j x_j \right\| \leq \sum_{j=1}^{n} |\alpha_{k_r j} - \alpha_j| \|x_j\| \to 0 \text{ as } r \to \infty$$

That is, $z_{k_r} \to z$ as $r \to \infty$. since K is closed, $z \in K$. Hence K is compact.

■

Theorem 11.2.21 — Riesz's Lemma. Let M be a closed proper linear subspace of a normed linear space $(X, \|\cdot\|)$. Then for each $0 < \epsilon < 1$, there is an element $z \in X$ such that $\|z\| = 1$ and

$$\|y - z\| > 1 - \epsilon \text{ for all } y \in M$$

Proof. Choose $x \in X \backslash M$ and define

$$d = d(x, M) = \inf_{m \in M} \|x - m\|$$

Since M is closed, $d > 0$. By definition of infimum, there is a $m \in M$ such that

$$d \leq \|x - m\| < d + \epsilon d = d(1 + \epsilon)$$

Take $z = -\left(\frac{m-x}{\|m-x\|}\right)$. Then $\|z\| = 1$ and for any $y \in M$

$$\begin{aligned}
\|y - z\| &= \left\| y + \left(\frac{m - x}{\|m - x\|}\right) \right\| = \frac{\|y(\|m - x\|) + m - x\|}{\|m - x\|} \\
&\geq \frac{d}{\|m - x\|} > \frac{d}{d(1 + \epsilon)} = \frac{1}{1 + \epsilon} = 1 - \frac{\epsilon}{1 + \epsilon} > 1 - \epsilon
\end{aligned}$$

■

Theorem 11.2.22 A normed linear space $(X, \|\cdot\|)$ is finite-dimensional if and only its closed unit ball $B(X) = \{x \in X | \|x\| \leq 1\}$ is compact.

Proof. Assume that $(X, \|\cdot\|)$ is finite-dimensional normed linear space. Since the ball $B(X)$ is closed and bounded, it is compact.

Assume that the closed unit ball $B(X) = \{x \in X | \|x\| \leq 1\}$ is compact. Then $B(X)$ is totally bounded. Hence there is a finite $\frac{1}{2}$-net $\{x_1, x_2, \ldots, x_n\}$ in $B(X)$. Let $M = \text{lin}\{x_1, x_2, \ldots, x_n\}$. Then M is a finite-dimensional linear subspace of X and hence closed.

Claim: $M = X$. If M is a proper subspace of X, then, by Riesz's Lemma11.2.21 there is an element $x_0 \in B(X)$ such that $d(x_0, M) > \frac{1}{2}$. In particular, $\|x_0 - x_k\| > \frac{1}{2}$ for all $k = 1, 2, \ldots, n$. However this contradicts the fact that $\{x_1, x_2, \ldots, x_n\}$ is a $\frac{1}{2}$-net in $B(X)$. Hence $M = X$ and, consequently, X is finite-dimensional.

■

R We now give another argument to show that boundedness does not imply total boundedness. Let $X = \ell_2$ and $B(X) = \{x \in X | \|x\|_2 \leq 1\}$. It is obvious that $B(X)$ is bounded. We show that $B(X)$ is not totally bounded. since X is complete and

$B(X)$ is a closed subset of $X, B(X)$ is complete. If $B(X)$ were totally bounded, then $B(X)$ would, according to Theorem 2.26, be compact. By Theorem 2.6.4, X would be finite-dimensional. But this is false since X is infinite-dimensional.

11.2.6 Separable Spaces and Schauder Bases

Definition 11.2.18 — Dense. A subset S of a normed linear space $(X, \|\cdot\|)$ is said to be dense in X if $\overline{S} = X$; i.e., for each $x \in X$ and $\epsilon > 0$, there is a $y \in S$ such that $\|x - y\| < \epsilon$

Definition 11.2.19 — Separable. A normed linear space (X, $\|\cdot\|$) is said to be separable if it contains a countable dense subset.

■ **Example 11.8** The real line $\mathbb{R}$ is separable since the set $\mathbb{Q}$ of rational numbers is a countable dense subset of $\mathbb{R}$.

The complex plane $\mathbb{C}$ is separable since the set of all complex numbers with rational real and imaginary parts is a countable dense subset of $\mathbb{C}$.

The sequence space ℓ_p, where $1 \leq p < \infty$, is separable. Take M to be the set of all sequences with rational entries such that all but a finite number of the entries are zero. (If the entries are complex, take for M the set of finitely nonzero sequences with rational real and imaginary parts.) It is clear that M is countable. We show that M is dense in ℓ_p. Let $\epsilon > 0$ and $x = (x_n) \in \ell_p$. Then there is an N such that

$$\sum_{k=N+1}^{\infty} |x_k|^p < \frac{\epsilon}{2}$$

Now, for each $1 \leq k \leq N$, there is a rational number q_k such that $|x_k - q_k|^p < \frac{\epsilon}{2N}$. Set $q = (q_1, q_2, \ldots, q_N, 0, 0, \ldots)$. Then $q \in M$ and

$$\|x - q\|_p^p = \sum_{k=1}^{N} |x_k - q_k|^p + \sum_{k=N+1}^{\infty} |x_k|^p < \epsilon$$

Hence M is dense in ℓ_p. ■

■ **Example 11.9** The sequence space ℓ_∞, with the supremum norm, is not separable.

To see this, consider the set M of elements $x = (x_n)$, in which x_n is either 0 or 1. This set is uncountable since we may consider each element of M as a binary representation of a number in the interval $[0,1]$. Hence there are uncountably many sequences of zeroes and ones. For any two distinct elements $x, y \in M$, $\|x - y\|_\infty = 1$. Let each of the elements of M be a centre of a ball of radius $\frac{1}{4}$. Then we get uncountably many nonintersecting balls. If A is any dense subset of ℓ_∞, then each of these balls contains a point of A. Hence A cannot be countable and, consequently, ℓ_∞ is not separable. ■

Theorem 11.2.23 A normed linear space $(X, \|\cdot\|)$ is separable if and only if it contains a countable set B such that $\overline{\text{lin}}(B) = X$

Proof. Assume that X is separable and let A be a countable dense subset of X. Since the linear hull of A, $\text{lin}(A)$, contains A and A is dense in X, we have that $\text{lin}(A)$ is dense in X, that is, $\overline{\text{lin}}(A) = X$

Conversely, assume that X contains a countable set B such that $\overline{\text{lin}}(B) = X$. Let $B = \{x_n | n \in \mathbb{N}\}$. Assume first that $= \mathbb{R}$, and put

$$C = \left\{ \sum_{j=1}^{n} \lambda_j x_j | \lambda_j \in \mathbb{Q}, j = 1, 2, \ldots, n, n \in \mathbb{N} \right\}$$

We first show that C is a countable subset of X. The set $\mathbb{Q} \times B$ is countable and consequently, the family $\mathcal{F}$ of all finite subsets of $\mathbb{Q} \times B$ is also countable. The mapping

$$\{(\lambda_1, x_1), (\lambda_2, x_2), \ldots, (\lambda_n, x_n)\} \mapsto \sum_{j=1}^{n} \lambda_j x_j$$

maps $\mathcal{F}$ onto C. Hence C is countable. Next, we show that C is dense in X. Let $x \in X$ and $\epsilon > 0$. Since $\overline{\text{lin}}(B) = X$, we can find an $n \in \mathbb{N}$ points $x_1, x_2, \ldots, x_n \in B$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$ such that

$$\left\| x - \sum_{j=1}^{n} \lambda_j x_j \right\| < \frac{\epsilon}{2}$$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for each $\lambda_i \in \mathbb{R}$, we can find a $\mu_i \in \mathbb{Q}$ such that

$$|\lambda_i - \mu_i| < \frac{\epsilon}{2n\,(1 + \|x_i\|)} \text{ for all } i = 1, 2, \ldots, n$$

Hence.

$$\begin{aligned}
\left\| x - \sum_{j=1}^{n} \mu_j x_j \right\| &\leq \left\| x - \sum_{j=1}^{n} \lambda_j x_j \right\| + \left\| \sum_{j=1}^{n} \lambda_j x_j - \sum_{j=1}^{n} \mu_j x_j \right\| \\
&< \frac{\epsilon}{2} + \sum_{j=1}^{n} |\lambda_j - \mu_j| \|x_j\| \\
&< \frac{\epsilon}{2} + \sum_{j=1}^{n} \frac{\epsilon \|x_j\|}{2n\,(1 + \|x_j\|)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

This shows that C is dense in X. If $\mathbb{F} = \mathbb{C}$, the set C is that of finite linear combinations with coefficients being those complex numbers with rational real and imaginary parts.

∎

Theorem 11.2.24 We now give another argument based on Theorem11.2.23 to show that the sequence space ℓ_p, where $1 \leq p < \infty$, is separable. Let $e_n = (\delta_{nm})_{m\in\mathbb{N}}$, where

$$\delta_{nm} = \begin{cases} 1 \text{ if } n = m \\ 0 \text{ otherwise} \end{cases}$$

Clearly, $e_n \in \ell_p$. Let $\epsilon > 0$ and $x = (x_n) \in \ell_p$. Then there is a natural number N such that

$$\sum_{k=n+1}^{\infty} |x_k|^p < \epsilon^p \text{ for all } n \geq N$$

Now, if $n \geq N$, then

$$\left\| x - \sum_{j=1}^{n} x_j e_j \right\|_p = \left(\sum_{k=n+1}^{\infty} |x_k|^p \right)^{1/p} < \epsilon$$

Hence $\overline{\text{lin}}\,(\{e_n | n \in \mathbb{N}\}) = \ell_p$. Of course, the set $\{e_n | n \in \mathbb{N}\}$ is countable.

Definition 11.2.20 — Schauder basis. A sequence (b_n) in a Banach space $(X, \|\cdot\|)$ is called a Schauder basis if for any $x \in X$, there is a unique sequence (α_n) of scalars such that

$$\lim_{n\to\infty} \left| x - \sum_{j=1}^{n} \alpha_j b_j \right| | = 0$$

In this case we write $x = \sum_{j=1}^{\infty} \alpha_j b_j$

(R) It is clear from Definition 11.2.20 that (b_n) is a Schauder basis if and only if $X = \overline{\text{lin}}\,\{b_n | n \in \mathbb{N}\}$ and every $x \in X$ has a unique expansion $x = \sum_{i=1}^{\infty} \alpha_j b_j$. Uniqueness of this expansion clearly implies that the set $\{b_n | n \in \mathbb{N}\}$ is linearly independent.

■ **Example 11.10**

1. For $1 \leq p < \infty$, the sequence (e_n), where $e_n = (\delta_{nm})_{m\in\mathbb{N}}$, is a Schauder basis for ℓ_p
2. (e_n) is a Schauder basis for c_0
3. $(e_n) \cup \{e\}$, where $e = (1,1,1,\ldots)$ (the constant 1 sequence), is a Schauder basis for c
4. ℓ_∞ has no Schauder basis.

Proposition 11.2.25 If a Banach space $(X, \|\cdot\|)$ has a Schauder basis, then it is separable.

Proof. Let (b_n) be a Schauder basis for X. Then $\{b_n | n \in \mathbb{N}\}$ is countable and $\overline{\text{lin}}\,(\{b_n | n \in \mathbb{N}\}) = X$ ■

(R) Schauder bases have been constructed for most of the well-known Banach spaces. Schauder conjectured that every separable Banach space has a Schauder basis. This conjecture, known as the Basis Problem, remained unresolved for a long time until Per Enflo in 1973 answered it in the negative. He constructed a separable reflexive Banach space with no basis.

11.3 Hilbert Spaces

Definition 11.3.1 — Inner product space. Let X be a linear space over a field $\mathbb{F}$. An inner product on X is a scalar-valued function $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{F}$ such that for all $x,y,z \in X$ and for all $\alpha,\beta \in \mathbb{F}$, we have

1. $\langle x,x\rangle \geq 0$
2. $\langle x,x\rangle = 0 \iff x = 0$
3. $\langle x,y\rangle = \overline{\langle y,x\rangle}$ (The bar denotes complex conjugation.)
4. $\langle \alpha x,y\rangle = \alpha\langle x,y\rangle$
5. $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$

An inner product space $(X,\langle\cdot,\cdot\rangle)$ is a linear space X together with an inner $\langle\cdot,\cdot\rangle$ product defined on it. An inner product space is also called pre-Hilbert space.

■ **Example 11.11 — Euclidean n -space.** Fix a positive integer n. Let $X = \mathbb{F}^n$. For $x = (x_1,x_2,\ldots,x_n)$ and $y = (y_1,y_2,\ldots,y_n)$ in X define

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i\overline{y_i}$$

since this is a finite sum, $\langle\cdot,\cdot\rangle$ is well-defined. It is easy to show that $(X,\langle\cdot,\cdot\rangle)$ is an inner product space. The space $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) with this inner product is called the Euclidean n -space (resp. unitary n -space) and will be denoted by $\ell_2(n)$

■ **Example 11.12** Let $X = \ell_2$, the space of all sequences $x = (x_1,x_2,\ldots)$ of real or complex numbers with $\sum_1^\infty |x_i|^2 < \infty$. For $x = (x_1,x_2,\ldots)$ and $y = (y_1,y_2,\ldots)$ in X, define

$$\langle x,y\rangle = \sum_{i=1}^{\infty} x_i\overline{y_i}$$

In order to show that $\langle\cdot,\cdot\rangle$ is well-defined we first observe that if a and b are real numbers, then

$$0 \leq (a-b)^2, \quad \text{whence} \quad ab \leq \frac{1}{2}\left(a^2+b^2\right)$$

Using this fact, we have that

$$\left|x_i\overline{y_i}\right| = \left|x_i\right|\left|\overline{y_i}\right| \leq \frac{1}{2}\left(|x_i|^2+|y_i|^2\right) \Rightarrow \sum_{i=1}^{\infty}|x_i\overline{y_i}| \leq \frac{1}{2}\left(\sum_{i=1}^{\infty}|x_i|^2+\sum_{i=1}^{\infty}|y_i|^2\right) < \infty$$

Hence, $\langle\cdot,\cdot\rangle$ is well-defined (i.e., the series converges).

■ **Example 11.13** Let $X = \mathcal{L}(\mathbb{C}^n)$ be the linear space of all $n \times n$ complex matrices. For $A \in \mathcal{L}(\mathbb{C}^n)$, let $\tau(A) = \sum_{i=1}^{n} (A)_{ii}$ be the trace of A. For $A, B \in \mathcal{L}(\mathbb{C}^n)$, define

$$\langle A, B \rangle = \tau(B^* A),$$

where B^* denotes conjugate transpose of matrix B.

Show that $(\mathcal{L}(\mathbb{C}^n), \langle \cdot, \cdot \rangle)$ is an inner product space.

Theorem 11.3.1 — Cauchy-Bunyakowsky-Schwarz Inequality. Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space over a field $\mathbb{F}$ Then for all $x, y \in X$

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle}$$

Moreover, given any $x, y \in X$, the equality

$$|\langle x, y \rangle| = \sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle}$$

holds if and only if x and y are linearly dependent.

Proof. Proof. If $x = 0$ or $y = 0$, then the result holds vacuously. Assume that $x \neq 0$ and $y \neq 0$. For any $\alpha \in \mathbb{F}$ we have

$$\begin{aligned} 0 \leq \langle x - \alpha y, x - \alpha y \rangle &= \langle x, x \rangle - \alpha \langle y, x \rangle - \bar{\alpha} \langle x, y \rangle + \alpha \bar{\alpha} \langle y, y \rangle \\ &= \langle x, x \rangle - \bar{\alpha} \langle x, y \rangle - \alpha [\langle y, x \rangle - \bar{\alpha} \langle y, y \rangle] \end{aligned}$$

Now choosing $\alpha = \frac{\langle x, y \rangle}{\langle y, y \rangle}$, we have

$$0 \leq \langle x, x \rangle - \frac{\langle y, x \rangle \langle x, y \rangle}{\langle y, y \rangle} = \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}$$

whence

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle}$$

Assume that $|\langle x, y \rangle| = \sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle}$. We show that x and y are linearly dependent. If $x = 0$ or $y = 0$, then x and y are obviously linearly dependent. We therefore assume that $x \neq 0$ and $y \neq 0$. Then $\langle y, y \rangle \neq 0$. With $\alpha = \frac{\langle x, y \rangle}{\langle y, y \rangle}$, we have that

$$\langle x - \alpha y, x - \alpha y \rangle = \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} = 0$$

That is,

$$\langle x - \alpha y, x - \alpha y \rangle = 0, \quad \Rightarrow \quad x = \alpha y$$

That is, x and y are linearly dependent. Conversely, assume that x and y are linearly dependent. Without loss of generality, $x = \lambda y$ for some $\lambda \in \mathbb{F}$. Then

$$\begin{aligned}
|\langle x,y\rangle| &= |\langle \lambda y,y\rangle| = |\lambda||\langle y,y\rangle| = |\lambda|\langle y,y\rangle \\
&= |\lambda|\sqrt{\langle y,y\rangle}\sqrt{\langle y,y\rangle} = \sqrt{|\lambda|^2\langle y,y\rangle}\sqrt{\langle y,y\rangle} = \sqrt{\langle \lambda y \lambda y\rangle}\sqrt{\langle y,y\rangle} \\
&= \sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}
\end{aligned}$$

■

Definition 11.3.2 — norm. Let $(X,\langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. For each $x \in X$, define

$$\|x\| := \sqrt{\langle x,x\rangle}$$

Then $\|\cdot\|$ defines a **norm** on X. That is, $(X,\|\cdot\|)$ is a normed linear space over $\mathbb{F}$.

The norm is called the inner product norm or a norm induced or generated by the inner product.

Proof.

$$\begin{aligned}
\|x+y\|^2 = \langle x+y,x+y\rangle &= \langle x,x\rangle + \langle x,y\rangle + (y,x) + \langle y,y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\
&= \|x\|^2 + 2\Re e\langle x,y\rangle + \|y\|^2 \le \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\
&\le \|x\|^2 + 2\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle} + \|y\|^2 (\text{ by Theorem 11.3.1}) \\
&= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2
\end{aligned}$$

Taking the positive square root both sides yields

$$\|x+y\| \le \|x\| + \|y\|$$

■

R The Cauchy-Bunyakowsky-Schwarz Inequality now becomes

$$|\langle x,y\rangle| \le \|x\|\|y\|$$

And Cauchy-Bunyakowsky-Schwarz Inequality implies that any inner product is continuous.

Theorem 11.3.2 — Polarization Identity. Let $(X,\langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Then for all $x,y \in X$

$$\langle x,y\rangle = \frac{\|x+y\|^2}{4} - \frac{\|x-y\|^2}{4} \quad \text{if } \mathbb{F} = \mathbb{R}, \quad \text{and}$$
$$\langle x,y\rangle = \frac{\|x+y\|^2}{4} - \frac{\|x-y\|^2}{4} + i\left(\frac{\|x+yi\|^2}{4} - \frac{\|x-yi\|^2}{4}\right) \text{ if } \mathbb{F} = \mathbb{C}$$

Proof. Assume that $\mathbb{F} = \mathbb{R}$. Then

$$\begin{aligned}\|x+y\|^2 - \|x-y\|^2 &= \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle - \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle - \langle y,y\rangle \\ &= 4\langle x,y\rangle, \quad \text{since } \langle x,y\rangle = \langle y,x\rangle\end{aligned}$$

■

Theorem 11.3.3 — Parallelogram Identity. Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Then for all $x,y \in X$

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

Proof.

$$\begin{aligned}\|x-y\|^2 + \|x+y\|^2 &= \langle x-y, x-y\rangle + \langle x+y, x+y\rangle \\ &= \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle + \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= 2\|x\|^2 + 2\|y\|^2\end{aligned}$$

■

R The geometric interpretation of the Parallelogram Identity is evident: the sum of the squares of the lengths of the diagonals of a parallelogram is equal to the sum of the squares of the lengths of the four.

Theorem 11.3.4 A normed linear space X over a field $\mathbb{F}$ is an inner product space if and only if the Parallelogram Identity

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

holds for all $x,y \in X$

R A normed linear space is an inner product space if and only if its norm satisfies the Parallelogram Identity.

Proof. $'' \Leftarrow ''$. Let X be a normed linear space in which the parallelogram identity (PI) holds. We shall only consider the case $\mathbb{F} = \mathbb{R}$. The polarization identity gives us a hint as to how we should define an inner product: For all $x,y \in X$, define

$$\langle x,y\rangle = \left\|\frac{x+y}{2}\right\|^2 - \left\|\frac{x-y}{2}\right\|^2$$

We claim that $\langle\cdot,\cdot\rangle$ is an inner product on X

1. $\langle x,x\rangle = \|\frac{x+x}{2}\|^2 - \|\frac{x-x}{2}\|^2 = \|x\|^2 \geq 0$
2. $\langle x,x\rangle = 0 \Longleftrightarrow \|x\|^2 = 0 \Longleftrightarrow x = 0$
3. $\langle x,y\rangle = \left\|\frac{x+y}{2}\right\|^2 - \left\|\frac{x-y}{2}\right\|^2 = \left\|\frac{y+x}{2}\right\|^2 - \left\|\frac{y-x}{2}\right\|^2 = \langle y,x\rangle = \langle y,x\rangle$ since $\mathbb{F} = \mathbb{R}$
4. Replace x by $u+v$ and y by $w+v$ in the parallelogram identity:

$$\|u+w+2v\|^2 + \|u-w\|^2 = 2\|u+v\|^2 + 2\|w+v\|^2 \tag{11.2}$$

Replace x by $u-v$ and y by $w-v$ in the parallelogram identity:

$$\|u+w-2v\|^2 + \|u-w\|^2 = 2\|u-v\|^2 + 2\|w-v\|^2 \tag{11.3}$$

Subtract 11.3 from 11.2

$$\|u+w+2v\|^2 - \|u+w-2v\|^2 = 2\left[\|u+v\|^2 - \|u-v\|^2 + \|v+w\|^2 - \|v-w\|^2\right]$$

Use the definition of $\langle\cdot,\cdot\rangle$

$$4\langle u+w,2v\rangle = 8[\langle u,v\rangle + \langle w,v\rangle] \quad \Rightarrow \quad \frac{1}{2}\langle u+w,2v\rangle = \langle u,v\rangle + \langle w,v\rangle \tag{11.4}$$

Take $w = 0$

$$\frac{1}{2}\langle u,2v\rangle = \langle u,v\rangle \tag{11.5}$$

Now replace u by $x+y$ and v by z in 11.5 and use 11.4 to get

$$\langle x+y,z\rangle = \frac{1}{2}\langle x+y,2z\rangle = \langle x,z\rangle + \langle y,z\rangle$$

5. We show that $\langle\lambda x,y\rangle = \lambda\langle x,y\rangle$ for all $\lambda \in \mathbb{R}$ and all $x,y \in X$.
If $\lambda = n$ is a nonzero integer, then using 4,

$$\langle nx,y\rangle = n\langle x,y\rangle \quad \Rightarrow \quad n\left\langle \frac{x}{n},y\right\rangle = \left\langle \frac{nx}{n},y\right\rangle = \langle x,y\rangle$$

That is,

$$\left\langle \frac{x}{n},y\right\rangle = \frac{1}{n}\langle x,y\rangle$$

If λ is a rational number, $\lambda = \frac{p}{q}$, say. Then

$$\left\langle \frac{p}{q}x,y\right\rangle = p\left\langle \frac{x}{q},y\right\rangle = \frac{p}{q}\langle x,y\rangle$$

If $\lambda \in \mathbb{R}$, then there is a sequence (r_k) of rational numbers such that $r_k \to \lambda$ as $k \to \infty$. Using continuity of the norm, we have that

$$\begin{aligned}
\langle \lambda x, y\rangle = \left\langle \lim_{k\to\infty} r_k x, y \right\rangle &= \frac{1}{4}\left\|\lim_{k\to\infty} r_k x + y\right\|^2 - \frac{1}{4}\left\|\lim_{k\to\infty} r_k x - y\right\|^2 \\
&= \frac{1}{4}\lim_{k\to\infty} \|r_k x + y\|^2 - \frac{1}{4}\lim_{k\to\infty}\|r_k x - y\|^2 \\
&= \lim_{k\to\infty}\left(\left\|\frac{r_k x + y}{2}\right\|^2 - \left\|\frac{r_k x - y}{2}\right\|^2\right) \\
&= \lim_{k\to\infty}\langle r_k x, y\rangle \\
&= \lim_{k\to\infty} r_k \langle x, y\rangle = \lambda\langle x, y\rangle
\end{aligned}$$

Thus, $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ for all $\lambda \in \mathbb{R}$ and all $x, y \in X$.

■

Corollary 11.3.5 Let $(X, \|\cdot\|)$ be a normed linear space over a field $\mathbb{F}$. If every two-dimensional linear subspace of X is an inner product space over $\mathbb{F}$, then X is an inner product space.

■ **Example 11.14 — Not inner product space.** 1. Let $X = \ell_p$, for $p \neq 2$. Then X is not an inner product space. We show that the norm on ℓ_p, $\quad p \neq 2$ does not satisfy the parallelogram identity. Take $x = (1,1,0,0,\dots)$ and $y = (1,-1,0,0,\dots)$ in ℓ_p. Then

$$\|x\| = 2^{\frac{1}{p}} = \|y\| \quad \text{and} \quad \|x+y\| = 2 = \|x-y\|$$

Thus,

$$\|x+y\|^2 + \|x-y\|^2 = 8 \neq 2\|x\|^2 + 2\|y\|^2 = 4 \cdot 2^{\frac{2}{p}}$$

2. The normed linear space $X = \mathcal{C}[a,b]$, with the supremum norm $\|\cdot\|_\infty$ is not an inner product space. We show that the norm

$$\|x\|_\infty = \max_{a\le t\le b} |x(t)|$$

does not satisfy the parallelogram identity. To that end, take

$$x(t) = 1 \quad \text{and} \quad y(t) = \frac{t-a}{b-a}$$

since

$$x(t) + y(t) = 1 + \frac{t-a}{b-a} \quad \text{and} \quad x(t) - y(t) = 1 - \frac{t-a}{b-a}$$

we have that

$$\|x\| = 1 = \|y\|, \quad \text{and} \quad \|x+y\| = 2, \quad \|x-y\| = 1$$

Thus,

$$\|x+y\|^2 + \|x-y\|^2 = 5 \neq 2\|x\|^2 + 2\|y\|^2 = 4$$

Definition 11.3.3 — complete. Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. If X is **complete** with respect to the norm induced by the inner product $\langle\cdot,\cdot\rangle$, then we say that X is a Hilbert space.

■ **Example 11.15**
1. The classical space ℓ_2 is a Hilbert space.
2. ℓ_0 is an incomplete inner product space.
3. The space $\mathcal{C}[-1,1]$ is an incomplete inner product space.

11.3.1 Orthogonal

Definition 11.3.4 — Orthogonal. Two elements x and y in an inner product space $(X, \langle\cdot,\cdot\rangle)$ are said to be **orthogonal**, denoted by $x \perp y$, if

$$\langle x,y\rangle = 0$$

The set $M \subset X$ is called **orthogonal** if it consists of non-zero pairwise orthogonal elements.

If M is a subset of X such that $\langle x,m\rangle = 0$ for all $m \in M$, then we say that x is orthogonal to M and write $x \perp M$. We shall denote by

$$M^{\perp} = \{x \in X : \langle x,m\rangle = 0 \ \ \forall m \in M\}$$

the set of all elements in X that are orthogonal to M. The set $M^{\perp}$ is called the **orthogonal complement** of M.

Proposition 11.3.6 Let M and N be subsets of an inner product space $(X, \langle\cdot,\cdot\rangle)$. Then
1. $\{0\}^{\perp} = X$ and $X^{\perp} = \{0\}$.
2. $M^{\perp}$ is a closed linear subspace of X.
3. $M \subset \left(M^{\perp}\right)^{\perp} = M^{\perp\perp}$
4. If M is a linear subspace, then $M \cap M^{\perp} = \{0\}$
5. If $M \subset N$, then $N^{\perp} \subset M^{\perp}$
6. $M^{\perp} = (\operatorname{lin} M)^{\perp} = (\overline{\operatorname{lin}} M)^{\perp}$

Proof.
1. Exercise
2. Let $x,y \in M^{\perp}$, and $\alpha, \beta \in \mathbb{F}$. Then for each $z \in M$

$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle = 0$$

Hence, $\alpha x + \beta y \in M^{\perp}$. That is, $M^{\perp}$ is a subspace of X. To show that $M^{\perp}$ is closed, let $x \in \overline{M^{\perp}}$ Then there exists a sequence (x_n) in $M^{\perp}$ such that $x_n \to x$ as $n \to \infty$. Thus, for all $y \in M$

$$\langle x,y\rangle = \lim_{n} \langle x_n, y\rangle = 0$$

whence $x \in M^{\perp}$.

3. Exercise
4. Exercise
5. Let $x \in N^{\perp}$. Then $\langle x,y \rangle = 0$ for all $y \in N$. In particular, $\langle x,y \rangle = 0$ for all $y \in M$ since $M \subset N$ Thus, $x \in M^{\perp}$
6. Since $M \subset \operatorname{lin} M \subset \overline{\operatorname{lin}} M$, we have, by [5], that $(\overline{\operatorname{lin}} M)^{\perp} \subset (\operatorname{lin} M)^{\perp} \subset M^{\perp}$. It remains to show that $M^{\perp} \subset (\overline{\operatorname{lin}} M)^{\perp}$. To that end, let $x \in M^{\perp}$. Then $\langle x,y \rangle = 0$ for all $y \in M$, and consequently $\langle x,y \rangle = 0$ for all $y \in \operatorname{lin} M$. If $z \in \overline{\operatorname{lin}} M$, then there exists a sequence (z_n) in $\operatorname{lin} M$ such that $z_n \to z$ as $n \to \infty$. Thus,

$$\langle x,z \rangle = \lim_n \langle x,z_n \rangle = 0$$

whence $x \in (\overline{\operatorname{lin}} M)^{\perp}$.

■

Theorem 11.3.7 — Pythagoras. Let $(X, \langle \cdot,\cdot \rangle)$ be an inner product space over a field $\mathbb{F}$ and let $x,y \in X$

1. If $\mathbb{F} = \mathbb{R}$, then $x \perp y$ if and only if

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2$$

2. If $\mathbb{F} = \mathbb{C}$, then $x \perp y$ if and only if

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \text{ and } \|x+iy\|^2 = \|x\|^2 + \|y\|^2$$

Corollary 11.3.8 If $M = \{x_1, x_2, \ldots, x_n\}$ is an orthogonal set in an inner product space $(X, \langle \cdot,\cdot \rangle)$ then

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

11.3.2 Best approximations

Definition 11.3.5 — Best approximation. Let K be a closed subset of an inner product space $(X, \langle \cdot,\cdot \rangle)$. For a given $x \in X \backslash K$, a best approximation or nearest point to x from K is any element $y_0 \in K$ such that

$$\|x - y_0\| \leq \|x - y\| \quad \text{for all} \quad y \in K$$

Equivalently, $y_0 \in K$ is a best approximation to x from K if

$$\|x - y_0\| = \inf_{y \in K} \|x - y\| = d(x,K)$$

The (possibly empty) set of all best approximations to x from K is denoted by $P_K(x)$.

That is

$$P_K(x) = \{y \in K : \|x - y\| = d(x, K)\}$$

Definition 11.3.6 — Metric projection. The (generally set-valued) map P_K which associates each x in X with its best approximations in K is called the **metric projection** or the **nearest point map**. The set K is called

1. **proximinal** if each $x \in X$ has a best approximation in K; i.e., $P_K(x) \neq \emptyset$ for each $x \in X$
2. **Chebyshev** if each $x \in X$ has a unique best approximation in K; i.e., the set $P_K(x)$ consists of a single point

Theorem 11.3.9 Every nonempty complete convex subset K of an inner product space $(X, \langle \cdot, \cdot \rangle)$ is a Chebyshev set.

Proof. Existence: Without loss of generality, $x \in X \backslash K$. Let

$$\delta = \inf_{y \in K} \|x - y\|$$

By definition of the infimum, there exists a sequence $(y_n)_1^\infty$ in K such that

$$\|x - y_n\| \to \delta \quad \text{as} \quad n \to \infty$$

We show that $(y_n)_1^\infty$ is a Cauchy sequence. By the Parallelogram Identity,

$$\begin{aligned} \|y_m - y_n\|^2 &= \|(x - y_n) - (x - y_m)\|^2 \\ &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - (y_n + y_m)\|^2 \\ &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \left(\frac{y_n + y_m}{2}\right)\right\|^2 \\ &\leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2 \end{aligned}$$

Since $\frac{y_n + y_m}{2} \in K$ by convexity of K. Thus,

$$\|y_m - y_n\|^2 \leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2 \to 0 \quad \text{as} \quad n, m \to \infty$$

That is, $(y_n)_1^\infty$ is a Cauchy sequence in K. Since K is complete, there exists $y \in K$ such that $y_n \to y \quad \text{as} \quad n \to \infty$. since the norm is continuous,

$$\|x - y\| = \left\|x - \lim_{n \to \infty} y_n\right\| = \left\|\lim_{n \to \infty} (x - y_n)\right\| = \lim_{n \to \infty} \|x - y_n\| = \delta$$

Thus

$$\|x - y\| = \delta = d(x, K)$$

Uniqueness: Assume that $y, y_0 \in K$ are two best approximations to x from K. That is,

$$\|x-y_0\| = \|x-y\| = \delta = d(x,K)$$

By the Parallelogram Identity,

$$\begin{aligned}
0 \leq \|y-y_0\|^2 &= \|(y-x)+(x-y_0)\|^2 \\
&= 2\|x-y\|^2 + 2\|x-y_0\|^2 - \|2x-(y+y_0)\|^2 \\
&= 2\delta^2 + 2\delta^2 - 4\left\|x - \left(\frac{y+y_0}{2}\right)\right\|^2 \\
&\leq 4\delta^2 - 4\delta^2 = 0
\end{aligned}$$

Thus, $y_0 = y$. ■

Corollary 11.3.10 Every nonempty closed convex subset of a Hilbert space is Chebyshev.

Theorem 11.3.11 Let K be a nonempty closed convex subset of a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle), x \in \mathcal{H} \backslash K$ and $y_0 \in K$. Then y_0 is the best approximation to x from K if and only if

$$\Re_{\mathfrak{e}} \langle x-y_0, y-y_0\rangle \leq 0 \text{ for all } y \in K$$

Proof. The existence and uniqueness of the best approximation to x in K are guaranteed by Theorem 3.4 .1 Let y_0 be the best approximation to x in K. Then, for any $y \in K$ and any $0<\lambda<1, \lambda y+(1-\lambda)y_0 \in K$ since K is convex. Thus,

$$\begin{aligned}
\|x-y_0\|^2 \leq \|x-[\lambda y+(1-\lambda)y_0]\|^2 &= \|(x-y_0)-\lambda(y-y_0)\|^2 \\
&= \langle (x-y_0)-\lambda(y-y_0), (x-y_0)-\lambda(y-y_0)\rangle \\
&= \langle x-y_0, x-y_0\rangle - \lambda[\langle x-y_0, y-y_0\rangle + \langle y-y_0, x-y_0\rangle] \\
&+ \lambda^2 \langle y-y_0, y-y_0\rangle \\
&= \|x-y_0\|^2 - 2\lambda \mathfrak{Re}(\langle x-y_0, y-y_0\rangle) + \lambda^2\|y-y_0\|^2
\end{aligned}$$

$$\Rightarrow \quad 2\lambda \Re_{\mathfrak{e}}(\langle x-y_0, y-y_0\rangle) \leq \lambda^2 \|y-y_0\|^2$$

$$\Rightarrow \quad \mathfrak{Re}(\langle x-y_0, y-y_0\rangle) \leq \frac{\lambda}{2}\|y-y_0\|^2$$

As $\lambda \to 0, \frac{\lambda}{2}\|y-y_0\|^2 \to 0$, and consequently $\mathfrak{Re}\langle x-y_0, y-y_0\rangle \leq 0$ Conversely, assume that for each $y \in K, \Re\mathfrak{e} \langle x - y_0, y - y_0 \rangle \leq 0$. Then, for any $y \in K$

$$\begin{aligned}
\|x-y\|^2 &= \|(x-y_0)-(y-y_0)\|^2 \\
&= \langle (x-y_0)-(y-y_0), (x-y_0)-(y-y_0) \rangle \\
&= \langle x-y_0, x-y_0 \rangle - \langle x-y_0, y-y_0 \rangle - \langle y-y_0, x-y_0 \rangle + \langle y-y_0, y-y_0 \rangle \\
&= \langle x-y_0, x-y_0 \rangle - [\langle x-y_0, y-y_0 \rangle + \langle y-y_0, x-y_0 \rangle] + \langle y-y_0, y-y_0 \rangle \\
&= \langle x-y_0, x-y_0 \rangle - [\langle x-y_0, y-y_0 \rangle + \langle x-y_0, y-y_0 \rangle] + \langle y-y_0, y-y_0 \rangle \\
&= \|x-y_0\|^2 - 2\Re_{\mathfrak{e}} \langle x-y_0, y-y_0 \rangle + \|y-y_0\|^2 \\
&\geq \|x-y_0\|^2
\end{aligned}$$

Taking the positive square root both sides, we have that $\|x - y_0\| \leq \|x - y\|$ for all $y \in K$. ■

Corollary 11.3.12 — Characterization of Best Approximations from closed subspaces. Let M be a closed subspace of a Hilbert space $\mathcal{H}$ and let $x \in \mathcal{H} \backslash M$. Then an element $y_0 \in M$ is the best approximation to x from M if and only if $\langle x - y_0, y \rangle = 0$ for all $y \in M$ (i.e., $x - y_0 \in M^{\perp}$).

R Corollary11.3.12 says that if M is a closed linear subspace of a Hilbert space $\mathcal{H}$, then $y_0 = P_M(x)$ (i.e., y_0 is the best approximation to x from M) if and only if $x - P_M(x) \perp M$. That is, the unique best approximation is obtained by "dropping the perpendicular from x onto M" It is for this reason that the map $P_M : x \to P_M(x)$ is also called the orthogonal projection of $\mathcal{H}$ onto M.

■ **Example 11.16** Let $X = \mathcal{C}_2[-1,1], M = \mathbb{P}_2 = \text{lin}\left\{1, t, t^2\right\}$, and $x(t) = t^3$. Find $P_M(x)$ ■

Proof. Note that $\mathcal{C}_2[-1,1]$ is an incomplete inner product space. since M is finite-dimensional, it is complete, and consequently proximinal in $\mathcal{C}_2[-1,1]$. Uniqueness of best approximations follows from the Parallelogram Identity.

Let $y_0 = \sum_{i=0}^{2} \alpha_i t^i \in M$

$$
\begin{aligned}
y_0 = P_M(x) &\iff x - y_0 \in M^{\perp} \\
&\iff \left\langle x - y_0, t^j \right\rangle = 0 \quad \text{for all } j = 0,1,2 \\
&\iff \left\langle t^3 - \sum_{i=0}^{2} \alpha_i t^i, t^j \right\rangle = 0 \quad \text{for all } j = 0,1,2 \\
&\iff \sum_{i=0}^{2} \alpha_i \left\langle t^i, t^j \right\rangle = \left\langle t^3, t^j \right\rangle \text{ for all } j = 0,1,2 \\
&\iff \sum_{i=0}^{2} \alpha_i \int_{-1}^{1} t^i \cdot t^j dt = \int_{-1}^{1} t^3 \cdot t^j dt \quad \text{for all } j = 0,1,2 \\
&\iff \sum_{i=0}^{2} \alpha_i \int_{-1}^{1} t^{i+j} dt = \int_{-1}^{1} t^{3+j} dt \quad \text{for all } j = 0,1,2 \\
&\iff \sum_{i=0}^{2} \alpha_i \frac{t^{i+j+1}}{i+j+1}\Bigg]_{-1}^{1} = \frac{t^{j+4}}{j+4}\Bigg]_{-1}^{1} \quad \text{for all } j = 0,1,2 \\
&\iff \sum_{i=0}^{2} \alpha_i \frac{1}{i+j+1}\left[1 - (-1)^{i+j+1}\right] = \frac{1}{j+4}\left[1 - (-1)^{j+4}\right] \\
&\quad \text{for all } j = 0,1,2 \\
&\iff \begin{cases} 2\alpha_0 + 0\alpha_1 + \frac{2}{3}\alpha_2 = 0 \\ 0\alpha_0 + \frac{2}{3}\alpha_1 + 0\alpha_2 = \frac{2}{5} \\ \frac{2}{3}\alpha_0 + 0\alpha_1 + \frac{2}{5}\alpha_2 = 0 \end{cases} \\
&\iff \quad \alpha_0 = 0, \quad \alpha_1 = \frac{3}{5}, \quad \alpha_2 = 0
\end{aligned}
$$

Thus, $P_M(x) = y_0 = \frac{3}{5}t$ ■

> **Theorem 11.3.13 — Projection Theorem.** Let $\mathcal{H}$ be a Hilbert space, M a closed subspace of $\mathcal{H}$. Then
>
> 1. $\mathcal{H} = M \oplus M^{\perp}$. That is, each $x \in \mathcal{H}$ can be uniquely decomposed in the form
>
> $$x = y + z \text{ with } y \in M \text{ and } z \in M^{\perp}$$
>
> 2. $M = M^{\perp\perp}$

Proof. 1. If $x \in M$, then $x = x + 0$, and we are done. Assume that $x \notin M$. Let $y = P_M(x)$ be the unique best approximation to x from M as advertised in Theorem 3.4.1. Then $z = x - P_M(x) \in M^{\perp}$, and

$$x = P_M(x) + (x - P_M(x)) = y + z$$

is the unique representation of x as a sum of an element of M and an element of $M^{\perp}$.

2. Since the containment $M \subset M^{\perp\perp}$ is clear, we only show that $M^{\perp\perp} \subset M$. To that end, let $x \in M^{\perp\perp}$. Then by [1] above

$$x = y + z, \quad \text{where } y \in M \quad \text{and } z \in M^{\perp}$$

Since $M \subset M^{\perp\perp}$ and $M^{\perp\perp}$ is a subspace, $z = x - y \in M^{\perp\perp}$. But $z \in M^{\perp}$ implies that $z \in M^{\perp} \cap M^{\perp\perp}$ which, in turn, implies that $z = 0$. Thus, $x = y \in M$.

■

Corollary 11.3.14 If M is a closed subspace of a Hilbert space $\mathcal{H}$, and if $M \neq \mathcal{H}$, then there exists $z \in \mathcal{H} \backslash \{0\}$ such that $z \perp M$.

Proof. Let $x \in \mathcal{H} \backslash M$. Then by the Projection Theorem,

$$x = y + z, \text{ where } y \in M \text{ and } z \in M^{\perp}$$

Hence $z \neq 0$ and $z \perp M$. ■

Proposition 11.3.15 Let S be a nonempty subset of a Hilbert space $\mathcal{H}$. Then

1. $S^{\perp\perp} = \overline{\operatorname{lin}} S$
2. $S^{\perp} = \{0\}$ if and only if $\overline{\operatorname{lin}} S = \mathcal{H}$

Proof. 1. Since $S^{\perp} = (\overline{\operatorname{lin}} S)^{\perp}$ by Proposition 3.3.2, we have, by the Projection Theorem, that

$$\overline{\operatorname{lin}} S = (\overline{\operatorname{lin}} S)^{\perp\perp} = S^{\perp\perp}$$

2. If $S^{\perp} = \{0\}$, then by [1]

$$\overline{\operatorname{lin}} S = S^{\perp\perp} = \{0\}^{\perp} = \mathcal{H}$$

On the other hand, if $\mathcal{H} = \overline{\operatorname{lin}} S$, then $\mathcal{H} = S^{\perp\perp}$ by [1], and so

$$S^{\perp} = S^{\perp\perp\perp} = \mathcal{H}^{\perp} = \{0\}$$

■

11.3.3 Orthonormal Sets and Orthonormal Bases

Definition 11.3.7 — Orthonormal set. Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{F}$. A set $S = \{x_\alpha : \alpha \in \Lambda\}$ of elements of X is called an **orthonormal set** if

1. $\langle x_\alpha, x_\beta \rangle = 0$ for all $\alpha \neq \beta$ (i.e., S is an orthogonal set), and
2. $\|x_\alpha\| = 1 \quad$ for all $\quad \alpha \in \Lambda$

Definition 11.3.8 — Fourier series. If $S = \{x_\alpha : \alpha \in \Lambda\}$ is an orthonormal set and $x \in X$, then the numbers $\langle x, x_\alpha \rangle$ are called the **Fourier coefficients** of x with respect to S and the formal series $\sum_{\alpha \in \Lambda} \langle x, x_\alpha \rangle x_\alpha$ the **Fourier series** of x.

Theorem 11.3.16 An orthonormal set S in a separable inner product space $(X, \langle \cdot, \cdot \rangle)$ is at most countable.

Proof. If S is finite, then there is nothing to prove. Assume that S is infinite. Observe that if $x, y \in S$ then $\|x - y\| = \sqrt{2}$ (since x and y are orthonormal). Let $D = \{y_n | n \in \mathbb{N}\}$ be a countable dense subset of X. Then to each $x \in S$ corresponds an element $y_n \in D$ such that $\|x - y_n\| < \frac{\sqrt{2}}{4}$. This defines a map $f : S \to \mathbb{N}$ given by $f(x) = n$, where n corresponds to the y_n as indicated above. Now, if x and y are distinct elements of S, then there are distinct elements y_n and y_m in D such that

$$\|x - y_n\| < \frac{\sqrt{2}}{4} \text{ and } \|y - y_m\| < \frac{\sqrt{2}}{4}$$

Hence,

$$\sqrt{2} = \|x - y\| \leq \|x - y_n\| + \|y_n - y_m\| + \|y_m - y\| < \frac{\sqrt{2}}{2} + \|y_n - y_m\| \Longleftrightarrow \frac{\sqrt{2}}{2} < \|y_n - y_m\|$$

and so $y_n \neq y_m$. In particular, $n \neq m$. Thus, we have a one-to-one correspondence between the elements of S and a subset of $\mathbb{N}$. ■

Definition 11.3.9 — Complete. An orthonormal set S in an inner product space $(X, \langle \cdot, \cdot \rangle)$ is said to be complete in X if $S \subset T$ and T is an orthonormal set in X, then $S = T$.

R Simply put, a complete orthonormal set S in an inner product space is an orthonormal set that is not properly contained in any other orthonormal set in X; in other words, S is complete if it is a maximal orthonormal set in X

It is easy exercise to show that a set S is complete in an inner product $(X, \langle \cdot, \cdot \rangle)$ if and only if $S^\perp = \{0\}$

Theorem 11.3.17 Let $(X, \langle \cdot, \cdot \rangle)$ be a separable inner product space over $\mathbb{F}$

1. (Best Fit). If $\{x_1, x_2, \ldots, x_n\}$ is a finite orthonormal set in X and

$$M = \text{lin}\{x_1, x_2, \ldots, x_n\}$$

then for each $x \in X$ there exists $y_0 \in M$ such that

$$\|x - y_0\| = d(x, M)$$

In fact, $y_0 = \sum_{k=1}^{n} \langle x, x_k \rangle x_k$

2. (Bessel's Inequality). Let $(x_n)_{n=1}^{\infty}$ be an orthonormal sequence in X. Then for any

$x \in X$

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2$$

Proof. 1. For any choice of scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$

$$\begin{aligned}
\left\| x - \sum_{k=1}^{n} \lambda_k x_k \right\|^2 &= \left\langle x - \sum_{i=1}^{n} \lambda_i x_i, x - \sum_{j=1}^{n} \lambda_j x_j \right\rangle \\
&= \|x\|^2 - \sum_{i=1}^{n} \lambda_i \langle x_i, x \rangle - \sum_{j=1}^{n} \overline{\lambda_j} \langle x, x_j \rangle + \sum_{i=1}^{n} \lambda_i \overline{\lambda_i} \\
&= \|x\|^2 - \sum_{i=1}^{n} \lambda_i \overline{\langle x, x_i \rangle} - \sum_{j=1}^{n} \overline{\lambda_j} \langle x, x_j \rangle + \sum_{i=1}^{n} \lambda_i \overline{\lambda_i} \\
&= \|x\|^2 + \sum_{i=1}^{n} \left[\lambda_i \overline{\lambda_i} - \lambda_i \overline{\langle x, x_i \rangle} - \overline{\lambda_i} \langle x, x_i \rangle + \langle x, x_i \rangle \overline{\langle x, x_i \rangle} \right] - \sum_{i=1}^{n} |\langle x, x_i \rangle|^2 \\
&= \|x\|^2 + \sum_{i=1}^{n} \left[(\lambda_i - \langle x, x_i \rangle) \left(\overline{\lambda_i} - \overline{\langle x, x_i \rangle} \right) \right] - \sum_{i=1}^{n} |\langle x, x_i \rangle|^2 \\
&= \|x\|^2 + \sum_{i=1}^{n} \left[(\lambda_i - \langle x, x_i \rangle) \left(\overline{\lambda_i - \langle x, x_i \rangle} \right) \right] - \sum_{i=1}^{n} |\langle x, x_i \rangle|^2 \\
&= \|x\|^2 - \sum_{i=1}^{n} |\langle x, x_i \rangle|^2 + \sum_{i=1}^{n} |\lambda_i - \langle x, x_i \rangle|^2
\end{aligned}$$

Therefore, $\|x - \sum_{k=1}^{n} \lambda_k x_k\|^2$ is minimal if and only if $\lambda_k = \langle x, x_k \rangle$ for each $k = 1, 2, \ldots, n$

2. For each positive integer n, and with $\lambda_k = \langle x, x_k \rangle$, the above argument shows that

$$0 \leq \left\| x - \sum_{k=1}^{n} \lambda_k x_k \right\|^2 = \|x\|^2 - \sum_{i=1}^{n} |\langle x, x_i \rangle|^2$$

Thus,

$$\sum_{k=1}^{n} |\langle x, x_k \rangle|^2 \leq \|x\|^2$$

Taking the limit as $n \to \infty$, we get

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2$$

■

Theorem 11.3.18 — Riesz-Fischer Theorem. Let $(x_n)_1^{\infty}$ be an orthonormal sequence in a separable Hilbert space $\mathcal{H}$ and let $(c_n)_1^{\infty}$ be a sequence of scalars. Then the series

$\sum_{k=1}^{\infty} c_k x_k$ converges in $\mathcal{H}$ if and only if $c = (c_n)_1^{\infty} \in \ell_2$. In this case

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\| = \left(\sum_{k=1}^{\infty} |c_k|^2 \right)^{\frac{1}{2}}$$

Proof. Assume that the series $\sum_{k=1}^{\infty} c_k x_k$ converges to x. Then for each $j, n \in \mathbb{N}$

$$\left\langle \sum_{k=1}^{n} c_k x_k, x_j \right\rangle = \sum_{k=1}^{n} c_k \langle x_k, x_j \rangle = c_j$$

Using continuity of the inner product

$$\langle x, x_j \rangle = \left\langle \sum_{k=1}^{\infty} c_k x_k, x_j \right\rangle = \lim_{n\to\infty} \left\langle \sum_{k=1}^{n} c_k x_k, x_j \right\rangle = \lim_{n\to\infty} c_j = c_j$$

By Bessel's Inequality, we have that

$$\sum_{k=1}^{\infty} |c_k|^2 = \sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2 < \infty$$

That is, $c = (c_n)_1^{\infty} \in \ell_2$.

Conversely, assume that $c = (c_n)_1^{\infty} \in \ell_2$. Set $z_n = \sum_{k=1}^{n} c_k x_k$. Then for $1 \leq n \leq m$

$$\|z_n - z_m\|^2 = \left\| \sum_{k=n+1}^{m} c_k x_k \right\|^2 = \sum_{k=n+1}^{m} |c_k|^2 \to 0 \text{ as } n \to \infty$$

Hence, $(z_n)_1^{\infty}$ is a Cauchy sequence in $\mathcal{H}$. Since $\mathcal{H}$ is complete, the sequence $(z_n)_1^{\infty}$ converges to some $x \in \mathcal{H}$. Hence the series $\sum_{k=1}^{\infty} c_k x_k$ converges to some element in $\mathcal{H}$. Also,

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\|^2 = \lim_{n\to\infty} \left\| \sum_{k=1}^{n} c_k x_k \right\|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} |c_k|^2$$

whence,

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\| = \left(\sum_{k=1}^{\infty} |c_k|^2 \right)^{\frac{1}{2}}$$

■

R Note that Bessel's Inequality says that

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2 < \infty$$

That is, $(\langle x, x_n \rangle)_1^{\infty} \in \ell_2$. Hence, by Theorem11.3.18 , the series $\sum_{k=1}^{\infty} \langle x, x_k \rangle x_k$ converges.

There is however no reason why this series should converge to x. In fact, the following example shows that this series may not converge to x.

■ **Example 11.17** Let $(e_n) \in \ell_2$, where $e_n = (\delta_{1n}, \delta_{2n}, \dots)$ with

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

For each $n \in \mathbb{N}$, let $f_n = e_{n+1}$. Then $(f_n)_{n=1}^{\infty}$ is an orthonormal sequence in ℓ_2. For any $x = (x_n)_1^{\infty} \in \ell_2$

$$\sum_{k=1}^{\infty} \langle x, f_k \rangle f_k = \sum_{k=1}^{\infty} \langle x, e_{k+1} \rangle e_{k+1} = (0, x_2, x_3, \dots) \neq (x_1, x_2, x_3, \dots) = x$$

■

Definition 11.3.10 — Orthonormal basis. Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{F}$. An orthonormal set $\{x_n\}$ is called an **orthonormal basis** for X if for each $x \in X$

$$x = \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k$$

That is, the sequence of partial sums (s_n), where $s_n = \sum_{k=1}^{n} \langle x, x_k \rangle x_k$, converges to x

Theorem 11.3.19 Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space and assume that $S = \{x_n\}$ is an orthonormal set in $\mathcal{H}$. Then the following statements are equivalent:

1. S is complete in $\mathcal{H}$; i.e., $S^{\perp} = \{0\}$
2. $\overline{\operatorname{lin} S} = \mathcal{H}$; i.e., the linear span of S is norm-dense in $\mathcal{H}$
3. (Fourier Series Expansion.) For any $x \in \mathcal{H}$, we have

$$x = \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i$$

   That is, S is an orthonormal basis for $\mathcal{H}$
4. (Parseval's Identity.) For all $x, y \in \mathcal{H}$

$$\langle x, y \rangle = \sum_{k=1}^{\infty} \langle x, x_k \rangle \overline{\langle y, x_k \rangle}$$

5. For any $x \in \mathcal{H}$

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2$$

Proof. 1 $\iff$ [2]". This equivalence was proved in Proposition $S^{\perp} = \{0\}$ if and only if $\overline{\operatorname{lin}} S = \mathcal{H}$.

1. [1] $\Rightarrow$ [3]". Let $x \in \mathcal{H}$ and $s_n = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$. Then for all $n > m$

$$\|s_n - s_m\|^2 = \left\| \sum_{i=m+1}^{n} \langle x, x_i \rangle x_i \right\|^2 = \sum_{m+1}^{n} |\langle x, x_i \rangle|^2 \leq \|x\|^2$$

Thus, (s_n) is a Cauchy sequence in $\mathcal{H}$. since $\mathcal{H}$ is complete, this sequence converges to some element which we denote by $\sum_{i=1}^{\infty} \langle x, x_i \rangle x_i$. We show that $x = \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i$. Indeed, for each fixed $j \in \mathbb{N}$

$$\begin{aligned}
\left\langle x - \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i, x_j \right\rangle &= \left\langle x - \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, x_i \rangle x_i, x_j \right\rangle \\
&= \lim_{n\to\infty} \left\langle x - \sum_{i=1}^{n} \langle x, x_i \rangle x_i, x_j \right\rangle \\
&= \lim_{n\to\infty} \left(\langle x, x_j \rangle - \sum_{i=1}^{n} \langle x, x_i \rangle \langle x_i, x_j \rangle \right) \\
&= \lim_{n\to\infty} \left(\langle x, x_j \rangle - \langle x, x_j \rangle \right) = 0
\end{aligned}$$

Thus, by [1], $x - \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i = 0$, whence

$$x = \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i$$

2. "[3] $\Rightarrow$ [4]". Let $x, y \in \mathcal{H}$. Then

$$\begin{aligned}
\langle x, y \rangle &= \lim_{n\to\infty} \left\langle \sum_{i=1}^{n} \langle x, x_i \rangle x_i, \sum_{j=1}^{n} \langle y, x_j \rangle x_j \right\rangle \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \langle x, x_i \rangle \overline{\langle y, x_j \rangle} \langle x_i, x_j \rangle \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, x_i \rangle \overline{\langle y, x_i \rangle} = \sum_{i=1}^{\infty} \langle x, x_i \rangle \overline{\langle y, x_i \rangle}
\end{aligned}$$

3. "[4] $\Rightarrow$ [5]". Take $x = y$ in [4]
4. "[5] $\Rightarrow$ [1]" since $\|x\|^2 = \sum_k \langle x, x_k \rangle \overline{\langle x, x_k \rangle}$, if $x \perp S$ then $\langle x, x_k \rangle = 0$ for all k. Thus, $\|x\|^2 = 0$ whence $x = 0$. That is, $S^{\perp} = \{0\}$

■

■ **Example 11.18** The set $S = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos nt}{\sqrt{\pi}}, \frac{\sin nt}{\sqrt{\pi}} \right\}_{n=1}^{\infty}$ is an orthonormal basis for the real $L_2[-\pi, \pi]$.

Hence, if $x \in L_2[-\pi, \pi]$, then by Theorem11.3.19 [3]

$$\begin{aligned}
x(t) &= \left\langle x(t), \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{n=1}^{\infty} \left[\left\langle x(t), \frac{\cos nt}{\sqrt{\pi}} \right\rangle \frac{\cos nt}{\sqrt{\pi}} + \left\langle x(t), \frac{\sin nt}{\sqrt{\pi}} \right\rangle \frac{\sin nt}{\sqrt{\pi}} \right] \\
&= \frac{1}{2\pi} \langle x(t), 1 \rangle + \sum_{n=1}^{\infty} \left[\frac{1}{\pi} \langle x(t), \cos nt \rangle \cos nt + \frac{1}{\pi} \langle x(t), \sin nt \rangle \sin nt \right] \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t) dt \\
&+ \sum_{n=1}^{\infty} \left[\left(\frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \cos nt dt \right) \cos nt + \left(\frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \sin nt dt \right) \sin nt \right] \\
&= a_0 + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt)
\end{aligned}$$

where

$$\left.\begin{aligned} a_0 &= \frac{1}{\sqrt{2\pi}}\left\langle x(t), \frac{1}{\sqrt{2\pi}}\right\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)dt \\ a_n &= \frac{1}{\sqrt{\pi}}\left\langle x(t), \frac{\cos nt}{\sqrt{\pi}}\right\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos ntdt, \quad \text{and} \\ b_n &= \frac{1}{\sqrt{\pi}}\left\langle x(t), \frac{\sin nt}{\sqrt{\pi}}\right\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin ntdt \end{aligned}\right\} \quad n = 1, 2, \ldots$$

That is, the Fourier series expansion of x is

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt) \tag{11.6}$$

It is clear from above that for all $n = 1, 2, \ldots$

$$2\pi|a_0|^2 = \left|\left\langle x(t), \tfrac{1}{\sqrt{2\pi}}\right\rangle\right|^2, \pi|a_n|^2 = \left|\left\langle x(t), \tfrac{\cos nt}{\sqrt{\pi}}\right||^2, \text{ and}$$
$$\pi|b_n|^2 = \left|\left\langle x(t), \tfrac{\sin nt}{\sqrt{\pi}}\right\rangle\right|^2$$

By Theorem 11.3.19 [5] we have that

$$\begin{aligned} \int_{-\pi}^{\pi}|x(t)|^2dt = \|x\|^2 &= \left|\left\langle x(t), \frac{1}{\sqrt{2\pi}}\right\rangle\right|^2 \\ &+ \sum_{n=1}^{\infty}\left(\left|\left\langle x(t), \frac{\cos nt}{\sqrt{\pi}}\right\rangle\right|^2 + \left|\left\langle x(t), \frac{\sin nt}{\sqrt{\pi}}\right||^2\right) \\ &= 2\pi|a_0|^2 + \sum_{n=1}^{\infty}\left(\pi|a_n|^2 + \pi|b_n|^2\right) \\ &= \pi\left(2|a_0|^2 + \sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right)\right) \end{aligned}$$

■

■ **Example 11.19** We now apply the above results to a particular function: Let $x(t) = t$. Then $a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} tdt = 0 \quad$ since $x(t) = t$ is an odd function For $n = 1, 2, \ldots, \quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos ntdt = 0 \quad$ since $t\cos nt$ is an odd function

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin ntdt = \frac{2}{\pi}\int_0^{\pi} t\sin ntdt \\ &= \frac{2}{\pi}\left[\left.\frac{-t\cos nt}{n}\right|_0^{\pi} + \frac{1}{n}\int_0^{\pi}\cos ntdt\right] \\ &= \frac{2}{\pi}\left[\frac{-\pi}{n}\cos n\pi\right] = \frac{2(-1)^{n+1}}{n} \end{aligned}$$

Hence, by Theorem 11.3.19 [3]

$$x(t) = \sum_{n=1}^{\infty}\frac{2(-1)^{n+1}}{n}\sin nt = \sum_{n=1}^{\infty}\frac{2(-1)^{n+1}\sqrt{\pi}}{n}\frac{\sin nt}{\sqrt{\pi}}$$

It now follows that

$$\frac{2(-1)^{n+1}\sqrt{\pi}}{n} = \left\langle t, \frac{\sin nt}{\sqrt{\pi}} \right\rangle$$

Now

$$\|x\|_2^2 = \int_{-\pi}^{\pi} t^2 dt = 2\int_0^{\pi} t^2 dt = \frac{2}{3}t^3\Big|_0^{\pi} = \frac{2\pi^3}{3}$$

Also, by Theorem 11.3.19 [5]

$$\|x\|_2^2 = \sum_{n=1}^{\infty} \left|\left\langle t, \frac{\sin nt}{\sqrt{\pi}} \right|\right|^2 = \sum_{n=1}^{\infty} \left|\frac{2(-1)^{n+1}\sqrt{\pi}}{n}\right|^2 = \sum_{n=1}^{\infty} \frac{4\pi}{n^2}$$

Thus,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

■

Theorem 11.3.20 We can express the Fourier Series Expansion 11.6 of $x \in L_2[-\pi, \pi]$ in exponential form. Recall that

$$e^{i\theta} = \cos\theta + i\sin\theta \qquad \text{(Euler's Formula).}$$

Therefore

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ and } \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Equation 11.6 now becomes

$$\begin{aligned}
x(t) &= a_0 + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt) \\
&= a_0 + \sum_{n=1}^{\infty} \left[a_n \left(\frac{e^{int} + e^{-int}}{2} \right) + b_n \left(\frac{e^{int} - e^{-int}}{2i} \right) \right] \\
&= a_0 + \sum_{n=1}^{\infty} \left[\left(\frac{a_n - ib_n}{2} \right) e^{int} + \left(\frac{a_n + ib_n}{2} \right) e^{-int} \right] \\
&= a_0 + \sum_{n=1}^{\infty} \left(\frac{a_n - ib_n}{2} \right) e^{int} + \sum_{n=1}^{\infty} \left(\frac{a_n + ib_n}{2} \right) e^{-int}
\end{aligned}$$

For each $n = 1, 2, 3, \ldots$, let $c_n = \frac{1}{2}(a_n + ib_n)$. Then $\overline{c_n} = \frac{1}{2}(a_n - ib_n)$ for each $n = 1, 2, 3, \ldots$, and so last equation becomes

$$x(t) = a_0 + \sum_{n=1}^{\infty} \overline{c_n} e^{int} + \sum_{n=1}^{\infty} c_n e^{-int}$$

Re-index the first sum by letting $n = -k$. Then

$$x(t) = a_0 + \sum_{k=-1}^{-\infty} \overline{c_{-k}} e^{-ikt} + \sum_{n=1}^{\infty} c_n e^{-int}$$

For $n = -1, -2, -3, \ldots$, define

$$c_n = \overline{c_{-n}}$$

and let $c_0 = a_0$. The we can rewrite equation as

$$x(t) = \sum_{-\infty}^{\infty} c_n e^{-int}$$

This is the complex exponential form of the Fourier Series of $x \in L_2[-\pi, \pi]$.

Note that,

$$c_0 = a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t)dt$$

and for $n = 1,2,3,\ldots$

$$\begin{aligned} c_n &= \frac{1}{2}(a_n + ib_n) = \frac{1}{2}\left[\frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos nt dt + i\frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin nt dt\right] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)(\cos nt + i \sin nt)dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{int}dt \end{aligned}$$

and

$$c_{-n} = \overline{c_n} = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-int}dt} = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\overline{e^{-int}}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{int}dt$$

Therefore, for all $n \in \mathbb{Z}$

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{int}dt$$

Now, for $n = 1,2,3,\ldots$

$$\begin{aligned} |c_n|^2 = c_n \cdot \overline{c_n} &= \frac{1}{2}(a_n + ib_n) \cdot \frac{1}{2}(a_n - ib_n) \\ &= \frac{1}{4}(a_n^2 + b_n^2) = \frac{1}{4}\left(|a_n|^2 + |b_n|^2\right) \end{aligned}$$

Therefore

$$\sum_{n=1}^{\infty} |c_n|^2 = \frac{1}{4}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right) \tag{11.7}$$

Since for $n = 1,2,3,\ldots, c_{-n} = \overline{c_n}$, it follows that

$$|c_{-n}|^2 = c_{-n} \cdot \overline{c_{-n}} = \overline{c_n} \cdot \overline{\overline{c_n}} = \overline{c_n} \cdot c_n = |c_n|^2$$

Hence, for $n = 1,2,3,\ldots$

$$\sum_{n=1}^{\infty} |c_{-n}|^2 = \sum_{n=1}^{\infty} |c_n|^2 = \frac{1}{4} \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right) \tag{11.8}$$

From 11.7 and 11.8 , we have that

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} |c_n|^2 &= \sum_{n=1}^{\infty} |c_{-n}|^2 + |c_0|^2 + \sum_{n=1}^{\infty} |c_n|^2 \\
&= \frac{1}{4} \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right) + |a_0|^2 + \frac{1}{4} \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right) \\
&= |a_0|^2 + \frac{1}{2} \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right) \\
&= \frac{1}{2\pi} \cdot \pi \left(2\,|a_0|^2 + \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right)\right) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} |x(t)|^2 dt
\end{aligned}$$

That is,

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |x(t)|^2 dt$$

Let $\{x_1, x_2, \ldots, x_n\}$ be a basis of an n -dimensional linear subspace M of an inner product space $(X, \langle\cdot,\cdot\rangle)$. We have seen in Theorem 3.5 .2 that if the set $\{x_1, x_2, \ldots, x_n\}$ is orthonormal, then the orthogonal projection (=best approximation) of any $x \in X$ onto M is given by

$$P_M(x) = \sum_{k=1}^{n} \langle x, x_k \rangle\, x_k$$

It is clearly easy to compute orthogonal projections from a linear subspace that has an orthonormal basis: the coefficients in the orthogonal projection of $x \in X$ are just the "Fourier coefficients" of x. If the basis of M is not orthogonal, it may be advantageous to find an orthonormal basis for M and express the orthogonal projection as a linear combination of the new orthonormal basis. The process of finding an orthonormal basis from a given (non-orthonormal) basis is known as the Gram-Schmidt Orthonormalisation Procedure.

Theorem 11.3.21 — Gram-Schmidt Orthonormalisation Procedure. If $\{x_k\}_1^{\infty}$ is a linearly independent set in an inner product space $(X, \langle\cdot,\cdot\rangle)$ then there exists an orthonormal set $\{e_k\}_1^\infty$ in X such that

$$\operatorname{lin}\{x_1, x_2, \ldots, x_n\} = \operatorname{lin}\{e_1, e_2, \ldots, e_n\} \quad \text{for all} \quad n$$

Proof. Set $e_1 = \frac{x_1}{\|x_1\|}$. Then $\operatorname{lin}\{x_1\} = \operatorname{lin}\{e_1\}$. Next, let $y_2 = x_2 - \langle x_2, e_1\rangle e_1$. Then

$$\langle y_2, e_1\rangle = \langle x_2 - \langle x_2, e_1\rangle e_1, e_1\rangle = \langle x_2, e_1\rangle - \langle x_2, e_1\rangle \langle e_1, e_1\rangle = \langle x_2, e_1\rangle - \langle x_2, e_1\rangle = 0$$

That is, $e_1 \perp y_2$. Set $e_2 = \frac{y_2}{\|y_2\|}$. Then $\{e_1, e_2\}$ is an orthonormal set with the property that $\operatorname{lin}\{x_1, x_2\} = \operatorname{lin}\{e_1, e_2\}$. In general, for each $k = 2, 3, \ldots,$ we let

$$y_k = x_k - \sum_{i=1}^{k-1} \langle x_k, e_i\rangle e_i$$

Then for $k = 2, 3, \ldots$

$$\langle y_k, e_1\rangle = \langle y_k, e_2\rangle = \langle y_k, e_3\rangle = \cdots = \langle y_k, e_{k-1}\rangle = 0$$

Set $e_k = \frac{y_k}{\|y_k\|}$. Then $\{e_1, e_2, \ldots, e_k\}$ is an orthonormal set in X with the property that

$$\operatorname{lin}\{e_1, e_2, \ldots, e_k\} = \operatorname{lin}\{x_1, x_2, \ldots, x_k\}$$

■

Definition 11.3.11 — Isomorphic. Two linear spaces X and Y over the same field $\mathbb{F}$ are said to be **isomorphic** it there is a one-to-one map T from X onto Y such that for all $x_1, x_2 \in X$ and all $\alpha, \beta \in \mathbb{F}$

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2)$$

Any map that satisfies previous condition is called a **linear operator**. Clearly, the linear structures of the two linear spaces X and Y are preserved under the map T.

Definition 11.3.12 — Isometry. Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be two normed linear spaces and $T : X \to Y$. Then T is called an isometry if

$$\|Tx\| = \|x\| \text{ for all } x \in X$$

Simply put, an isometry is a map that preserves lengths.

It is implicit in the above definition that the norm on the left of 11.3.12 is in Y and that on the right is in X. In order to avoid possible confusion, we should perhaps have labelled the norms as $\|\cdot\|_X$ and $\|\cdot\|_Y$ for the norms in X and Y respectively. This notation is however cumbersome and will therefore be avoided.

R Normed linear spaces that are isometrically isomorphic are essentially identical.

Theorem 11.3.22 Let $M = \text{lin}\{x_n\}$ be a linear subspace of X. Then there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ which has the following properties:

1. $\text{lin}\{x_{n_k}\} = M$
2. $\{x_{n_k}\}$ is linearly independent.

Proof. We define the subsequence inductively as follows: Let x_{n_1} be the first nonzero element of the sequence $\{x_n\}$. Therefore $x_n = 0 \cdot x_{n_1}$ for all $n < n_1$. If there is an $\alpha \in \mathbb{F}$ such that $x_n = \alpha x_{n_1}$ for all $n > n_1$, then we are done. Otherwise, let x_{n_2} be the first element of the sequence $\{x_n\}_{n>n_1}$ that is not a multiple of x_{n_1}. Thus there is an $\alpha \in \mathbb{F}$ such that $x_n = \alpha x_{n_1} + 0 x_{n_2}$ for all $n < n_2$. If $x_n = \alpha x_{n_1} + \beta x_{n_2}$ for some $\alpha, \beta \in \mathbb{F}$ and all $n > n_2$ then we are done. Otherwise let x_{n_3} be the first element of the sequence $\{x_n\}$ which is not a linear combination of x_{n_1} and x_{n_2}. Then $x_n = \alpha x_{n_1} + \beta x_{n_2} + 0 x_{n_3}$ for all $n < n_3$. If $x_n = \alpha x_{n_1} + \beta x_{n_2} + \gamma x_{n_3}$ for all $n > n_3$, then we are done. Otherwise let x_{n_4} be the first element of the sequence $\{x_n\}$ that is not in $\text{lin}\{x_{n_1}, x_{n_2}, x_{n_3}\}$. Continue in this fashion to obtain elements $x_{n_1}, x_{n_2}, \ldots$. If $x \in \text{lin}\{x_1, x_2, \ldots, x_n\}$, then $x \in \text{lin}\{x_{n_1}, x_{n_2}, \ldots, x_{n_r}\}$ for r sufficiently large. That is $\text{lin}\{x_{n_k}\} = M$ The subsequence $\{x_{n_k}\}$ is, by its construction, linearly independent.

■

Theorem 11.3.23 Every separable Hilbert space $\mathcal{H}$ has a countable orthonormal basis.

Proof. By Theorem 11.2.23 there is a set $\{x_n | n \in \mathbb{N}\}$ such that $\overline{\text{lin}}\{x_n | n \in \mathbb{N}\} = \mathcal{H}$. Using 11.3.22 extract from $\{x_n | n \in \mathbb{N}\}$ a linearly independent subsequence $\{x_{n_k}\}$ such that $\text{lin}\{x_n\} = \text{lin}\{x_{n_k}\}$. Apply the Gram-Schmidt Orthonormalisation Procedure to the subsequence $\{x_{n_k}\}$ to obtain an orthonormal basis for $\mathcal{H}$ ■

Theorem 11.3.24 Every separable infinite-dimensional Hilbert space $\mathcal{H}$ is isometrically isomorphic to ℓ_2.

Proof. Let $\{x_n | n \in \mathbb{N}\}$ be an orthonormal basis for $\mathcal{H}$. Define $T : \mathcal{H} \to \ell_2$ by

$$Tx = (\langle x, x_n \rangle)_{n \in \mathbb{N}} \text{ for each } x \in \mathcal{H}$$

It follows from Bessel's Inequality that the right hand side is in ℓ_2. We must show that T is a surjective linear isometry. (One-to-oneness follows from isometry.) (i) $\underline{T}$ is linear: Let $x, y \in \mathcal{H}$ and $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} T(x+y) &= (\langle x+y, x_n \rangle)_{n \in \mathbb{N}} = (\langle x, x_n \rangle + \langle y, x_n \rangle)_{n \in \mathbb{N}} \\ &= (\langle x, x_n \rangle)_{n \in \mathbb{N}} + (\langle y, x_n \rangle)_{n \in \mathbb{N}} = Tx + Ty \end{aligned}$$

and

$$T(\lambda x) = (\langle \lambda x, x_n \rangle)_{n\in\mathbb{N}} = (\lambda \langle x, x_n \rangle)_{n\in\mathbb{N}} = \lambda (\langle x, x_n \rangle)_{n\in\mathbb{N}}$$

(ii) T is surjective: Let $(c_n)_{n\in\mathbb{N}} \in \ell_2$. By the Riesz-Fischer Theorem (Theorem 3.5.3), the series $\sum_{k=1}^{\infty} c_k x_k$ converges to some $x \in \mathcal{H}$. By continuity of the inner product, we have that for each $j \in \mathbb{N}$

$$\langle x, x_j \rangle = \lim_{n\to\infty} \left\langle \sum_{k=1}^{n} c_k x_k, x_j \right\rangle = \lim_{n\to\infty} c_j = c_j$$

Hence, $Tx = (\langle x, x_n \rangle)_{n\in\mathbb{N}} = (c_n)_{n\in\mathbb{N}}$ (iii) T is an isometry: For each $x \in \mathcal{H}$

$$\|Tx\|_2^2 = \sum_{n\in\mathbb{N}} |\langle x, x_n \rangle|^2 = \|x\|^2$$

where the second equality follows from the fact that $\{x_n\}_{n\in\mathrm{N}}$ is an orthonormal basis and Theorem 3.5 .4[5] ■

11.4 Bounded Linear Operators and Functionals

Definition 11.4.1 — Linear operator. Let X and Y be linear spaces over the same field $\mathbb{F}$. A **linear operator** from X into Y is a mapping $T: X \to Y$ such that

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 \text{ for all } x_1, x_2 \in X \text{ and all } \alpha, \beta \in \mathbb{F}$$

Simply put, a linear operator between linear spaces is a mapping that preserves the structure of the underlying linear space.

We shall denote by $\mathcal{L}(X,Y)$ the set of all linear operators from X into Y. We shall write $\mathcal{L}(X)$ for $\mathcal{L}(X,X)$

Definition 11.4.2 — range,null space. The **range** of a linear operator $T: X \to Y$ is the set

$$\operatorname{ran}(T) = \{y \in Y | y = Tx \text{ for some } x \in X\} = TX$$

and the **null space** or the **kernel** of $T \in \mathcal{L}(X,Y)$ is the set

$$\mathcal{N}(T) = \ker(T) = \{x \in X : Tx = 0\} = T^{-1}(0)$$

If $T \in \mathcal{L}(X,Y)$, then $\ker(T)$ is a linear subspace of X and $\operatorname{ran}(T)$ is a linear subspace of Y.

Theorem 11.4.1 An operator $T \in \mathcal{L}(X,Y)$ is **one-to-one** (or injective) if $\ker(T) = \{0\}$ and **onto** (or surjective) if $\operatorname{ran}(T) = Y$.

If $T \in \mathcal{L}(X,Y)$ is one-to-one, then there exists a map $T^{-1}: \operatorname{ran}(T) \to \operatorname{dom}(T)$ which maps each $y \in \operatorname{ran}(T)$ onto that $x \in \operatorname{dom}(T)$ for which $Tx = y$.

In this case we write $T^{-1}y = x$ and the map T^{-1} is called the inverse of the operator $T \in \mathcal{L}(X,Y)$. An operator $T : X \to Y$ is invertible if it has an inverse T^{-1}

Proposition 11.4.2 Let X and Y be linear spaces over $\mathbb{F}$. Suppose that $T \in \mathcal{L}(X,Y)$ is invertible. Then

1. T^{-1} is also invertible and $\left(T^{-1}\right)^{-1} = T$
2. $TT^{-1} = I_Y$ and $T^{-1}T = I_X$
3. T^{-1} is a linear operator.

Proof. (c) Linearity of T^{-1} : Let $x,y \in Y$ and $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} T^{-1}(\alpha x + y) &= T^{-1}\left(\alpha TT^{-1}x + TT^{-1}y\right) = T^{-1}\left[T\left(\alpha T^{-1}x + T^{-1}y\right)\right] \\ &= T^{-1}T\left(\alpha T^{-1}x + T^{-1}y\right) = \alpha T^{-1}x + T^{-1}y \end{aligned}$$

■

Definition 11.4.3 Let X and Y be linear spaces over $\mathbb{F}$. For all $T,S \in \mathcal{L}(X,Y)$ and $\alpha \in \mathbb{F}$, define the operations of addition and scalar multiplication as follows:

$$\begin{aligned} (T+S)(x) &= Tx + Sx \text{ and} \\ (\alpha T)(x) &= \alpha Tx \quad \text{for each} \quad x \in X \end{aligned}$$

Then $\mathcal{L}(X,Y)$ is a linear space over $\mathbb{F}$

Definition 11.4.4 — Bounded linear operators. Let X and Y be normed linear spaces over the same field $\mathbb{F}$. A linear operator $T : X \to Y$ is said to be **bounded** if there exists a constant $M > 0$ such that

$$\|Tx\| \leq M\|x\| \quad \text{for all} \quad x \in X$$

If should be emphasised that the norm on the left side is in Y and that on the right side is in X.)

An operator $T : X \to Y$ is said to be continuous at $x_0 \in X$ if given any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|Tx - Tx_0\| < \epsilon \quad \text{whenever } \|x - x_0\| < \delta$$

T is continuous on X if it is continuous at each point of X

We shall denote by $\mathcal{B}(X,Y)$ the set of all bounded linear operators from X into Y. We shall write $\mathcal{B}(X)$ for $\mathcal{B}(X,X)$

Definition 11.4.5 Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and let

$T \in \mathcal{B}(X,Y)$. The **operator norm** (or simply norm) of T, denoted by $\|T\|$, is defined as

$$\|T\| = \inf\{M : \|Tx\| \leq M\|x\|, \quad \text{for all } x \in X\}$$

Since T is bounded, $\|T\| < \infty$. Furthermore

$$\|Tx\| \leq \|T\|\|x\| \quad \text{for all} \quad x \in X$$

Theorem 11.4.3 Let X and Y be normed linear spaces over a field $\mathbb{F}$ and let $T \in \mathcal{B}(X,Y)$. Then

$$\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} = \sup\{\|Tx\| : \|x\| = 1\} = \sup\{\|Tx\| : \|x\| \leq 1\}$$

Proof. Let $\alpha = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\}$, $\beta = \sup\{\|Tx\| : \|x\| = 1\}$, and $\gamma = \sup\{\|Tx\| : \|x\| \leq 1\}$.

We first show that $\|T\| = \alpha$. Now, for all $x \in X\backslash\{0\}$ we have that $\frac{\|Tx\|}{\|x\|} \leq \alpha$, and therefore $\|Tx\| \leq \alpha\|x\|$. By definition of $\|T\|$ we have that $\|T\| \leq \alpha$. On the other hand, for all $x \in X$ we have that $\|Tx\| \leq \|T\|\|x\|$. In particular, for all $x \in X\backslash\{0\}$, $\frac{\|Tx\|}{\|x\|} \leq \|T\|$, and therefore

$$\alpha = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} \leq \|T\|. \text{ Thus, } \alpha = \|T\|$$

Next, we show that $\alpha = \beta = \gamma$. Now, for each $x \in X$

$$\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} = \left\{\left\|T\left(\frac{x}{\|x\|}\right)\right\| : x \neq 0\right\} \subset \{\|Tx\| : \|x\| = 1\} \subset \{\|Tx\| : \|x\| \leq 1\}$$

Thus,

$$\alpha = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} \leq \beta = \sup\{\|Tx\| : \|x\| = 1\} \leq \gamma = \sup\{\|Tx\| : \|x\| \leq 1\}$$

But for all $x \neq 0$ $\frac{\|Tx\|}{\|x\|} \leq \alpha \Rightarrow \|Tx\| \leq \alpha\|x\| \leq \alpha$ for all x such that $\|x\| \leq 1$.

Therefore,

$$\gamma = \sup\{\|Tx\| : \|x\| \leq 1\} \leq \alpha$$

That is, $\alpha \leq \beta \leq \gamma \leq \alpha$. Hence, $\alpha = \beta = \gamma$ ∎

Theorem 11.4.4 Let X and Y be normed linear spaces over a field $\mathbb{F}$. Then the function $\|\cdot\|$ defined above is a norm on $\mathcal{B}(X,Y)$

Proof. Properties N1 and N2 of a norm are easy to verify. We prove N3 and N4. Let $T \in \mathcal{B}(X,Y)$ and $\alpha \in \mathbb{F}$.

N3. $\|\alpha T\| = \sup\{\|\alpha Tx\| : \|x\| = 1\} = |\alpha|\sup\{\|Tx\| : \|x\| = 1\} = |\alpha|\|T\|$.

N4. Let $T, S \in \mathcal{B}(X,Y)$. Then for each $x \in X$

$$\|(T+S)(x)\| = \|Tx+Sx\| \leq \|Tx\| + \|Sx\| \leq (\|T\|+\|S\|)\|x\|$$

Thus, $\|T+S\| \leq \|T\| + \|S\|$ ■

■ **Example 11.20** Let $X = \mathbb{F}^n$ with the uniform norm $\|\cdot\|_\infty$. For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$, define $T : \mathbb{F}^n \to \mathbb{F}^n$ by

$$Tx = T(x_1, x_2, \ldots, x_n) = \left(\sum_{j=1}^{n} \alpha_{1j} x_j, \sum_{j=1}^{n} \alpha_{2j} x_j, \ldots, \sum_{j=1}^{n} \alpha_{nj} x_j \right)$$

It is easy to show that T is a linear operator on X. We show that T is bounded.

$$\begin{aligned}
\|Tx\|_\infty &= \sup_{1\leq i\leq n} \left| \sum_{j=1}^{n} \alpha_{ij} x_j \right| \leq \sup_{1\leq i\leq n} \sum_{j=1}^{n} |\alpha_{ij}||x_j| \\
&\leq \sup_{1\leq i\leq n} \sum_{j=1}^{n} |\alpha_{ij}| \sup_{1\leq j\leq n} |x_j| = M\|x\|_\infty
\end{aligned}$$

where $M = \sup_{1\leq i\leq n} \sum_{j=1}^{n} |\alpha_{ij}|$. Hence, $\|T\| \leq M$ We claim that $\|T\| = M$. We need to show that $\|Tx\|_\infty \geq M\|x\|_\infty$. To that end, choose an index k such that $\sum_{j=1}^{n} |\alpha_{kj}| = M = \sup_{1\leq i\leq n} \sum_{j=1}^{n} |\alpha_{ij}|$ and let x be the unit vector whose j-th component is $\frac{\overline{\alpha_{kj}}}{|\alpha_{kj}|}$. Then

$$\|Tx\|_\infty = \sup_{1\leq i\leq n} \left| \sum_{j=1}^{n} \alpha_{ij} x_j \right| \geq \left| \sum_{j=1}^{n} \alpha_{kj} x_j \right| = \sum_{j=1}^{n} |\alpha_{kj}| = M\|x\|_\infty$$

Thus $\|T\| = \sup_{1\leq i\leq n} \sum_{j=1}^{n} |\alpha_{ij}|$.

R According to 11.4.3, we can take x on the unit ball, and at this time $\|x\|_\infty = 1$.

■ **Example 11.21** Let $\{x_n | n \in \mathbb{N}\}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. For $(\lambda_i)_{i=1}^{\infty} \in \ell_\infty$, define $T : \mathcal{H} \to \mathcal{H}$ by

$$Tx = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i$$

Then T is a bounded linear operator on $\mathcal{H}$. Linearity is an immediate consequence of the inner product. Boundedness:

$$\begin{aligned}
\|Tx\|^2 &= \left\| \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i \right\|^2 = \sum_{i=1}^{\infty} |\lambda_i|^2 |\langle x, x_i \rangle|^2 \|x_i\|^2 \\
&\leq M^2 \sum_{i=1}^{\infty} |\langle x, x_i \rangle|^2, \quad \text{where } M = \sup_{i\in\mathbb{N}} |\lambda_i| \\
&\leq M^2 \|x\|^2 \text{ by Bessel's Inequality.}
\end{aligned}$$

Thus, $\|Tx\| \le M\|x\|$, and consequently $\|T\| \le M$. We show that $\|T\| = \sup_{i\in\mathbb{N}} |\lambda_i|$. Indeed, for any $\epsilon > 0$, there exists λ_k such that $|\lambda_k| > M - \epsilon$. Hence,

$$\|T\| \ge \|Tx_k\| = \|\lambda_k x_k\| = |\lambda_k| > M - \epsilon$$

since ϵ is arbitrary, we have that $\|T\| \ge M$.

■ **Example 11.22 — left-shift operator.** Define an operator $L : \ell_2 \to \ell_2$ by

$$Lx = L((x_1, x_2, x_3, \ldots)) = (x_2, x_3, \ldots)$$

The L is a bounded linear operator on ℓ_2 Linearity: Easy. Boundedness: For all $x = (x_1, x_2, x_3, \ldots) \in \ell_2$

$$\|Lx\|_2^2 = \sum_{i=2}^{\infty} |x_i|^2 \le \sum_{i=1}^{\infty} |x_i|^2 = \|x\|_2^2$$

That is, L is a bounded linear operator and $\|L\| \le 1$. We show that $\|L\| = 1$. To that end, consider $e_2 = (0,1,0,0,\ldots) \in \ell_2$. Then $\|e_2\|_2 = 1$ and $Le_2 = (1,0,0,\ldots)$ which implies that $\|Le_2\|_2 = 1$. Thus, $\|L\| = 1$. The operator L is called the left-shift operator

■ **Example 11.23** Let $\mathcal{BC}[0,\infty)$ be the linear space of all bounded continuous functions on the interval $[0,\infty)$ with the uniform norm $\|\cdot\|_\infty$. Define $T : \mathcal{BC}[0,\infty) \to \mathcal{BC}[0,\infty)$ by

$$(Tx)(t) = \frac{1}{t}\int_0^t x(\tau)d\tau$$

Then T is a bounded linear operator on $\mathcal{BC}[0,\infty)$.

Linearity: For all $x, y \in \mathcal{BC}[0,\infty)$ and all $\alpha, \beta \in \mathbb{F}$

$$(T(\alpha x + \beta y))(t) = \frac{1}{t}\int_0^t (x+y)(\tau)d\tau = \alpha\left(\frac{1}{t}\int_0^t x(\tau)d\tau\right) + \beta\left(\frac{1}{t}\int_0^t y(\tau)d\tau\right)$$

Boundedness: For each $x \in \mathcal{BC}[0,\infty)$

$$\begin{aligned}\|Tx\|_\infty &= \sup_t |(Tx)(t)| = \sup_t \left|\frac{1}{t}\int_0^t x(\tau)d\tau\right| \\ &\le \sup_t \frac{1}{t}\int_0^t |x(\tau)|d\tau \le \left(\sup_t \frac{1}{t}\int_0^t d\tau\right)\|x\|_\infty = \|x\|_\infty\end{aligned}$$

■ **Example 11.24** Let M be a closed subspace of a normed linear space X and $Q_M : X \to X/M$ the qoutient map. Then Q_M is bounded and $\|Q_M\| = 1$. Indeed, since $\|Q_M(x)\| = \|x + M\| \le \|x\|, Q_M$ is bounded and $\|Q_M\| \le 1$. But since Q_M maps the open unit ball in X onto the open unit ball in X/M, it follows that $\|Q_M\| = 1$

■ **Example 11.25** Let $X = \mathcal{P}[0,1]$ - the set of polynomials on the interval [0,1] with the uniform norm $\|x\|_\infty = \max_{0\le t\le 1} |x(t)|$. For each $x \in X$, define $T : X \to X$ by

$$Tx = x'(t) = \frac{dx}{dt} \quad \text{(differentiation with respect to } t\text{)}$$

Linearity: For $x, y \in X$ and all $\alpha, \beta \in \mathbb{F}$

$$T(\alpha x + \beta y) = (\alpha x + \beta y)'(t) = \alpha x'(t) + \beta y'(t) = \alpha Tx + \beta Ty$$

T is not bounded: Let $x_n(t) = t^n, n \in \mathbb{N}$. Then

$$\|x_n\| = 1, \quad Tx_n = x_n'(t) = nt^{n-1}, \quad \text{and} \quad \|Tx_n\| = \frac{\|Tx_n\|}{\|x_n\|} = n$$

Hence T is unbounded.

Theorem 11.4.5 Let X and Y be normed linear spaces over a field $\mathbb{F}$. Then $T \in \mathcal{L}(X, Y)$ is bounded if and only if T maps a bounded set into a bounded set.

Proof. Assume that T is bounded. That is, there exists a constant $M > 0$ such that $\|Tx\| \leq M\|x\|$ for all $x \in X$. If $\|x\| \leq k$ for some constant k, then $\|Tx\| \leq M\|x\| \leq kM$. That is, T maps a bounded set into a bounded set.

Now assume that T maps a bounded set into a bounded set. Then T maps the unit ball $B = \{x \in X \; \|x\| \leq 1\}$ into a bounded set. That is, there exists a constant $M > 0$ such that $\|Tx\| \leq M$ for all $x \in B$. Therefore, for any nonzero $x \in X$

$$\frac{\|Tx\|}{\|x\|} = \left\| T\left(\frac{x}{\|x\|}\right) \right\| \leq M$$

Hence, $\|Tx\| \leq M\|x\|$. That is, T is bounded. ∎

Exercise 11.1 Show that the inverse of a bounded linear operator is not necessarily bounded.

Proposition 11.4.6 Let $T \in \mathcal{B}(X, Y)$. Then T^{-1} exists and is bounded if and only if there is a constant $K > 0$ such that

$$\|Tx\| \geq K\|x\| \text{ for all } x \in X$$

Proof. Assume that there is a constant $K > 0$ such that $\|Tx\| \geq K\|x\|$ for all $x \in X$. If $x \neq 0$, then $Tx \neq 0$ and so T is one-to-one and hence T^{-1} exists. Also, given $y \in \operatorname{ran}(T)$, let $y = Tx$ for some $x \in X$. Then

$$\left\|T^{-1}y\right\| = \left\|T^{-1}(Tx)\right\| = \|x\| \leq \frac{1}{K}\|Tx\| = \frac{1}{K}\|y\|$$

i.e., $\|T^{-1}y\| \leq \frac{1}{K}\|y\|$ for all $y \in Y$. Thus T^{-1} is bounded. Assume that T^{-1} exists and is bounded. Then for each $x \in X$

$$\|x\| = \left\|T^{-1}(Tx)\right\| \leq \left\|T^{-1}\right\| \|Tx\| \iff \frac{1}{\|T^{-1}\|}\|x\| \leq \|Tx\| \iff K\|x\| \leq \|Tx\|$$

where $K = \frac{1}{\|T^{-1}\|}$

∎

Theorem 11.4.7 Let X and Y be normed linear spaces over a field $\mathbb{F}$ and $T \in \mathcal{L}(X,Y)$. The following statements are equivalent:

1. T is continuous on X
2. T is continuous at some point in X
3. T is bounded on X

Proof. $2 \to 1$

$$Ty_n - Ty = T(y_n - y) = T(y_n - y + x_0) - Tx_0$$

The implication $(1) \Rightarrow (2)$ is obvious. $(2) \Rightarrow (3)$: Assume that T is continuous at $x \in X$, but T is not bounded on X. Then there is a sequence (x_n) in X such that $\|Tx_n\| > n\|x_n\|$ for each $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, let

$$y_n = \frac{x_n}{n\|x_n\|} + x$$

Then

$$\|y_n - x\| = \frac{1}{n} \to 0 \text{ as } n \to \infty$$

i.e., $y_n \overset{n\to\infty}{\longrightarrow} x$, but

$$\|Ty_n - Tx\| = \frac{\|Tx_n\|}{n\|x_n\|} > \frac{n\|x_n\|}{n\|x_n\|} = 1$$

That is, $Ty_n \nrightarrow Tx$ as $n \to \infty$, contradicting (2)

$(3) \Rightarrow (1)$: Assume that T is bounded on X. Let (x_n) be a sequence in X which converges to $x \in X$ Then

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \le \|T\|\,\|x_n - x\| \to 0 \text{ as } n \to \infty$$

Thus, T is continuous on X ■

Theorem 11.4.8 Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be normed linear spaces with $\dim(X) < \infty$ and $T : X \to Y$ be a linear operator. Then T is continuous. That is, every linear operator on a finite-dimensional normed linear space is automatically continuous.

Proof. Define a new norm $\|\cdot\|_0$ on X by

$$\|x\|_0 = \|x\| + \|Tx\| \text{ for all } x \in X$$

Since X is finte-dimensional, the norms $\|\cdot\|_0$ and $\|\cdot\|$ on X are equivalent. Hence there are constants α and β such that

$$\alpha\|x\|_0 \le \|x\| \le \beta\|x\|_0 \text{ for all } x \in X$$

Hence,

$$\|Tx\| \leq \|x\|_0 \leq \frac{1}{\alpha}\|x\| = K\|x\|$$

where $K = \frac{1}{\alpha}$. Therefore T is bounded ■

Definition 11.4.6 — uniform convergence and strongly operator convergence. Let X and Y be normed linear spaces over a field $\mathbb{F}$

1. A sequence $(T_n)_1^\infty$ in $\mathcal{B}(X,Y)$ is said to be uniformly operator convergent to T if

$$\lim_{n\to\infty} \|T_n - T\| = 0$$

This is also referred to as **convergence in the uniform topology or convergence in the operator norm topology** of $\mathcal{B}(X,Y)$. In this case T is called the uniform operator limit of the sequence $(T_n)_1^\infty$

2. A sequence $(T_n)_1^\infty$ in $\mathcal{B}(X,Y)$ is said to be strongly operator convergent to T if

$$\lim_{n\to\infty} \|T_n x - Tx\| = 0 \quad \text{for each } x \in X$$

In this case T is called the strong operator limit of the sequence $(T_n)_1^\infty$

A sequence $(T_n)_1^\infty$ in $\mathcal{B}(X,Y)$ is said to be strongly operator convergent to T if

$$\lim_{n\to\infty} \|T_n x - Tx\| = 0 \quad \text{for each } x \in X$$

In this case T is called the **strong operator limit** of the sequence $(T_n)_1^\infty$. Of course, if T is the uniform operator limit of the sequence $(T_n)_1^\infty \subset \mathcal{B}(X,Y)$, then $T \in \mathcal{B}(X,Y)$. On the other hand, the strong operator limit T of a sequence $(T_n)_1^\infty \subset \mathcal{B}(X,Y)$ need not be bounded in general.

Theorem 11.4.9 If the sequence $(T_n)_1^\infty$ in $\mathcal{B}(X,Y)$ is uniformly convergent to $T \in \mathcal{B}(X,Y)$, then it is strongly convergent to T

Proof. Since, for each $x \in X$, $\|T_n x - Tx\| = \|(T_n - T)(x)\| \leq \|T_n - T\| \, \|x\|$, if $\|T_n - T\| \to 0$ as $n \to \infty$, then

$$\|(T_n - T)(x)\| \quad \to 0 \quad \text{as } n \to \infty$$

The converse does not hold. ■

■ **Example 11.26** Consider the sequence (T_n) of operators, where for each $n \in \mathbb{N}$ $T_n : \ell_2 \to \ell_2$ is given by

$$T_n(x_1, x_2, \ldots) = (0, 0, \ldots, 0, x_{n+1}, x_{n+2}, \ldots)$$

Let $\epsilon > 0$ be given. Then for each $x = (x_i)_{i=1}^{\infty} \in \ell_2$, there exists N such that

$$\sum_{n+1}^{\infty} |x_i|^2 < \epsilon^2, \qquad \text{for all } n \geq N$$

Hence, for all $n \geq N$

$$\|T_n x\|_2^2 = \sum_{n+1}^{\infty} |x_i|^2 < \epsilon^2$$

That is, for each $x \in \ell_2, T_n x \to 0$. Hence, $T_n \to 0$ strongly. Now, since

$$\|T_n x\|_2^2 = \sum_{n+1}^{\infty} |x_i|^2 \leq \sum_{1}^{\infty} |x_i|^2 = \|x\|_2^2$$

for $n \in \mathbb{N}$ and $x = (x_i)_{i=1}^{\infty} \in \ell_2$, it follows that $\|T_n\| \leq 1$ for each $n \in \mathbb{N}$. But $\|T_n\| \geq 1$ for all n. To see this, take $x = (0, 0, \ldots, 0, x_{n+1}, 0, \ldots) \in \ell_2$, where $x_{n+1} \neq 0$ Then $T_n x = x$ and hence $\|T_n x\|_2 = |x_{n+1}|$, and consequently, $\|T_n\| \geq 1$. That is, (T_n) does not converge to zero in the uniform topology.

Theorem 11.4.10 Let X and Y be normed linear spaces over a field $\mathbb{F}$. Then $\mathcal{B}(X, Y)$ is a Banach space if Y is a Banach space.

Proof. We have shown that $\mathcal{B}(X, Y)$ is a normed linear space. It remains to show that it is complete if Y is complete. To that end, let $(T_n)_1^{\infty}$ be a Cauchy sequence in $\mathcal{B}(X, Y)$. Then given any $\epsilon > 0$ there exists a positive integer N such that

$$\|T_n - T_{n+r}\| < \epsilon \qquad \text{for all } n > N$$

whence

$$\|T_n x - T_{n+r} x\| \leq \|T_n - T_{n+r}\| \, \|x\| < \epsilon \|x\| \text{for all} x \in X \tag{11.9}$$

Hence, $(T_n x)_1^{\infty}$ is a Cauchy sequence in Y. Since Y is complete there exists $y \in Y$ such that $T_n x \to y$ as $n \to \infty$. Set $Tx = y$. We show that $T \in \mathcal{B}(X, Y)$ and $T_n \to T$. Let $x_1, x_2 \in X$, and $\alpha, \beta \in \mathbb{F}$. Then

$$\begin{aligned} T(\alpha x_1 + \beta x_2) &= \lim_{n\to\infty} T_n(\alpha x_1 + \beta x_2) = \lim_{n\to\infty} [\alpha T_n x_1 + \beta T_n x_2] \\ &= \alpha \lim_{n\to\infty} T_n x_1 + \beta \lim_{n\to\infty} T_n x_2 = \alpha T x_1 + \beta T x_2 \end{aligned}$$

That is, $T \in \mathcal{L}(X, Y)$. Taking the limit as $r \to \infty$ in 11.9 we get that $\|(T_n - T)x\| = \|T_n x - Tx\| \leq \epsilon \|x\|$ for all $n > N$, and all $x \in X$.

That is, $T_n - T$ is a bounded operator for all $n > N$. Since $\mathcal{B}(X, Y)$ is a linear space

$$T = T_n - (T_n - T) \in \mathcal{B}(X, Y)$$

Finally

$$\|T_n - T\| = \sup\{\|T_n x - Tx\| : \|x\| \leq 1\} \leq \sup\{\|x\|\epsilon : \|x\| \leq 1\} \leq \epsilon \text{ for all } n > N$$

That is, $T_n \to T$ as $n \to \infty$.

■

Definition 11.4.7 — Composition. Let $T : X \to Y$ and $S : Y \to Z$. We define the **composition** of T and S as the map $ST : X \to Z$ defined by

$$(ST)(x) = (S \circ T)(x) = S(Tx)$$

Theorem 11.4.11 Let X, Y and Z be normed linear spaces over a field $\mathbb{F}$ and let $T \in \mathcal{B}(X,Y)$ and $S \in \mathcal{B}(Y,Z)$. Then $ST \in \mathcal{B}(X,Z)$ and $\|ST\| \leq \|S\|\|T\|$

Proof. Since linearity is trivial, we only prove boundedness of ST. Let $x \in X$. Then

$$\|(ST)(x)\| = \|S(Tx)\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|$$

Thus, $\|ST\| \leq \|S\|\|T\|$

■

Definition 11.4.8 — Banach algebra. Let X be a normed linear space over $\mathbb{F}$. For $S, T_1, T_2 \in \mathcal{B}(X)$ it is easy to show that

$$\begin{aligned} (ST_1)T_2 &= S(T_1T_2) \\ S(T_1 + T_2) &= ST_1 + ST_2 \\ (T_1 + T_2)S &= T_1S + T_2S \end{aligned}$$

The operator I defined by $Ix = x$ for all $x \in X$ belongs to $\mathcal{B}(X), \|I\| = 1$, and it has the property that $IT = TI = T$ for all $T \in \mathcal{B}(X)$. We call I the **identity operator**. The set $\mathcal{B}(X)$ is therefore an algebra with an identity element. In fact, $\mathcal{B}(X)$ is a normed algebra with an identity element. If X is a Banach space then $\mathcal{B}(X)$ is a **Banach algebra**.

Definition 11.4.9 — Linear functional. Let X be a linear space over $\mathbb{F}$. A linear operator $f : X \to \mathbb{F}$ is called a **linear functional** on X. Of course, $\mathcal{L}(X, \mathbb{F})$ denotes the set of all linear functionals on X.

Since every linear functional is a linear operator, all of the foregoing discussion on linear operators applies equally well to linear functionals. For example, if X is a normed linear space then we say that $f \in \mathcal{L}(X, \mathbb{F})$ is bounded if there exists a constant $M > 0$ such that $|f(x)| \leq M\|x\|$ for all $x \in X$.

The norm of f is defined by

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

We shall denote by $X^* = \mathcal{B}(X, \mathbb{F})$ the set of all bounded (i.e., continuous) linear functionals on X. We call X^* the **dual** of X. It follows from Theorem 11.4.10 that X^* is always a Banach space under the above norm.

■ **Example 11.27** Let $X = \mathcal{C}[a,b]$. For each $x \in X$, define $f : X \to \mathbb{R}$ by

$$f(x) = \int_a^b x(t)dt$$

Then f is a bounded linear functional on X. Linearity: For any $x, y \in X$ and any $\alpha, \beta \in \mathbb{R}$

$$f(\alpha x + \beta y) = \int_a^b (\alpha x + \beta y)(t)dt = \alpha \int_a^b x(t)dt + \beta \int_a^b y(t)dt = \alpha f(x) + \beta f(y)$$

Boundedness:

$$|f(x)| = \left| \int_a^b x(t)dt \right| \le \int_a^b |x(t)|dt \le \max_{a \le t \le b} |x(t)|(b-a) = \|x\|_\infty (b-a)$$

Hence f is bounded and $\|f\| \le b-a$. We show that $\|f\| = b-a$. Take $x = 1$, the constant function 1. Then

$$f(1) = \int_a^b dt = b-a, \quad \text{i.e., } |f(1)| = b-a$$

Hence

$$b-a = \frac{|f(1)|}{1} \le \sup\left\{ \frac{|f(x)|}{\|x\|} : x \ne 0 \right\} = \|f\| \le b-a$$

That is, $\|f\| = b-a$

11.4.1 Examples of Dual Spaces

Definition 11.4.10 — isomorphic. Let X and Y be normed linear spaces over the same field $\mathbb{F}$. Then X and Y are said to be **isomorphic** to each other, denoted by $X \simeq Y$, if there is a bijective linear operator T from X onto Y. If, in addition, T is an **isometry**, i.e., $\|Tx\| = \|x\|$ for each $x \in X$, then we say that T is an **isometric isomorphism**. In this case, X and Y are are said to be isometrically isomorphic and we write $X \cong Y$

Two normed linear spaces which are isometrically isomorphic can be regarded as identical, the isometry merely amounting to a relabelling of the elements.

Theorem 11.4.12 Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and T a linear operator from X onto Y. Then T is an isometry if and only if

1. T is one-to-one
2. T is continuous on X
3. T has a continuous inverse (in fact, $\|T^{-1}\| = \|T\| = 1$);
4. T is distance-preserving: For all $x, y \in X, \|Tx - Ty\| = \|x - y\|$

Proof. If T satisfies (iv), then, taking $y=0$, we have that $\|Tx\|=\|x\|$ for each $x\in X$; i.e., T is an isometry. Conversely, assume that T is an isometry. If $x\neq y$, then

$$\|Tx-Ty\|=\|T(x-y)\|=\|x-y\|>0$$

Hence, $Tx\neq Ty$. This shows that T is one-to-one and distance-preserving. Since $\|Tx\|=\|x\|$ for each $x\in X$, it follows that T is bounded and $\|T\|=1$. By Theorem 11.4.7 , T is continuous on X. Let $y_1,y_2\in Y$ and $\alpha_1,\alpha_2\in\mathbb{F}$. Then there exist $x_1,x_2\in X$ such that $Tx_i=y_i$ for $i=1,2$. Therefore

$$\alpha_1y_1+\alpha_2y_2=\alpha_1Tx_1+\alpha_2Tx_2=T\left(\alpha_1x_1+\alpha_2x_2\right) \text{ or}$$
$$T^{-1}\left(\alpha_1y_1+\alpha_2y_2\right)=\alpha_1x_1+\alpha_2x_2=\alpha_1T^{-1}y_1+\alpha_2T^{-1}y_2$$

That is, T^{-1} is linear. Furthermore, for $y\in Y$, let $x=T^{-1}y$. Then

$$\left\|T^{-1}y\right\|=\|x\|=\|Tx\|=\|y\|$$

Therefore T^{-1} is bounded and $\left\|T^{-1}\right\|=1$ ■

■ **Example 11.28** The dual of ℓ_1 is (isometrically isomorphic to) ℓ_∞; i.e., $\ell_1^*\cong\ell_\infty$

Proof. Let $y=(y_n)\in\ell_\infty$ and define $\Phi:\ell_\infty\to\ell_1^*$ by

$$(\Phi y)(x)=\sum_{j=1}^{\infty}x_jy_j \text{ for } x=(x_n)\in\ell_1$$

Claim 1 : $\Phi y\in\ell_1^*$ Linearity of Φy : Let $x=(x_n),z=(z_n)\in\ell_1$ and $\alpha\in\mathbb{F}$. Then

$$\begin{aligned}(\Phi y)(\alpha x+z)&=\sum_{j=1}^{\infty}\left(\alpha x_j+z_j\right)y_j=\sum_{j=1}^{\infty}\alpha x_jy_j+\sum_{j=1}^{\infty}z_jy_j\\&=\alpha\sum_{j=1}^{\infty}x_jy_j+\sum_{j=1}^{\infty}z_jy_j\\&=\alpha(\Phi y)(x)+(\Phi y)(z)\end{aligned}$$

Boundedness of Φy : For any $x=(x_n)\in\ell_1$

$$|(\Phi y)(x)|=\left|\sum_{j=1}^{\infty}x_jy_j\right|\leq\sum_{j=1}^{\infty}\left|x_jy_j\right|\leq\|y\|_\infty\sum_{j=1}^{\infty}\left|x_j\right|=\|y\|_\infty\|x\|_1$$

That is, $\Phi y\in\ell_1^*$ and

$$\|\Phi y\|\leq\|y\|_\infty$$

Claim 2 : Φ is a surjective linear isometry.

(i) Φ is a surjective: A basis for ℓ_1 is (e_n), where $e_n=(\delta_{nm})$ has 1 in the n -th position and zeroes elsewhere. Let $f\in\ell_1^*$ and $x=(x_n)\in\ell_1$. Then $x=\sum_{n=1}^{\infty}x_ne_n$ and therefore

$$f(x)=\sum_{n=1}^{\infty}x_nf\left(e_n\right)=\sum_{n=1}^{\infty}x_nz_n$$

where, for each $n \in \mathbb{N}, z_n = f(e_n)$. We show that $z = (z_n) \in \ell_\infty$. Indeed, for each $n \in \mathbb{N}$

$$|z_n| = |f(e_n)| \leq \|f\| \|e_n\| = \|f\|$$

Hence, $z = (z_n) \in \ell_\infty$. Also, for any $x = (x_n) \in \ell_1$

$$(\Phi z)(x) = \sum_{n=1}^{\infty} x_n z_n = \sum_{n=1}^{\infty} x_n f(e_n) = f(x)$$

That is, $\Phi z = f$ and so Φ is surjective. Furthermore,

$$\|z\|_\infty = \sup_{n \in \mathbb{N}} |z_n| = \sup_{n \in \mathbb{N}} |f(e_n)| \leq \|f\| = \|\Phi z\|$$

(ii) Φ is linear: Let $y = (y_n), z = (z_n) \in \ell_\infty$ and $\beta \in \mathbb{F}$. Then, for any $x = (x_n) \in \ell_1$

$$\begin{aligned}[\Phi(\beta y + z)](x) &= \sum_{j=1}^{\infty} x_j (\beta y_j + z_j) = \beta \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} x_j z_j \\ &= \beta(\Phi y)(x) + (\Phi z)(x) = [\beta \Phi y + \Phi z](x)\end{aligned}$$

Hence, $\Phi(\beta y + z) = \beta \Phi y + \Phi$, which proves linearity of Φ

(iii) Φ is an isometry: This follows from Claim 1 and Claim 2(i). ■

■ **Example 11.29** The dual of c_0 is (isometrically isomorphic to) ℓ_1, i.e., $c_0^* \cong \ell_1$

Proof. Let $y = (y_n) \in \ell_1$ and define $\Phi : \ell_1 \to c_0^*$ by

$$(\Phi y)(x) = \sum_{j=1}^{\infty} x_j y_j \text{ for } x = (x_n) \in c_0$$

Proceeding as in Example 1 above, one shows that Φy is a bounded linear functional on c_0 and

$$\|\Phi y\| \leq \|y\|_1$$

Claim: Φ is a surjective linear isometry. (i) Φ is a surjective: A basis for c_0 is (e_n), where $e_n = (\delta_{nm})$ has 1 in the n -th position and zeroes elsewhere. Let $f \in c_0^*$ and $x = (x_n) \in c_0$. Then $x = \sum_{n=1}^{\infty} x_n e_n$ and therefore

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n) = \sum_{n=1}^{\infty} x_n w_n$$

where, for each $n \in \mathbb{N}, w_n = f(e_n)$. For $n, k \in \mathbb{N}$, let

$$z_{nk} = \begin{cases} \frac{|w_k|}{w_k} & \text{if } w_k \neq 0 \text{ and } k \leq n \\ 0 & \text{if } w_k = 0 \text{ or } k > n \end{cases}$$

and let

$$z_n = (z_{n1}, z_{n2}, \ldots, z_{nn}, 0, 0, \ldots)$$

Then $z_n \in c_0$ and

$$\|z_n\|_\infty = \sup_{k \in \mathbb{N}} |z_{nk}| = 1$$

Also,

$$f(z_n) = \sum_{k=1}^{\infty} z_{nk} w_k = \sum_{k=1}^{n} |w_k|$$

Hence, for each $n \in \mathbb{N}$

$$\sum_{k=1}^{n} |w_k| = |f(z_n)| \le \|f\| \, \|z_n\| \le \|f\|$$

since the right hand side is independent of n, it follows that $\sum_{k=1}^{\infty} |w_k| \le \|f\|$ Hence, $w = (w_n) \in \ell_1$. Also, for any $x = (x_n) \in c_0$

$$(\Phi w)(x) = \sum_{n=1}^{\infty} x_n w_n = \sum_{n=1}^{\infty} x_n f(e_n) = f(x)$$

That is, $\Phi w = f$ and so Φ is surjective. Furthermore,

$$\|w\|_1 = \sum_{k=1}^{\infty} |w_k| \le \|f\| = \|\Phi w\|$$

(ii) Φ is linear: Let $y = (y_n), z = (z_n) \in \ell_1$ and $\beta \in \mathbb{F}$. Then, for any $x = (x_n) \in c_0$

$$\begin{aligned}[\Phi(\beta y + z)](x) &= \sum_{j=1}^{\infty} x_j (\beta y_j + z_j) = \beta \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} x_j z_j \\ &= \beta(\Phi y)(x) + (\Phi z)(x) = [\beta \Phi y + \Phi z](x)\end{aligned}$$

Hence, $\Phi(\beta y + z) = \beta \Phi y + \Phi z$, which proves linearity of Φ (iii) Φ is an isometry.

■

■ **Example 11.30** Let $1 < p < \infty, \frac{1}{p} + \frac{1}{q} = 1$. Then the dual of ℓ_p is (isometrically isomorphic to) ℓ_q, i.e., $\ell_p^* \cong \ell_q$

Proof. Let $y = (y_n) \in \ell_q$ and define $\Phi : \ell_q \to \ell_p^*$ by

$$(\Phi y)(x) = \sum_{j=1}^{\infty} x_j y_j \text{ for } x = (x_n) \in \ell_p$$

It is straightforward to show that Φy is linear. We show that Φy is bounded. By Hölder's Inequality,

$$|(\Phi y)(x)| = \left| \sum_{j=1}^{\infty} x_j y_j \right| \le \sum_{j=1}^{\infty} |x_j y_j| \le \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^{\infty} |y_j|^q \right)^{\frac{1}{q}} = \|x\|_p \|y\|_q$$

That is, $\Phi y \in \ell_p^*$ and

$$\|\Phi y\| \leq \|y\|_q$$

Claim: Φ is a surjective linear isometry. (i) Φ is a surjective: A basis for ℓ_p is (e_n), where $e_n = (\delta_{nm})$ has 1 in the n -th position and zeroes elsewhere. Let $f \in \ell_p^*$ and $x = (x_n) \in \ell_p$. Then $x = \sum_{n=1}^{\infty} x_n e_n$ and therefore

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n) = \sum_{n=1}^{\infty} x_n w_n$$

where, for each $n \in \mathbb{N}, w_n = f(e_n)$. For $n, k \in \mathbb{N}$, let

$$z_{nk} = \begin{cases} \frac{|w_k|^q}{w_k} & \text{if } k \leq n \text{ and } w_k \neq 0 \\ 0 & \text{if } w_k = 0 \text{ or } k > n \end{cases}$$

and let

$$z_n = (z_{n1}, z_{n2}, \ldots, z_{nn}, 0, 0, \ldots)$$

Then $z_n \in \ell_p$ and

$$f(z_n) = \sum_{k=1}^{\infty} z_{nk} w_k = \sum_{k=1}^{n} |w_k|^q$$

Hence, for each $n \in \mathbb{N}$

$$\sum_{k=1}^{n} |w_k|^q = |f(z_n)| \leq \|f\| \|z_n\|_p$$

Since

$$\begin{aligned} \|z_n\|_p &= \left(\sum_{k=1}^{\infty} |z_{nk}|^p \right)^{1/p} = \left(\sum_{k=1}^{n} |z_{nk}|^p \right)^{1/p} \\ &= \left(\sum_{k=1}^{n} |w_k|^{p(q-1)} \right)^{1/p} = \left(\sum_{k=1}^{n} |w_k|^q \right)^{1/p} \end{aligned}$$

it follows that, for each $n \in \mathbb{N}$

$$\begin{aligned} \sum_{k=1}^{n} |w_k|^q \leq \|f\| \|z_n\|_p \quad &\Longleftrightarrow \sum_{k=1}^{n} |w_k|^q \leq \|f\| \left(\sum_{k=1}^{n} |w_k|^q \right)^{1/p} \\ &\Longleftrightarrow \left(\sum_{k=1}^{n} |w_k|^q \right)^{1-1/p} \leq \|f\| \end{aligned}$$

$$\Longleftrightarrow \left(\sum_{k=1}^{n} |w_k|^q \right)^{1/q} \leq \|f\|$$

since the right hand side is independent of n, it follows that $\left(\sum_{k=1}^{\infty}|w_k|^q\right)^{1/q} \le \|f\|$, and so $w = (w_n) \in \ell_q$. Also, for any $x = (x_n) \in \ell_p$

$$(\Phi w)(x) = \sum_{n=1}^{\infty} x_n w_n = \sum_{n=1}^{\infty} x_n f(e_n) = f(x)$$

That is, $\Phi w = f$ and so Φ is surjective. Furthermore,

$$\|w\|_q = \left(\sum_{k=1}^{\infty}|w_k|^q\right)^{1/q} \le \|f\| = \|\Phi w\|$$

(ii) Φ is linear: Let $y = (y_n), z = (z_n) \in \ell_q$ and $\beta \in \mathbb{F}$. Then, for any $x = (x_n) \in \ell_p$

$$\begin{aligned}[\Phi(\beta y + z)](x) &= \sum_{j=1}^{\infty} x_j (\beta y_j + z_j) = \beta \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} x_j z_j \\ &= \beta(\Phi y)(x) + (\Phi z)(x) = [\beta \Phi y + \Phi z](x)\end{aligned}$$

Hence, $\Phi(\beta y + z) = \beta\Phi y + \Phi z$, which proves linearity of Φ

(iii) Φ is an isometry. ■

Theorem 11.4.13 Every linear functional on a finite-dimensional normed linear space is continuous.

Proposition 11.4.14 Let X be a normed linear space over $\mathbb{F}$. If X is finite-dimensional, then X^* is also finite-dimensional and $\dim X = \dim X^*$

Proof. Let $\{x_1, x_2, \cdots, x_n\}$ be a basis for X. For each $j = 1,2,\ldots,n$, let x_j^* be defined by $x_j^*(x_k) = \delta_{jk}$ for $k = 1,2,\ldots,n$. Then each x_j^* is a bounded linear functional on X. We show that $\left\{x_j^* | j = 1,2,\ldots,n\right\}$ is a basis for X^*. Let x^* be an element of X^* and define $\lambda_j = x^*(x_j)$ for each $j = 1,2,\ldots,n$. Then for any $k = 1,2,\ldots,n$

$$\left(\sum_{j=1}^{n} \lambda_j x_j^*\right)(x_k) = \sum_{j=1}^{n} \lambda_j \delta_{jk} = \lambda_k = x^*(x_k)$$

Hence $x^* = \sum_{j=1}^{n} \lambda_j x_j^*$; i.e., $\left\{x_j^* | j = 1,2,\ldots,n\right\}$ spans X^*. It remains to show that $\left\{x_j^* | j = 1,2,\ldots,n\right\}$ is linearly independent. Suppose that $\sum_{j=1}^{n} \alpha_j x_j^* = 0$. Then, for each $k = 1,2,\ldots,n$

$$0 = \left(\sum_{j=1}^{n} \alpha_j x_j^*\right)(x_k) = \sum_{j=1}^{n} \alpha_j \delta_{jk} = \alpha_k$$

Hence $\left\{x_j^* | j = 1,2,\ldots,n\right\}$ is a linearly independent set.

■

11.4.2 The Dual Space of a Hilbert Space

Theorem 11.4.15 Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Choose and fix $y \in X \backslash \{0\}$. Define a map $f_y : X \to \mathbb{F}$ by $f_y(x) = \langle x, y\rangle$. We claim that f_y is a bounded (= continuous) linear functional on X

Linearity: Let $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{F}$. Then

$$f_y(\alpha x_1 + \beta x_2) = \langle \alpha x_1 + \beta x_2, y\rangle = \alpha \langle x_1, y\rangle + \beta \langle x_2, y\rangle = \alpha f_y(x_1) + \beta f_y(x_2)$$

Boundedness: For any $x \in X$ $|f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$ (by the CBS Inequality) That is, f_y is bounded and $\|f_y\| \leq \|y\|$. since

$$f_y(y) = \langle y, y\rangle = \|y\|^2 \Rightarrow \frac{|f_y(y)|}{\|y\|} = \|y\|$$

we have that $\|f_y\| = \|y\|$.

The above observation simply says that each element y in an inner product space $(X, \langle\cdot,\cdot\rangle)$ determines a bounded linear functional on X

The following theorem asserts that if $\mathcal{H}$ is a Hilbert space then the converse of this statement is true.

Theorem 11.4.16 — Riesz-Fréchet Theorem. Let $\mathcal{H}$ be a Hilbert space over $\mathbb{F}$. If $f : \mathcal{H} \to \mathbb{F}$ is a bounded linear functional on $\mathcal{H}$ (i.e., $f \in \mathcal{H}^*$) then there exists one and only one $y \in \mathcal{H}$ such that

$$f(x) = \langle x, y\rangle \quad \text{for all } x \in \mathcal{H}$$

Moreover, $\|f\| = \|y\|$.

Proof. Existence: If $f = 0$ then take $y = 0$. Assume that $f \neq 0$. Let $N = \{x \in \mathcal{H} | f(x) = 0\}$, the kernel of f. Then N is a closed proper subspace of $\mathcal{H}$. By Corollary 3.4 .5 there exists $z \in N^{\perp} \backslash \{0\}$ Without loss of generality, $\|z\| = 1$. Put $u = f(x)z - f(z)x$. Then

$$f(u) = f(f(x)z - f(z)x) = f(x)f(z) - f(z)f(x) = 0, \quad \text{i.e., } u \in N$$

Thus,

$$0 = \langle u, z\rangle = \langle f(x)z - f(z)x, z\rangle = f(x)\langle z, z\rangle - f(z)\langle x, z\rangle = f(x) - f(z)\langle x, z\rangle$$

whence, $f(x) = f(z)\langle x, z\rangle = \langle x, \overline{f(z)}z\rangle$. Take $y = \overline{f(z)}z$. Then $f(x) = \langle x, y\rangle$.

Uniqueness: Assume that $f(x) = \langle x, y\rangle = \langle x, y_0\rangle$ for each $x \in \mathcal{H}$. Then

$$0 = \langle x, y\rangle - \langle x, y_0\rangle = \langle x, y - y_0\rangle, \quad \text{for all} \quad x \in \mathcal{H}$$

In particular, take $x = y - y_0$

$$0 = \langle y - y_0, y - y_0\rangle = \|y - y_0\|^2 \Rightarrow y - y_0 = 0 \Rightarrow y = y_0$$

Finally, for any $x \in \mathcal{H}$

$$|f(x)| = |\langle x, y \rangle| \leq \|x\| \|y\| \quad \text{(by the CBS Inequality)}$$

That is, $\|f\| \leq \|y\|$. Since

$$f(y) = \langle y, y \rangle = \|y\|^2 \Rightarrow \quad \frac{|f(y)|}{\|y\|} = \|y\|$$

we have that $\|f\| = \|y\|$

■

R

1. Every bounded linear functional on a Hilbert space $\mathcal{H}$ is, in fact, determined by some element $y \in \mathcal{H}$.
2. The element $y \in \mathcal{H}$ as advertised in Theorem 11.4.16 is called the representer of the functional f
3. The conclusion of Theorem 11.4.16 may fail if $(X, \langle \cdot, \cdot \rangle)$ is an incomplete inner product space.

■ **Example 11.31** Let X be the linear space of polynomials over $\mathbb{R}$ with the inner product defined by

$$\langle x, y \rangle = \int_0^1 x(t) y(t) dt$$

For each $x \in X$, let $f : X \to \mathbb{R}$ be defined by $f(x) = x(0)$, (i.e., f is a point evaluation at 0) Then f is a bounded linear functional on X. We show that there does not exist an element $y \in X$ such that

$$f(x) = \langle x, y \rangle \quad \text{for all } x \in X$$

Assume that such an element exists. Then for each $x \in X$

$$f(x) = x(0) = \int_0^1 x(t) y(t) dt$$

Since for any $x \in X$ the functional f maps the polynomial $tx(t)$ onto zero, we have that

$$\int_0^1 tx(t) y(t) dt = 0 \quad \text{for all } x \in X$$

In particular, with $x(t) = ty(t)$ we have that

$$\int_0^1 t^2 [y(t)]^2 dt = 0$$

whence $y \equiv 0$, i.e. y is the zero polynomial. Hence, for all $x \in X$

$$f(x) = \langle x, y \rangle = \langle x, 0 \rangle = 0$$

That is, f is the zero functional, a contradiction since f maps a polynomial with a nonzero constant term to that constant term.

Theorem 11.4.17 Let $\mathcal{H}$ be a Hilbert space.

1. If $\mathcal{H}$ is a real Hilbert space, then $\mathcal{H} \cong \mathcal{H}^*$
2. If $\mathcal{H}$ is a complex Hilbert space, then $\mathcal{H}$ is isometrically embedded onto $\mathcal{H}^*$

Proof. For each $y \in \mathcal{H}$, define $\Lambda : \mathcal{H} \to \mathcal{H}^*$ by $\Lambda y = f_y$, where $f_y(x) = \langle x,y \rangle$ for each $x \in \mathcal{H}$ Let $y,z \in \mathcal{H}$. Then, for each $x \in \mathcal{H}$

$$y \neq z \iff \langle x,y \rangle \neq \langle x,z \rangle \iff f_y \neq f_z \iff \Lambda y \neq \Lambda z$$

Hence, Λ is well defined and one-to-one. Furthermore, since

$$\|\Lambda y\| = \|f_y\| = \|y\|$$

for each $y \in \mathcal{H}, \Lambda$ is an isometry. If $f \in \mathcal{H}^*$, then by Riesz-Fréchet Theorem (Theorem 4.3 .1), there is a unique $y_f \in \mathcal{H}$ such that $f(x) = \langle x, y_f \rangle$. Hence $\Lambda y_f = f$, i.e., Λ is onto. The inverse Λ^{-1} of Λ is given by $\Lambda^{-1} f = y$, where $f(x) = \langle x,y \rangle$ for all $x \in \mathcal{H}$ since $\|\Lambda^{-1} f\| = \|y\| = \|f\|$ for each $f \in \mathcal{H}^*, \Lambda^{-1}$ is bounded (in fact an isometry). If $\mathcal{H}$ is real, then Λ is linear. Indeed, for all $x,y,z \in \mathcal{H}$ and all $\alpha \in \mathbb{R}$, then

$$\begin{aligned}
(\Lambda(\alpha y + z))(x) = f_{(\alpha y+z)}(x) &= \langle x, \alpha y + z \rangle \\
&= \langle x, \alpha y \rangle + \langle x,z \rangle = \alpha \langle x,y \rangle + \langle x,z \rangle \\
&= (\alpha \Lambda y)(x) + (\Lambda z)(x) = (\alpha \Lambda y + \Lambda z)(x)
\end{aligned}$$

Hence, $\Lambda(\alpha y + z) = \alpha \Lambda y + \Lambda z$ If $\mathcal{H}$ is complex, then Λ is conjugate-linear; i.e., $\Lambda(\alpha y + z) = \bar{\alpha} \Lambda y + \Lambda z$.

■

11.5 The Hahn-Banach Theorem and its Consequences

Definition 11.5.1 — partial order. A binary relation $\preccurlyeq$ on a set P is a **partial order** if it satisfies the following properties: For all $x,y,z \in P$

1. $\preccurlyeq$ is reflexive: $x \preccurlyeq x$
2. $\preccurlyeq$ is antisymmetric: if $x \preccurlyeq y$ and $y \preccurlyeq x$, then $x = y$
3. $\preccurlyeq$ is transitive: if $x \preccurlyeq y$ and $y \preccurlyeq z$, then $x \preccurlyeq z$

A partially ordered set is a pair $(P, \preccurlyeq)$, where P is a set $\preccurlyeq$ is a partial order on P

■ **Example 11.32**

1. Let $P = \mathbb{R}$ and take $\preccurlyeq$ to be $\leq$, the usual less than or equal to relation on $\mathbb{R}$.
2. Let $P = \mathcal{P}(X)$ the power set of a set X and take $\preccurlyeq$ to be $\subseteq$, the usual set inclusion relation.
3. Let $P = \mathcal{C}[0,1]$, the space of continuous real-valued functions on the interval [0,1] and take $\preccurlyeq$ to be the relation $\leq$ given by $f \leq g$ if and only if $f(x) \leq g(x)$ for each $x \in [0,1]$

Definition 11.5.2 Let C be a subset of a partially ordered set $(P, \preccurlyeq)$

1. An element $u \in P$ is an upper bound of C if $x \preccurlyeq u$ for every $x \in C$
2. An element $m \in C$ is said to be maximal if for any element $y \in C$, the relation $m \preccurlyeq y$ implies that $m = y$

Definition 11.5.3 — linear order (or a total order). Let $(P, \preccurlyeq)$ be a partially ordered set and $x, y \in P$. We say that x and y are **comparable** if either $x \preccurlyeq y$ or $y \preccurlyeq x$. Otherwise, x and y are incomparable. A partial order $\preccurlyeq$ is called a **linear order (or a total order)** if any two elements of P are comparable. In this case we say that $(P, \preccurlyeq)$ is a linearly ordered (or totally ordered) set. A linearly ordered set is also called a **chain**.

Theorem 11.5.1 — Zorn's Lemma. Let $(P, \preccurlyeq)$ be a partially ordered set. If each linearly ordered subset of P has an upper bound, then P has a maximal element.

Definition 11.5.4 — extension. Let M and N be linear subspaces of a linear space X with $M \subset N$ and let f be a linear functional on M. A linear functional F on N is called an **extension** of f to N if $F|_M = f$; i.e., $F(x) = f(x)$ for each $x \in M$

Definition 11.5.5 — sublinear functional. Let X be a linear space. A function $p : X \to \mathbb{R}$ is called a **sublinear functional** provided that:

1. $p(x+y) \leq p(x) + p(y) \quad$ for $x, y \in X$
2. $p(\lambda x) = \lambda p(x), \quad \lambda \geq 0$

Observe that any linear functional or any norm on X is a sublinear functional. Also, every positive scalar multiple of a sublinear functional is again a sublinear functional

Theorem 11.5.2 Let M be a proper linear subspace of a real linear space $X, x_0 \in X \backslash M$, and $N = \{m + \alpha x_0 | m \in M, \alpha \in \mathbb{R}\}$. Suppose that $p : X \to \mathbb{R}$ a sublinear functional defined on X, and f a linear functional defined on M such that $f(x) \leq p(x)$ for all $x \in M$. Then f can be extended to a linear functional F defined on N such that $F(x) \leq p(x)$ for all $x \in N$.

Proof. Since $x_0 \notin M$, it is readily verified that $N = M \oplus \operatorname{lin}\{x_0\}$. Therefore each $x \in N$ has a unique representation of the form $x = m + \lambda x_0$ for some unique $m \in M$ and $\lambda \in \mathbb{R}$. Define a functional F on N by

$$F(x) = f(m) + \lambda c \text{ for some } c \in \mathbb{R}$$

This functional F is well defined since each $x \in N$ is uniquely determined. Furthermore F is linear and $F(y) = f(y)$ for all $y \in M$. It remains to show that it is possible to choose a $c \in \mathbb{R}$ such that for each $x \in N$

$$F(x) \leq p(x)$$

Let $y_1, y_2 \in M$. Since $f(y) \leq p(y)$ for all $y \in M$, we have that

$$\begin{aligned} f(y_1) - f(y_2) = f(y_1 - y_2) \leq p(y_1 - y_2) &= p(y_1 + x_0 - y_2 - x_0) \\ &\leq p(y_1 + x_0) + p(-y_2 - x_0) \\ \iff -f(y_2) - p(-y_2 - x_0) \leq p(y_1 + x_0) - f(y_1) & \end{aligned}$$

Therefore, for fixed $y_1 \in M$, the set of real numbers $\{-f(y_2) - p(-y_2 - x_0) \,|\, y_2 \in M\}$ is bounded above and hence has the least upper bound. Let

$$a = \sup\{-f(y_2) - p(-y_2 - x_0) \,|\, y_2 \in M\}$$

Similarly, for fixed $y_2 \in M$, the set $\{p(y_1 + x_0) - f(y_1) \,|\, y_1 \in M\}$ is bounded below. Let

$$b = \inf\{p(y_1 + x_0) - f(y_1) \,|\, y_1 \in M\}$$

Of course, $a \leq b$. Hence there is a real number c such that $a \leq c \leq b$. Therefore

$$-f(y) - p(-y - x_0) \leq c \leq p(y + x_0) - f(y)$$

for each $y \in M$. Now, let $x = y + \lambda x_0 \in N$. If $\lambda = 0$, then $F(x) = f(x) \leq p(x)$. If $\lambda > 0$, then

$$\begin{aligned} c \leq p\left(\frac{y}{\lambda} + x_0\right) - f(y/\lambda) &\iff \lambda c \leq p(y + \lambda x_0) - f(y) \\ &\iff f(y) + \lambda c \leq p(y + \lambda x_0) \\ &\iff F(x) \leq p(x) \end{aligned}$$

Finally, if $\lambda < 0$, then

$$\begin{aligned} -f(y/\lambda) - p(-y/\lambda - x_0) \leq c &\iff -\frac{1}{\lambda} f(y) + \frac{1}{\lambda} p(y + \lambda x_0) \leq c \\ &\iff f(y) - p(y + \lambda x_0) \leq -\lambda c \\ &\iff f(y) + \lambda c \leq p(y + \lambda x_0) \\ &\iff F(x) \leq p(x) \end{aligned}$$

We now state our main result. What this theorem essentially states is that there are enough bounded (continuous) linear functionals for a rich theory and as mentioned before it is used ubiquitously thoughout functional analysis.

■

Theorem 11.5.3 — Hahn-Banach Extension Theorem for real linear spaces. Let p be a sublinear functional on a real linear space X and let M be a subspace of X. If f is a linear functional on M such that $f(x) \leq p(x)$ for all $x \in M$, then f has an extension F to X such that $F(x) \leq p(x)$ for all $x \in X$

Proof. Let $\mathcal{F}$ be the set of all pairs (M_α, f_α), where M_α is a subspace of X containing $M, f_\alpha(y) = f(y)$ for each $y \in M$, i.e., f_α is an extension of f, and $f_\alpha(x) \le p(x)$ for all $x \in M_\alpha$. Clearly, $\mathcal{F} \neq \emptyset$ since $(M, f) \in \mathcal{F}$. Define a partial order on $\mathcal{F}$ by:

$$(M_\alpha, f_\alpha) \prec (M_\beta, f_\beta) \Longleftrightarrow M_\alpha \subset M_\beta \text{ and } f_\beta\big|_{M_\alpha} = f_\alpha$$

Let $\mathcal{T}$ be a totally ordered subset of $\mathcal{F}$ and let

$$X_0 = \bigcup \{M_\alpha | (M_\alpha, f_\alpha) \in \mathcal{T}\}$$

Then X_0 is a linear subspace of X since $\mathcal{T}$ is totally ordered. Define a functional $f_0 : X_0 \to \mathbb{R}$ by $f_0(x) = f_\alpha(x)$ for all $x \in M_\alpha$. Then f_0 is well-defined, since if $x \in M_\alpha \cap M_\beta$, then $x \in M_\alpha$ and $x \in M_\beta$. Therefore $f_0(x) = f_\alpha(x)$ and $f_0(x) = f_\beta(x)$. By total ordering of $\mathcal{T}$, either f_α extends f_β or vice versa. Hence $f_\alpha(x) = f_\beta(x)$. It is clear that f_0 is a linear extension of f. Furthermore $f_0(x) \le p(x)$ for all $x \in X_0$ and $(M_\alpha, f_\alpha) \prec (X_0, f_0)$ for all $(M_\alpha, f_\alpha) \in \mathcal{T}$, i.e., (X_0, f_0) is an upper bound for $\mathcal{T}$.

By Zorn's lemma11.5.1, $\mathcal{F}$ has a maximal element (X_1, F). To complete the proof, it suffices to show that $X_1 = X$. If $X_1 \neq X$, then choose $y \in X \backslash X_1$. By Lemma 11.5.2, we can extend F to a linear functional $\widetilde{F}$ defined on $\tilde{X} = X_1 \oplus \text{lin}\{y\}$ and extending f such that $\widetilde{F}(x) \le p(x)$ for all $x \in \widetilde{X}$. Thus $(\widetilde{X}, \widetilde{F}) \in \mathcal{F}$ and $(X_1, F) \prec (\widetilde{X}, \widetilde{F})$, which contradicts the maximality of (X_1, F). ■

Definition 11.5.6 — seminorm. A seminorm p on a (complex) linear space X is a function $p : X \to \mathbb{R}$ such that for all $x, y \in X$ and $\lambda \in \mathbb{C}$

1. $p(x) \ge 0$ and $p(0) = 0$
2. $p(x+y) \le p(x) + p(y)$, and
3. $p(\lambda x) = |\lambda| p(x)$

Theorem 11.5.4 — Hahn-Banach Extension Theorem for (complex) linear spaces. Let X be a real or complex linear space, p be a seminorm on X and f a linear functional on a linear subspace M of X such that $|f(x)| \le p(x)$ for all $x \in M$. Then there is a linear functional F on X such that $F|_M = f$ and $|F(x)| \le p(x)$ for all $x \in X$

Proof. Assume first that X is a real linear space. Then, by Theorem 11.5.3 there is an extension F of f such that $F(x) \le p(x)$ for all $x \in X$. Since

$$-F(x) = F(-x) \le p(-x) = p(x) \text{ for all } x \in X$$

it follows that $-p(x) \le F(x) \le p(x)$, or $|F(x)| \le p(x)$ for all $x \in X$. Now assume that X is a complex linear space. Then we may regard X as a real linear space by restricting the scalar field to $\mathbb{R}$. We denote the resulting real linear space by $X_\mathbb{R}$ and the real linear subspace by $M_\mathbb{R}$. Write f as $f = f_1 + i f_2$, where f_1 and f_2 are real linear functionals given by $f_1(x) = \Re e[f(x)]$ and $f_2(x) = \Im[f(x)]$. Then f_1 is a real linear functional of $M_{\mathbb{R}}$ and $f_1(x) \leq |f(x)| \leq p(x)$ for all $x \in M_{\mathbb{R}}$. Hence, by Theorem 11.5.3, f_1 has a real linear extension F_1 such that $F_1(x) \leq p(x)$ for all $x \in X_{\mathbb{R}}$. Since

$$\begin{aligned} f(ix) = if(x) &\Longleftrightarrow f_1(ix) + if_2(ix) = if_1(x) - f_2(x) \\ &\Longleftrightarrow f_2(x) = -f_1(ix) \text{ and } f_2(ix) = f_1(x) \end{aligned}$$

we can write $f(x) = f_1(x) - if_1(ix)$. Set

$$F(x) = F_1(x) - iF_1(ix) \quad \text{for all } x \in X$$

Then F is a real linear extension of f and, for all $x, y \in X$

$$\begin{aligned} F(x+y) &= F_1(x+y) - iF_1(ix+iy) = F_1(x) - iF_1(ix) + F_1(y) - iF_1(iy) \\ &= F(x) + F(y) \end{aligned}$$

For all $x \in X$

$$F(ix) = F_1(ix) - iF_1(-x) = F_1(ix) + iF_1(x) = i\left(F_1(x) - iF_1(ix)\right) = iF(x)$$

If $\alpha = a + bi$ for $a, b \in \mathbb{R}$, and $x \in X$, then

$$\begin{aligned} F(\alpha x) = F((a+bi)x) &= F(ax + bix) = F(ax) + F(bix) \\ &= aF(x) + bF(ix) = aF(x) + biF(x) = (a+bi)F(x) \\ &= \alpha F(x) \end{aligned}$$

Hence, F is also complex linear. Finally, for $x \in X$, write $F(x) = |F(x)|e^{i\theta}$. Then, since $\Re e F = F_1$,

$$|F(x)| = F(x)e^{-i\theta} = F\left(xe^{-i\theta}\right) = F_1\left(xe^{-i\theta}\right) \leq p\left(xe^{-i\theta}\right) = \left|e^{-i\theta}\right| p(x) = p(x)$$

Suppose that M is a subspace of a normed linear space X and f is a bounded linear functional on M. If F is any extension of f to X, then the norm of F is at least as large as $\|f\|$ because

$$\begin{aligned} \|F\| = \sup\{|F(x)| : x \in X, \|x\| \leq 1\} &\geq \sup\{|F(x)| : x \in M, \|x\| \leq 1\} \\ &= \sup\{|f(x)| : x \in M, \|x\| \leq 1\} = \|f\| \end{aligned}$$

■

The following consequence of the Hahn-Banach theorem states that it is always possible to find a bounded extension of f to the whole space which has the same, i.e., smallest possible, norm.

Theorem 11.5.5 — Hahn-Banach Extension Theorem for Normed linear spaces. Let M be a linear subspace of the normed linear space $(X, \|\cdot\|)$ and let $f \in M^*$. Then there exists an extension $x^* \in X^*$ of f such that $\|x^*\| = \|f\|$

Proof. Define p on X by $p(x) = \|f\|\|x\|$. Then p is a seminorm on X and $|f(x)| \leq p(x)$ for all $x \in M$. By Theorem 11.5.4, f has an extension F to X such that $|F(x)| \leq p(x)$ for all $x \in X$. That is $|F(x)| \leq \|f\|\|x\|$. This shows that F is bounded and $\|F\| \leq \|f\|$. Since F must have norm at least as large as $\|f\|$, $\|F\| = \|f\|$ and the result follows with $x^* = F$ ■

11.5.1 Consequences of the Hahn-Banach Extension Theorem

Theorem 11.5.6 Let M be a linear subspace of a normed linear space $(X, \|\cdot\|)$ and $x \in X$ such that

$$d = d(x, M) := \inf_{y \in M} \|x - y\| > 0$$

Then there is an $x^* \in X^*$ such that

1. $\|x^*\| = 1$
2. $x^*(x) = d$
3. $x^*(m) = 0$ for all $m \in M$

Proof. Let $Y = M + \text{lin}\{x\} := \{m + \alpha x, m \in M, \alpha \in \mathbb{F}\}$. Then each y in Y is uniquely expressible in the form $y = m + \alpha x$ for some $m \in M$ and some scalar α. Indeed, if

$$y = m_1 + \alpha x = m_2 + \beta x$$

for some $m_1, m_2 \in M$ and some scalars α and β, then $(\beta - \alpha)x = m_1 - m_2 \in M$.

Claim: $\alpha = \beta$. If $\alpha \neq \beta$, then since M is a subspace

$$x = \frac{1}{\alpha - \beta}(m_1 - m_2) \in M$$

a contradiction since $x \notin M$. Hence, $\alpha = \beta$ and consequently $m_1 = m_2$. Define $f : Y \to \mathbb{F}$ by

$$f(y) = f(m + \alpha x) = \alpha d$$

Since the scalar α is uniquely determined, f is well defined.

Claim: f is a bounded linear functional on Y. Linearity: Let $y_1 = m_1 + \alpha_1 x$ and $y_2 = m_2 + \alpha_2 x$ be any two elements of Y and $\lambda \in \mathbb{F}$. Then

$$f(\lambda y_1 + y_2) = f((\lambda m_1 + m_2) + (\lambda \alpha_1 + \alpha_2)x) = (\lambda \alpha_1 + \alpha_2)d = \lambda \alpha_1 d + \alpha_2 d = \lambda f(y_1) + f(y_2)$$

Boundedness: Let $y = m + \alpha x \in Y$. Then

$$\|y\| = \|m + \alpha x\| = |\alpha| \left\| \frac{m}{\alpha} + x \right\| = |\alpha| \left\| x - \left(-\frac{m}{\alpha} \right) \right\| \geq |\alpha| d = |f(y)|$$

i.e., $|f(y)| \leq \|y\|$ for all $y \in Y$. Thus, f is bounded and $\|f\| \leq 1$.

We show next that $\|f\| = 1$. By definition of infimum, given any $\epsilon > 0$, there is an element $m_\epsilon \in M$ such that $\|x - m_\epsilon\| < d + \epsilon$. Let $z = \frac{x - m_\epsilon}{\|x - m_\epsilon\|}$. Then $z \in Y, \|z\| = 1$ and

$$|f(z)| = \frac{d}{\|x - m_\epsilon\|} > \frac{d}{d + \epsilon}$$

Since ϵ is arbitrary, it follows that $|f(z)| \geq 1$. Thus

$$1 \leq |f(z)| \leq \|f\|\|z\| = \|f\|$$

Thus, $\|f\| = 1$. It is clear that and $f(m) = 0$ for all $m \in M$ and $f(x) = d$. By Theorem 11.5.5, there is an $x^* \in X^*$ such that $x^*(y) = f(y)$ for all $y \in Y$ and $\|x^*\| = \|f\|$. Hence, $\|x^*\| = 1$ and $x^*(m) = 0$ for all $m \in M$ and $x^*(x) = d$. ■

Corollary 11.5.7 Let $(X, \|\cdot\|)$ be a normed linear space and $x_0 \in X \backslash \{0\}$. Then there exists an $x^* \in X^*$, such that $x^*(x_0) = \|x_0\|$ and $\|x^*\| = 1$

Proof. Consider $M = \{0\}$. Since $x_0 \in X \backslash \{0\}$, it follows that $x_0 \notin M$ and so $d = d(x_0, M) = \|x_0\| > 0$. By Theorem 11.5.3, there is an $x^* \in X^*$ such that $x^*(x_0) = \|x_0\|$ and $\|x^*\| = 1$ ■

The following result asserts that X^* is big enough to distinguish the points of X.

Corollary 11.5.8 Let $(X, \|\cdot\|)$ be a normed linear space and $y, z \in X$. If $y \neq z$, then there exists an $x^* \in X^*$, such that

$$x^*(y) \neq x^*(z)$$

Proof. Consider $M = \{0\}$. Since $y \neq z$, it follows that $y - z \notin M$ and consequently

$$d = d(y - z, M) = \|y - z\| > 0$$

By Theorem 11.5.6, there is an $x^* \in X^*$ such that $x^*(y - z) = d > 0$. Hence $x^*(y) \neq x^*(z)$ ■

Corollary 11.5.9 For each x in a normed linear space $(X, \|\cdot\|)$

$$\|x\| = \sup\{|x^*(x)| \,|\, x^* \in X^*, \|x^*\| = 1\}$$

Proof. If $x = 0$, then the result holds vacuously. Assume $x \in X \backslash \{0\}$. For any $x^* \in X^*$ with $\|x^*\| = 1$

$$|x^*(x)| \leq \|x^*\| \, \|x\| = \|x\|$$

Hence, $\sup\{|x^*(x)|\,|x^* \in X^*, \|x^*\| = 1\} \leq \|x\|$. By Corollary 11.5.7, there is a $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \|x\|$. Therefore

$$\|x\| = |x^*(x)| \leq \sup\{|x^*(x)|\,|x^* \in X^*, \|x^*\| = 1\}$$

whence $\|x\| = \sup\{|x^*(x)|\,|x^* \in X^*, \|x^*\| = 1\}$ ■

Theorem 11.5.10 If the dual X^* of a normed linear space $(X, \|\cdot\|)$ is separable, then X is also separable.

Proof. Let $S = S(X^*) = \{x^* \in X^* |\, \|x^*\| = 1\}$. Since any subset of a separable space is separable, S is separable. Let $\{x_n^* | n \in \mathbb{N}\}$ be a countable dense subset of S. Since $x_n^* \in S$ for each $n \in \mathbb{N}$, we have that $\|x_n^*\| = 1$. Hence, for each $n \in \mathbb{N}$ there is an element $x_n \in X$ such that $\|x_n\| = 1$ and $|x_n^*(x_n)| > \frac{1}{2}$ (Otherwise $|x_n^*(x)| \leq \frac{1}{2}$ for all $x \in X$ and so $\|x_n^*\| \leq \frac{1}{2}$, a contradiction.) Let

$$M = \overline{\text{lin}}(\{x_n | n \in \mathbb{N}\})$$

Then M is separable since M contains a countable dense subset comprising all linear combinations of the x_n 's with coefficients whose real and imaginary parts are rational. Claim: $M = X$. If $M \neq X$, then there is an element $x_0 \in X \backslash M$ such that $d = d(x_0, M) > 0$. By Theorem 11.5.6, there is an $x^* \in X^*$ such that $\|x^*\| = 1$, i.e. $x^* \in S$, and $x^*(y) = 0$ for all $y \in M$. In particular, $x^*(x_n) = 0$ for all $n \in \mathbb{N}$. Now, for each $n \in \mathbb{N}$

$$\frac{1}{2} < |x_n^*(x_n)| = |x_n^*(x_n) - x^*(x_n)| = |(x_n^* - x^*)(x_n)| \leq \|x_n^* - x^*\|$$

But this contradicts the fact that the set $\{x_n^* | n \in \mathbb{N}\}$ is dense in S. Hence $M = X$ and, consequently, X is separable. ■

R The converse of Theorem 11.5.10 does not hold. That is, if X separable, it does not follow that its dual X^* is also be separable. Take, for example, ℓ_1. Its dual is (isometrically isomorphic to) ℓ_∞. The space ℓ_1 is separable whereas ℓ_∞ is not. This also shows that the dual of ℓ_∞ is not (isometrically isomorphic to) ℓ_1

Definition 11.5.7 Let M be a subset of a normed linear space X. The annihilator of M, denoted by $M^\perp$, is the set

$$M^\perp = \{x^* \in X^* | x^*(y) = 0 \text{ for all } y \in M\}$$

It is easy to show that $M^\perp$ is a closed linear subspace of X^*

Theorem 11.5.11 Let M be a linear subspace of a normed linear space X. Then

$$X^*/M^\perp \cong M^*$$

Proof. Define $\Phi : X^*/M^\perp \to M^*$ by

$$\Phi\left(x^* + M^\perp\right)(m) = x^*(m)$$

for all $x^* \in X^*$ and all $m \in M$. We show that Φ is well-defined. Let $x^*, y^* \in X^*$ such that $x^* + M^\perp = y^* + M^\perp$. Then $x^* - y^* \in M^\perp$ and so $x^*(m) = y^*(m)$ for all $m \in M$. Thus $\Phi\left(x^* + M^\perp\right) = \Phi\left(y^* + M^\perp\right)$; i.e., Φ is well-defined. Clearly, $\Phi\left(x^* + M^\perp\right)$ is a linear functional on M. We show that Φ is linear. Let $x^*, y^* \in X^*$ and $\lambda \in \mathbb{F}$. Then, for all $m \in M$

$$\begin{aligned}\Phi\left(\left(x^* + M^\perp\right) + \lambda\left(y^* + M^\perp\right)\right)(m) &= \Phi\left(x^* + \lambda y^* + M^\perp\right)(m) = \left(x^* + \lambda y^*\right)(m) \\ &= x^*(m) + \lambda y^*(m) \\ &= \Phi\left(x^* + M^\perp\right)(m) + \lambda\Phi\left(y^* + M^\perp\right)(m) \\ &= \left(\Phi\left(x^* + M^\perp\right) + \lambda\Phi\left(y^* + M^\perp\right)\right)(m)\end{aligned}$$

Hence, $\Phi\left(\left(x^* + M^\perp\right) + \lambda\left(y^* + M^\perp\right)\right) = \Phi\left(x^* + M^\perp\right) + \lambda\Phi\left(y^* + M^\perp\right)$. We now show that Φ is surjective. Let $y^* \in M^*$. Then, by Theorem 11.5.5, there is an $x^* \in X^*$ such that $y^*(m) = x^*(m)$ for all $m \in M$ and $\|y^*\| = \|x^*\|$. Hence, for all $m \in M$

$$\Phi\left(x^* + M^\perp\right)(m) = x^*(m) = y^*(m)$$

Thus $\Phi\left(x^* + M^\perp\right) = y^*$. Furthermore

$$\left\|x^* + M^\perp\right\| \leq \|x^*\| = \|y^*\| = \left\|\Phi\left(x^* + M^\perp\right)\right\|$$

But for any $y^* \in M^\perp, x^* + M^\perp = (x^* + y^*) + M^\perp$. Hence, for all $m \in M$

$$\left|\Phi\left(x^* + M^\perp\right)(m)\right| = \left|(x^* + y^*)(m)\right| \leq \|x^* + y^*\| \, \|m\|$$

That is, $\Phi\left(x^* + M^\perp\right)$ is a bounded linear functional on M and $\left\|\Phi\left(x^* + M^\perp\right)\right\| \leq \|x^* + y^*\|$ for all $y^* \in M^\perp$. Thus

$$\left\|\Phi\left(x^* + M^\perp\right)\right\| \leq \inf_{y^* \in M^\perp} \|x^* + y^*\| = \left\|x^* + M^\perp\right\|$$

It now follows that $\left\|\Phi\left(x^* + M^\perp\right)\right\| = \left\|x^* + M^\perp\right\|$ ∎

11.5.2 Bidual of a normed linear space and Reflexivity

Definition 11.5.8 — Second dual space or bidual space. Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$ and $x \in X$. Define a functional $\Phi_x : X^* \to \mathbb{F}$ by

$$\Phi_x(x^*) = x^*(x) \text{ for all } x^* \in X^*$$

It is easy to verify that Φ_x is linear and for each $x^* \in X^*$

$$|\Phi_x(x^*)| = |x^*(x)| \leq \|x^*\| \|x\|$$

That is, Φ_x is bounded and $\|\Phi_x\| \leq \|x\|$. By Corollary 11.5.9

$$\|x\| = \sup\{|x^*(x)| \, |x^* \in X^*, \|x^*\| = 1\} = \sup\{|\Phi_x(x)| \, |x^* \in X^*, \|x^*\| = 1\} = \|\Phi_x\|$$

This shows that Φ_x is a bounded linear functional on X^*, i.e., $\Phi_x \in (X^*)^* = X^{**}$ and $\|\Phi_x\| = \|x\|$. The space X^{**} is called the **second dual space or bidual space** of X. It now follows that we can define a map

$$J_X : X \to X^{**} \text{ by } J_X x = \Phi_x, \text{ for } x \in X$$

that is

$$(J_X x)(x^*) = x^*(x) \text{ for } x \in X \text{ and } x^* \in X^*$$

It is easy to show that J_X is linear and $\|x\| = \|\Phi_x\| = \|J_X x\|$. That is, J_X is a linear isometry of X into its bidual X^{**}. The map J_X as defined above is called the canonical or natural embedding of X into its bidual X^{**}. This shows that we can identify X with the subspace $J_X X = \{J_X x | x \in X\}$ of X^{**}

Definition 11.5.9 — Reflexive. Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$. Then X is said to be **reflexive** if the canonical embedding $J_X : X \to X^{**}$ of X into its bidual X^{**} is surjective. In this case $X \cong X^{**}$.

If X is reflexive, we customarily write $X = X^{**}$. The equality simply means that X is isometrically isomorphic to X^{**}. Reflexivity of X basically means that each bounded linear functional on X^* is an evaluation functional. Since dual spaces are complete, a reflexive normed linear space is necessarily a Banach space. It is therefore appropriate to speak of a reflexive Banach space rather than a reflexive normed linear space.

Theorem 11.5.12
1. *Every finite-dimensional normed linear space is reflexive.*
2. *A closed linear subspace of a reflexive space is reflexive.*

Proof.
1. Proof. (1). If $\dim X < \infty$, then Proposition 11.4.14 implies that $\dim X = \dim X^* = \dim X^{**}$. Since $J_X X$ is isometrically isomorphic to $X, \dim(J_X X) = \dim X = \dim X^{**}$. Since $J_X X$ is a subspace of X^{**}, it must equal to X^{**}
2. Let X be reflexive and M a closed linear subspace of X. Given $y^{**} \in M^{**}$, it must be shown that there exists $y \in M$ such that $J_M y(y^*) = y^*(y) = y^{**}(y^*)$ for all $y^* \in M^*$. Define a functional ψ on X^* by

$$\psi(x^*) = y^{**}(x^*|_M), \quad x^* \in X^*$$

Clearly, ψ is linear and

$$|\psi(x^*)| \leq \|y^{**}\| \, \|x^*|_M\| \leq \|y^{**}\| \, \|x^*\|$$

so $\psi \in X^{**}$. By reflexivity of X, there exists $y \in X$ such that $J_X y = \psi$. That is, $\psi(x^*) = x^*(y)$ for each $x^* \in X^*$. If $y \notin M$, then by Theorem 11.5.6, there exists an $x_0^* \in X^*$ such that $x_0^*(y) \neq 0$ and $x_0^*(m) = 0$ for all $m \in M$. Then

$$0 \neq x_0^*(y) = \psi(x_0^*) = y^{**}(x_0^*|_M) = y^{**}(0) = 0$$

which is absurd. Thus $y \in M$ and $x^*(y) = \psi(x^*) = y^{**}(x^*|_M), \quad x^* \in X^*$. By Theorem 11.5.5, every $y^* \in M^*$ is of the form $y^* = x^*|_M$ for some $x^* \in X^*$. Thus $(J_M y)(y^*) = y^*(y) = y^{**}(y^*), \quad y^* \in M^*$, and the proof is complete.

■

Theorem 11.5.13 A Banach space X is reflexive if and only if its dual X^* is reflexive.

Proof. Assume that X is reflexive. Let $J_X : X \to X^{**}$ and $J_{X^*} : X^* \to (X^*)^{**} = X^{***}$ be the canonical embeddings of X and X^* respectively. We must show that J_{X^*} is surjective. To that end, let $x^{***} \in X^{***} = (X^{**})^*$ and consider the following diagram:

$$X \xrightarrow{J_X} X^{**} \xrightarrow{x^{***}} \mathbb{F}$$

Define a functional x^* on X by $x^* = x^{***} J_X$. It is obvious that x^* is linear since both x^{***} and J_X are linear. Also, for each $x \in X$

$$|x^*(x)| = |x^{***} J_X(x)| \leq \|x^{***}\| \, \|J_X x\| = \|x^{***}\| \, \|x\|$$

i.e., x^* is bounded and $\|x^*\| \leq \|x^{***}\|$. Hence $x^* \in X^*$. We now show that $J_{X^*}(x^*) = x^{***}$. Let $x^{**} \in X^{**}$ be any element. Since J_X is surjective, there is an $x \in X$ such that $x^{**} = J_X x$. Hence

$$x^{***}(x^{**}) = x^{***}(J_X x) = x^*(x) = J_X x(x^*) = J_{X^*} x^*(J_X x) = J_{X^*} x^*(x^{**})$$

and therefore $J_{X^*} x^* = x^{***}$. That is, J_{X^*} is surjective.

Assume that X^* is reflexive. Then the canonical embedding $J_{X^*} : X^* \to X^{***}$ is surjective. If $J_X X \neq X^{**}$, let $x^{**} \in X^{**} \setminus J_X X$. since $J_X X$ is a closed linear subspace of X^{**}, it follows from Theorem 11.5.6 that there is a functional $\phi \in X^{***}$ such that $\|\phi\| = 1, \phi(x^{**}) = d(x^{**}, J_X X)$, and $\phi(J_X x) = 0$ for all $x \in X$. since J_{X^*} is surjective, there is an $x^* \in X^*$ such that $J_{X^*} x^* = \phi$. Hence, for each $x \in X$

$$0 = \phi(J_X x) = J_{X^*} x^*(J_X x) = (J_X x)(x^*) = x^*(x)$$

i.e., $x^*(x) = 0$ for all $x \in X$. This implies that $x^* = 0$. But then $0 = J_{X^*} x^* = \phi$, a contradiction since $\phi \neq 0$. Hence $J_X X = X^{**}$; i.e., J_X is surjective. ■

We showed earlier (Theorem 11.5.10) that if the dual space X^* of a normed linear space X is separable, then X is also separable, but not conversely. However, if X is reflexive, then the converse holds

Theorem 11.5.14 If X is a reflexive separable Banach space, then its dual X^* is also separable.

Proof. Since X is reflexive and separable, its bidual $X^{**} = J_X X$ is also separable. Hence, by Theorem 11.5.10, X^* is separable. ■

■ **Example 11.33** 1. For $1 < p < \infty$, the sequence space ℓ_p is reflexive.

2. The spaces c_0, c, ℓ_1, and ℓ_∞ are non-reflexive.
3. Every Hilbert space $\mathcal{H}$ is reflexive.

11.5.3 The Adjoint Operator

Definition 11.5.10 — Adjoint. Let X and Y be normed linear spaces and $T \in \mathcal{B}(X,Y)$. The Banach space adjoint (or simply adjoint) of T, denoted by T^*, is the operator $T^* : Y^* \to X^*$ defined by

$$(T^* y^*)(x) = y^*(Tx) \text{ for all } y^* \in Y^* \text{ and all } x \in X$$

The following diagram helps make sense of the above definition.

$$X \xrightarrow{T} Y$$
$$X^* \xleftarrow{T^*} Y^*$$

It is straightforward to show that for any $y^* \in Y^*, T^* y^*$ is a linear functional on X. Furthermore, for any $y^* \in Y^*$ and $x \in X$

$$|T^* y^*(x)| = |y^*(Tx)| \leq \|y^*\| \, \|T\| \|x\|$$

i.e., $T^* y^*$ is a bounded linear functional on X and $\|T^* y^*\| \leq \|T\| \, \|y^*\|$

■ **Example 11.34** Let $X = \ell_1 = Y$ and define $T : \ell_1 \to \ell_1$ by

$$Tx = T(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots), \text{ where } x = (x_n) \in \ell_1$$

the right-shift operator. Then the adjoint of T is $T^* : \ell_\infty \to \ell_\infty$ is given by

$$T^* y = T^*(y_1, y_2, y_3, \ldots) = (y_2, y_3, \ldots), \text{ where } y = (y_n) \in \ell_\infty$$

the left-shift operator

Theorem 11.5.15 Let X and Y be normed linear spaces over $\mathbb{F}$ and let $T \in \mathcal{B}(X,Y)$

1. T^* is a bounded linear operator on Y^*
2. The map $\Lambda : \mathcal{B}(X,Y) \to \mathcal{B}(Y^*,X^*)$ defined by $\Lambda T = T^*$ is an isometric isomorphism of $\mathcal{B}(X,Y)$ into $\mathcal{B}(Y^*,X^*)$

Proof. 1. Let $y_1^*, y_2^* \in Y^*$ and $\alpha \in \mathbb{F}$. Then, for all $x \in X$

$$\begin{aligned} T^*(\alpha y_1^* + y_2^*)(x) &= (\alpha y_1^* + y_2^*)(Tx) = \alpha y_1^*(Tx) + y_2^*(Tx) \\ &= \alpha T^* y_1^*(x) + T^* y_2^*(x) = (\alpha T^* y_1^* + T^* y_2^*)(x) \end{aligned}$$

Hence, $T^*(\alpha y_1^* + y_2^*) = \alpha T^* y_1^* + T^* y_2^*$. Furthermore, as shown above, $\|T^* y^*\| \le \|T\| \|y^*\|$. Hence, $T^* \in \mathcal{B}(Y^*,X^*)$ and $\|T^*\| \le \|T\|$

2. We show that $\|T^*\| = \|T\|$, whence $\|\Lambda T^*\| = \|T^*\|$. Indeed

$$\begin{aligned} \|T\| &= \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|=1} \left(\sup_{\|y^*\|=1} |y^*(Tx)| \right) \quad (\text{ by Corollary 11.5.9}) \\ &= \sup_{\|y^*\|=1} \left(\sup_{\|x\|=1} |y^*(Tx)| \right) = \sup_{\|y^*\|=1} \|T^* y^*\| \\ &= \|T^*\|. \end{aligned}$$

■

11.5.4 Weak Topologies

Definition 11.5.11 — weak topology. Let $(X, \|\cdot\|)$ be a normed linear space and $\mathcal{F} \subset X^*$. The weak topology on X induced by the family $\mathcal{F}$ denoted by $\sigma(X,\mathcal{F})$, is the weakest topology on X with respect to which each $x^* \in \mathcal{F}$ is continuous.

R The weak topology on X induced by the dual space X^* is simply referred to as "the weak topology on X'' and is denoted by $\sigma(X,X^*)$

Definition 11.5.12 What do the basic open sets for the weak topology $\sigma(X,X^*)$ look like? Unless otherwise indicated, we shall denote by $\Phi, \Phi_1, \Phi_2 \ldots$ finite subsets of X^*. Let $x_0 \in X, \Phi$ and $\epsilon > 0$ be given. Consider all sets of the form

$$\begin{aligned} V(x_0;\Phi;\epsilon) &:= \{x \in X \,|\, |x^*(x) - x^*(x_0)| < \epsilon, x^* \in \Phi\} \\ &= \bigcap_{x^* \in \Phi} \{x \in X \,|\, |x^*(x) - x * (x_0)| < \epsilon\} \end{aligned}$$

Proposition 11.5.16 1. $x_0 \in V(x_0;\Phi;\epsilon)$

2. Given $V(x_0;\Phi_1;\epsilon_1)$ and $V(x_0;\Phi_2;\epsilon_2)$, we have

$$V(x_0;\Phi_1 \cup \Phi_2; \min\{\epsilon_1,\epsilon_2\}) \subset V(x_0;\Phi_1;\epsilon_1) \cap V(x_0;\Phi_2;\epsilon_2)$$

3. If $x \in V(x_0;\Phi;\epsilon)$, then there is a $\delta > 0$ such that $V(x;\Phi;\delta) \subset V(x_0;\Phi;\epsilon)$

Proof. 1. It is obvious that $x_0 \in V(x_0;\Phi;\epsilon)$

2. Let $x \in V(x_0;\Phi_1 \cup \Phi_2;\min\{\epsilon_1,\epsilon_2\})$. Then for each $x^* \in \Phi_1$

$$|x^*(x) - x^*(x_0)| < \min\{\epsilon_1,\epsilon_2\} \leq \epsilon_1$$

Hence $x \in V(x_0;\Phi_1;\epsilon_1)$. Similarly, $x \in V(x_0;\Phi_2;\epsilon_2)$. It now follows that $x \in V(x_0;\Phi_1;\epsilon_1) \cap V(x_0;\Phi_2;\epsilon_2)$ and, consequently

$$V(x_0;\Phi_1 \cup \Phi_2;\min\{\epsilon_1,\epsilon_2\}) \subset V(x_0;\Phi_1;\epsilon_1) \cap V(x_0;\Phi_2;\epsilon_2)$$

3. Let $x \in V(x_0;\Phi;\epsilon)$ and $\gamma = \max\{|x^*(x) - x^*(x_0)|\,|x^* \in \Phi\}$. Then $0 \leq \gamma < \epsilon$. Choose δ such that $0 < \delta < \epsilon - \gamma$. Then, for any $y \in V(x;\Phi;\delta)$ and any $x^* \in \Phi$, we have

$$|x^*(y) - x^*(x_0)| \leq |x^*(y) - x^*(x)| + |x^*(x) - x^*(x_0)| < \delta + \gamma < \epsilon$$

■

Definition 11.5.13 Recall that a collection $\mathcal{B}$ of subsets of a set X is a base for a topology on X if and only if

1. $X = \bigcup\{B|B \in \mathcal{B}\}$; i.e., each $x \in X$ belongs to some $B \in \mathcal{B}$, and
2. if $x \in B_1 \cap B_2$ for some B_1 and B_2 in $\mathcal{B}$, then there is a $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$

Theorem 11.5.17 Let $\mathcal{B} = \{V(x;\Phi;\epsilon)|x \in X, \Phi(\text{finite}) \subset X^*, \epsilon > 0\}$. Then $\mathcal{B}$ is a base for a Hausdorff topology on X

Proof. 1. It is clear that $x \in V(x;\Phi;\epsilon)$ for each $x \in X$

2. Let $x \in V(x_1;\Phi_1;\epsilon_1) \cap V(x_2;\Phi_2;\epsilon_2)$. Then $x \in V(x_1;\Phi_1;\epsilon_1)$ and $x \in V(x_2;\Phi_2;\epsilon_2)$. By Proposition11.5.16 (3), there are $\delta_1 > 0$ and $\delta_2 > 0$ such that $V(x;\Phi_1;\delta_1) \subset V(x_1;\Phi_1;\epsilon_1)$ and $V(x;\Phi_2;\delta_2) \subset V(x_2;\Phi_2;\epsilon_2)$. By Proposition11.5.16 (2)

$$V(x;\Phi_1 \cup \Phi_2;\min\{\delta_1,\delta_2\}) \subset V(x;\Phi_1;\delta_1) \cap V(x;\Phi_2;\delta_2) \subset V(x_1;\Phi_1;\epsilon_1) \cap V(x_2;\Phi_2;\epsilon_2)$$

Hence, $\mathcal{B}$ is a base for a topology on X. Finally, we show that the topology generated by $\mathcal{B}$ is Hausdorff. Let x and y be distinct elements of X. By Corollary 11.5.8 , there is an $x^* \in X^*$ such that $x^*(x) \neq x^*(y)$. Let $0 < \epsilon < |x^*(x) - x^*(y)|$. Then $V\left(x;x^*;\frac{\epsilon}{2}\right)$ and $V\left(y;x^*;\frac{\epsilon}{2}\right)$ are disjoint neighbourhoods of x and y respectively.

■

R It is easy to see that each $x^* \in X^*$ is continuous with respect to the topology generated by $\mathcal{B}$. Indeed, let $x_0 \in X, x^* \in X^*$ and $\epsilon > 0$. Since x^* is continuous with respect to the norm topology on X, there is a norm neighbourhood U of x_0 such $|x^*(x) - x^*(x_0)| < \epsilon$ for all $x \in U$. It now follows $V(x_0; x^*; \epsilon)$ is a neighbourhood of x_0 in the topology generated by $\mathcal{B}$ and $|x^*(x) - x^*(x_0)| < \epsilon$ for all $x \in V(x_0; x^*; \epsilon)$.

Definition 11.5.14 One shows quite easily that the topology generated by

$$\mathcal{B} = \{V(x; \Phi; \epsilon) | x \in X, \Phi(\text{ finite }) \subset X^*, \epsilon > 0\}$$

is precisely $\sigma(X, X^*)$, the weak topology on X induced by X^*. Therefore, a set G is open in the topology $\sigma(X, X^*)$ if and only if for each $x \in G$ there is a finite set $\Phi = \{x^*, x_1^*, x_2^*, \ldots, x_n^*\} \subset X^*$ and an $\epsilon > 0$ such that $V(x; \Phi; \epsilon) \subset G$. It now follows that a normed linear space X carries two natural topologies: the norm topology induced by the norm on X and the weak topology induced by its dual space X^*.

Topological concepts that are associated with the weak topology are usually preceded by the word "**weak**"; for example, weak compactness, weak closure, etc. Those topological concepts pertaining to the topology generated by the norm on X are usually preceded by the word "**norm**", e.g. norm-closure or by the word "**strong**", e.g. strongly open set.

Theorem 11.5.18 Let$\{x^*, x_1^*, x_2^*, \ldots, x_n^*\} \subset X^*$. Then

1. $x^* \in \operatorname{lin}\{x_1^*, x_2^*, \ldots, x_n^*\}$ if and only if $\bigcap_{i=1}^n \ker(x_i^*) \subset \ker(x^*)$
2. If $\{x_1^*, x_2^*, \ldots, x_n^*\}$ is a linearly independent set, then for any set of scalars $\{c_1, c_2, \ldots, c_n\}$ $\bigcap_{i=1}^n \{x \in X | x_i^*(x) = c_i\} \neq \emptyset$

Proof. 1. If $x^* \in \operatorname{lin}\{x_1^*, x_2^*, \ldots, x_n^*\}$, then $x^* = \sum_{i=1}^n \alpha_i x_i^*$ for some scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$. Let $x \in \bigcap_{i=1}^n \ker(x_i^*)$. Then $x_i^*(x) = 0$ for each $i = 1, 2, \ldots, n$. Hence, $\sum_{i=1}^n \alpha_i x_i^*(x) = 0$ and consequently $x^*(x) = 0$; i.e., $x \in \ker(x^*)$. Therefore $\bigcap_{i=1}^n \ker(x_i^*) \subset \ker(x^*)$. Conversely, assume that $\bigcap_{i=1}^n \ker(x_i^*) \subset \ker(x^*)$. We use induction on n. Let us first show that if $\ker(x_1^*) = \ker(x^*)$, then $x^* = \alpha x_1^*$ for some nonzero scalar α. Let $K = \ker(x_1^*)$ and $z \in X \backslash K$. Then proceeding as in Theorem11.5.6 , each $x \in X$ is uniquely expressible as $x = y + \lambda z$, where $y \in K$ and $\lambda \in \mathbb{F}$. Hence, since $x^*(y) = 0 = x_1^*(y)$

$$x^*(x) = \lambda x^*(z) = \frac{\lambda x^*(z)}{x_1^*(z)} x_1^*(z) = \left(\frac{x^*(z)}{x_1^*(z)}\right) \lambda x_1^*(z) = \left(\frac{x^*(z)}{x_1^*(z)}\right) x_1^*(x) = \alpha x_1^*(x)$$

where $\alpha = \frac{x^*(z)}{x_1^*(z)}$ Assume that the result is true for $n-1$. For each $i = 1, 2, \ldots, n, x_i^*$ is not a linear combination of the x_j^* 's for $j = 1, 2, \ldots, n$ and $i \neq j$. Hence, $\bigcap_{j \neq i} \ker\left(x_j^*\right)$ is not contained in $\ker(x_i^*)$. Therefore there is an $x_i \in \bigcap_{j \neq i} \ker\left(x_j^*\right)$ such that $x_i^*(x_i) = 1$. Let $\alpha_i = x^*(x_i)$ for each $i = 1, 2, \ldots, n$. Let $x \in X$ and $y = x - \sum_{i=1}^n x_i^*(x) x_i$.

Then, for each $j = 1,2,\dots,n$

$$x_j^*(y) = x_j^*(x) - \sum_{i=1}^{n} x_i^*(x) x_j^*(x_i) = x_j^*(x) - x_j^*(x) = 0$$

Thus, $y \in \bigcap_{i=1}^{n} \ker(x_i^*)$. By the assumption, $y \in \ker(x^*)$. Therefore

$$0 = x^*(y) = x^*(x) - \sum_{i=1}^{n} x_i^*(x) x^*(x_i) = x^*(x) - \sum_{i=1}^{n} \alpha_i x_i^*(x) \Longleftrightarrow x^*(x) = \sum_{i=1}^{n} \alpha_i x_i^*(x)$$

whence $x^* = \sum_{i=1}^{n} \alpha_i x_i^*$

2. Let $H_i = \{x \in X | x_i^*(x) = c_i\}$ for each $i = 1,2,\dots,n$. We want to show that $\bigcap_{i=1}^{n} H_i \neq \emptyset$ The proof is by induction on n. If $n = 1$, then, since $x_1^* \neq 0$, it follows that $H_1 \neq \emptyset$. Assume true for $n = k$ and let $H = \bigcap_{i=1}^{k+1} H_i$. By the linear independence of $\{x_1^*, x_2^*, \dots, x_{k+1}^*\}$, $\bigcap_{i=1}^{k} \ker(x_i^*) \not\subseteq \ker(x_{k+1}^*)$ Hence, there is an $x_0 \in \bigcap_{i=1}^{k} \ker(x_i^*)$ such that $x_{k+1}^*(x_0) \neq 0$. Take any $x \in \bigcap_{i=1}^{k} \ker(x_i^*)$ and set $y = x + \alpha x_0$, where $\alpha = c_{k+1} - \frac{x_{k+1}^*(x)}{x_{k+1}^*(x_0)}$. Then $x_i^*(y) = x_i^*(x) = c_i$ for each $i = 1,2,\dots,k$ and $x_{k+1}^*(y) = c_{k+1}$. That is, $y \in H$

■

Theorem Let τ denote the norm topology on X. Then (a) $\sigma(X, X^*) \subset \tau$ (b) $\sigma(X, X^*) = \tau$ if and only if X is finite-dimensional. Thus, if X is infinite-dimensional, then the weak topology $\sigma(X, X^*)$ is strictly weaker than the norm topology.

Proof. (a) The topology $\sigma(X, X^*)$ is the weakest topology on X making each $x^* \in X^*$ continuous. Hence, $\sigma(X, X^*)$ is weaker than the norm topology τ.

(b) Assume that $\sigma(X, X^*) = \tau$ and let $x^* \in X^*$. Then, since x^* is continuous when X is equipped with the norm topology and, by the hypothesis, it is continuous in the weak topology $\sigma(X, X^*)$, it is continous at 0. Therefore there is a finite set $\Phi = \{x_1^*, x_2^*, \dots, x_n^*\} \subset X^*$ and an $\epsilon > 0$ such that $|x^*(x)| < 1$ for all $x \in V(0;\Phi;\epsilon)$. Let $z \in \bigcap_{i=1}^{n} \ker(x_i^*)$. Then $x_i^*(z) = 0$ and so $|x_i^*(z)| < \epsilon$ for each $i = 1,2,\dots,n$ That is, $z \in V(0;\Phi;\epsilon)$. If $x \in \bigcap_{i=1}^{n} \ker(x_i^*)$, then $mx \in \bigcap_{i=1}^{n} \ker(x_i^*)$ for each $m \in \mathbb{Z}^+$ since $\bigcap_{i=1}^{n} \ker(x_i^*)$ is a linear subspace of X. It now follows that $mx \in V(0;\Phi;\epsilon)$ for each $m \in \mathbb{Z}^+$. This, in turn, implies that

$$1 > |x^*(mx)| = m|x^*(x)| \Longleftrightarrow |x^*(x)| < \frac{1}{m}$$

since m is arbitrary, $x^*(x) = 0$; i.e., $x \in \ker(x^*)$. Hence $\bigcap_{i=1}^{n} \ker(x_i^*) \subset \ker(x^*)$. By Lemma 11.5.18 $x^* \in X^*$ is expressible as $x^* = \sum_{i=1}^{n} \alpha_i x_i^*$ for some scalars $\alpha_1, \alpha_2, \dots, \alpha_n$. Hence X^* is spanned by the set $\{x_1^*, x_2^*, \dots, x_n^*\}$. This shows that X^* is finite-dimensional. By Proposition 4.2.4, X is also finite-dimensional.

Conversely, assume that X is finite-dimensional. Let $\{x_1, x_2, \dots, x_n\}$ be a basis for X such that $\|x_k\| = 1$ for each $k = 1,2,\dots,n$. Let $U \subset X$ be open in the norm topology of X. We want to show that U is open in the weak topology of X. Let $x_0 \in U$. Then there is an

$r > 0$ such that $B(x_0, r) \subset U$. For any $x \in X, x = \sum_{k=1}^{n} \alpha_k x_k$. Define $x_i^* : X \to \mathbb{F}$ by $x_i^*(x) = \alpha_i$ for each $i = 1, 2, \ldots, n$. Since the α_i 's are uniquely determined, x_i^* is well-defined. One shows quite easily that $x_i^* \in X^*$ for each $i = 1, 2, \ldots, n$ Let $\Phi = \{x_1^*, x_2^*, \ldots, x_n^*\}$ and $\epsilon = \frac{r}{n}$. Then, for any $x \in V(x_0; \Phi; \epsilon)$, we have $|x_i^*(x) - x_i^*(x_0)| < \epsilon$ for each $i = 1, 2, \ldots, n$. Hence, if $x \in V(x_0; \Phi; \epsilon)$, then

$$\|x - x_0\| = \left\| \sum_{k=1}^{n} x_k^*(x - x_0) x_k \right\| \le \sum_{k=1}^{n} |x_k^*(x - x_0)| < n\epsilon = r$$

That is, $x \in B(x_0, r) \subset U$. It now follows that for each $x \in U$, there is a $V(x; \Phi; \epsilon)$ such that $V(x; \Phi; \epsilon) \subset U$. Hence U is open in the weak topology $\sigma(X, X^*)$. Thus, $\sigma(X, X^*) = \tau$

11.6 Baire's Category Theorem and its Applications

Definition 11.6.1 — dense. Recall that a subset S of a metric space (X, d) is **dense** in X if $\bar{S} = X$; i.e., for each $x \in X$ and each $\epsilon > 0$, there is an element $y \in S$ such that $d(x, y) < \epsilon$, or equivalently, $S \cap B(x, \epsilon) \neq \emptyset$.

Theorem 11.6.1 Let (X, d) be a complete metric space. If (G_n) is a sequence of nonempty, open and dense subsets of X then $G = \bigcap_{n \in \mathbb{N}} G_n$ is dense in X

Proof. Let $x \in X$ and $\epsilon > 0$. since G_1 is dense in X, there is a point x_1 in the open set $G_1 \cap B(x, \epsilon)$ Let r_1 be a number such that $0 < r_1 < \frac{\epsilon}{2}$ and

$$\overline{B(x_1, r_1)} \subset G_1 \cap B(x, \epsilon)$$

Since G_2 is dense in X, there is a point x_2 in the open set $G_2 \cap B(x_1, r_1)$. Let r_2 be a number such that $0 < r_2 < \frac{\epsilon}{2^2}$ and

$$\overline{B(x_2, r_2)} \subset G_2 \cap B(x_1, r_1)$$

Since G_3 is dense in X, there is a point x_3 in the open set $G_3 \cap B(x_2, r_2)$. Let r_3 be a number such that $0 < r_3 < \frac{\epsilon}{2^3}$ and

$$\overline{B(x_3, r_3)} \subset G_3 \cap B(x_2, r_2)$$

Continuing in this fashion, we obtain a sequence (x_n) in X and a sequence (r_n) of radii such that for each $n = 1, 2, 3, \ldots$

$$0 < r_n < \frac{\epsilon}{2^n}, \overline{B(x_{n+1}, r_{n+1})} \subset G_{n+1} \cap B(x_n, r_n) \text{ and } \overline{B(x_1, r_1)} \subset G_1 \cap B(x, \epsilon)$$

It is clear that

$$\overline{B(x_{n+1}, r_{n+1})} \subset B(x_n, r_n) \subset B(x_{n-1}, r_{n-1}) \subset \cdots \subset B(x_1, r_1) \subset B(x, \epsilon)$$

Let $N \in \mathbb{N}$. If $k > N$ and $\ell > N$, then both x_k and x_ℓ lie in $B(x_N, r_N)$. By the triangle inequality

$$d(x_k, x_\ell) \leq d(x_k, x_N) + d(x_N, x_\ell) < 2r_N < \frac{2\epsilon}{2^N} = \frac{\epsilon}{2^{N-1}}$$

Hence, (x_n) is a Cauchy sequence in X. Since X is complete, there is a $y \in X$ such that $x_n \to y$ as $n \to \infty$. since x_k lies in the closed set $\overline{B(x_n, r_n)}$ if $k > n$, it follows that y lies in each $\overline{B(x_n, r_n)}$. Hence y lies in each G_n. That is, $G = \bigcap G_n \neq \emptyset$. It is also clear that $y \in B(x, \epsilon)$ ■

Definition 11.6.2 A subset S of metric space (X, d) is said to be nowhere dense in X if the set $X \backslash \bar{S}$ is dense in X; i.e.

$$\overline{X \backslash \bar{S}} = X$$

Proposition 11.6.2 A subset S of a metric space (X, d) is nowhere dense in X if and only if the closure $\bar{S}$ of S contains no interior points.

Proof. Assume that S is nowhere dense in X and that $(\bar{S})^\circ \neq \emptyset$. Then there is an $\epsilon > 0$ and an $x \in \bar{S}$ such that $B(x, \epsilon) \subset \bar{S}$. But then $X \backslash \bar{S} \subset X \backslash B(x, \epsilon)$. Since $X \backslash B(x, \epsilon)$ is closed, $X \backslash B(x, \epsilon) = \overline{X \backslash B(x, \epsilon)}$ Therefore

$$\overline{X \backslash \bar{S}} \subset X \backslash B(x, \epsilon) \subset X$$

where the second containment is proper. This is a contradiction. Hence, $(\bar{S})^\circ = \emptyset$. Conversely, assume that $(\bar{S})^\circ = \emptyset$. Then, for each $x \in \bar{S}$ and each $\epsilon > 0$

$$B(x, \epsilon) \cap X \backslash \bar{S} \neq \emptyset$$

This means that each $x \in \bar{S}$ is a limit point of the set $X \backslash \bar{S}$. That is, $\bar{S} \subset \overline{X \backslash \bar{S}}$. Thus,

$$X = \bar{S} \cup (X \backslash \bar{S}) \subset \overline{X \backslash \bar{S}} \cup X \backslash \bar{S} = \overline{X \backslash \bar{S}} \subset X$$

Hence $X = \overline{X \backslash \bar{S}}$ and so S is nowhere dense in X. ■

■ **Example 11.35** Each finite subset of $\mathbb{R}$ is nowhere dense in $\mathbb{R}$

Definition 11.6.3 A subset S of a metric space (X, d) is said to be

1. of first category or meagre in X if S can be written as a countable union of sets which are nowhere dense in X. Such sets are also called thin.
2. of second category or nonmeagre in X if it is not of first category in X. Such sets are also called fat or thick.

It is clear that a subset of a set of first category is itself a set of first category. Also, a countable union of sets of first category is again a set of first category.

■ **Example 11.36** The set $\mathbb{Q}$ of rationals is of first category in $\mathbb{R}$.

Theorem 11.6.3 — Baire's Category Theorem. A complete metric space (X,d) is of second category in itself.

Proof. Assume that X is of first category. Then there is a sequence (G_n) of sets which are nowhere dense in X such that $X = \bigcup_n G_n$. Replacing each G_n by its closure, we get $X = \bigcup_n \overline{G_n}$. The sets $\overline{G_n}$ are closed and nowhere dense in X. It follows that the sets $U_n = X \backslash \overline{G_n}$ are open and dense in X. Since X is complete it follows, by Theorem 6.1 .1 , that $U = \bigcap_n U_n$ is dense in X and therefore nonempty since X is nonempty. However $X = \bigcup_n \overline{G_n}$ implies that

$$\emptyset \neq \bigcap_n U_n = \bigcap_n (X \backslash \overline{G_n}) = X \backslash \bigcup_n \overline{G_n} = \emptyset$$

which is absurd. ■

11.6.1 Uniform Boundedness Principle

Definition 11.6.4 — norm (or uniformly) bounded/ pointwise bounded. A subset $\mathcal{F}$ of $\mathcal{B}(X,Y)$ is said to be

1. **norm (or uniformly) bounded** if

$$\sup\{\|T\| | T \in \mathcal{F}\} < \infty$$

2. **pointwise bounded** on X if

$$\sup\{\|Tx\| | T \in \mathcal{F}\} < \infty$$

for each $x \in X$

Clearly, a norm bounded set is pointwise bounded on X. Uniform Boundedness Principle (or Banach-Steinhaus Theorem) says that if X is a Banach space, then the converse also holds.

Theorem 11.6.4 — Uniform Boundedness Principle/ Banach-Steinhaus. Let X be a Banach space, Y a normed linear space and let $\mathcal{F}$ be subset of $\mathcal{B}(X,Y)$ such that $\sup\{\|Tx\| | T \in \mathcal{F}\} < \infty$ for each $x \in X$. Then $\sup\{ \|T\| | T \in \mathcal{F}\} < \infty$

Proof. For each $k \in \mathbb{N}$, let

$$A_k = \{x \in X | \|Tx\| \leq k \text{ for all } T \in \mathcal{F}\}$$

Since T is continuous, A_k is closed. Indeed, let $x \in \overline{A_k}$. Then there is a sequence $(x_n) \subset A_k$ such that $\lim_{n\to\infty} x_n = x$. Since $x_n \in A_k$ for each n, $\|Tx_n\| \leq k$ for all $T \in \mathcal{F}$. Hence

$$\|Tx\| \leq \|Tx - Tx_n\| + \|Tx_n\| \leq \|T\| \, \|x_n - x\| + k \to k \text{ as } n \to \infty$$

That is, $\|Tx\| \leq k$ and consequently $x \in A_k$. By the hypothesis, $X = \bigcup_{k=1}^{\infty} A_k$. By Baire's Category Theorem, there is an index k_0 such that $(\overline{A_{k_0}})^\circ \neq \emptyset$. That is, there is an $x_0 \in \overline{A_{k_0}}$ and an $\epsilon > 0$ such that $B(x_0,\epsilon) \subset \overline{A_{k_0}} = A_{k_0}$ Let $x \in X \backslash \{0\}$ and set $z = x_0 + \lambda x$, where $\lambda = \frac{\epsilon}{2\|x\|}$. Then $\|z - x_0\| = \lambda \|x\| = \frac{\epsilon}{2} < \epsilon$. Hence $z \in B(x_0,\epsilon) \subset A_{k_0}$ and, consequently, $\|Tz\| \leq k_0$ for all $T \in \mathcal{F}$. It now follows that

$$\|Tx\| = \frac{1}{\lambda}\|Tz - Tx_0\| \leq \frac{1}{\lambda}(\|Tz\| + \|Tx_0\|) \leq \frac{2k_0}{\lambda} = \frac{4k_0}{\epsilon}\|x\|$$

Hence $\|T\| \leq \frac{4k_0}{\epsilon}$ for all $T \in \mathcal{F}$. ∎

R It is essential that X be complete in Theorem 11.6.4 Consider the subset $\ell_0 \subset \ell_1$ of finitely nonzero sequences in ℓ_1. The set ℓ_0 is dense but not closed in ℓ_1. For each $n \in \mathbb{N}$, let $T_n x = n x_n$, where $x = (x_n) \in \ell_0$. For each $x \in \ell_0, T_n x = 0$ for sufficiently large n. Clearly, (T_n) is pointwise bounded on ℓ_0. On the other hand, for $(e_n) \in \ell_0, \|e_n\| = 1$ and $\|T_n\| \geq T_n e_n = n$ for all $n \in \mathbb{N}$. Thus (T_n) is not norm bounded.

Corollary 11.6.5 — Corollary. Let S be a subset of a normed linear space $(X, \|\cdot\|)$ such that the set $\{x^*(x) | x \in S\}$ is bounded for each $x^* \in X^*$. Then the set S is bounded.

Proof. Let J_X be the canonical embedding of X into X^{**}. By the hypothesis, the set $\{J_X x(x^*) | x \in S\}$ is bounded for each $x^* \in X^*$. Since X^* is a Banach space, it follows from the Uniform Boundedness Principle that the set $\{J_X x | x \in S\}$ is bounded. since $\|J_X x\| = \|x\|$, the set S is also bounded. ∎

R Let X and Y be normed linear spaces. We remarked earlier that the strong operator limit T of a sequence $(T_n) \subset \mathcal{B}(X,Y)$ need not be bounded. However, if X is complete, then T is also bounded. This is a consequence of the Uniform Boundedness Principle.

Corollary 11.6.6 Let (T_n) be a sequence of bounded linear operators from a Banach space X into a normed linear space Y. If T is the strong operator limit of the sequence (T_n), then $T \in \mathcal{B}(X,Y)$

Proof. The proof of linearity of T is straightforward. We show that T is bounded. Since for each $x \in X, T_n x \to Tx$ as $n \to \infty$, the sequence $(T_n x)$ is bounded for each $x \in X$. By the Uniform Boundedness Principle, we have that the sequence $(\|T_n\|)$ is bounded. That is, there is a constant $M > 0$ such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Therefore

$$\|T_n x\| \leq \|T_n\| \|x\| \leq M\|x\| \text{ for all } n \in \mathbb{N}$$

By continuity of the norm,

$$\|Tx\| \leq \|Tx - T_n x\| + \|T_n x\| \leq \|Tx - T_n x\| + M\|x\| \to M\|x\| \text{ as } n \to \infty$$

Hence, $\|Tx\| \leq M\|x\|$ for each $x \in X$, i.e., $T \in \mathcal{B}(X,Y)$ ■

11.6.2 The Open Mapping Theorem

Definition 11.6.5 Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and let $T : X \to Y$. Then we say that T is an open mapping if TU is open in Y whenever U is open in X

Theorem 11.6.7 Let X and Y be Banach spaces over the field $\mathbb{F}$ and let T be a bounded linear operator from X onto Y. Then there is a constant $r > 0$ such that

$$B_Y(0,2r) := \{y \in Y | \|y\| < 2r\} \subset \overline{TB_X(0,1)}$$

Proof. It is easy to see that $X = \bigcup_{n=1}^{\infty} nB_X(0,1)$. Indeed, if $x \in X$, then there is an $n \in \mathbb{N}$ such that $\|x\| < n$. Hence, $x \in nB_X(0,1)$. Since T is surjective

$$Y = TX = T\left(\bigcup_{n=1}^{\infty} nB_X(0,1)\right) = \bigcup_{n=1}^{\infty} nTB_X(0,1) = \bigcup_{n=1}^{\infty} n\overline{TB_X(0,1)}$$

By Baire's Category Theorem, there is a positive integer n_0 such that $\left(n_0\overline{TB_X(0,1)}\right)^{\circ} \neq \emptyset$. This implies that $(\overline{TB_X(0,1)})^{\circ} \neq \emptyset$. Hence, there is a constant $r > 0$ and an element $y_0 \in Y$ such that $B_Y(y_0,4r) \subset \overline{TB_X(0,1)}$. since $y_0 \in \overline{TB_X(0,1)}$, it follows, by symmetry, that $-y_0 \in \overline{TB_X(0,1)}$. Therefore

$$B_Y(0,4r) = B_Y(y_0,4r) - y_0 \subset \overline{TB_X(0,1)} + \overline{TB_X(0,1)}$$

since $\overline{TB_X(0,1)}$ is a convex set, $\overline{TB_X(0,1)} + \overline{TB_X(0,1)} = 2\overline{TB_X(0,1)}$. Hence, $B_Y(0,4r) \subset 2\overline{TB_X(0,1)}$ and, consequently, $B_Y(0,2r) \subset \overline{TB_X(0,1)}$

■

Theorem 11.6.8 Let X and Y be Banach spaces over the field $\mathbb{F}$ and let T be a bounded linear operator from X onto Y. Then there is a constant $r > 0$ such that

$$B_Y(0,r) := \{y \in Y | \|y\| < r\} \subset TB_X(0,1)$$

Proof. By Lemma11.6.7, there is a constant $r > 0$ such that $B_Y(0,2r) \subset \overline{TB_X(0,1)}$. Let $y \in B_Y(0,r)$ i.e., $y \in Y$ and $\|y\| < r$. Then, with $\epsilon = \frac{r}{2}$, there is an element $z_1 \in X$ such that

$$\|z_1\| < \frac{1}{2} \text{ and } \|y - Tz_1\| < \frac{r}{2}$$

Since $y - Tz_1 \in Y$ and $\|y - Tz_1\| < \frac{r}{2} < r$, it follows that $y - Tz_1 \in B_Y(0,r)$. Therefore there is an element $z_2 \in X$ such that

$$\|z_2\| < \frac{1}{2^2} \text{ and } \|(y - Tz_1) - Tz_2\| < \frac{r}{2^2}$$

In general, having chosen elements $z_k \in X, 1 \le k \le n$, such that $\|z_k\| < \frac{1}{2^k}$ and

$$\|y - (Tz_1 + Tz_2 + \cdots + Tz_n)\| < \frac{r}{2^n}$$

pick $z_{n+1} \in X$ such that $\|z_{n+1}\| < \frac{1}{2^{n+1}}$ and

$$\|y - T(z_1 + z_2 + \cdots + z_n + z_{n+1})\| = \|y - (Tz_1 + Tz_2 + \cdots + Tz_n + Tz_{n+1})\| < \frac{r}{2^{n+1}}$$

Claim: The series $\sum_{k=1}^{\infty} z_k$ converges to a point $x \in B_X(0,1)$ and $Tx = y$. Proof of Claim: since X is complete, it suffices to show that $\sum_{k=1}^{\infty} \|z_k\| < \infty$. But this is obviously true since

$$\sum_{k=1}^{\infty} \|z_k\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1$$

Hence, the series $\sum_{k=1}^{\infty} z_k$ converges to some $x \in X$ with $\|x\| < 1$, i.e., $x \in B_X(0,1)$. since

$$\lim_{n\to\infty} \left\| y - T\left(\sum_{k=1}^{n} z_k\right) \right\| = \lim_{n\to\infty} \frac{r}{2^n} = 0$$

continuity of T implies that

$$Tx = \lim_{n\to\infty} T\left(\sum_{k=1}^{n} z_k\right) = y$$

That is, $Tx = y$ ■

Theorem 11.6.9 — Open Mapping Theorem. Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X,Y)$. If T maps X onto Y, then T is an open mapping.

Proof. Let U be an open set in X. We need to show that TU is open in Y. Let $y \in TU$. since T is surjective, there is an $x \in U$ such that $Tx = y$. since U is open, there is an $\epsilon > 0$ such that $B_X(x,\epsilon) = x + B_X(0,\epsilon) \subset U$. But then $y + TB_X(0,\epsilon) \subset TU$. By Lemma 11.6.8, there is a constant $r > 0$ such that $B_Y(0,r) \subset TB_X(0,1)$. Hence $B_Y(0,r\epsilon) \subset TB_X(0,\epsilon)$. Therefore

$$B(y,r\epsilon) = y + B_Y(0,r\epsilon) \subset y + TB_X(0,\epsilon) \subset TU$$

Hence TU is open in Y ■

Theorem 11.6.10 — Banach's Theorem. Let X and Y be Banach spaces and assume $T \in \mathcal{B}(X,Y)$ is bijective. Then T^{-1} is a bounded linear operator from Y onto X, i.e., $T^{-1} \in \mathcal{B}(Y,X)$

Proof. We have shown that T^{-1} is linear. It remains to show that T^{-1} is bounded. By Theorem 4.1.4, it suffices to show that T^{-1} is continuous on Y. To that end, let U be an open set in X. By Theorem 6.3 .1 $\left(T^{-1}\right)^{-1}(U) = TU$ is open in Y. Hence T^{-1} is continuous on Y ■

11.6.3 Closed Graph Theorem

Definition 11.6.6 Let X and Y be linear spaces over a field $\mathbb{F}$ and $T : X \to Y$. The graph of T, denoted by $\mathcal{G}(T)$, is the subset of $X \times Y$ given by

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

Since T is linear, $\mathcal{G}(T)$ is a linear subspace of $X \times Y$. Let $\|\cdot\|_X$ and $\|\cdot\|_Y$ be norms on X and Y respectively. Then, for $x \in X$ and $y \in Y$, $\|(x,y)\| := \|x\|_X + \|y\|_Y$ defines a norm on $X \times Y$. If X and Y are Banach spaces, then so is $X \times Y$

Definition 11.6.7 Let X and Y be normed linear spaces over $\mathbb{F}$. A linear operator $T : X \to Y$ is closed if its graph $\mathcal{G}(T)$ is a closed linear subspace of $X \times Y$

Theorem 11.6.11 — Closed Graph Theorem. Let X and Y be Banach spaces and $T : X \to Y$ a closed linear operator. Then T is bounded

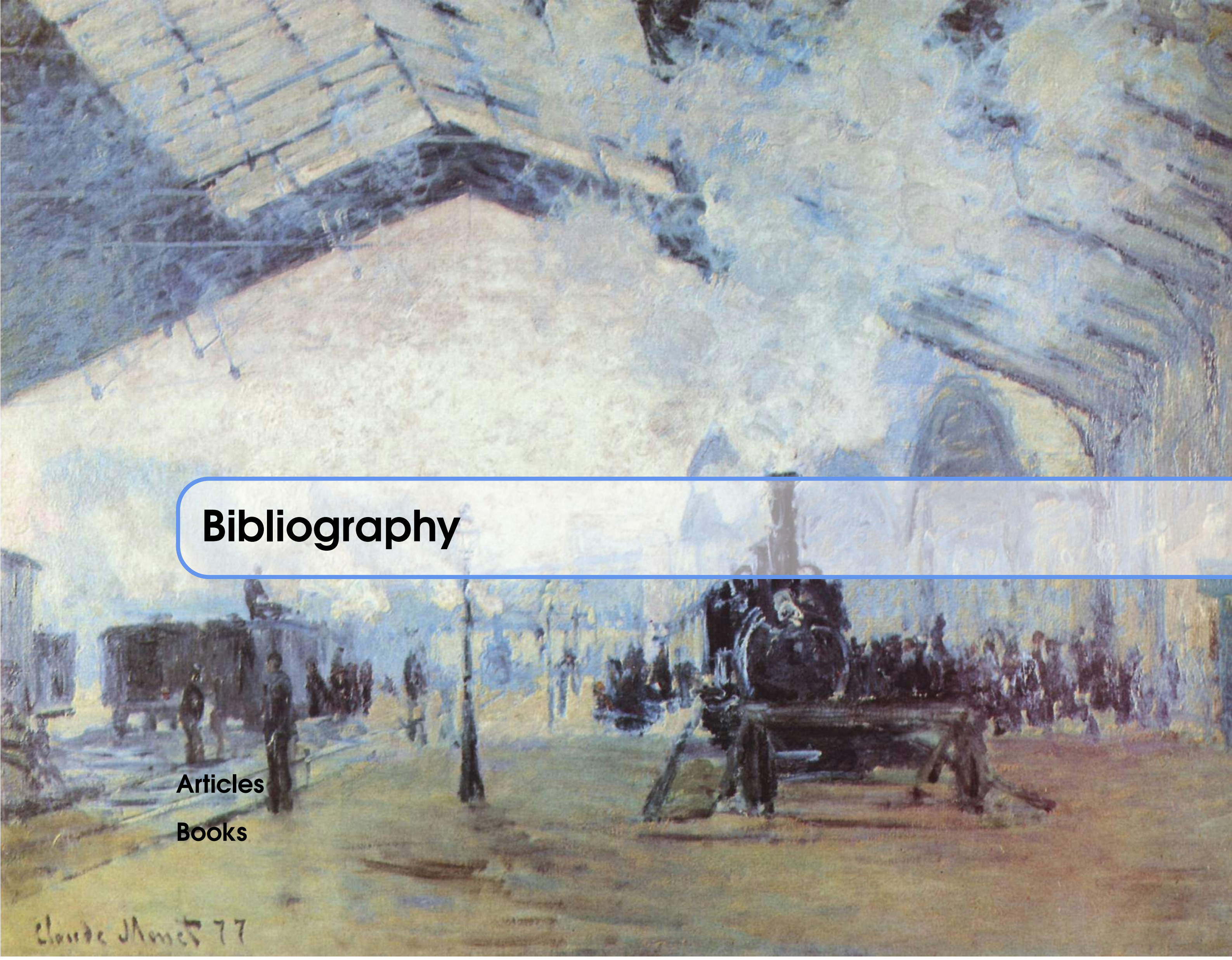
Proof. Since $X \times Y$, with the norm defined above, is a Banach space, and by the hypothesis $\mathcal{G}(T)$ is closed it follows that $\mathcal{G}(T)$ is also a Banach space. Consider the map $P : \mathcal{G}(T) \to X$ given by $P(x, Tx) = x$ Then P is linear and bijective. It is also bounded since

$$\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|$$

That is, P is bounded and $\|P\| \leq 1$. By Banach's Theorem (Corollary 6.3 .4), it follows that $P^{-1} : X \to \mathcal{G}(T)$ given by $P^{-1}x = (x, Tx)$ for $x \in X$, is also bounded. Hence $\|(x, Tx)\| = \|P^{-1}x\| \leq \|P^{-1}\| \|x\|$ Therefore

$$\|(x, Tx)\| = \|x\| + \|Tx\| \leq \left\|P^{-1}\right\| \|x\| \iff \|Tx\| \leq \left\|P^{-1}\right\| \|x\|$$

That is, T is bounded and $\|T\| \leq \|P^{-1}\|$ ■



Bibliography

Articles

Books