6

Metric Spaces

Definition 6.1. *A function $d : X \times X \to [0, \infty)$ is called a metric if*

1. *(Symmetry) $d(x, y) = d(y, x)$ for all $x, y \in X$*
2. *(Non-degenerate) $d(x, y) = 0$ if and only if $x = y \in X$*
3. *(Triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.*

As primary examples, any normed space $(X, \|\cdot\|)$ (see Definition 5.1) is a metric space with $d(x, y) := \|x - y\|$. Thus the space $\ell^p(\mu)$ (as in Theorem 5.2) is a metric space for all $p \in [1, \infty]$. Also any subset of a metric space is a metric space. For example a surface Σ in $\mathbb{R}^3$ is a metric space with the distance between two points on Σ being the usual distance in $\mathbb{R}^3$.

Definition 6.2. *Let (X, d) be a metric space. The* ***open ball*** *$B(x, \delta) \subset X$ centered at $x \in X$ with radius $\delta > 0$ is the set*

$$B(x, \delta) := \{y \in X : d(x, y) < \delta\}.$$

We will often also write $B(x, \delta)$ as $B_x(\delta)$. We also define the ***closed ball*** *centered at $x \in X$ with radius $\delta > 0$ as the set $C_x(\delta) := \{y \in X : d(x, y) \leq \delta\}$.*

Definition 6.3. *A sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space (X, d) is said to be* ***convergent*** *if there exists a point $x \in X$ such that $\lim_{n\to\infty} d(x, x_n) = 0$. In this case we write $\lim_{n\to\infty} x_n = x$ of $x_n \to x$ as $n \to \infty$.*

Exercise 6.1. Show that x in Definition 6.3 is necessarily unique.

Definition 6.4. *A set $E \subset X$ is* ***bounded*** *if $E \subset B(x, R)$ for some $x \in X$ and $R < \infty$. A set $F \subset X$ is* ***closed*** *iff every convergent sequence $\{x_n\}_{n=1}^{\infty}$ which is contained in F has its limit back in F. A set $V \subset X$ is* ***open*** *iff V^c is closed. We will write $F \sqsubset X$ to indicate the F is a closed subset of X and $V \subset_o X$ to indicate the V is an open subset of X. We also let τ_d denote the collection of open subsets of X relative to the metric d.*

Definition 6.5. *A subset $A \subset X$ is a **neighborhood** of x if there exists an open set $V \subset_o X$ such that $x \in V \subset A$. We will say that $A \subset X$ is an **open neighborhood** of x if A is open and $x \in A$.*

Exercise 6.2. Let $\mathcal{F}$ be a collection of closed subsets of X, show $\cap\mathcal{F} := \cap_{F\in\mathcal{F}} F$ is closed. Also show that finite unions of closed sets are closed, i.e. if $\{F_k\}_{k=1}^n$ are closed sets then $\cup_{k=1}^n F_k$ is closed. (By taking complements, this shows that the collection of open sets, τ_d, is closed under finite intersections and arbitrary unions.)

The following "continuity" facts of the metric d will be used frequently in the remainder of this book.

Lemma 6.6. *For any non empty subset $A \subset X$, let $d_A(x) := \inf\{d(x,a)|a \in A\}$, then*

$$|d_A(x) - d_A(y)| \le d(x,y)\ \forall\, x,y \in X \tag{6.1}$$

and in particular if $x_n \to x$ in X then $d_A(x_n) \to d_A(x)$ as $n \to \infty$. Moreover the set $F_\varepsilon := \{x \in X | d_A(x) \ge \varepsilon\}$ is closed in X.

Proof. Let $a \in A$ and $x, y \in X$, then

$$d(x,a) \le d(x,y) + d(y,a).$$

Take the inf over a in the above equation shows that

$$d_A(x) \le d(x,y) + d_A(y) \quad \forall x,y \in X.$$

Therefore, $d_A(x) - d_A(y) \le d(x,y)$ and by interchanging x and y we also have that $d_A(y) - d_A(x) \le d(x,y)$ which implies Eq. (6.1). If $x_n \to x \in X$, then by Eq. (6.1),

$$|d_A(x) - d_A(x_n)| \le d(x,x_n) \to 0 \text{ as } n \to \infty$$

so that $\lim_{n\to\infty} d_A(x_n) = d_A(x)$. Now suppose that $\{x_n\}_{n=1}^\infty \subset F_\varepsilon$ and $x_n \to x$ in X, then

$$d_A(x) = \lim_{n\to\infty} d_A(x_n) \ge \varepsilon$$

since $d_A(x_n) \ge \varepsilon$ for all n. This shows that $x \in F_\varepsilon$ and hence F_ε is closed. ∎

Corollary 6.7. *The function d satisfies,*

$$|d(x,y) - d(x',y')| \le d(y,y') + d(x,x').$$

In particular $d : X \times X \to [0,\infty)$ is "continuous" in the sense that $d(x,y)$ is close to $d(x',y')$ if x is close to x' and y is close to y'. (The notion of continuity will be developed shortly.)

Proof. By Lemma 6.6 for single point sets and the triangle inequality for the absolute value of real numbers,

$$\begin{aligned}|d(x,y)-d(x',y')| &\le |d(x,y)-d(x,y')|+|d(x,y')-d(x',y')| \\ &\le d(y,y')+d(x,x').\end{aligned}$$

■

Example 6.8. Let $x \in X$ and $\delta > 0$, then $C_x(\delta)$ and $B_x(\delta)^c$ are closed subsets of X. For example if $\{y_n\}_{n=1}^{\infty} \subset C_x(\delta)$ and $y_n \to y \in X$, then $d(y_n, x) \le \delta$ for all n and using Corollary 6.7 it follows $d(y,x) \le \delta$, i.e. $y \in C_x(\delta)$. A similar proof shows $B_x(\delta)^c$ is open, see Exercise 6.3.

Exercise 6.3. Show that $V \subset X$ is open iff for every $x \in V$ there is a $\delta > 0$ such that $B_x(\delta) \subset V$. In particular show $B_x(\delta)$ is open for all $x \in X$ and $\delta > 0$. **Hint:** by definition V is not open iff V^c is not closed.

Lemma 6.9 (Approximating open sets from the inside by closed sets). *Let A be a closed subset of X and $F_\varepsilon := \{x \in X | d_A(x) \ge \varepsilon\} \sqsubset X$ be as in Lemma 6.6. Then $F_\varepsilon \uparrow A^c$ as $\varepsilon \downarrow 0$.*

Proof. It is clear that $d_A(x) = 0$ for $x \in A$ so that $F_\varepsilon \subset A^c$ for each $\varepsilon > 0$ and hence $\cup_{\varepsilon>0} F_\varepsilon \subset A^c$. Now suppose that $x \in A^c \subset_o X$. By Exercise 6.3 there exists an $\varepsilon > 0$ such that $B_x(\varepsilon) \subset A^c$, i.e. $d(x,y) \ge \varepsilon$ for all $y \in A$. Hence $x \in F_\varepsilon$ and we have shown that $A^c \subset \cup_{\varepsilon>0} F_\varepsilon$. Finally it is clear that $F_\varepsilon \subset F_{\varepsilon'}$ whenever $\varepsilon' \le \varepsilon$. ■

Definition 6.10. *Given a set A contained a metric space X, let $\bar{A} \subset X$ be the **closure of** A defined by*

$$\bar{A} := \{x \in X : \exists \{x_n\} \subset A \ni x = \lim_{n\to\infty} x_n\}.$$

*That is to say $\bar{A}$ contains all **limit points** of A. We say A is **dense in** X if $\bar{A} = X$, i.e. every element $x \in X$ is a limit of a sequence of elements from A.*

Exercise 6.4. Given $A \subset X$, show $\bar{A}$ is a closed set and in fact

$$\bar{A} = \cap\{F : A \subset F \subset X \text{ with } F \text{ closed}\}. \tag{6.2}$$

That is to say $\bar{A}$ is the smallest closed set containing A.

6.1 Continuity

Suppose that (X,ρ) and (Y,d) are two metric spaces and $f : X \to Y$ is a function.

Definition 6.11. *A function $f : X \to Y$ is continuous at $x \in X$ if for all $\varepsilon > 0$ there is a $\delta > 0$ such that*

$$d(f(x), f(x')) < \varepsilon \text{ provided that } \rho(x, x') < \delta. \tag{6.3}$$

The function f is said to be continuous if f is continuous at all points $x \in X$.

The following lemma gives two other characterizations of continuity of a function at a point.

Lemma 6.12 (Local Continuity Lemma). *Suppose that (X, ρ) and (Y, d) are two metric spaces and $f : X \to Y$ is a function defined in a neighborhood of a point $x \in X$. Then the following are equivalent:*

1. *f is continuous at $x \in X$.*
2. *For all neighborhoods $A \subset Y$ of $f(x)$, $f^{-1}(A)$ is a neighborhood of $x \in X$.*
3. *For all sequences $\{x_n\}_{n=1}^{\infty} \subset X$ such that $x = \lim_{n\to\infty} x_n$, $\{f(x_n)\}$ is convergent in Y and*

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

Proof. $1 \implies 2$. If $A \subset Y$ is a neighborhood of $f(x)$, there exists $\varepsilon > 0$ such that $B_{f(x)}(\varepsilon) \subset A$ and because f is continuous there exists a $\delta > 0$ such that Eq. (6.3) holds. Therefore

$$B_x(\delta) \subset f^{-1}\left(B_{f(x)}(\varepsilon)\right) \subset f^{-1}(A)$$

showing $f^{-1}(A)$ is a neighborhood of x.

$2 \implies 3$. Suppose that $\{x_n\}_{n=1}^{\infty} \subset X$ and $x = \lim_{n\to\infty} x_n$. Then for any $\varepsilon > 0$, $B_{f(x)}(\varepsilon)$ is a neighborhood of $f(x)$ and so $f^{-1}\left(B_{f(x)}(\varepsilon)\right)$ is a neighborhood of x which must containing $B_x(\delta)$ for some $\delta > 0$. Because $x_n \to x$, it follows that $x_n \in B_x(\delta) \subset f^{-1}\left(B_{f(x)}(\varepsilon)\right)$ for a.a. n and this implies $f(x_n) \in B_{f(x)}(\varepsilon)$ for a.a. n, i.e. $d(f(x), f(x_n)) < \varepsilon$ for a.a. n. Since $\varepsilon > 0$ is arbitrary it follows that $\lim_{n\to\infty} f(x_n) = f(x)$.

$3. \implies 1$. We will show not $1. \implies$ not 3. If f is not continuous at x, there exists an $\varepsilon > 0$ such that for all $n \in \mathbb{N}$ there exists a point $x_n \in X$ with $\rho(x_n, x) < \frac{1}{n}$ yet $d(f(x_n), f(x)) \geq \varepsilon$. Hence $x_n \to x$ as $n \to \infty$ yet $f(x_n)$ does not converge to $f(x)$. ■

Here is a global version of the previous lemma.

Lemma 6.13 (Global Continuity Lemma). *Suppose that (X, ρ) and (Y, d) are two metric spaces and $f : X \to Y$ is a function defined on all of X. Then the following are equivalent:*

1. *f is continuous.*
2. *$f^{-1}(V) \in \tau_\rho$ for all $V \in \tau_d$, i.e. $f^{-1}(V)$ is open in X if V is open in Y.*
3. *$f^{-1}(C)$ is closed in X if C is closed in Y.*

4. For all convergent sequences $\{x_n\} \subset X$, $\{f(x_n)\}$ is convergent in Y and

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

Proof. Since $f^{-1}(A^c) = \left[f^{-1}(A)\right]^c$, it is easily seen that 2. and 3. are equivalent. So because of Lemma 6.12 it only remains to show 1. and 2. are equivalent. If f is continuous and $V \subset Y$ is open, then for every $x \in f^{-1}(V)$, V is a neighborhood of $f(x)$ and so $f^{-1}(V)$ is a neighborhood of x. Hence $f^{-1}(V)$ is a neighborhood of all of its points and from this and Exercise 6.3 it follows that $f^{-1}(V)$ is open. Conversely if $x \in X$ and $A \subset Y$ is a neighborhood of $f(x)$, then there exists $V \subset_o X$ such that $f(x) \in V \subset A$. Hence $x \in f^{-1}(V) \subset f^{-1}(A)$ and by assumption $f^{-1}(V)$ is open showing $f^{-1}(A)$ is a neighborhood of x. Therefore f is continuous at x and since $x \in X$ was arbitrary, f is continuous. ■

Example 6.14. The function d_A defined in Lemma 6.6 is continuous for each $A \subset X$. In particular, if $A = \{x\}$, it follows that $y \in X \to d(y,x)$ is continuous for each $x \in X$.

Exercise 6.5. Use Example 6.14 and Lemma 6.13 to recover the results of Example 6.8.

The next result shows that there are lots of continuous functions on a metric space (X, d).

Lemma 6.15 (Urysohn's Lemma for Metric Spaces). *Let (X, d) be a metric space and suppose that A and B are two disjoint closed subsets of X. Then*

$$f(x) = \frac{d_B(x)}{d_A(x) + d_B(x)} \text{ for } x \in X \tag{6.4}$$

defines a continuous function, $f : X \to [0,1]$, such that $f(x) = 1$ for $x \in A$ and $f(x) = 0$ if $x \in B$.

Proof. By Lemma 6.6, d_A and d_B are continuous functions on X. Since A and B are closed, $d_A(x) > 0$ if $x \notin A$ and $d_B(x) > 0$ if $x \notin B$. Since $A \cap B = \emptyset$, $d_A(x) + d_B(x) > 0$ for all x and $(d_A + d_B)^{-1}$ is continuous as well. The remaining assertions about f are all easy to verify. ■

Sometimes Urysohn's lemma will be use in the following form. Suppose $F \subset V \subset X$ with F being closed and V being open, then there exists $f \in C(X, [0,1]))$ such that $f = 1$ on F while $f = 0$ on V^c. This of course follows from Lemma 6.15 by taking $A = F$ and $B = V^c$.

6.2 Completeness in Metric Spaces

Definition 6.16 (Cauchy sequences). *A sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space (X, d) is **Cauchy** provided that*

$$\lim_{m,n\to\infty} d(x_n, x_m) = 0.$$

Exercise 6.6. Show that convergent sequences are always Cauchy sequences. The converse is not always true. For example, let $X = \mathbb{Q}$ be the set of rational numbers and $d(x,y) = |x-y|$. Choose a sequence $\{x_n\}_{n=1}^\infty \subset \mathbb{Q}$ which converges to $\sqrt{2} \in \mathbb{R}$, then $\{x_n\}_{n=1}^\infty$ is $(\mathbb{Q}, d)$ – Cauchy but not $(\mathbb{Q}, d)$ – convergent. The sequence does converge in $\mathbb{R}$ however.

Definition 6.17. *A metric space (X, d) is* ***complete*** *if all Cauchy sequences are convergent sequences.*

Exercise 6.7. Let (X, d) be a complete metric space. Let $A \subset X$ be a subset of X viewed as a metric space using $d|_{A\times A}$. Show that $(A, d|_{A\times A})$ is complete iff A is a closed subset of X.

Example 6.18. Examples 2. – 4. of complete metric spaces will be verified in Chapter 7 below.

1. $X = \mathbb{R}$ and $d(x,y) = |x-y|$, see Theorem 3.8 above.
2. $X = \mathbb{R}^n$ and $d(x,y) = \|x-y\|_2 = \sum_{i=1}^n (x_i - y_i)^2$.
3. $X = \ell^p(\mu)$ for $p \in [1, \infty]$ and any weight function $\mu : X \to (0, \infty)$.
4. $X = C([0,1], \mathbb{R})$ – the space of continuous functions from $[0,1]$ to $\mathbb{R}$ and
$$d(f,g) := \max_{t\in[0,1]} |f(t) - g(t)|.$$
This is a special case of Lemma 7.3 below.
5. Let $X = C([0,1], \mathbb{R})$ and
$$d(f,g) := \int_0^1 |f(t) - g(t)|\, dt.$$
You are asked in Exercise 7.14 to verify that (X, d) is a metric space which is **not** complete.

Exercise 6.8 (Completions of Metric Spaces). Suppose that (X, d) is a (not necessarily complete) metric space. Using the following outline show there exists a complete metric space $\left(\bar{X}, \bar{d}\right)$ and an isometric map $i : X \to \bar{X}$ such that $i(X)$ is dense in $\bar{X}$, see Definition 6.10.

1. Let $\mathcal{C}$ denote the collection of Cauchy sequences $a = \{a_n\}_{n=1}^\infty \subset X$. Given two element $a, b \in \mathcal{C}$ show
$$d_{\mathcal{C}}(a,b) := \lim_{n\to\infty} d(a_n, b_n) \text{ exists,}$$
$d_{\mathcal{C}}(a,b) \geq 0$ for all $a, b \in \mathcal{C}$ and $d_{\mathcal{C}}$ satisfies the triangle inequality,
$$d_{\mathcal{C}}(a,c) \leq d_{\mathcal{C}}(a,b) + d_{\mathcal{C}}(b,c) \text{ for all } a, b, c \in \mathcal{C}.$$
Thus $(\mathcal{C}, d_{\mathcal{C}})$ would be a metric space if it were true that $d_{\mathcal{C}}(a,b) = 0$ iff $a = b$. This however is false, for example if $a_n = b_n$ for all $n \geq 100$, then $d_{\mathcal{C}}(a,b) = 0$ while a need not equal b.

2. Define two elements $a, b \in \mathcal{C}$ to be equivalent (write $a \sim b$) whenever $d_{\mathcal{C}}(a, b) = 0$. Show " $\sim$ " is an equivalence relation on $\mathcal{C}$ and that $d_{\mathcal{C}}(a', b') = d_{\mathcal{C}}(a, b)$ if $a \sim a'$ and $b \sim b'$. (**Hint:** see Corollary 6.7.)
3. Given $a \in \mathcal{C}$ let $\bar{a} := \{b \in \mathcal{C} : b \sim a\}$ denote the equivalence class containing a and let $\bar{X} := \{\bar{a} : a \in \mathcal{C}\}$ denote the collection of such equivalence classes. Show that $\bar{d}\left(\bar{a}, \bar{b}\right) := d_{\mathcal{C}}(a, b)$ is well defined on $\bar{X} \times \bar{X}$ and verify $\left(\bar{X}, \bar{d}\right)$ is a metric space.
4. For $x \in X$ let $i(x) = \bar{a}$ where a is the constant sequence, $a_n = x$ for all n. Verify that $i : X \to \bar{X}$ is an isometric map and that $i(X)$ is dense in $\bar{X}$.
5. Verify $\left(\bar{X}, \bar{d}\right)$ is complete. **Hint:** if $\{\bar{a}(m)\}_{m=1}^{\infty}$ is a Cauchy sequence in $\bar{X}$ choose $b_m \in X$ such that $\bar{d}(i(b_m), \bar{a}(m)) \le 1/m$. Then show $\bar{a}(m) \to \bar{b}$ where $b = \{b_m\}_{m=1}^{\infty}$.

6.3 Supplementary Remarks

6.3.1 Word of Caution

Example 6.19. Let (X, d) be a metric space. It is always true that $\overline{B_x(\varepsilon)} \subset C_x(\varepsilon)$ since $C_x(\varepsilon)$ is a closed set containing $B_x(\varepsilon)$. However, it is not always true that $\overline{B_x(\varepsilon)} = C_x(\varepsilon)$. For example let $X = \{1, 2\}$ and $d(1, 2) = 1$, then $B_1(1) = \{1\}$, $\overline{B_1(1)} = \{1\}$ while $C_1(1) = X$. For another counter example, take

$$X = \left\{(x, y) \in \mathbb{R}^2 : x = 0 \text{ or } x = 1\right\}$$

with the usually Euclidean metric coming from the plane. Then

$$\begin{aligned} B_{(0,0)}(1) &= \left\{(0, y) \in \mathbb{R}^2 : |y| < 1\right\}, \\ \overline{B_{(0,0)}(1)} &= \left\{(0, y) \in \mathbb{R}^2 : |y| \le 1\right\}, \text{ while} \\ C_{(0,0)}(1) &= \overline{B_{(0,0)}(1)} \cup \{(0,1)\}. \end{aligned}$$

In spite of the above examples, Lemmas 6.20 and 6.21 below shows that for certain metric spaces of interest it is true that $\overline{B_x(\varepsilon)} = C_x(\varepsilon)$.

Lemma 6.20. *Suppose that $(X, |\cdot|)$ is a normed vector space and d is the metric on X defined by $d(x, y) = |x - y|$. Then*

$$\begin{aligned} \overline{B_x(\varepsilon)} &= C_x(\varepsilon) \text{ and} \\ \mathrm{bd}(B_x(\varepsilon)) &= \{y \in X : d(x, y) = \varepsilon\}. \end{aligned}$$

where the boundary operation, $\mathrm{bd}(\cdot)$ *is defined in Definition 10.30 below.*

Proof. We must show that $C := C_x(\varepsilon) \subset \overline{B_x(\varepsilon)} =: \bar{B}$. For $y \in C$, let $v = y - x$, then

$$|v| = |y - x| = d(x, y) \le \varepsilon.$$

Let $\alpha_n = 1 - 1/n$ so that $\alpha_n \uparrow 1$ as $n \to \infty$. Let $y_n = x + \alpha_n v$, then $d(x, y_n) = \alpha_n d(x, y) < \varepsilon$, so that $y_n \in B_x(\varepsilon)$ and $d(y, y_n) = 1 - \alpha_n \to 0$ as $n \to \infty$. This shows that $y_n \to y$ as $n \to \infty$ and hence that $y \in \bar{B}$. ■

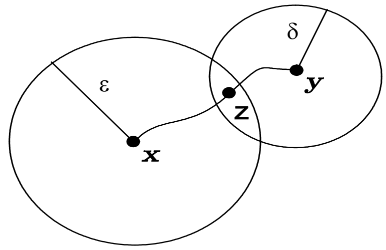

Fig. 6.1. An almost length minimizing curve joining x to y.

6.3.2 Riemannian Metrics

This subsection is not completely self contained and may safely be skipped.

Lemma 6.21. *Suppose that X is a Riemannian (or sub-Riemannian) manifold and d is the metric on X defined by*

$$d(x,y) = \inf\{\ell(\sigma) : \sigma(0) = x \text{ and } \sigma(1) = y\}$$

where $\ell(\sigma)$ is the length of the curve σ. We define $\ell(\sigma) = \infty$ if σ is not piecewise smooth.

Then

$$\overline{B_x(\varepsilon)} = C_x(\varepsilon) \text{ and}$$
$$\mathrm{bd}(B_x(\varepsilon)) = \{y \in X : d(x,y) = \varepsilon\}$$

where the boundary operation, $\mathrm{bd}(\cdot)$ *is defined in Definition 10.30 below.*

Proof. Let $C := C_x(\varepsilon) \subset \overline{B_x(\varepsilon)} =: \bar{B}$. We will show that $C \subset \bar{B}$ by showing $\bar{B}^c \subset C^c$. Suppose that $y \in \bar{B}^c$ and choose $\delta > 0$ such that $B_y(\delta) \cap \bar{B} = \emptyset$. In particular this implies that

$$B_y(\delta) \cap B_x(\varepsilon) = \emptyset.$$

We will finish the proof by showing that $d(x,y) \geq \varepsilon + \delta > \varepsilon$ and hence that $y \in C^c$. This will be accomplished by showing: if $d(x,y) < \varepsilon + \delta$ then $B_y(\delta) \cap B_x(\varepsilon) \neq \emptyset$.

If $d(x,y) < \max(\varepsilon,\delta)$ then either $x \in B_y(\delta)$ or $y \in B_x(\varepsilon)$. In either case $B_y(\delta) \cap B_x(\varepsilon) \neq \emptyset$. Hence we may assume that $\max(\varepsilon,\delta) \leq d(x,y) < \varepsilon + \delta$. Let $\alpha > 0$ be a number such that

$$\max(\varepsilon,\delta) \leq d(x,y) < \alpha < \varepsilon + \delta$$

and choose a curve σ from x to y such that $\ell(\sigma) < \alpha$. Also choose $0 < \delta' < \delta$ such that $0 < \alpha - \delta' < \varepsilon$ which can be done since $\alpha - \delta < \varepsilon$. Let $k(t) = d(y, \sigma(t))$ a continuous function on $[0,1]$ and therefore $k([0,1]) \subset \mathbb{R}$ is a connected set which contains 0 and $d(x,y)$. Therefore there exists $t_0 \in [0,1]$ such that $d(y,\sigma(t_0)) = k(t_0) = \delta'$. Let $z = \sigma(t_0) \in B_y(\delta)$ then

$$d(x,z) \le \ell(\sigma|_{[0,t_0]}) = \ell(\sigma) - \ell(\sigma|_{[t_0,1]}) < \alpha - d(z,y) = \alpha - \delta' < \varepsilon$$

and therefore $z \in B_x(\varepsilon) \cap B_x(\delta) \neq \emptyset$. ■

Remark 6.22. Suppose again that X is a Riemannian (or sub-Riemannian) manifold and

$$d(x,y) = \inf\{\ell(\sigma) : \sigma(0) = x \text{ and } \sigma(1) = y\}.$$

Let σ be a curve from x to y and let $\varepsilon = \ell(\sigma) - d(x,y)$. Then for all $0 \le u < v \le 1$,

$$d(\sigma(u),\sigma(v)) \le \ell(\sigma|_{[u,v]}) + \varepsilon.$$

So if σ is within ε of a length minimizing curve from x to y that $\sigma|_{[u,v]}$ is within ε of a length minimizing curve from $\sigma(u)$ to $\sigma(v)$. In particular if $d(x,y) = \ell(\sigma)$ then $d(\sigma(u),\sigma(v)) = \ell(\sigma|_{[u,v]})$ for all $0 \le u < v \le 1$, i.e. if σ is a length minimizing curve from x to y that $\sigma|_{[u,v]}$ is a length minimizing curve from $\sigma(u)$ to $\sigma(v)$.

To prove these assertions notice that

$$\begin{aligned} d(x,y) + \varepsilon = \ell(\sigma) &= \ell(\sigma|_{[0,u]}) + \ell(\sigma|_{[u,v]}) + \ell(\sigma|_{[v,1]}) \\ &\ge d(x,\sigma(u)) + \ell(\sigma|_{[u,v]}) + d(\sigma(v),y) \end{aligned}$$

and therefore

$$\begin{aligned} \ell(\sigma|_{[u,v]}) &\le d(x,y) + \varepsilon - d(x,\sigma(u)) - d(\sigma(v),y) \\ &\le d(\sigma(u),\sigma(v)) + \varepsilon. \end{aligned}$$

6.4 Exercises

Exercise 6.9. Let (X,d) be a metric space. Suppose that $\{x_n\}_{n=1}^\infty \subset X$ is a sequence and set $\varepsilon_n := d(x_n, x_{n+1})$. Show that for $m > n$ that

$$d(x_n, x_m) \le \sum_{k=n}^{m-1} \varepsilon_k \le \sum_{k=n}^{\infty} \varepsilon_k.$$

Conclude from this that if

$$\sum_{k=1}^{\infty} \varepsilon_k = \sum_{n=1}^{\infty} d(x_n, x_{n+1}) < \infty$$

then $\{x_n\}_{n=1}^\infty$ is Cauchy. Moreover, show that if $\{x_n\}_{n=1}^\infty$ is a convergent sequence and $x = \lim_{n\to\infty} x_n$ then

$$d(x, x_n) \le \sum_{k=n}^{\infty} \varepsilon_k.$$

Exercise 6.10. Show that (X,d) is a complete metric space iff every sequence $\{x_n\}_{n=1}^\infty \subset X$ such that $\sum_{n=1}^\infty d(x_n, x_{n+1}) < \infty$ is a convergent sequence in X. You may find it useful to prove the following statements in the course of the proof.

1. If $\{x_n\}$ is Cauchy sequence, then there is a subsequence $y_j := x_{n_j}$ such that $\sum_{j=1}^\infty d(y_{j+1}, y_j) < \infty$.
2. If $\{x_n\}_{n=1}^\infty$ is Cauchy and there exists a subsequence $y_j := x_{n_j}$ of $\{x_n\}$ such that $x = \lim_{j\to\infty} y_j$ exists, then $\lim_{n\to\infty} x_n$ also exists and is equal to x.

Exercise 6.11. Suppose that $f : [0,\infty) \to [0,\infty)$ is a C^2 – function such that $f(0) = 0$, $f' > 0$ and $f'' \le 0$ and (X,ρ) is a metric space. Show that $d(x,y) = f(\rho(x,y))$ is a metric on X. In particular show that

$$d(x,y) := \frac{\rho(x,y)}{1+\rho(x,y)}$$

is a metric on X. (Hint: use calculus to verify that $f(a+b) \le f(a) + f(b)$ for all $a, b \in [0,\infty)$.)

Exercise 6.12. Let $\{(X_n, d_n)\}_{n=1}^\infty$ be a sequence of metric spaces, $X := \prod_{n=1}^\infty X_n$, and for $x = (x(n))_{n=1}^\infty$ and $y = (y(n))_{n=1}^\infty$ in X let

$$d(x,y) = \sum_{n=1}^\infty 2^{-n} \frac{d_n(x(n), y(n))}{1 + d_n(x(n), y(n))}.$$

Show:

1. (X,d) is a metric space,
2. a sequence $\{x_k\}_{k=1}^\infty \subset X$ converges to $x \in X$ iff $x_k(n) \to x(n) \in X_n$ as $k \to \infty$ for each $n \in \mathbb{N}$ and
3. X is complete if X_n is complete for all n.

Exercise 6.13. Suppose (X,ρ) and (Y,d) are metric spaces and A is a dense subset of X.

1. Show that if $F : X \to Y$ and $G : X \to Y$ are two continuous functions such that $F = G$ on A then $F = G$ on X. **Hint:** consider the set $C := \{x \in X : F(x) = G(x)\}$.
2. Suppose $f : A \to Y$ is a function which is uniformly continuous, i.e. for every $\varepsilon > 0$ there exists a $\delta > 0$ such that
$$d(f(a), f(b)) < \varepsilon \text{ for all } a, b \in A \text{ with } \rho(a,b) < \delta.$$
Show there is a unique continuous function $F : X \to Y$ such that $F = f$ on A. **Hint:** each point $x \in X$ is a limit of a sequence consisting of elements from A.
3. Let $X = \mathbb{R} = Y$ and $A = \mathbb{Q} \subset X$, find a function $f : \mathbb{Q} \to \mathbb{R}$ which is continuous on $\mathbb{Q}$ but does **not** extend to a continuous function on $\mathbb{R}$.

7

Banach Spaces

Let $(X, \|\cdot\|)$ be a normed vector space and $d(x, y) := \|x - y\|$ be the associated metric on X. We say $\{x_n\}_{n=1}^{\infty} \subset X$ **converges to** $x \in X$ (and write $\lim_{n\to\infty} x_n = x$ or $x_n \to x$) if

$$0 = \lim_{n\to\infty} d(x, x_n) = \lim_{n\to\infty} \|x - x_n\|.$$

Similarly $\{x_n\}_{n=1}^{\infty} \subset X$ is said to be a **Cauchy sequence** if

$$0 = \lim_{m,n\to\infty} d(x_m, x_n) = \lim_{m,n\to\infty} \|x_m - x_n\|.$$

Definition 7.1 (Banach space). *A normed vector space $(X, \|\cdot\|)$ is a **Banach space** if the associated metric space (X, d) is complete, i.e. all Cauchy sequences are convergent.*

Remark 7.2. Since $\|x\| = d(x, 0)$, it follows from Lemma 6.6 that $\|\cdot\|$ is a continuous function on X and that

$$|\|x\| - \|y\|| \leq \|x - y\| \text{ for all } x, y \in X.$$

It is also easily seen that the vector addition and scalar multiplication are continuos on any normed space as the reader is asked to verify in Exercise 7.7. These facts will often be used in the sequel without further mention.

7.1 Examples

Lemma 7.3. *Suppose that X is a set then the bounded functions, $\ell^\infty(X)$, on X is a Banach space with the norm*

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

Moreover if X is a metric space (more generally a topological space, see Chapter 10) the set $BC(X) \subset \ell^\infty(X) = B(X)$ is closed subspace of $\ell^\infty(X)$ and hence is also a Banach space.

Proof. Let $\{f_n\}_{n=1}^\infty \subset \ell^\infty(X)$ be a Cauchy sequence. Since for any $x \in X$, we have

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty \tag{7.1}$$

which shows that $\{f_n(x)\}_{n=1}^\infty \subset \mathbb{F}$ is a Cauchy sequence of numbers. Because $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) is complete, $f(x) := \lim_{n\to\infty} f_n(x)$ exists for all $x \in X$. Passing to the limit $n \to \infty$ in Eq. (7.1) implies

$$|f(x) - f_m(x)| \le \lim\inf_{n\to\infty} \|f_n - f_m\|_\infty$$

and taking the supremum over $x \in X$ of this inequality implies

$$\|f - f_m\|_\infty \le \lim\inf_{n\to\infty} \|f_n - f_m\|_\infty \to 0 \text{ as } m \to \infty$$

showing $f_m \to f$ in $\ell^\infty(X)$.

For the second assertion, suppose that $\{f_n\}_{n=1}^\infty \subset BC(X) \subset \ell^\infty(X)$ and $f_n \to f \in \ell^\infty(X)$. We must show that $f \in BC(X)$, i.e. that f is continuous. To this end let $x, y \in X$, then

$$\begin{aligned}|f(x) - f(y)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\le 2\|f - f_n\|_\infty + |f_n(x) - f_n(y)|\,.\end{aligned}$$

Thus if $\varepsilon > 0$, we may choose n large so that $2\|f - f_n\|_\infty < \varepsilon/2$ and then for this n there exists an open neighborhood V_x of $x \in X$ such that $|f_n(x) - f_n(y)| < \varepsilon/2$ for $y \in V_x$. Thus $|f(x) - f(y)| < \varepsilon$ for $y \in V_x$ showing the limiting function f is continuous. ■

Here is an application of this theorem.

Theorem 7.4 (Metric Space Tietze Extension Theorem). *Let (X,d) be a metric space, D be a closed subset of X, $-\infty < a < b < \infty$ and $f \in C(D,[a,b])$. (Here we are viewing D as a metric space with metric $d_D := d_{D\times D}$.) Then there exists $F \in C(X,[a,b])$ such that $F|_D = f$.*

Proof. 1. By scaling and translation (i.e. by replacing f by $(b-a)^{-1}(f-a)$), it suffices to prove Theorem 7.4 with $a = 0$ and $b = 1$.

2. Suppose $\alpha \in (0,1]$ and $f : D \to [0,\alpha]$ is continuous function. Let $A := f^{-1}([0,\frac{1}{3}\alpha])$ and $B := f^{-1}([\frac{2}{3}\alpha,\alpha])$. By Lemma 6.15 there exists a function $\tilde{g} \in C(X,[0,\alpha/3])$ such that $\tilde{g} = 0$ on A and $\tilde{g} = 1$ on B. Letting $g := \frac{\alpha}{3}\tilde{g}$, we have $g \in C(X,[0,\alpha/3])$ such that $g = 0$ on A and $g = \alpha/3$ on B. Further notice that
$$0 \le f(x) - g(x) \le \frac{2}{3}\alpha \text{ for all } x \in D.$$

3. Now suppose $f : D \to [0,1]$ is a continuous function as in step 1. Let $g_1 \in C(X,[0,1/3])$ be as in step 2, see Figure 7.1. with $\alpha = 1$ and let $f_1 := f - g_1|_D \in C(D,[0,2/3])$. Apply step 2. with $\alpha = 2/3$ and $f = f_1$ to

find $g_2 \in C(X, [0, \frac{1}{3}\frac{2}{3}])$ such that $f_2 := f - (g_1 + g_2)|_D \in C(D, [0, \left(\frac{2}{3}\right)^2])$. Continue this way inductively to find $g_n \in C(X, [0, \frac{1}{3}\left(\frac{2}{3}\right)^{n-1}])$ such that

$$f - \sum_{n=1}^{N} g_n|_D =: f_N \in C(D, [0, \left(\frac{2}{3}\right)^N]). \tag{7.2}$$

4. Define $F := \sum_{n=1}^{\infty} g_n$. Since

$$\sum_{n=1}^{\infty} \|g_n\|_\infty \le \sum_{n=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{n-1} = \frac{1}{3}\frac{1}{1-\frac{2}{3}} = 1,$$

the series defining F is uniformly convergent so $F \in C(X, [0,1])$. Passing to the limit in Eq. (7.2) shows $f = F|_D$.

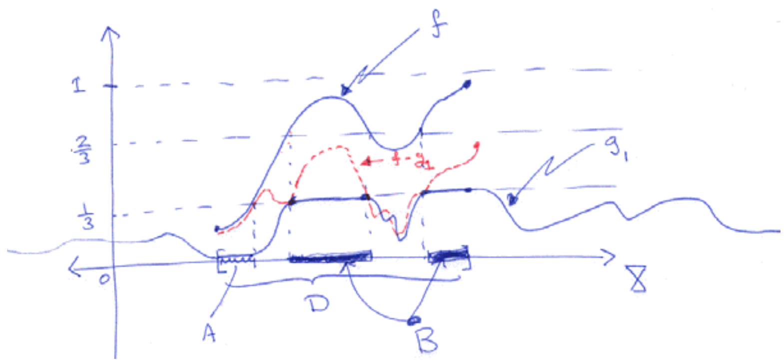

Fig. 7.1. Reducing f by subtracting off a globally defined function $g_1 \in C\left(X, [0, \frac{1}{3}]\right)$.

■

Theorem 7.5 (Completeness of $\ell^p(\mu)$). *Let X be a set and $\mu : X \to (0,\infty)$ be a given function. Then for any $p \in [1,\infty]$, $(\ell^p(\mu), \|\cdot\|_p)$ is a Banach space.*

Proof. We have already proved this for $p = \infty$ in Lemma 7.3 so we now assume that $p \in [1,\infty)$. Let $\{f_n\}_{n=1}^{\infty} \subset \ell^p(\mu)$ be a Cauchy sequence. Since for any $x \in X$,

$$|f_n(x) - f_m(x)| \le \frac{1}{\mu(x)} \|f_n - f_m\|_p \to 0 \text{ as } m,n \to \infty$$

it follows that $\{f_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence of numbers and $f(x) := \lim_{n\to\infty} f_n(x)$ exists for all $x \in X$. By Fatou's Lemma,

$$\|f_n - f\|_p^p = \sum_X \mu \cdot \lim_{m\to\infty} \inf |f_n - f_m|^p \le \lim_{m\to\infty} \inf \sum_X \mu \cdot |f_n - f_m|^p$$
$$= \lim_{m\to\infty} \inf \|f_n - f_m\|_p^p \to 0 \text{ as } n \to \infty.$$

This then shows that $f = (f - f_n) + f_n \in \ell^p(\mu)$ (being the sum of two ℓ^p – functions) and that $f_n \xrightarrow{\ell^p} f$. ■

Remark 7.6. Let X be a set, Y be a Banach space and $\ell^\infty(X, Y)$ denote the bounded functions $f : X \to Y$ equipped with the norm

$$\|f\| = \|f\|_\infty = \sup_{x\in X} \|f(x)\|_Y .$$

If X is a metric space (or a general topological space, see Chapter 10), let $BC(X, Y)$ denote those $f \in \ell^\infty(X, Y)$ which are continuous. The same proof used in Lemma 7.3 shows that $\ell^\infty(X, Y)$ is a Banach space and that $BC(X, Y)$ is a closed subspace of $\ell^\infty(X, Y)$. Similarly, if $1 \le p < \infty$ we may define

$$\ell^p(X, Y) = \left\{ f : X \to Y : \|f\|_p = \left(\sum_{x\in X} \|f(x)\|_Y^p \right)^{1/p} < \infty \right\}.$$

The same proof as in Theorem 7.5 would then show that $\left(\ell^p(X, Y), \|\cdot\|_p\right)$ is a Banach space.

7.2 Bounded Linear Operators Basics

Definition 7.7. *Let X and Y be normed spaces and $T : X \to Y$ be a linear map. Then T is said to be bounded provided there exists $C < \infty$ such that $\|T(x)\| \le C\|x\|_X$ for all $x \in X$. We denote the best constant by $\|T\|$, i.e.*

$$\|T\| = \sup_{x\neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{x\neq 0} \{\|T(x)\| : \|x\| = 1\} .$$

The number $\|T\|$ is called the operator norm of T.

Proposition 7.8. *Suppose that X and Y are normed spaces and $T : X \to Y$ is a linear map. The the following are equivalent:*

(a) T is continuous.
(b) T is continuous at 0.
(c) T is bounded.

Proof. (a) $\Rightarrow$ (b) trivial. (b) $\Rightarrow$ (c) If T continuous at 0 then there exist $\delta > 0$ such that $\|T(x)\| \le 1$ if $\|x\| \le \delta$. Therefore for any $x \in X$, $\|T(\delta x/\|x\|)\| \le 1$

which implies that $\|T(x)\| \leq \frac{1}{\delta}\|x\|$ and hence $\|T\| \leq \frac{1}{\delta} < \infty$. (c) $\Rightarrow$ (a) Let $x \in X$ and $\varepsilon > 0$ be given. Then

$$\|Ty - Tx\| = \|T(y-x)\| \leq \|T\| \, \|y-x\| < \varepsilon$$

provided $\|y-x\| < \varepsilon/\|T\| := \delta$. ∎

For the next three exercises, let $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ and $T : X \to Y$ be a linear transformation so that T is given by matrix multiplication by an $m \times n$ matrix. Let us identify the linear transformation T with this matrix.

Exercise 7.1. Assume the norms on X and Y are the ℓ^1 – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \sum_{j=1}^n |x_j|$. Then the operator norm of T is given by

$$\|T\| = \max_{1\leq j\leq n} \sum_{i=1}^m |T_{ij}| \, .$$

Exercise 7.2. Suppose that norms on X and Y are the ℓ^∞ – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \max_{1\leq j\leq n} |x_j|$. Then the operator norm of T is given by

$$\|T\| = \max_{1\leq i\leq m} \sum_{j=1}^n |T_{ij}| \, .$$

Exercise 7.3. Assume the norms on X and Y are the ℓ^2 – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\|^2 = \sum_{j=1}^n x_j^2$. Show $\|T\|^2$ is the largest eigenvalue of the matrix $T^{tr}T : \mathbb{R}^n \to \mathbb{R}^n$. **Hint:** Use the spectral theorem for orthogonal matrices.

Notation 7.9 *Let $L(X,Y)$ denote the bounded linear operators from X to Y and $L(X) = L(X,X)$. If $Y = \mathbb{F}$ we write X^* for $L(X,\mathbb{F})$ and call X^* the (continuous)* ***dual space*** *to X.*

Lemma 7.10. *Let X, Y be normed spaces, then the operator norm $\|\cdot\|$ on $L(X,Y)$ is a norm. Moreover if Z is another normed space and $T : X \to Y$ and $S : Y \to Z$ are linear maps, then $\|ST\| \leq \|S\|\|T\|$, where $ST := S \circ T$.*

Proof. As usual, the main point in checking the operator norm is a norm is to verify the triangle inequality, the other axioms being easy to check. If $A, B \in L(X,Y)$ then the triangle inequality is verified as follows:

$$\begin{aligned}\|A+B\| &= \sup_{x\neq 0} \frac{\|Ax+Bx\|}{\|x\|} \leq \sup_{x\neq 0} \frac{\|Ax\|+\|Bx\|}{\|x\|} \\ &\leq \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} + \sup_{x\neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\| \, .\end{aligned}$$

For the second assertion, we have for $x \in X$, that

$$\|STx\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|.$$

From this inequality and the definition of $\|ST\|$, it follows that $\|ST\| \leq \|S\|\|T\|$. ■

The reader is asked to prove the following continuity lemma in Exercise 7.12.

Lemma 7.11. *Let X, Y and Z be normed spaces. Then the maps*

$$(S,x) \in L(X,Y) \times X \longrightarrow Sx \in Y$$

and

$$(S,T) \in L(X,Y) \times L(Y,Z) \longrightarrow ST \in L(X,Z)$$

are continuous relative to the norms

$$\begin{aligned} \|(S,x)\|_{L(X,Y)\times X} &:= \|S\|_{L(X,Y)} + \|x\|_X \quad \textit{and} \\ \|(S,T)\|_{L(X,Y)\times L(Y,Z)} &:= \|S\|_{L(X,Y)} + \|T\|_{L(Y,Z)} \end{aligned}$$

on $L(X,Y) \times X$ and $L(X,Y) \times L(Y,Z)$ respectively.

Proposition 7.12. *Suppose that X is a normed vector space and Y is a Banach space. Then $(L(X,Y), \|\cdot\|_{op})$ is a Banach space. In particular the dual space X^* is always a Banach space.*

Proof. Let $\{T_n\}_{n=1}^\infty$ be a Cauchy sequence in $L(X,Y)$. Then for each $x \in X$,

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \, \|x\| \to 0 \text{ as } m,n \to \infty$$

showing $\{T_n x\}_{n=1}^\infty$ is Cauchy in Y. Using the completeness of Y, there exists an element $Tx \in Y$ such that

$$\lim_{n\to\infty} \|T_n x - Tx\| = 0.$$

The map $T : X \to Y$ is linear map, since for $x, x' \in X$ and $\lambda \in \mathbb{F}$ we have

$$T(x + \lambda x') = \lim_{n\to\infty} T_n(x + \lambda x') = \lim_{n\to\infty} [T_n x + \lambda T_n x'] = Tx + \lambda T x',$$

wherein we have used the continuity of the vector space operations in the last equality. Moreover,

$$\|Tx - T_n x\| \leq \|Tx - T_m x\| + \|T_m x - T_n x\| \leq \|Tx - T_m x\| + \|T_m - T_n\| \, \|x\|$$

and therefore

$$\begin{aligned} \|Tx - T_n x\| &\leq \lim \inf_{m\to\infty} (\|Tx - T_m x\| + \|T_m - T_n\| \, \|x\|) \\ &= \|x\| \cdot \lim \inf_{m\to\infty} \|T_m - T_n\| \,. \end{aligned}$$

Hence

$$\|T - T_n\| \leq \lim \inf_{m\to\infty} \|T_m - T_n\| \to 0 \text{ as } n \to \infty.$$

Thus we have shown that $T_n \to T$ in $L(X,Y)$ as desired. ■

The following characterization of a Banach space will sometimes be useful in the sequel.

Theorem 7.13. *A normed space* $(X, \|\cdot\|)$ *is a Banach space iff for every sequence* $\{x_n\}_{n=1}^{\infty}$ *such that* $\sum\limits_{n=1}^{\infty}\|x_n\| < \infty$ *implies* $\lim_{N\to\infty}\sum\limits_{n=1}^{N} x_n = s$ *exists in* X *(that is to say every absolutely convergent series is a convergent series in* X*.) As usual we will denote* s *by* $\sum\limits_{n=1}^{\infty} x_n$.

Proof. This is very similar to Exercise 6.10.

($\Rightarrow$)If X is complete and $\sum\limits_{n=1}^{\infty}\|x_n\| < \infty$ then sequence $s_N := \sum\limits_{n=1}^{N} x_n$ for $N \in \mathbb{N}$ is Cauchy because (for $N > M$)

$$\|s_N - s_M\| \leq \sum_{n=M+1}^{N}\|x_n\| \to 0 \text{ as } M, N \to \infty.$$

Therefore $s = \sum\limits_{n=1}^{\infty} x_n := \lim_{N\to\infty}\sum\limits_{n=1}^{N} x_n$ exists in X.

($\Longleftarrow$) Suppose that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence and let $\{y_k = x_{n_k}\}_{k=1}^{\infty}$ be a subsequence of $\{x_n\}_{n=1}^{\infty}$ such that $\sum\limits_{n=1}^{\infty}\|y_{n+1} - y_n\| < \infty$. By assumption

$$y_{N+1} - y_1 = \sum_{n=1}^{N}(y_{n+1} - y_n) \to s = \sum_{n=1}^{\infty}(y_{n+1} - y_n) \in X \text{ as } N \to \infty.$$

This shows that $\lim_{N\to\infty} y_N$ exists and is equal to $x := y_1 + s$. Since $\{x_n\}_{n=1}^{\infty}$ is Cauchy,

$$\|x - x_n\| \leq \|x - y_k\| + \|y_k - x_n\| \to 0 \text{ as } k, n \to \infty$$

showing that $\lim_{n\to\infty} x_n$ exists and is equal to x. ■

Example 7.14. Here is another proof of Theorem 7.12 which makes use of Proposition 7.12. Suppose that $T_n \in L(X, Y)$ is a sequence of operators such that $\sum\limits_{n=1}^{\infty}\|T_n\| < \infty$. Then

$$\sum_{n=1}^{\infty}\|T_n x\| \leq \sum_{n=1}^{\infty}\|T_n\|\,\|x\| < \infty$$

and therefore by the completeness of Y, $Sx := \sum\limits_{n=1}^{\infty} T_n x = \lim_{N\to\infty} S_N x$ exists in Y, where $S_N := \sum\limits_{n=1}^{N} T_n$. The reader should check that $S : X \to Y$ so defined is linear. Since,

$$\|Sx\| = \lim_{N\to\infty} \|S_N x\| \le \lim_{N\to\infty} \sum_{n=1}^{N} \|T_n x\| \le \sum_{n=1}^{\infty} \|T_n\| \, \|x\| \, ,$$

S is bounded and

$$\|S\| \le \sum_{n=1}^{\infty} \|T_n\|. \tag{7.3}$$

Similarly,

$$\begin{aligned}\|Sx - S_M x\| &= \lim_{N\to\infty} \|S_N x - S_M x\| \\ &\le \lim_{N\to\infty} \sum_{n=M+1}^{N} \|T_n\| \, \|x\| = \sum_{n=M+1}^{\infty} \|T_n\| \, \|x\|\end{aligned}$$

and therefore,

$$\|S - S_M\| \le \sum_{n=M}^{\infty} \|T_n\| \to 0 \text{ as } M \to \infty.$$

7.3 General Sums in Banach Spaces

Definition 7.15. *Suppose X is a normed space.*

1. *Suppose that $\{x_n\}_{n=1}^{\infty}$ is a sequence in X, then we say $\sum_{n=1}^{\infty} x_n$* ***converges*** *in X and $\sum_{n=1}^{\infty} x_n = s$ if*

$$\lim_{N\to\infty} \sum_{n=1}^{N} x_n = s \text{ in } X.$$

2. *Suppose that $\{x_\alpha : \alpha \in A\}$ is a given collection of vectors in X. We say the sum $\sum_{\alpha\in A} x_\alpha$* ***converges*** *in X and write $s = \sum_{\alpha\in A} x_\alpha \in X$ if for all $\varepsilon > 0$ there exists a finite set $\Gamma_\varepsilon \subset A$ such that $\left\| s - \sum_{\alpha\in\Lambda} x_\alpha \right\| < \varepsilon$ for any $\Lambda \subset\subset A$ such that $\Gamma_\varepsilon \subset \Lambda$.*

Warning: As usual if $\sum_{\alpha\in A} \|x_\alpha\| < \infty$ then $\sum_{\alpha\in A} x_\alpha$ exists in X, see Exercise 7.16. However, unlike the case of real valued sums the existence of $\sum_{\alpha\in A} x_\alpha$ does **not** imply $\sum_{\alpha\in A} \|x_\alpha\| < \infty$. See Proposition 29.19 below, from which one may manufacture counter-examples to this false premise.

Lemma 7.16. *Suppose that $\{x_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in a normed space, X.*

1. *If $s = \sum_{\alpha\in A} x_\alpha \in X$ exists and $T : X \to Y$ is a bounded linear map between normed spaces, then $\sum_{\alpha\in A} T x_\alpha$ exists in Y and*

$$Ts = T \sum_{\alpha\in A} x_\alpha = \sum_{\alpha\in A} T x_\alpha.$$

2. *If* $s = \sum_{\alpha \in A} x_\alpha$ *exists in* X *then for every* $\varepsilon > 0$ *there exists* $\Gamma_\varepsilon \subset\subset A$ *such that* $\left\| \sum_{\alpha \in \Lambda} x_\alpha \right\| < \varepsilon$ *for all* $\Lambda \subset\subset A \setminus \Gamma_\varepsilon$.
3. *If* $s = \sum_{\alpha \in A} x_\alpha$ *exists in* X, *the set* $\Gamma := \{\alpha \in A : x_a \neq 0\}$ *is at most countable. Moreover if* Γ *is infinite and* $\{\alpha_n\}_{n=1}^\infty$ *is an enumeration of* Γ, *then*
$$s = \sum_{n=1}^\infty x_{\alpha_n} := \lim_{N \to \infty} \sum_{n=1}^N x_{\alpha_n}. \tag{7.4}$$
4. *If we further assume that* X *is a Banach space and suppose for all* $\varepsilon > 0$ *there exists* $\Gamma_\varepsilon \subset\subset A$ *such that* $\left\| \sum_{\alpha \in \Lambda} x_\alpha \right\| < \varepsilon$ *whenever* $\Lambda \subset\subset A \setminus \Gamma_\varepsilon$, *then* $\sum_{\alpha \in A} x_\alpha$ *exists in* X.

Proof.

1. Let Γ_ε be as in Definition 7.15 and $\Lambda \subset\subset A$ such that $\Gamma_\varepsilon \subset \Lambda$. Then
$$\left\| Ts - \sum_{\alpha \in \Lambda} Tx_\alpha \right\| \leq \|T\| \left\| s - \sum_{\alpha \in \Lambda} x_\alpha \right\| < \|T\| \, \varepsilon$$
which shows that $\sum_{\alpha \in \Lambda} Tx_\alpha$ exists and is equal to Ts.
2. Suppose that $s = \sum_{\alpha \in A} x_\alpha$ exists and $\varepsilon > 0$. Let $\Gamma_\varepsilon \subset\subset A$ be as in Definition 7.15. Then for $\Lambda \subset\subset A \setminus \Gamma_\varepsilon$,
$$\begin{aligned} \left\| \sum_{\alpha \in \Lambda} x_\alpha \right\| &= \left\| \sum_{\alpha \in \Gamma_\varepsilon \cup \Lambda} x_\alpha - \sum_{\alpha \in \Gamma_\varepsilon} x_\alpha \right\| \\ &\leq \left\| \sum_{\alpha \in \Gamma_\varepsilon \cup \Lambda} x_\alpha - s \right\| + \left\| \sum_{\alpha \in \Gamma_\varepsilon} x_\alpha - s \right\| < 2\varepsilon. \end{aligned}$$
3. If $s = \sum_{\alpha \in A} x_\alpha$ exists in X, for each $n \in \mathbb{N}$ there exists a finite subset $\Gamma_n \subset A$ such that $\left\| \sum_{\alpha \in \Lambda} x_\alpha \right\| < \frac{1}{n}$ for all $\Lambda \subset\subset A \setminus \Gamma_n$. Without loss of generality we may assume $x_\alpha \neq 0$ for all $\alpha \in \Gamma_n$. Let $\Gamma_\infty := \cup_{n=1}^\infty \Gamma_n$ – a countable subset of A. Then for any $\beta \notin \Gamma_\infty$, we have $\{\beta\} \cap \Gamma_n = \emptyset$ and therefore
$$\|x_\beta\| = \left\| \sum_{\alpha \in \{\beta\}} x_\alpha \right\| \leq \frac{1}{n} \to 0 \text{ as } n \to \infty.$$
Let $\{\alpha_n\}_{n=1}^\infty$ be an enumeration of Γ and define $\gamma_N := \{\alpha_n : 1 \leq n \leq N\}$. Since for any $M \in \mathbb{N}$, γ_N will eventually contain Γ_M for N sufficiently large, we have
$$\limsup_{N \to \infty} \left\| s - \sum_{n=1}^N x_{\alpha_n} \right\| \leq \frac{1}{M} \to 0 \text{ as } M \to \infty.$$
Therefore Eq. (7.4) holds.

4. For $n \in \mathbb{N}$, let $\Gamma_n \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} x_\alpha\right\| < \frac{1}{n}$ for all $\Lambda \subset\subset A \setminus \Gamma_n$. Define $\gamma_n := \cup_{k=1}^n \Gamma_k \subset A$ and $s_n := \sum_{\alpha\in\gamma_n} x_\alpha$. Then for $m > n$,
$$\|s_m - s_n\| = \left\|\sum_{\alpha\in\gamma_m\setminus\gamma_n} x_\alpha\right\| \leq 1/n \to 0 \text{ as } m, n \to \infty.$$
Therefore $\{s_n\}_{n=1}^\infty$ is Cauchy and hence convergent in X, because X is a Banach space. Let $s := \lim_{n\to\infty} s_n$. Then for $\Lambda \subset\subset A$ such that $\gamma_n \subset \Lambda$, we have
$$\left\|s - \sum_{\alpha\in\Lambda} x_\alpha\right\| \leq \|s - s_n\| + \left\|\sum_{\alpha\in\Lambda\setminus\gamma_n} x_\alpha\right\| \leq \|s - s_n\| + \frac{1}{n}.$$
Since the right side of this equation goes to zero as $n \to \infty$, it follows that $\sum_{\alpha\in A} x_\alpha$ exists and is equal to s.

■

Exercise 7.4. Prove Theorem 8.4. BRUCE: Delete

7.4 Inverting Elements in $L(X)$

Definition 7.17. *A linear map $T : X \to Y$ is an **isometry** if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$. T is said to be **invertible** if T is a bijection and T^{-1} is bounded.*

Notation 7.18 *We will write $GL(X, Y)$ for those $T \in L(X, Y)$ which are invertible. If $X = Y$ we simply write $L(X)$ and $GL(X)$ for $L(X, X)$ and $GL(X, X)$ respectively.*

Proposition 7.19. *Suppose X is a Banach space and $\Lambda \in L(X) := L(X, X)$ satisfies $\sum_{n=0}^\infty \|\Lambda^n\| < \infty$. Then $I - \Lambda$ is invertible and*
$$(I - \Lambda)^{-1} = \text{“}\frac{1}{I - \Lambda}\text{”} = \sum_{n=0}^\infty \Lambda^n \text{ and } \left\|(I - \Lambda)^{-1}\right\| \leq \sum_{n=0}^\infty \|\Lambda^n\|.$$
In particular if $\|\Lambda\| < 1$ then the above formula holds and
$$\left\|(I - \Lambda)^{-1}\right\| \leq \frac{1}{1 - \|\Lambda\|}.$$

Proof. Since $L(X)$ is a Banach space and $\sum_{n=0}^\infty \|\Lambda^n\| < \infty$, it follows from Theorem 7.13 that

$$S := \lim_{N\to\infty} S_N := \lim_{N\to\infty} \sum_{n=0}^{N} \Lambda^n$$

exists in $L(X)$. Moreover, by Lemma 7.11,

$$\begin{aligned}(I-\Lambda)\,S &= (I-\Lambda)\lim_{N\to\infty} S_N = \lim_{N\to\infty}(I-\Lambda)\,S_N \\ &= \lim_{N\to\infty}(I-\Lambda)\sum_{n=0}^{N}\Lambda^n = \lim_{N\to\infty}(I-\Lambda^{N+1}) = I\end{aligned}$$

and similarly $S\,(I-\Lambda) = I$. This shows that $(I-\Lambda)^{-1}$ exists and is equal to S. Moreover, $(I-\Lambda)^{-1}$ is bounded because

$$\left\|(I-\Lambda)^{-1}\right\| = \|S\| \le \sum_{n=0}^{\infty}\|\Lambda^n\|.$$

If we further assume $\|\Lambda\| < 1$, then $\|\Lambda^n\| \le \|\Lambda\|^n$ and

$$\sum_{n=0}^{\infty}\|\Lambda^n\| \le \sum_{n=0}^{\infty}\|\Lambda\|^n = \frac{1}{1-\|\Lambda\|} < \infty.$$

■

Corollary 7.20. *Let X and Y be Banach spaces. Then $GL(X,Y)$ is an open (possibly empty) subset of $L(X,Y)$. More specifically, if $A \in GL(X,Y)$ and $B \in L(X,Y)$ satisfies*

$$\|B-A\| < \|A^{-1}\|^{-1} \tag{7.5}$$

then $B \in GL(X,Y)$

$$B^{-1} = \sum_{n=0}^{\infty}\left[I_X - A^{-1}B\right]^n A^{-1} \in L(Y,X), \tag{7.6}$$

$$\left\|B^{-1}\right\| \le \|A^{-1}\|\frac{1}{1-\|A^{-1}\|\,\|A-B\|} \tag{7.7}$$

and

$$\left\|B^{-1}-A^{-1}\right\| \le \frac{\|A^{-1}\|^2\,\|A-B\|}{1-\|A^{-1}\|\,\|A-B\|}. \tag{7.8}$$

In particular the map

$$A \in GL(X,Y) \to A^{-1} \in GL(Y,X) \tag{7.9}$$

is continuous.

Proof. Let A and B be as above, then

$$B = A - (A - B) = A\left[I_X - A^{-1}(A - B))\right] = A(I_X - \Lambda)$$

where $\Lambda : X \to X$ is given by

$$\Lambda := A^{-1}(A - B) = I_X - A^{-1}B.$$

Now

$$\|\Lambda\| = \left\|A^{-1}(A - B))\right\| \le \|A^{-1}\| \, \|A - B\| < \|A^{-1}\| \|A^{-1}\|^{-1} = 1.$$

Therefore $I - \Lambda$ is invertible and hence so is B (being the product of invertible elements) with

$$B^{-1} = (I_X - \Lambda)^{-1} A^{-1} = \left[I_X - A^{-1}(A - B))\right]^{-1} A^{-1}.$$

Taking norms of the previous equation gives

$$\begin{aligned} \left\|B^{-1}\right\| &\le \left\|(I_X - \Lambda)^{-1}\right\| \|A^{-1}\| \le \|A^{-1}\| \frac{1}{1 - \|\Lambda\|} \\ &\le \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \, \|A - B\|} \end{aligned}$$

which is the bound in Eq. (7.7). The bound in Eq. (7.8) holds because

$$\begin{aligned} \left\|B^{-1} - A^{-1}\right\| &= \left\|B^{-1}(A - B) A^{-1}\right\| \le \left\|B^{-1}\right\| \left\|A^{-1}\right\| \|A - B\| \\ &\le \frac{\|A^{-1}\|^2 \, \|A - B\|}{1 - \|A^{-1}\| \, \|A - B\|}. \end{aligned}$$

■

For an application of these results to linear ordinary differential equations, see Section 8.3.

7.5 Hahn Banach Theorem

Our next goal is to show that continuous dual X^* of a Banach space X is always large. This will be the content of the Hahn – Banach Theorem 7.24 below.

Proposition 7.21. *Let X be a complex vector space over $\mathbb{C}$ and let $X_{\mathbb{R}}$ denote X thought of as a real vector space. If $f \in X^*$ and $u = \mathrm{Re} f \in X_{\mathbb{R}}^*$ then*

$$f(x) = u(x) - iu(ix). \tag{7.10}$$

Conversely if $u \in X_{\mathbb{R}}^*$ and f is defined by Eq. (7.10), then $f \in X^*$ and $\|u\|_{X_{\mathbb{R}}^*} = \|f\|_{X^*}$. More generally if p is a semi-norm on X, then

$$|f| \le p \text{ iff } u \le p.$$

Proof. Let $v(x) = \text{Im } f(x)$, then

$$v(ix) = \text{Im } f(ix) = \text{Im}(if(x)) = \text{Re}f(x) = u(x).$$

Therefore

$$f(x) = u(x) + iv(x) = u(x) + iu(-ix) = u(x) - iu(ix).$$

Conversely for $u \in X_{\mathbb{R}}^*$ let $f(x) = u(x) - iu(ix)$. Then

$$\begin{aligned} f((a+ib)x) &= u(ax + ibx) - iu(iax - bx) \\ &= au(x) + bu(ix) - i(au(ix) - bu(x)) \end{aligned}$$

while

$$(a + ib)f(x) = au(x) + bu(ix) + i(bu(x) - au(ix)).$$

So f is complex linear.

Because $|u(x)| = |\text{Re}f(x)| \le |f(x)|$, it follows that $\|u\| \le \|f\|$. For $x \in X$ choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$ so

$$|f(x)| = f(\lambda x) = u(\lambda x) \le \|u\| \; \|\lambda x\| = \|u\|\|x\|.$$

Since $x \in X$ is arbitrary, this shows that $\|f\| \le \|u\|$ so $\|f\| = \|u\|$.[1]

For the last assertion, it is clear that $|f| \le p$ implies that $u \le |u| \le |f| \le p$. Conversely if $u \le p$ and $x \in X$, choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$. Then

$$|f(x)| = \lambda f(x) = f(\lambda x) = u(\lambda x) \le p(\lambda x) = p(x)$$

holds for all $x \in X$. ∎

Definition 7.22 (Minkowski functional). *A function $p : X \to \mathbb{R}$ is a Minkowski functional if*

[1]

Proof. To understand better why $\|f\| = \|u\|$, notice that

$$\|f\|^2 = \sup_{\|x\|=1} |f(x)|^2 = \sup_{\|x\|=1} (|u(x)|^2 + |u(ix)|^2).$$

Supppose that $M = \sup_{\|x\|=1} |u(x)|$ and this supremum is attained at $x_0 \in X$ with $\|x_0\| = 1$. Replacing x_0 by $-x_0$ if necessary, we may assume that $u(x_0) = M$. Since u has a maximum at x_0,

$$\begin{aligned} 0 &= \frac{d}{dt}\bigg|_0 u\left(\frac{x_0 + itx_0}{\|x_0 + itx_0\|}\right) \\ &= \frac{d}{dt}\bigg|_0 \left\{\frac{1}{|1+it|} \; (u(x_0) + tu(ix_0))\right\} = u(ix_0) \end{aligned}$$

since $\frac{d}{dt}|_0 |1 + it| = \frac{d}{dt}|_0 \sqrt{1 + t^2} = 0$.This explains why $\|f\| = \|u\|$. ∎

1. *$p(x+y) \le p(x) + p(y)$ for all $x, y \in X$ and*
2. *$p(cx) = cp(x)$ for all $c \ge 0$ and $x \in X$.*

Example 7.23. Suppose that $X = \mathbb{R}$ and

$$p(x) = \inf\{\lambda \ge 0 : x \in \lambda[-1,2] = [-\lambda, 2\lambda]\}.$$

Notice that if $x \ge 0$, then $p(x) = x/2$ and if $x \le 0$ then $p(x) = -x$, i.e.

$$p(x) = \begin{cases} x/2 \text{ if } x \ge 0 \\ |x| \text{ if } x \le 0. \end{cases}$$

From this formula it is clear that $p(cx) = cp(x)$ for all $c \ge 0$ but not for $c < 0$. Moreover, p satisfies the triangle inequality, indeed if $p(x) = \lambda$ and $p(y) = \mu$, then $x \in \lambda[-1,2]$ and $y \in \mu[-1,2]$ so that

$$x + y \in \lambda[-1,2] + \mu[-1,2] \subset (\lambda + \mu)[-1,2]$$

which shows that $p(x+y) \le \lambda + \mu = p(x) + p(y)$. To check the last set inclusion let $a, b \in [-1,2]$, then

$$\lambda a + \mu b = (\lambda + \mu)\left(\frac{\lambda}{\lambda+\mu} a + \frac{\mu}{\lambda+\mu} b\right) \in (\lambda+\mu)[-1,2]$$

since $[-1,2]$ is a convex set and $\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1$.

BRUCE: Add in the relationship to convex sets and separation theorems, see Reed and Simon Vol. 1. for example.

Theorem 7.24 (Hahn-Banach). *Let X be a real vector space, $M \subset X$ be a subspace $f : M \to \mathbb{R}$ be a linear functional such that $f \le p$ on M. Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F|_M = f$ and $F \le p$.*

Proof. Step (1) We show for all $x \in X \setminus M$ there exists and extension F to $M \oplus \mathbb{R}x$ with the desired properties. If F exists and $\alpha = F(x)$, then for all $y \in M$ and $\lambda \in \mathbb{R}$ we must have $f(y) + \lambda\alpha = F(y + \lambda x) \le p(y + \lambda x)$ i.e. $\lambda\alpha \le p(y + \lambda x) - f(y)$. Equivalently put we must find $\alpha \in \mathbb{R}$ such that

$$\alpha \le \frac{p(y+\lambda x) - f(y)}{\lambda} \text{ for all } y \in M \text{ and } \lambda > 0$$
$$\alpha \ge \frac{p(z - \mu x) - f(z)}{\mu} \text{ for all } z \in M \text{ and } \mu > 0.$$

So if $\alpha \in \mathbb{R}$ is going to exist, we have to prove, for all $y, z \in M$ and $\lambda, \mu > 0$ that

$$\frac{f(z) - p(z - \mu x)}{\mu} \le \frac{p(y + \lambda x) - f(y)}{\lambda}$$

or equivalently

$$
\begin{aligned}
f(\lambda z+\mu y) &\leq \mu p(y+\lambda x)+\lambda p(z-\mu x) \\
&= p(\mu y+\mu\lambda x)+p(\lambda z-\lambda\mu x).
\end{aligned} \tag{7.11}
$$

But by assumption and the triangle inequality for p,

$$
\begin{aligned}
f(\lambda z+\mu y) \leq p(\lambda z+\mu y) &= p(\lambda z+\mu\lambda x+\lambda z-\lambda\mu x) \\
&\leq p(\lambda z+\mu\lambda x)+p(\lambda z-\lambda\mu x)
\end{aligned}
$$

which shows that Eq. (7.11) is true and by working backwards, there exist an $\alpha\in\mathbb{R}$ such that $f(y)+\lambda\alpha\leq p(y+\lambda x)$. Therefore $F(y+\lambda x):=f(y)+\lambda\alpha$ is the desired extension.

Step (2) Let us now write $F:X\to\mathbb{R}$ to mean F is defined on a linear subspace $D(F)\subset X$ and $F:D(F)\to\mathbb{R}$ is linear. For $F,G:X\to\mathbb{R}$ we will say $F\prec G$ if $D(F)\subset D(G)$ and $F=G|_{D(F)}$, that is G is an extension of F. Let

$$\mathcal{F}=\{F:X\to\mathbb{R}:f\prec F\text{ and }F\leq p\text{ on }D(F)\}.$$

Then $(\mathcal{F},\prec)$ is a partially ordered set. If $\Phi\subset\mathcal{F}$ is a chain (i.e. a linearly ordered subset of $\mathcal{F}$) then Φ has an upper bound $G\in\mathcal{F}$ defined by $D(G)=\bigcup\limits_{F\in\Phi}D(F)$ and $G(x)=F(x)$ for $x\in D(F)$. Then it is easily checked that $D(G)$ is a linear subspace, $G\in\mathcal{F}$, and $F\prec G$ for all $F\in\Phi$. We may now apply Zorn's Lemma[2] (see Theorem B.7) to conclude there exists a maximal element $F\in\mathcal{F}$. Necessarily, $D(F)=X$ for otherwise we could extend F by step (1), violating the maximality of F. Thus F is the desired extension of f. ■

Corollary 7.25. *Suppose that X is a complex vector space, $p:X\to[0,\infty)$ is a semi-norm, $M\subset X$ is a linear subspace, and $f:M\to\mathbb{C}$ is linear functional such that $|f(x)|\leq p(x)$ for all $x\in M$. Then there exists $F\in X'$ (X' is the* ***algebraic*** *dual of X) such that $F|_M=f$ and $|F|\leq p$.*

Proof. Let $u=\mathrm{Re}f$ then $u\leq p$ on M and hence by Theorem 7.24, there exists $U\in X'_{\mathbb{R}}$ such that $U|_M=u$ and $U\leq p$ on M. Define $F(x)=U(x)-iU(ix)$ then as in Proposition 7.21, $F=f$ on M and $|F|\leq p$. ■

Theorem 7.26. *Let X be a normed space $M\subset X$ be a closed subspace and $x\in X\setminus M$. Then there exists $f\in X^*$ such that $\|f\|=1$, $f(x)=\delta=d(x,M)$ and $f=0$ on M.*

[2] The use of Zorn's lemma in this step may be avoided in the case that $p(x)$ is a norm and X may be written as $\overline{M\oplus\mathrm{span}(\beta)}$ where $\beta:=\{x_n\}_{n=1}^\infty$ is a countable subset of X. In this case, by step (1) and induction, $f:M\to\mathbb{R}$ may be extended to a linear functional $F:M\oplus\mathrm{span}(\beta)\to\mathbb{R}$ with $F(x)\leq p(x)$ for $x\in M\oplus\mathrm{span}(\beta)$. This function F then extends by continuity to X and gives the desired extension of f.

Proof. Define $h : M \oplus \mathbb{C}x \to \mathbb{C}$ by $h(m + \lambda x) \equiv \lambda\delta$ for all $m \in M$ and $\lambda \in \mathbb{C}$. Then

$$\|h\| := \sup_{m\in M \text{ and } \lambda\neq 0} \frac{|\lambda|\,\delta}{\|m+\lambda x\|} = \sup_{m\in M \text{ and } \lambda\neq 0} \frac{\delta}{\|x+m/\lambda\|} = \frac{\delta}{\delta} = 1$$

and by the Hahn – Banach theorem there exists $f \in X^*$ such that $f|_{M\oplus\mathbb{C}x} = h$ and $\|f\| \le 1$. Since $1 = \|h\| \le \|f\| \le 1$, it follows that $\|f\| = 1$. ■

Corollary 7.27. *The linear map $x \in X \to \hat{x} \in X^{**}$ where $\hat{x}(f) = f(x)$ for all $x \in X$ is an isometry.*

Proof. Since $|\hat{x}(f)| = |f(x)| \le \|f\|_{X^*}\|x\|_X$ for all $f \in X^*$, it follows that $\|\hat{x}\|_{X^{**}} \le \|x\|_X$. Now applying Theorem 7.26 with $M = \{0\}$, there exists $f \in X^*$ such that $\|f\| = 1$ and $|\hat{x}(f)| = f(x) = \|x\|$, which shows that $\|\hat{x}\|_{X^{**}} \ge \|x\|_X$. This shows that $x \in X \to \hat{x} \in X^{**}$ is an isometry. Since isometries are necessarily injective, we are done. ■

Definition 7.28. *A Banach space X is reflexive if the map $x \in X \to \hat{x} \in X^{**}$ is surjective. (BRUCE: this is defined again in Definition 33.44 below.)*

Exercise 7.5. Show all finite dimensional Banach spaces are reflexive.

Definition 7.29. *For $M \subset X$ and $N \subset X^*$ let*

$$M^0 := \{f \in X^* : f|_M = 0\} \text{ and}$$
$$N^\perp := \{x \in X : f(x) = 0 \text{ for all } f \in N\}.$$

Lemma 7.30. *Let $M \subset X$ and $N \subset X^*$, then*

1. *M^0 and $N^\perp$ are always closed subspace of X^* and X respectively.*
2. *$\left(M^0\right)^\perp = \bar{M}$.*

Proof. The first item is an easy consequence of the assumed continuity off all linear functionals involved.

If $x \in M$, then $f(x) = 0$ for all $f \in M^0$ so that $x \in \left(M^0\right)^\perp$. Therefore $\bar{M} \subset \left(M^0\right)^\perp$. If $x \notin \bar{M}$, then there exists $f \in X^*$ such that $f|_M = 0$ while $f(x) \neq 0$, i.e. $f \in M^0$ yet $f(x) \neq 0$. This shows $x \notin \left(M^0\right)^\perp$ and we have shown $\left(M^0\right)^\perp \subset \bar{M}$. ■

Proposition 7.31. *Suppose X is a Banach space, then $X^{***} = \widehat{(X^*)} \oplus \left(\hat{X}\right)^0$ where*

$$\left(\hat{X}\right)^0 = \{\lambda \in X^{***} : \lambda(\hat{x}) = 0 \text{ for all } x \in X\}.$$

In particular X is reflexive iff X^* is reflexive.

Proof. Let $\psi \in X^{***}$ and define $f_\psi \in X^*$ by $f_\psi(x) := \psi(\hat{x})$ for all $x \in X$ and set $\psi' := \psi - \hat{f}_\psi$. For $x \in X$ (so $\hat{x} \in X^{**}$) we have

$$\psi'(\hat{x}) = \psi(\hat{x}) - \hat{f}_\psi(\hat{x}) = f_\psi(x) - \hat{x}(f_\psi) = f_\psi(x) - f_\psi(x) = 0.$$

This shows $\psi' \in \hat{X}^0$ and we have shown $X^{***} = \widehat{X^*} + \hat{X}^0$. If $\psi \in \widehat{X^*} \cap \hat{X}^0$, then $\psi = \hat{f}$ for some $f \in X^*$ and $0 = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$, i.e. $f = 0$ so $\psi = 0$. Therefore $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ as claimed.

If X is reflexive, then $\hat{X} = X^{**}$ and so $\hat{X}^0 = \{0\}$ showing $(X^*)^{**} = X^{***} = \widehat{(X^*)}$, i.e. X^* is reflexive. Conversely if X^* is reflexive we conclude that $\left(\hat{X}\right)^0 = \{0\}$ and therefore

$$X^{**} = \{0\}^\perp = \left(\hat{X}^0\right)^\perp = \hat{X},$$

which shows $\hat{X}$ is reflexive. Here we have used

$$\left(\hat{X}^0\right)^\perp = \overline{\hat{X}} = \hat{X}$$

since $\hat{X}$ is a closed subspace of X^{**}. ■

For the remainder of this section let X be an infinite set, $\mu : X \to (0, \infty)$ be a given function and $p, q \in [1, \infty]$ such that $q = p/(p-1)$. it will also be convenient to define $\delta_x : X \to \mathbb{R}$ for $x \in X$ by

$$\delta_x(y) = \begin{cases} 1 \text{ if } y = x \\ 0 \text{ if } y \neq x. \end{cases}$$

Notation 7.32 *Let $c_0(X)$ denote those functions $f \in \ell^\infty(X)$ which "vanish at ∞," i.e. for every $\varepsilon > 0$ there exists a finite subset $\Lambda_\varepsilon \subset X$ such that $|f(x)| < \varepsilon$ whenever $x \notin \Lambda_\varepsilon$. Also let $c_f(X)$ denote those functions $f : X \to \mathbb{F}$ with finite support, i.e.*

$$c_f(X) := \{f \in \ell^\infty(X) : \#(\{x \in X : f(x) \neq 0\}) < \infty\}.$$

Exercise 7.6. Show $c_f(X)$ is a dense subspace of the Banach spaces $\left(\ell^p(\mu), \|\cdot\|_p\right)$ for $1 \le p < \infty$, while the closure of $c_f(X)$ inside the Banach space, $(\ell^\infty(X), \|\cdot\|_\infty)$ is $c_0(X)$. Note from this it follows that $c_0(X)$ is a closed subspace of $\ell^\infty(X)$.

Theorem 7.33. *Let X be an infinite set, $\mu : X \to (0, \infty)$ be a function, $p \in [1, \infty]$, $q := p/(p-1)$ be the conjugate exponent and for $f \in \ell^q(\mu)$ define $\phi_f : \ell^p(\mu) \to \mathbb{F}$ by*

$$\phi_f(g) := \sum_{x \in X} f(x)\, g(x)\, \mu(x).$$

Then

1. *$\phi_f(g)$ is well defined and $\phi_f \in \ell^p(\mu)^*$.*
2. *The map*
$$f \in \ell^q(\mu) \xrightarrow{\phi} \phi_f \in \ell^p(\mu)^* \tag{7.12}$$
is a isometric linear map of Banach spaces.
3. *If $p \in [1,\infty)$, then the map in Eq. (7.12) is also surjective and hence, $\ell^p(\mu)^*$ is isometrically isomorphic to $\ell^q(\mu)$. When $p=\infty$, the map*
$$f \in \ell^1(\mu) \to \phi_f \in c_0^*$$
is an isometric and surjective, i.e. $\ell^1(\mu)$ is isometrically isomorphic to c_0^*.
4. *$\ell^p(\mu)$ is reflexive for $p \in (1,\infty)$.*
5. *The map $\phi : \ell^1(\mu) \to \ell^\infty(X)^*$ is not surjective.*
6. *$\ell^1(\mu)$ and $\ell^\infty(X)$ are **not** reflexive.*

Proof.

1. By Holder's inequality,
$$\sum_{x\in X} |f(x)|\,|g(x)|\,\mu(x) \le \|f\|_q \|g\|_p$$
which shows that ϕ_f is well defined. The $\phi_f : \ell^p(\mu) \to \mathbb{F}$ is linear by the linearity of sums and since
$$|\phi_f(g)| = \left|\sum_{x\in X} f(x)\,g(x)\,\mu(x)\right| \le \sum_{x\in X} |f(x)|\,|g(x)|\,\mu(x) \le \|f\|_q \|g\|_p,$$
we learn that
$$\|\phi_f\|_{\ell^p(\mu)^*} \le \|f\|_q. \tag{7.13}$$
Therefore $\phi_f \in \ell^p(\mu)^*$.
2. The map ϕ in Eq. (7.12) is linear in f by the linearity properties of infinite sums.
For $p \in (1,\infty)$, define $g(x) = \overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}$ where
$$\operatorname{sgn}(z) := \begin{cases} \frac{z}{|z|} & \text{if } z \ne 0 \\ 0 & \text{if } z = 0. \end{cases}$$
Then
$$\begin{aligned}\|g\|_p^p &= \sum_{x\in X} |f(x)|^{(q-1)p}\,\mu(x) = \sum_{x\in X} |f(x)|^{\left(\frac{p}{p-1}-1\right)p}\,\mu(x) \\ &= \sum_{x\in X} |f(x)|^q\,\mu(x) = \|f\|_q^q\end{aligned}$$
and

$$\phi_f(g) = \sum_{x\in X} f(x)\,\overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}\mu(x) = \sum_{x\in X} |f(x)|\,|f(x)|^{q-1}\mu(x)$$
$$= \|f\|_q^{q\left(\frac{1}{q}+\frac{1}{p}\right)} = \|f\|_q \|f\|_q^{\frac{q}{p}} = \|f\|_q \|g\|_p .$$

Hence $\|\phi_f\|_{\ell^p(\mu)^*} \geq \|f\|_q$ which combined with Eq. (7.13) shows $\|\phi_f\|_{\ell^p(\mu)^*} = \|f\|_q$.

For $p=\infty$, let $g(x) = \overline{\operatorname{sgn}(f(x))}$, then $\|g\|_\infty = 1$ and

$$|\phi_f(g)| = \sum_{x\in X} f(x)\,\overline{\operatorname{sgn}(f(x))}\mu(x)$$
$$= \sum_{x\in X} |f(x)|\,\mu(x) = \|f\|_1 \|g\|_\infty$$

which shows $\|\phi_f\|_{\ell^\infty(\mu)^*} \geq \|f\|_{\ell^1(\mu)}$. Combining this with Eq. (7.13) shows $\|\phi_f\|_{\ell^\infty(\mu)^*} = \|f\|_{\ell^1(\mu)}$.

For $p=1$,

$$|\phi_f(\delta_x)| = \mu(x)\,|f(x)| = |f(x)|\,\|\delta_x\|_1$$

and therefore $\|\phi_f\|_{\ell^1(\mu)^*} \geq |f(x)|$ for all $x \in X$. Hence $\|\phi_f\|_{\ell^1(\mu)^*} \geq \|f\|_\infty$ which combined with Eq. (7.13) shows $\|\phi_f\|_{\ell^1(\mu)^*} = \|f\|_\infty$.

3. Suppose that $p \in [1,\infty)$ and $\lambda \in \ell^p(\mu)^*$ or $p=\infty$ and $\lambda \in c_0^*$. We wish to find $f \in \ell^q(\mu)$ such that $\lambda = \phi_f$. If such an f exists, then $\lambda(\delta_x) = f(x)\mu(x)$ and so we must define $f(x) := \lambda(\delta_x)/\mu(x)$. As a preliminary estimate,

$$|f(x)| = \frac{|\lambda(\delta_x)|}{\mu(x)} \leq \frac{\|\lambda\|_{\ell^p(\mu)^*}\|\delta_x\|_{\ell^p(\mu)}}{\mu(x)}$$
$$= \frac{\|\lambda\|_{\ell^p(\mu)^*}[\mu(x)]^{\frac{1}{p}}}{\mu(x)} = \|\lambda\|_{\ell^p(\mu)^*}[\mu(x)]^{-\frac{1}{q}} .$$

When $p=1$ and $q=\infty$, this implies $\|f\|_\infty \leq \|\lambda\|_{\ell^1(\mu)^*} < \infty$. If $p \in (1,\infty]$ and $\Lambda \subset\subset X$, then

$$\|f\|_{\ell^q(\Lambda,\mu)}^q := \sum_{x\in\Lambda} |f(x)|^q \mu(x) = \sum_{x\in\Lambda} f(x)\,\overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}\mu(x)$$
$$= \sum_{x\in\Lambda} \frac{\lambda(\delta_x)}{\mu(x)}\overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}\mu(x)$$
$$= \sum_{x\in\Lambda} \lambda(\delta_x)\,\overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}$$
$$= \lambda\left(\sum_{x\in\Lambda} \overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}\delta_x\right)$$
$$\leq \|\lambda\|_{\ell^p(\mu)^*}\left\|\sum_{x\in\Lambda} \overline{\operatorname{sgn}(f(x))}\,|f(x)|^{q-1}\delta_x\right\|_p .$$

Since

$$\left\| \sum_{x \in \Lambda} \overline{\operatorname{sgn}(f(x))} \left| f(x) \right|^{q-1} \delta_x \right\|_p = \left(\sum_{x \in \Lambda} \left| f(x) \right|^{(q-1)p} \mu(x) \right)^{1/p}$$

$$= \left(\sum_{x \in \Lambda} \left| f(x) \right|^q \mu(x) \right)^{1/p} = \|f\|_{\ell^q(\Lambda,\mu)}^{q/p}$$

which is also valid for $p = \infty$ provided $\|f\|_{\ell^1(\Lambda,\mu)}^{1/\infty} := 1$. Combining the last two displayed equations shows

$$\|f\|_{\ell^q(\Lambda,\mu)}^q \leq \|\lambda\|_{\ell^p(\mu)^*} \|f\|_{\ell^q(\Lambda,\mu)}^{q/p}$$

and solving this inequality for $\|f\|_{\ell^q(\Lambda,\mu)}^q$ (using $q - q/p = 1$) implies $\|f\|_{\ell^q(\Lambda,\mu)} \leq \|\lambda\|_{\ell^p(\mu)^*}$ Taking the supremum of this inequality on $\Lambda \subset\subset X$ shows $\|f\|_{\ell^q(\mu)} \leq \|\lambda\|_{\ell^p(\mu)^*}$, i.e. $f \in \ell^q(\mu)$. Since $\lambda = \phi_f$ agree on $c_f(X)$ and $c_f(X)$ is a dense subspace of $\ell^p(\mu)$ for $p < \infty$ and $c_f(X)$ is dense subspace of c_0 when $p = \infty$, it follows that $\lambda = \phi_f$.

4. This basically follows from two applications of item 3. More precisely if $\lambda \in \ell^p(\mu)^{**}$, let $\tilde{\lambda} \in \ell^q(\mu)^*$ be defined by $\tilde{\lambda}(g) = \lambda(\phi_g)$ for $g \in \ell^q(\mu)$. Then by item 3., there exists $f \in \ell^p(\mu)$ such that, for all $g \in \ell^q(\mu)$,

$$\lambda(\phi_g) = \tilde{\lambda}(g) = \phi_f(g) = \phi_g(f) = \hat{f}(\phi_g).$$

Since $\ell^p(\mu)^* = \{\phi_g : g \in \ell^q(\mu)\}$, this implies that $\lambda = \hat{f}$ and so $\ell^p(\mu)$ is reflexive.

5. Let $\mathbf{1} \in \ell^\infty(X)$ denote the constant function 1 on X. Notice that $\|\mathbf{1} - f\|_\infty \geq 1$ for all $f \in c_0$ and therefore there exists $\lambda \in \ell^\infty(X)^*$ such that $\lambda(\mathbf{1}) = 0$ while $\lambda|_{c_0} \equiv 0$. Now if $\lambda = \phi_f$ for some $f \in \ell^1(\mu)$, then $\mu(x) f(x) = \lambda(\delta_x) = 0$ for all x and f would have to be zero. This is absurd.

6. As we have seen $\ell^1(\mu)^* \cong \ell^\infty(X)$ while $\ell^\infty(X)^* \cong c_0^* \neq \ell^1(\mu)$. Let $\lambda \in \ell^\infty(X)^*$ be the linear functional as described above. We view this as an element of $\ell^1(\mu)^{**}$ by using

$$\tilde{\lambda}(\phi_g) := \lambda(g) \text{ for all } g \in \ell^\infty(X).$$

Suppose that $\tilde{\lambda} = \hat{f}$ for some $f \in \ell^1(\mu)$, then

$$\lambda(g) = \tilde{\lambda}(\phi_g) = \hat{f}(\phi_g) = \phi_g(f) = \phi_f(g).$$

But λ was constructed in such a way that $\lambda \neq \phi_f$ for any $f \in \ell^1(\mu)$. It now follows from Proposition 7.31 that $\ell^1(\mu)^* \cong \ell^\infty(X)$ is also not reflexive.

■

7.6 Exercises

Exercise 7.7. Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Show the map

$$(\lambda, x, y) \in \mathbb{F} \times X \times X \to x + \lambda y \in X$$

is continuous relative to the norm on $\mathbb{F} \times X \times X$ defined by

$$\|(\lambda, x, y)\|_{\mathbb{F}\times X\times X} := |\lambda| + \|x\| + \|y\| .$$

(See Exercise 10.21 for more on the metric associated to this norm.) Also show that $\|\cdot\| : X \to [0, \infty)$ is continuous.

Exercise 7.8. Let $X = \mathbb{N}$ and for $p, q \in [1, \infty)$ let $\|\cdot\|_p$ denote the $\ell^p(\mathbb{N})$ – norm. Show $\|\cdot\|_p$ and $\|\cdot\|_q$ are inequivalent norms for $p \neq q$ by showing

$$\sup_{f\neq 0} \frac{\|f\|_p}{\|f\|_q} = \infty \text{ if } p < q.$$

Exercise 7.9. Suppose that $(X, \|\cdot\|)$ is a normed space and $S \subset X$ is a linear subspace.

1. Show the closure $\bar{S}$ of S is also a linear subspace.
2. Now suppose that X is a Banach space. Show that S with the inherited norm from X is a Banach space iff S is closed.

Exercise 7.10. Folland Problem 5.9. Showing $C^k([0, 1])$ is a Banach space.

Exercise 7.11. (Do not use.) Folland Problem 5.11. Showing Holder spaces are Banach spaces.

Exercise 7.12. Suppose that X, Y and Z are Banach spaces and $Q : X \times Y \to Z$ is a bilinear form, i.e. we are assuming $x \in X \to Q(x, y) \in Z$ is linear for each $y \in Y$ and $y \in Y \to Q(x, y) \in Z$ is linear for each $x \in X$. Show Q is continuous relative to the product norm, $\|(x, y)\|_{X\times Y} := \|x\|_X + \|y\|_Y$, on $X \times Y$ iff there is a constant $M < \infty$ such that

$$\|Q(x, y)\|_Z \leq M \|x\|_X \cdot \|y\|_Y \text{ for all } (x, y) \in X \times Y. \tag{7.14}$$

Then apply this result to prove Lemma 7.11.

Exercise 7.13. Let $d : C(\mathbb{R}) \times C(\mathbb{R}) \to [0, \infty)$ be defined by

$$d(f, g) = \sum_{n=1}^{\infty} 2^{-n} \frac{\|f - g\|_n}{1 + \|f - g\|_n},$$

where $\|f\|_n := \sup\{|f(x)| : |x| \leq n\} = \max\{|f(x)| : |x| \leq n\}$.

1. Show that d is a metric on $C(\mathbb{R})$.

2. Show that a sequence $\{f_n\}_{n=1}^{\infty} \subset C(\mathbb{R})$ converges to $f \in C(\mathbb{R})$ as $n \to \infty$ iff f_n converges to f uniformly on bounded subsets of $\mathbb{R}$.
3. Show that $(C(\mathbb{R}), d)$ is a complete metric space.

Exercise 7.14. Let $X = C([0,1], \mathbb{R})$ and for $f \in X$, let

$$\|f\|_1 := \int_0^1 |f(t)|\, dt.$$

Show that $(X, \|\cdot\|_1)$ is normed space and show by example that this space is **not** complete. Hint: For the last assertion find a sequence of $\{f_n\}_{n=1}^{\infty} \subset X$ which is "trying" to converge to the function $f = 1_{[\frac{1}{2},1]} \notin X$.

Exercise 7.15. Let $(X, \|\cdot\|_1)$ be the normed space in Exercise 7.14. Compute the closure of A when

1. $A = \{f \in X : f(1/2) = 0\}$.
2. $A = \left\{f \in X : \sup_{t\in[0,1]} f(t) \le 5\right\}$.
3. $A = \left\{f \in X : \int_0^{1/2} f(t)\, dt = 0\right\}$.

Exercise 7.16. Suppose $\{x_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in a Banach space X. Show $\sum_{\alpha\in A} x_\alpha$ exists in X and

$$\left\|\sum_{\alpha\in A} x_\alpha\right\| \le \sum_{\alpha\in A} \|x_\alpha\|$$

if $\sum_{\alpha\in A} \|x_\alpha\| < \infty$. That is to say **"absolute convergence"** implies convergence in a Banach space.

Exercise 7.17 (Dominated Convergence Theorem Again). Let X be a Banach space, A be a set and suppose $f_n : A \to X$ is a sequence of functions. Further assume there exists a summable function $g : A \to [0,\infty)$ such that $\|f_n(\alpha)\| \le g(\alpha)$ for all $\alpha \in A$. Show $\sum_{\alpha\in A} f(\alpha)$ exists in X and

$$\lim_{n\to\infty} \sum_{\alpha\in A} f_n(\alpha) = \sum_{\alpha\in A} f(\alpha)$$

where $f(\alpha) := \lim_{n\to\infty} f_n(\alpha)$.

7.6.1 Hahn – Banach Theorem Problems

Exercise 7.18. Folland 5.20, p. 160.

Exercise 7.19. Folland 5.21, p. 160.

Exercise 7.20. Let X be a Banach space such that X^* is separable. Show X is separable as well. (The converse is not true as can be seen by taking $X = \ell^1(\mathbb{N})$.) **Hint:** use the greedy algorithm, i.e. suppose $D \subset X^* \setminus \{0\}$ is a countable dense subset of X^*, for $\ell \in D$ choose $x_\ell \in X$ such that $\|x_\ell\| = 1$ and $|\ell(x_\ell)| \geq \frac{1}{2}\|\ell\|$.

Exercise 7.21. Folland 5.26.

8

The Riemann Integral

In this Chapter, the Riemann integral for Banach space valued functions is defined and developed. Our exposition will be brief, since the Lebesgue integral and the Bochner Lebesgue integral will subsume the content of this chapter. In Definition 11.1 below, we will give a general notion of a compact subset of a "topological" space. However, by Corollary 11.9 below, when we are working with subsets of $\mathbb{R}^d$ this definition is equivalent to the following definition.

Definition 8.1. *A subset $A \subset \mathbb{R}^d$ is said to be **compact** if A is closed and **bounded.***

Theorem 8.2. *Suppose that $K \subset \mathbb{R}^d$ is a compact set and $f \in C(K, X)$. Then*

1. *Every sequence $\{u_n\}_{n=1}^{\infty} \subset K$ has a convergent subsequence.*
2. *The function f is uniformly continuous on K, namely for every $\varepsilon > 0$ there exists a $\delta > 0$ only depending on ε such that $\|f(u) - f(v)\| < \varepsilon$ whenever $u, v \in K$ and $|u - v| < \delta$ where $|\cdot|$ is the standard Euclidean norm on $\mathbb{R}^d$.*

Proof.

1. (This is a special case of Theorem 11.7 and Corollary 11.9 below.) Since K is bounded, $K \subset [-R, R]^d$ for some sufficiently large d. Let t_n be the first component of u_n so that $t_n \in [-R, R]$ for all n. Let $J_1 = [0, R]$ if $t_n \in J_1$ for infinitely many n otherwise let $J_1 = [-R, 0]$. Similarly split J_1 in half and let $J_2 \subset J_1$ be one of the halves such that $t_n \in J_2$ for infinitely many n. Continue this way inductively to find a nested sequence of intervals $J_1 \supset J_2 \supset J_3 \supset J_4 \supset \dots$ such that the length of J_k is $2^{-(k-1)}R$ and for each k, $t_n \in J_k$ for infinitely many n. We may now choose a subsequence, $\{n_k\}_{k=1}^{\infty}$ of $\{n\}_{n=1}^{\infty}$ such that $\tau_k := t_{n_k} \in J_k$ for all k. The sequence $\{\tau_k\}_{k=1}^{\infty}$ is Cauchy and hence convergent. Thus by replacing $\{u_n\}_{n=1}^{\infty}$ by a subsequence if necessary we may assume the first component of $\{u_n\}_{n=1}^{\infty}$ is

convergent. Repeating this argument for the second, then the third and all the way through the d^{th} – components of $\{u_n\}_{n=1}^{\infty}$, we may, by passing to further subsequences, assume all of the components of u_n are convergent. But this implies $\lim u_n = u$ exists and since K is closed, $u \in K$.

2. (This is a special case of Exercise 11.5 below.) If f were not uniformly continuous on K, there would exists an $\varepsilon > 0$ and sequences $\{u_n\}_{n=1}^{\infty}$ and $\{v_n\}_{n=1}^{\infty}$ in K such that

$$\|f(u_n) - f(v_n)\| \geq \varepsilon \text{ while } \lim_{n\to\infty} |u_n - v_n| = 0.$$

By passing to subsequences if necessary we may assume that $\lim_{n\to\infty} u_n$ and $\lim_{n\to\infty} v_n$ exists. Since $\lim_{n\to\infty} |u_n - v_n| = 0$, we must have

$$\lim_{n\to\infty} u_n = u = \lim_{n\to\infty} v_n$$

for some $u \in K$. Since f is continuous, vector addition is continuous and the norm is continuous, we may now conclude that

$$\varepsilon \leq \lim_{n\to\infty} \|f(u_n) - f(v_n)\| = \|f(u) - f(u)\| = 0$$

which is a contraction.

■

For the remainder of the chapter, let $[a, b]$ be a fixed compact interval and X be a Banach space. The collection $\mathcal{S} = \mathcal{S}([a, b], X)$ of **step functions,** $f : [a, b] \to X$, consists of those functions f which may be written in the form

$$f(t) = x_0 1_{[a,t_1]}(t) + \sum_{i=1}^{n-1} x_i 1_{(t_i,t_{i+1}]}(t), \tag{8.1}$$

where $\pi := \{a = t_0 < t_1 < \cdots < t_n = b\}$ is a partition of $[a, b]$ and $x_i \in X$. For f as in Eq. (8.1), let

$$I(f) := \sum_{i=0}^{n-1} (t_{i+1} - t_i) x_i \in X. \tag{8.2}$$

Exercise 8.1. Show that $I(f)$ is well defined, independent of how f is represented as a step function. (**Hint:** show that adding a point to a partition π of $[a, b]$ does not change the right side of Eq. (8.2).) Also verify that $I : \mathcal{S} \to X$ is a linear operator.

Notation 8.3 *Let $\bar{\mathcal{S}}$ denote the closure of $\mathcal{S}$ inside the Banach space, $\ell^\infty([a, b], X)$ as defined in Remark 7.6.*

The following simple "Bounded Linear Transformation" theorem will often be used in the sequel to define linear transformations.

Theorem 8.4 (B. L. T. Theorem). *Suppose that Z is a normed space, X is a Banach space, and $\mathcal{S} \subset Z$ is a dense linear subspace of Z. If $T : \mathcal{S} \to X$ is a bounded linear transformation (i.e. there exists $C < \infty$ such that $\|Tz\| \le C\|z\|$ for all $z \in \mathcal{S}$), then T has a unique extension to an element $\bar{T} \in L(Z, X)$ and this extension still satisfies*

$$\left\|\bar{T}z\right\| \le C\|z\| \text{ for all } z \in \bar{\mathcal{S}}.$$

Exercise 8.2. Prove Theorem 8.4.

Proposition 8.5 (Riemann Integral). *The linear function $I : \mathcal{S} \to X$ extends uniquely to a continuous linear operator $\bar{I}$ from $\bar{\mathcal{S}}$ to X and this operator satisfies,*

$$\|\bar{I}(f)\| \le (b-a)\|f\|_\infty \text{ for all } f \in \bar{\mathcal{S}}. \tag{8.3}$$

Furthermore, $C([a,b], X) \subset \bar{\mathcal{S}} \subset \ell^\infty([a,b], X)$ and for $f \in$, $\bar{I}(f)$ may be computed as

$$\bar{I}(f) = \lim_{|\pi| \to 0} \sum_{i=0}^{n-1} f(c_i^\pi)(t_{i+1} - t_i) \tag{8.4}$$

where $\pi := \{a = t_0 < t_1 < \cdots < t_n = b\}$ denotes a partition of $[a,b]$, $|\pi| = \max\{|t_{i+1} - t_i| : i = 0, \ldots, n-1\}$ is the mesh size of π and c_i^π may be chosen arbitrarily inside $[t_i, t_{i+1}]$. See Figure 8.1.

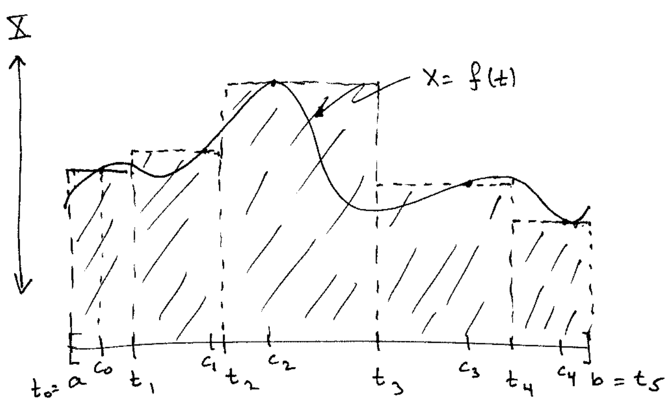

Fig. 8.1. The usual picture associated to the Riemann integral.

Proof. Taking the norm of Eq. (8.2) and using the triangle inequality shows,

$$\|I(f)\| \le \sum_{i=0}^{n-1}(t_{i+1} - t_i)\|x_i\| \le \sum_{i=0}^{n-1}(t_{i+1} - t_i)\|f\|_\infty \le (b-a)\|f\|_\infty. \tag{8.5}$$

The existence of $\bar{I}$ satisfying Eq. (8.3) is a consequence of Theorem 8.4.

Given $f \in C([a,b],X)$, $\pi := \{a = t_0 < t_1 < \cdots < t_n = b\}$ a partition of $[a,b]$, and $c_i^\pi \in [t_i, t_{i+1}]$ for $i = 0,1,2\dots,n-1$, let $f_\pi \in \mathcal{S}$ be defined by

$$f_\pi(t) := f(c_0)_0 1_{[t_0,t_1]}(t) + \sum_{i=1}^{n-1} f(c_i^\pi) 1_{(t_i,t_{i+1}]}(t).$$

Then by the uniform continuity of f on $[a,b]$ (Theorem 8.2), $\lim_{|\pi|\to 0} \|f - f_\pi\|_\infty = 0$ and therefore $f \in \bar{\mathcal{S}}$. Moreover,

$$I(f) = \lim_{|\pi|\to 0} I(f_\pi) = \lim_{|\pi|\to 0} \sum_{i=0}^{n-1} f(c_i^\pi)(t_{i+1} - t_i)$$

which proves Eq. (8.4). ■

If $f_n \in \mathcal{S}$ and $f \in \bar{\mathcal{S}}$ such that $\lim_{n\to\infty} \|f - f_n\|_\infty = 0$, then for $a \le \alpha < \beta \le b$, then $1_{(\alpha,\beta]} f_n \in \mathcal{S}$ and $\lim_{n\to\infty} \left\|1_{(\alpha,\beta]} f - 1_{(\alpha,\beta]} f_n\right\|_\infty = 0$. This shows $1_{(\alpha,\beta]} f \in \bar{\mathcal{S}}$ whenever $f \in \bar{\mathcal{S}}$.

Notation 8.6 *For $f \in \bar{\mathcal{S}}$ and $a \le \alpha \le \beta \le b$ we will write denote $\bar{I}(1_{(\alpha,\beta]} f)$ by $\int_\alpha^\beta f(t)\,dt$ or $\int_{(\alpha,\beta]} f(t)dt$. Also following the usual convention, if $a \le \beta \le \alpha \le b$, we will let*

$$\int_\alpha^\beta f(t)\,dt = -\bar{I}(1_{(\beta,\alpha]} f) = -\int_\beta^\alpha f(t)\,dt.$$

The next Lemma, whose proof is left to the reader contains some of the many familiar properties of the Riemann integral.

Lemma 8.7. *For $f \in \bar{\mathcal{S}}([a,b],X)$ and $\alpha, \beta, \gamma \in [a,b]$, the Riemann integral satisfies:*

1. $\left\|\int_\alpha^\beta f(t)\,dt\right\|_X \le (\beta - \alpha) \sup\{\|f(t)\| : \alpha \le t \le \beta\}$.
2. $\int_\alpha^\gamma f(t)\,dt = \int_\alpha^\beta f(t)\,dt + \int_\beta^\gamma f(t)\,dt$.
3. *The function $G(t) := \int_a^t f(\tau)d\tau$ is continuous on $[a,b]$.*
4. *If Y is another Banach space and $T \in L(X,Y)$, then $Tf \in \bar{\mathcal{S}}([a,b],Y)$ and*
$$T\left(\int_\alpha^\beta f(t)\ dt\right) = \int_\alpha^\beta Tf(t)\ dt.$$
5. *The function $t \to \|f(t)\|_X$ is in $\bar{\mathcal{S}}([a,b],\mathbb{R})$ and*
$$\left\|\int_a^b f(t)\ dt\right\|_X \le \int_a^b \|f(t)\|_X\ dt.$$

6. If $f, g \in \bar{\mathcal{S}}([a,b], \mathbb{R})$ and $f \le g$, then

$$\int_a^b f(t)\ dt \le \int_a^b g(t)\ dt.$$

Exercise 8.3. Prove Lemma 8.7.

8.1 The Fundamental Theorem of Calculus

Our next goal is to show that our Riemann integral interacts well with differentiation, namely the fundamental theorem of calculus holds. Before doing this we will need a couple of basic definitions and results of differential calculus, more details and the next few results below will be done in greater detail in Chapter 16.

Definition 8.8. *Let $(a,b) \subset \mathbb{R}$. A function $f : (a,b) \to X$ is differentiable at $t \in (a,b)$ iff*

$$L := \lim_{h\to 0} \left[f(t+h) - f(t)\right] h^{-1} = \lim_{h\to 0} \text{“}\frac{f(t+h) - f(t)}{h}\text{”}$$

exists in X. The limit L, if it exists, will be denoted by $\dot{f}(t)$ or $\frac{df}{dt}(t)$. We also say that $f \in C^1((a,b) \to X)$ if f is differentiable at all points $t \in (a,b)$ and $\dot{f} \in C((a,b) \to X)$.

As for the case of real valued functions, the derivative operator $\frac{d}{dt}$ is easily seen to be linear. The next two results have proves very similar to their real valued function analogues.

Lemma 8.9 (Product Rules). *Suppose that $t \to U(t) \in L(X)$, $t \to V(t) \in L(X)$ and $t \to x(t) \in X$ are differentiable at $t = t_0$, then*

1. *$\frac{d}{dt}|_{t_0} [U(t)\, x(t)] \in X$ exists and*

$$\frac{d}{dt}|_{t_0} [U(t)\, x(t)] = \left[\dot{U}(t_0)\, x(t_0) + U(t_0)\, \dot{x}(t_0)\right]$$

   and

2. *$\frac{d}{dt}|_{t_0} [U(t)\, V(t)] \in L(X)$ exists and*

$$\frac{d}{dt}|_{t_0} [U(t)\, V(t)] = \left[\dot{U}(t_0)\, V(t_0) + U(t_0)\, \dot{V}(t_0)\right].$$

3. *If $U(t_0)$ is invertible, then $t \to U(t)^{-1}$ is differentiable at $t = t_0$ and*

$$\frac{d}{dt}|_{t_0} U(t)^{-1} = -U(t_0)^{-1}\, \dot{U}(t_0)\, U(t_0)^{-1}. \tag{8.6}$$

Proof. The reader is asked to supply the proof of the first two items in Exercise 8.10. Before proving item 3., let us assume that $U(t)^{-1}$ is differentiable, then using the product rule we would learn

$$0 = \frac{d}{dt}|_{t_0} I = \frac{d}{dt}|_{t_0} \left[U(t)^{-1} U(t)\right] = \left[\frac{d}{dt}|_{t_0} U(t)^{-1}\right] U(t_0) + U(t_0)^{-1} \dot{U}(t_0).$$

Solving this equation for $\frac{d}{dt}|_{t_0} U(t)^{-1}$ gives the formula in Eq. (8.6). The problem with this argument is that we have not yet shown $t \to U(t)^{-1}$ is invertible at t_0. Here is the formal proof.

Since $U(t)$ is differentiable at t_0, $U(t) \to U(t_0)$ as $t \to t_0$ and by Corollary 7.20, $U(t_0 + h)$ is invertible for h near 0 and

$$U(t_0 + h)^{-1} \to U(t_0)^{-1} \text{ as } h \to 0.$$

Therefore, using Lemma 7.11, we may let $h \to 0$ in the identity,

$$\frac{U(t_0 + h)^{-1} - U(t_0)^{-1}}{h} = U(t_0 + h)^{-1} \left(\frac{U(t_0) - U(t_0 + h)}{h}\right) U(t_0)^{-1},$$

to learn

$$\lim_{h \to 0} \frac{U(t_0 + h)^{-1} - U(t_0)^{-1}}{h} = -U(t_0)^{-1} \dot{U}(t_0) U(t_0)^{-1}.$$

■

Proposition 8.10 (Chain Rule). *Suppose $s \to x(s) \in X$ is differentiable at $s = s_0$ and $t \to T(t) \in \mathbb{R}$ is differentiable at $t = t_0$ and $T(t_0) = s_0$, then $t \to x(T(t))$ is differentiable at t_0 and*

$$\frac{d}{dt}|_{t_0} x(T(t)) = x'(T(t_0)) T'(t_0).$$

The proof of the chain rule is essentially the same as the real valued function case, see Exercise 8.11.

Proposition 8.11. *Suppose that $f : [a, b] \to X$ is a continuous function such that $\dot{f}(t)$ exists and is equal to zero for $t \in (a, b)$. Then f is constant.*

Proof. Let $\varepsilon > 0$ and $\alpha \in (a, b)$ be given. (We will later let $\varepsilon \downarrow 0$.) By the definition of the derivative, for all $\tau \in (a, b)$ there exists $\delta_\tau > 0$ such that

$$\|f(t) - f(\tau)\| = \left\|f(t) - f(\tau) - \dot{f}(\tau)(t - \tau)\right\| \le \varepsilon |t - \tau| \text{ if } |t - \tau| < \delta_\tau. \tag{8.7}$$

Let

$$A = \{t \in [\alpha, b] : \|f(t) - f(\alpha)\| \le \varepsilon(t - \alpha)\} \tag{8.8}$$

and t_0 be the least upper bound for A. We will now use a standard argument which is referred to as **continuous induction** to show $t_0 = b$.

Eq. (8.7) with $\tau = \alpha$ shows $t_0 > \alpha$ and a simple continuity argument shows $t_0 \in A$, i.e.

$$\|f(t_0) - f(\alpha)\| \leq \varepsilon(t_0 - \alpha). \tag{8.9}$$

For the sake of contradiction, suppose that $t_0 < b$. By Eqs. (8.7) and (8.9),

$$\begin{aligned}\|f(t) - f(\alpha)\| &\leq \|f(t) - f(t_0)\| + \|f(t_0) - f(\alpha)\| \\ &\leq \varepsilon(t_0 - \alpha) + \varepsilon(t - t_0) = \varepsilon(t - \alpha)\end{aligned}$$

for $0 \leq t - t_0 < \delta_{t_0}$ which violates the definition of t_0 being an upper bound. Thus we have shown $b \in A$ and hence

$$\|f(b) - f(\alpha)\| \leq \varepsilon(b - \alpha).$$

Since $\varepsilon > 0$ was arbitrary we may let $\varepsilon \downarrow 0$ in the last equation to conclude $f(b) = f(\alpha)$. Since $\alpha \in (a, b)$ was arbitrary it follows that $f(b) = f(\alpha)$ for all $\alpha \in (a, b]$ and then by continuity for all $\alpha \in [a, b]$, i.e. f is constant. ∎

Remark 8.12. The usual real variable proof of Proposition 8.11 makes use Rolle's theorem which in turn uses the extreme value theorem. This latter theorem is not available to vector valued functions. However with the aid of the Hahn Banach Theorem 7.24 and Lemma 8.7, it is possible to reduce the proof of Proposition 8.11 and the proof of the Fundamental Theorem of Calculus 8.13 to the real valued case, see Exercise 8.24.

Theorem 8.13 (Fundamental Theorem of Calculus). *Suppose that $f \in C([a, b], X)$, Then*

1. $\frac{d}{dt}\int_a^t f(\tau)\, d\tau = f(t)$ *for all* $t \in (a, b)$.
2. *Now assume that $F \in C([a, b], X)$, F is continuously differentiable on (a, b) (i.e. $\dot{F}(t)$ exists and is continuous for $t \in (a, b)$) and $\dot{F}$ extends to a continuous function on $[a, b]$ which is still denoted by $\dot{F}$. Then*

$$\int_a^b \dot{F}(t)\, dt = F(b) - F(a).$$

Proof. Let $h > 0$ be a small number and consider

$$\begin{aligned}\left\|\int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h\right\| &= \left\|\int_t^{t+h} (f(\tau) - f(t))\, d\tau\right\| \\ &\leq \int_t^{t+h} \|(f(\tau) - f(t))\|\, d\tau \leq h\varepsilon(h),\end{aligned}$$

where $\varepsilon(h) := \max_{\tau\in[t,t+h]} \|(f(\tau) - f(t))\|$. Combining this with a similar computation when $h < 0$ shows, for all $h \in \mathbb{R}$ sufficiently small, that

$$\| \int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h\| \leq |h|\varepsilon(h),$$

where now $\varepsilon(h) := \max_{\tau\in[t-|h|,t+|h|]} \|(f(\tau) - f(t))\|$. By continuity of f at t, $\varepsilon(h) \to 0$ and hence $\frac{d}{dt}\int_a^t f(\tau)\,d\tau$ exists and is equal to $f(t)$.

For the second item, set $G(t) := \int_a^t \dot{F}(\tau)\,d\tau - F(t)$. Then G is continuous by Lemma 8.7 and $\dot{G}(t) = 0$ for all $t \in (a,b)$ by item 1. An application of Proposition 8.11 shows G is a constant and in particular $G(b) = G(a)$, i.e. $\int_a^b \dot{F}(\tau)\,d\tau - F(b) = -F(a)$. ∎

Corollary 8.14 (Mean Value Inequality). *Suppose that $f : [a,b] \to X$ is a continuous function such that $\dot{f}(t)$ exists for $t \in (a,b)$ and $\dot{f}$ extends to a continuous function on $[a,b]$. Then*

$$\|f(b) - f(a)\| \leq \int_a^b \|\dot{f}(t)\|dt \leq (b-a)\cdot \left\|\dot{f}\right\|_\infty . \tag{8.10}$$

Proof. By the fundamental theorem of calculus, $f(b) - f(a) = \int_a^b \dot{f}(t)dt$ and then by Lemma 8.7,

$$\begin{aligned}\|f(b) - f(a)\| &= \left\|\int_a^b \dot{f}(t)dt\right\| \leq \int_a^b \|\dot{f}(t)\|dt \\ &\leq \int_a^b \left\|\dot{f}\right\|_\infty dt = (b-a)\cdot \left\|\dot{f}\right\|_\infty .\end{aligned}$$

∎

Corollary 8.15 (Change of Variable Formula). *Suppose that $f \in C([a,b],X)$ and $T : [c,d] \to (a,b)$ is a continuous function such that $T(s)$ is continuously differentiable for $s \in (c,d)$ and $T'(s)$ extends to a continuous function on $[c,d]$. Then*

$$\int_c^d f(T(s))\,T'(s)\,ds = \int_{T(c)}^{T(d)} f(t)\,dt.$$

Proof. For $s \in (a,b)$ define $F(t) := \int_{T(c)}^t f(\tau)\,d\tau$. Then $F \in C^1((a,b),X)$ and by the fundamental theorem of calculus and the chain rule,

$$\frac{d}{ds}F(T(s)) = F'(T(s))\,T'(s) = f(T(s))\,T'(s)\,.$$

Integrating this equation on $s \in [c,d]$ and using the chain rule again gives

$$\int_c^d f(T(s))\,T'(s)\,ds = F(T(d)) - F(T(c)) = \int_{T(c)}^{T(d)} f(t)\,dt.$$

∎

8.2 Integral Operators as Examples of Bounded Operators

In the examples to follow all integrals are the standard Riemann integrals and we will make use of the following notation.

Notation 8.16 *Given an open set $U \subset \mathbb{R}^d$, let $C_c(U)$ denote the collection of real valued continuous functions f on U such that*

$$\operatorname{supp}(f) := \overline{\{x \in U : f(x) \neq 0\}}$$

is a compact subset of U.

Example 8.17. Suppose that $K : [0,1] \times [0,1] \to \mathbb{C}$ is a continuous function. For $f \in C([0,1])$, let

$$Tf(x) = \int_0^1 K(x,y) f(y) dy.$$

Since

$$\begin{aligned} |Tf(x) - Tf(z)| &\le \int_0^1 |K(x,y) - K(z,y)| \, |f(y)| \, dy \\ &\le \|f\|_\infty \max_y |K(x,y) - K(z,y)| \end{aligned} \tag{8.11}$$

and the latter expression tends to 0 as $x \to z$ by uniform continuity of K. Therefore $Tf \in C([0,1])$ and by the linearity of the Riemann integral, $T : C([0,1]) \to C([0,1])$ is a linear map. Moreover,

$$|Tf(x)| \le \int_0^1 |K(x,y)| \, |f(y)| \, dy \le \int_0^1 |K(x,y)| \, dy \cdot \|f\|_\infty \le A \|f\|_\infty$$

where

$$A := \sup_{x \in [0,1]} \int_0^1 |K(x,y)| \, dy < \infty. \tag{8.12}$$

This shows $\|T\| \le A < \infty$ and therefore T is bounded. We may in fact show $\|T\| = A$. To do this let $x_0 \in [0,1]$ be such that

$$\sup_{x \in [0,1]} \int_0^1 |K(x,y)| \, dy = \int_0^1 |K(x_0,y)| \, dy.$$

Such an x_0 can be found since, using a similar argument to that in Eq. (8.11), $x \to \int_0^1 |K(x,y)| \, dy$ is continuous. Given $\varepsilon > 0$, let

$$f_\varepsilon(y) := \frac{\overline{K(x_0,y)}}{\sqrt{\varepsilon + |K(x_0,y)|^2}}$$

and notice that $\lim_{\varepsilon\downarrow 0}\|f_\varepsilon\|_\infty = 1$ and

$$\|Tf_\varepsilon\|_\infty \geq |Tf_\varepsilon(x_0)| = Tf_\varepsilon(x_0) = \int_0^1 \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}}dy.$$

Therefore,

$$\begin{aligned}\|T\| &\geq \lim_{\varepsilon\downarrow 0}\frac{1}{\|f_\varepsilon\|_\infty}\int_0^1 \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}}dy \\ &= \lim_{\varepsilon\downarrow 0}\int_0^1 \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}}dy = A\end{aligned}$$

since

$$\begin{aligned}0 &\leq |K(x_0,y)| - \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}} \\ &= \frac{|K(x_0,y)|}{\sqrt{\varepsilon + |K(x_0,y)|^2}}\left[\sqrt{\varepsilon + |K(x_0,y)|^2} - |K(x_0,y)|\right] \\ &\leq \sqrt{\varepsilon + |K(x_0,y)|^2} - |K(x_0,y)|\end{aligned}$$

and the latter expression tends to zero uniformly in y as $\varepsilon \downarrow 0$.

We may also consider other norms on $C([0,1])$. Let (for now) $L^1([0,1])$ denote $C([0,1])$ with the norm

$$\|f\|_1 = \int_0^1 |f(x)|\,dx,$$

then $T : L^1([0,1],dm) \to C([0,1])$ is bounded as well. Indeed, let $M = \sup\{|K(x,y)| : x,y \in [0,1]\}$, then

$$|(Tf)(x)| \leq \int_0^1 |K(x,y)f(y)|\,dy \leq M\|f\|_1$$

which shows $\|Tf\|_\infty \leq M\|f\|_1$ and hence,

$$\|T\|_{L^1\to C} \leq \max\{|K(x,y)| : x,y \in [0,1]\} < \infty.$$

We can in fact show that $\|T\| = M$ as follows. Let $(x_0,y_0) \in [0,1]^2$ satisfying $|K(x_0,y_0)| = M$. Then given $\varepsilon > 0$, there exists a neighborhood $U = I \times J$ of (x_0,y_0) such that $|K(x,y) - K(x_0,y_0)| < \varepsilon$ for all $(x,y) \in U$. Let $f \in C_c(I,[0,\infty))$ such that $\int_0^1 f(x)dx = 1$. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $\alpha K(x_0,y_0) = M$, then

$$\begin{aligned}|(T\alpha f)(x_0)| &= \left|\int_0^1 K(x_0,y)\alpha f(y)dy\right| = \left|\int_I K(x_0,y)\alpha f(y)dy\right| \\ &\geq \operatorname{Re}\int_I \alpha K(x_0,y)f(y)dy \\ &\geq \int_I (M-\varepsilon)\, f(y)dy = (M-\varepsilon)\, \|\alpha f\|_{L^1}\end{aligned}$$

and hence

$$\|T\alpha f\|_C \geq (M-\varepsilon)\, \|\alpha f\|_{L^1}$$

showing that $\|T\| \geq M-\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we learn that $\|T\| \geq M$ and hence $\|T\| = M$.

One may also view T as a map from $T : C([0,1]) \to L^1([0,1])$ in which case one may show

$$\|T\|_{L^1 \to C} \leq \int_0^1 \max_y |K(x,y)|\, dx < \infty.$$

8.3 Linear Ordinary Differential Equations

Let X be a Banach space, $J = (a,b) \subset \mathbb{R}$ be an open interval with $0 \in J$, $h \in C(J \to X)$ and $A \in C(J \to L(X))$. In this section we are going to consider the ordinary differential equation,

$$\dot{y}(t) = A(t)y(t) + h\,(t) \text{ where } y(0) = x \in X, \tag{8.13}$$

where y is an unknown function in $C^1(J \to X)$. This equation may be written in its equivalent (as the reader should verify) integral form, namely we are looking for $y \in C(J,X)$ such that

$$y(t) = x + \int_0^t h\,(\tau)\, d\tau + \int_0^t A(\tau)y(\tau)d\tau. \tag{8.14}$$

In what follows, we will abuse notation and use $\|\cdot\|$ to denote the operator norm on $L\,(X)$ associated to then norm, $\|\cdot\|$, on X and let $\|\phi\|_\infty := \max_{t\in J}\|\phi(t)\|$ for $\phi \in BC(J,X)$ or $BC(J, L\,(X))$.

Notation 8.18 *For $t \in \mathbb{R}$ and $n \in \mathbb{N}$, let*

$$\Delta_n(t) = \begin{cases} \{(\tau_1,\dots,\tau_n) \in \mathbb{R}^n : 0 \leq \tau_1 \leq \dots \leq \tau_n \leq t\} \text{ if } t \geq 0 \\ \{(\tau_1,\dots,\tau_n) \in \mathbb{R}^n : t \leq \tau_n \leq \dots \leq \tau_1 \leq 0\} \text{ if } t \leq 0 \end{cases}$$

and also write $d\tau = d\tau_1 \dots d\tau_n$ and

$$\int_{\Delta_n(t)} f(\tau_1,\dots\tau_n)d\tau := (-1)^{n\cdot 1_{t<0}} \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} \dots \int_0^{\tau_2} d\tau_1 f(\tau_1,\dots\tau_n).$$

Lemma 8.19. *Suppose that* $\psi \in C(\mathbb{R}, \mathbb{R})$, *then*

$$(-1)^{n \cdot 1_{t<0}} \int_{\Delta_n(t)} \psi(\tau_1) \dots \psi(\tau_n) d\tau = \frac{1}{n!} \left(\int_0^t \psi(\tau) d\tau \right)^n . \tag{8.15}$$

Proof. Let $\Psi(t) := \int_0^t \psi(\tau) d\tau$. The proof will go by induction on n. The case $n = 1$ is easily verified since

$$(-1)^{1 \cdot 1_{t<0}} \int_{\Delta_1(t)} \psi(\tau_1) d\tau_1 = \int_0^t \psi(\tau) d\tau = \Psi(t).$$

Now assume the truth of Eq. (8.15) for $n-1$ for some $n \geq 2$, then

$$\begin{aligned}
(-1)^{n \cdot 1_{t<0}} & \int_{\Delta_n(t)} \psi(\tau_1) \dots \psi(\tau_n) d\tau \\
&= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} \dots \int_0^{\tau_2} d\tau_1 \psi(\tau_1) \dots \psi(\tau_n) \\
&= \int_0^t d\tau_n \frac{\Psi^{n-1}(\tau_n)}{(n-1)!} \psi(\tau_n) = \int_0^t d\tau_n \frac{\Psi^{n-1}(\tau_n)}{(n-1)!} \dot{\Psi}(\tau_n) \\
&= \int_0^{\Psi(t)} \frac{u^{n-1}}{(n-1)!} du = \frac{\Psi^n(t)}{n!},
\end{aligned}$$

wherein we made the change of variables, $u = \Psi(\tau_n)$, in the second to last equality. ■

Remark 8.20. Eq. (8.15) is equivalent to

$$\int_{\Delta_n(t)} \psi(\tau_1) \dots \psi(\tau_n) d\tau = \frac{1}{n!} \left(\int_{\Delta_1(t)} \psi(\tau) d\tau \right)^n$$

and another way to understand this equality is to view $\int_{\Delta_n(t)} \psi(\tau_1) \dots \psi(\tau_n) d\tau$ as a multiple integral (see Section 20 below) rather than an iterated integral. Indeed, taking $t > 0$ for simplicity and letting S_n be the permutation group on $\{1, 2, \dots, n\}$ we have

$$[0, t]^n = \cup_{\sigma \in S_n} \{(\tau_1, \dots, \tau_n) \in \mathbb{R}^n : 0 \leq \tau_{\sigma 1} \leq \dots \leq \tau_{\sigma n} \leq t\}$$

with the union being "essentially" disjoint. Therefore, making a change of variables and using the fact that $\psi(\tau_1) \dots \psi(\tau_n)$ is invariant under permutations, we find

$$\begin{aligned}
\left(\int_0^t \psi(\tau)d\tau\right)^n &= \int_{[0,t]^n} \psi(\tau_1)\dots\psi(\tau_n)d\tau \\
&= \sum_{\sigma\in S_n} \int_{\{(\tau_1,\dots,\tau_n)\in\mathbb{R}^n:0\le\tau_{\sigma 1}\le\dots\le\tau_{\sigma n}\le t\}} \psi(\tau_1)\dots\psi(\tau_n)d\tau \\
&= \sum_{\sigma\in S_n} \int_{\{(s_1,\dots,s_n)\in\mathbb{R}^n:0\le s_1\le\dots\le s_n\le t\}} \psi(s_{\sigma^{-1}1})\dots\psi(s_{\sigma^{-1}n})d\mathbf{s} \\
&= \sum_{\sigma\in S_n} \int_{\{(s_1,\dots,s_n)\in\mathbb{R}^n:0\le s_1\le\dots\le s_n\le t\}} \psi(s_1)\dots\psi(s_n)d\mathbf{s} \\
&= n!\int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau.
\end{aligned}$$

Theorem 8.21. *Let $\phi\in BC(J,X)$, then the integral equation*

$$y(t) = \phi(t) + \int_0^t A(\tau)y(\tau)d\tau \tag{8.16}$$

has a unique solution given by

$$y(t) = \phi(t) + \sum_{n=1}^{\infty} (-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(\tau_1)d\tau \tag{8.17}$$

and this solution satisfies the bound

$$\|y\|_\infty \le \|\phi\|_\infty e^{\int_J \|A(\tau)\|d\tau}.$$

Proof. Define $\Lambda : BC(J,X)\to BC(J,X)$ by

$$(\Lambda y)(t) = \int_0^t A(\tau)y(\tau)d\tau.$$

Then y solves Eq. (8.14) iff $y=\phi+\Lambda y$ or equivalently iff $(I-\Lambda)y=\phi$.

An induction argument shows

$$\begin{aligned}
(\Lambda^n\phi)(t) &= \int_0^t d\tau_n A(\tau_n)(\Lambda^{n-1}\phi)(\tau_n) \\
&= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} A(\tau_n)A(\tau_{n-1})(\Lambda^{n-2}\phi)(\tau_{n-1}) \\
&\vdots \\
&= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1}\dots\int_0^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1)\phi(\tau_1) \\
&= (-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(\tau_1)d\tau.
\end{aligned}$$

Taking norms of this equation and using the triangle inequality along with Lemma 8.19 gives,

$$\begin{aligned}\|(\Lambda^n\phi)(t)\| \leq \|\phi\|_\infty \cdot \int_{\Delta_n(t)} \|A(\tau_n)\| \ldots \|A(\tau_1)\| d\tau \\ \leq \|\phi\|_\infty \cdot \frac{1}{n!}\left(\int_{\Delta_1(t)} \|A(\tau)\| d\tau\right)^n \\ \leq \|\phi\|_\infty \cdot \frac{1}{n!}\left(\int_J \|A(\tau)\| d\tau\right)^n.\end{aligned}$$

Therefore,

$$\|\Lambda^n\|_{op} \leq \frac{1}{n!}\left(\int_J \|A(\tau)\| d\tau\right)^n \tag{8.18}$$

and

$$\sum_{n=0}^{\infty} \|\Lambda^n\|_{op} \leq e^{\int_J \|A(\tau)\| d\tau} < \infty$$

where $\|\cdot\|_{op}$ denotes the operator norm on $L\left(BC(J,X)\right)$. An application of Proposition 7.19 now shows $(I-\Lambda)^{-1} = \sum_{n=0}^{\infty} \Lambda^n$ exists and

$$\left\|(I-\Lambda)^{-1}\right\|_{op} \leq e^{\int_J \|A(\tau)\| d\tau}.$$

It is now only a matter of working through the notation to see that these assertions prove the theorem. ■

Corollary 8.22. *Suppose $h \in C(J \to X)$ and $x \in X$, then there exits a unique solution, $y \in C^1(J,X)$, to the linear ordinary differential Eq. (8.13).*

Proof. Let

$$\phi(t) = x + \int_0^t h(\tau)\, d\tau.$$

By applying Theorem 8.21 with and J replaced by any open interval J_0 such that $0 \in J_0$ and $\bar{J}_0$ is a compact subinterval[1] of J, there exists a unique solution y_{J_0} to Eq. (8.13) which is valid for $t \in J_0$. By uniqueness of solutions, if J_1 is a subinterval of J such that $J_0 \subset J_1$ and $\bar{J}_1$ is a compact subinterval of J, we have $y_{J_1} = y_{J_0}$ on J_0. Because of this observation, we may construct a solution y to Eq. (8.13) which is defined on the full interval J by setting $y(t) = y_{J_0}(t)$ for any J_0 as above which also contains $t \in J$. ■

Corollary 8.23. *Suppose that $A \in L(X)$ is independent of time, then the solution to*

$$\dot{y}(t) = Ay(t) \text{ with } y(0) = x$$

[1] We do this so that $\phi|_{J_0}$ will be bounded.

is given by $y(t) = e^{tA}x$ *where*

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n. \tag{8.19}$$

Moreover,

$$e^{(t+s)A} = e^{tA}e^{sA} \text{ for all } s, t \in \mathbb{R}. \tag{8.20}$$

Proof. The first assertion is a simple consequence of Eq. 8.17 and Lemma 8.19 with $\psi = 1$. The assertion in Eq. (8.20) may be proved by explicit computation but the following proof is more instructive.

Given $x \in X$, let $y(t) := e^{(t+s)A}x$. By the chain rule,

$$\begin{aligned} \frac{d}{dt} y(t) &= \frac{d}{d\tau}|_{\tau = t+s} e^{\tau A}x = Ae^{\tau A}x|_{\tau = t+s} \\ &= Ae^{(t+s)A}x = Ay(t) \text{ with } y(0) = e^{sA}x. \end{aligned}$$

The unique solution to this equation is given by

$$y(t) = e^{tA}x(0) = e^{tA}e^{sA}x.$$

This completes the proof since, by definition, $y(t) = e^{(t+s)A}x$. ■

We also have the following converse to this corollary whose proof is outlined in Exercise 8.21 below.

Theorem 8.24. *Suppose that* $T_t \in L(X)$ *for* $t \geq 0$ *satisfies*

1. *(Semi-group property.)* $T_0 = Id_X$ *and* $T_tT_s = T_{t+s}$ *for all* $s, t \geq 0$.
2. *(Norm Continuity)* $t \to T_t$ *is continuous at* 0, *i.e.* $\|T_t - I\|_{L(X)} \to 0$ *as* $t \downarrow 0$.

Then there exists $A \in L(X)$ *such that* $T_t = e^{tA}$ *where* e^{tA} *is defined in Eq. (8.19).*

8.4 Classical Weierstrass Approximation Theorem

Definition 8.25 (Support). *Let* $f : X \to Y$ *be a function from a topological space* (X, τ_X) *to a vector space* Y. *Then we define the support of* f *by*

$$\operatorname{supp}(f) := \overline{\{x \in X : f(x) \neq 0\}},$$

a closed subset of X.

Example 8.26. For example if $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = \sin(x)1_{[0,4\pi]}(x) \in \mathbb{R}$, then

$$\{f \neq 0\} = (0, 4\pi) \setminus \{\pi, 2\pi, 3\pi\}$$

and therefore $\operatorname{supp}(f) = [0, 4\pi]$.

Definition 8.27 (Convolution). *For $f, g \in C(\mathbb{R})$ with either f or g having compact support, we define the* ***convolution*** *of f and g by*

$$f * g(x) = \int_{\mathbb{R}} f(x-y)g(y)dy = \int_{\mathbb{R}} f(y)g(x-y)dy.$$

Lemma 8.28 (Approximate δ − sequences). *Suppose that $\{q_n\}_{n=1}^{\infty}$ is a sequence non-negative continuous real valued functions on $\mathbb{R}$ with compact support that satisfy*

$$\int_{\mathbb{R}} q_n(x)\ dx = 1 \text{ and} \tag{8.21}$$

$$\lim_{n\to\infty} \int_{|x|\geq\varepsilon} q_n(x)dx = 0 \text{ for all } \varepsilon > 0. \tag{8.22}$$

If W is a compact subset of $\mathbb{R}^d$ and $f \in BC(\mathbb{R}\times W)$, then

$$q_n * f\,(x,w) := \int_{\mathbb{R}} q_n(y) f(x-y,w)dy$$

converges to f uniformly on compact subsets of $\mathbb{R}\times W \subset \mathbb{R}^{d+1}$.

Proof. Let $(x,w) \in \mathbb{R}\times W$, then because of Eq. (8.21),

$$\begin{aligned} |q_n * f(x,w) - f(x,w)| &= \left| \int_{\mathbb{R}} q_n(y)\left(f(x-y,w) - f(x,w)\right) dy \right| \\ &\leq \int_{\mathbb{R}} q_n(y) \left| f(x-y,w) - f(x,w) \right| dy. \end{aligned}$$

Let $M = \sup\{|f(x,w)| : (x,w) \in \mathbb{R}\times W\}$. Then for any $\varepsilon > 0$, using Eq. (8.21),

$$\begin{aligned} |q_n * f(x,w) - f(x,w)| &\leq \int_{|y|\leq\varepsilon} q_n(y) \left| f(x-y,w) - f(x,w) \right| dy \\ &\quad + \int_{|y|>\varepsilon} q_n(y) \left| f(x-y,w) - f(x,w) \right| dy \\ &\leq \sup_{|z|\leq\varepsilon} |f(x+z,w) - f(x,w)| + 2M \int_{|y|>\varepsilon} q_n(y)dy. \end{aligned}$$

So if K is a compact subset of $\mathbb{R}$ (for example a large interval) we have

$$\begin{aligned} &\sup_{(x,w)\in K\times W} |q_n * f(x,w) - f(x,w)| \\ &\qquad \leq \sup_{|z|\leq\varepsilon,(x,w)\in K\times W} |f(x+z,w) - f(x,w)| + 2M \int_{|y|>\varepsilon} q_n(y)dy \end{aligned}$$

and hence by Eq. (8.22),

$$\limsup_{n\to\infty} \sup_{(x,w)\in K\times W} |q_n * f(x,w) - f(x,w)| \le \sup_{|z|\le\varepsilon,(x,w)\in K\times W} |f(x+z,w) - f(x,w)|\,.$$

This finishes the proof since the right member of this equation tends to 0 as $\varepsilon \downarrow 0$ by uniform continuity of f on compact subsets of $\mathbb{R}\times W$. ■

Let $q_n : \mathbb{R} \to [0,\infty)$ be defined by

$$q_n(x) := \frac{1}{c_n}(1-x^2)^n 1_{|x|\le 1} \text{ where } c_n := \int_{-1}^{1}(1-x^2)^n dx. \tag{8.23}$$

Figure 8.2 displays the key features of the functions q_n.

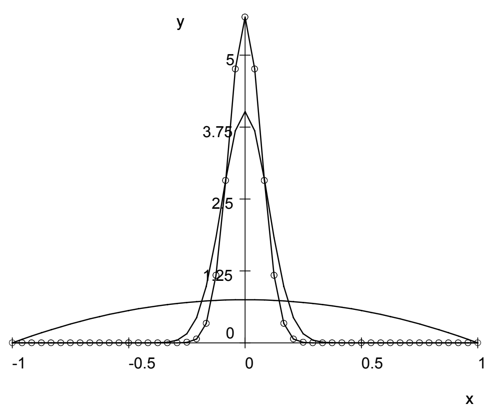

Fig. 8.2. A plot of q_1, q_{50}, and q_{100}. The most peaked curve is q_{100} and the least is q_1. The total area under each of these curves is one.

Lemma 8.29. *The sequence* $\{q_n\}_{n=1}^{\infty}$ *is an approximate* δ *– sequence, i.e. they satisfy Eqs. (8.21) and (8.22).*

Proof. By construction, $q_n \in C_c(\mathbb{R},[0,\infty))$ for each n and Eq. 8.21 holds. Since

$$\int_{|x|\geq\varepsilon} q_n(x)dx = \frac{2\int_\varepsilon^1 (1-x^2)^n dx}{2\int_0^\varepsilon (1-x^2)^n dx + 2\int_\varepsilon^1 (1-x^2)^n dx}$$
$$\leq \frac{\int_\varepsilon^1 \frac{x}{\varepsilon}(1-x^2)^n dx}{\int_0^\varepsilon \frac{x}{\varepsilon}(1-x^2)^n dx} = \frac{(1-x^2)^{n+1}|_\varepsilon^1}{(1-x^2)^{n+1}|_0^\varepsilon}$$
$$= \frac{(1-\varepsilon^2)^{n+1}}{1-(1-\varepsilon^2)^{n+1}} \to 0 \text{ as } n\to\infty,$$

the proof is complete. ■

Notation 8.30 *Let $\mathbb{Z}_+ := \mathbb{N}\cup\{0\}$ and for $x\in\mathbb{R}^d$ and $\alpha\in\mathbb{Z}_+^d$ let $x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$ and $|\alpha| = \sum_{i=1}^d \alpha_i$. A polynomial on $\mathbb{R}^d$ is a function $p:\mathbb{R}^d\to\mathbb{C}$ of the form*

$$p(x) = \sum_{\alpha:|\alpha|\leq N} p_\alpha x^\alpha \text{ with } p_\alpha\in\mathbb{C} \text{ and } N\in\mathbb{Z}_+.$$

If $p_\alpha \neq 0$ for some α such that $|\alpha| = N$, then we define $\deg(p) := N$ to be the degree of p. The function p has a natural extension to $z\in\mathbb{C}^d$, namely $p(z) = \sum_{\alpha:|\alpha|\leq N} p_\alpha z^\alpha$ where $z^\alpha = \prod_{i=1}^d z_i^{\alpha_i}$.

Theorem 8.31 (Weierstrass Approximation Theorem). *Suppose that $K\subset\mathbb{R}^d$ is a compact subset and $f\in C(K,\mathbb{C})$[2]. Then there exists polynomials p_n on $\mathbb{R}^d$ such that $p_n\to f$ uniformly on K.*

Proof. Choose $\lambda>0$ and $b\in\mathbb{R}^d$ such that

$$K_0 := \lambda K - b := \{\lambda x - b : x\in K\}\subset B_d$$

where $B_d := (0,1)^d$. The function $F(y) := f\left(\lambda^{-1}(y+b)\right)$ for $y\in K_0$ is in $C(K_0,\mathbb{C})$ and if $\hat{p}_n(y)$ are polynomials on $\mathbb{R}^d$ such that $\hat{p}_n\to F$ uniformly on K_0 then $p_n(x) := \hat{p}_n(\lambda x - b)$ are polynomials on $\mathbb{R}^d$ such that $p_n\to f$ uniformly on K. Hence we may now assume that K is a compact subset of B_d.

Let $g\in C(K\cup B_d^c)$ be defined by

$$g(x) = \begin{cases} f(x) \text{ if } x\in K \\ 0 \quad\text{ if } x\in B_d^c \end{cases}$$

and then use the Tietze extension Theorem 7.4 to find a continuous function $F\in C(\mathbb{R}^d,\mathbb{C})$ such that $F = g|_{K\cup B_d^c}$. If p_n are polynomials on $\mathbb{R}^d$ such

[2] Note that f is automatically bounded because if not there would exist $u_n\in K$ such that $\lim_{n\to\infty}|f(u_n)| = \infty$. Using Theorem 8.2 we may, by passing to a subsequence if necessary, assume $u_n\to u\in K$ as $n\to\infty$. Now the continuity of f would then imply

$$\infty = \lim_{n\to\infty}|f(u_n)| = |f(u)|$$

which is absurd since f takes values in $\mathbb{C}$.

that $p_n \to F$ uniformly on $[0,1]^d$ then p_n also converges to f uniformly on K. Hence, by replacing f by F, we may now assume that $f \in C(\mathbb{R}^d, \mathbb{C})$, $K = \bar{B}_d = [0,1]^d$, and $f \equiv 0$ on B_d^c.

With q_n defined as in Eq. (8.23), $x \in [0,1]$ and $w \in \mathbb{R}^{d-1}$, let

$$\begin{aligned} f_n(x,w) &:= (q_n * f)(x,w) = \int_{\mathbb{R}} q_n(x-y) f(y,w) dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y,w) \left[(1-(x-y)^2)^n 1_{|x-y|\le 1}\right] dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y,w)(1-(x-y)^2)^n dy = \sum_{k=0}^{2n} A_k^n(w)\, x^k \end{aligned}$$

where

$$A_k^n(w) = \int_{[0,1]} f(y,w)\, \rho_k^n(y)\, dy$$

and ρ_k is a polynomial function in y for each k. Then $A_k^n(w) = 0$ if $w \notin (0,1)^{d-1}$ and using the uniform continuity of f on $[0,1]^d$, one easily shows $A_k^n \in C\left(\mathbb{R}^{d-1}, \mathbb{C}\right)$. Moreover by Lemmas 8.28 and 8.29, $f_n(x,w) \to f(x,w)$ uniformly for $(x,w) \in [0,1]^d$ as $n \to \infty$. This completes the proof of $d = 1$ since then A_k^n are constants and $p_n(x) := f_n(x)$ is a polynomial in x.

The case of general d now follows by induction. Indeed, by the inductive hypothesis there exists polynomial functions a_k^n on $\mathbb{R}^{d-1}$ such that

$$\sup_{w \in [0,1]^{d-1}} |A_k^n(w) - a_k^n(w)| \le \frac{1}{2(n+1)n}.$$

Then

$$p_n(x,w) := \sum_{k=0}^{2n} a_k^n(w)\, x^k$$

is a polynomial function on $\mathbb{R}^d$ such that

$$\begin{aligned} |f(x,w) - p_n(x,w)| &\le |f(x,w) - f_n(x,w)| + |f_n(x,w) - p_n(x,w)| \\ &\le \varepsilon_n + \sum_{k=0}^{2n} |A_k^n(w) - a_k^n(w)|\, x^k \\ &\le \varepsilon_n + \frac{1}{n} \end{aligned}$$

where

$$\varepsilon_n := \sup_{(x,w) \in [0,1]^d} |f(x,w) - f_n(x,w)| \to 0 \text{ as } n \to \infty.$$

This shows

$$\sup_{(x,w) \in [0,1]^d} |f(x,w) - p_n(x,w)| \to 0 \text{ as } n \to \infty$$

which completes the proof. ■

Remark 8.32. The mapping $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d \to z = x+iy \in \mathbb{C}^d$ is an isomorphism of vector spaces. Letting $\bar{z} = x - iy$ as usual, we have $x = \frac{z+\bar{z}}{2}$ and $y = \frac{z-\bar{z}}{2i}$. Therefore under this identification any polynomial $p(x,y)$ on $\mathbb{R}^d \times \mathbb{R}^d$ may be written as a polynomial q in $(z,\bar{z})$, namely

$$q(z,\bar{z}) = p(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}).$$

Conversely a polynomial q in $(z,\bar{z})$ may be thought of as a polynomial p in (x,y), namely $p(x,y) = q(x+iy, x-iy)$.

Corollary 8.33 (Complex Weierstrass Approximation Theorem). *Suppose that $K \subset \mathbb{C}^d$ is a compact set and $f \in C(K,\mathbb{C})$. Then there exists polynomials $p_n(z,\bar{z})$ for $z \in \mathbb{C}^d$ such that $\sup_{z\in K} |p_n(z,\bar{z}) - f(z)| \to 0$ as $n \to \infty$.*

Proof. This is an immediate consequence of Theorem 8.31 and Remark 8.32. ■

Example 8.34. Let $K = S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and $\mathcal{A}$ be the set of polynomials in $(z,\bar{z})$ restricted to S^1. Then $\mathcal{A}$ is dense in $C(S^1)$.[3] Since $\bar{z} = z^{-1}$ on S^1, we have shown polynomials in z and z^{-1} are dense in $C(S^1)$. This example generalizes in an obvious way to $K = \left(S^1\right)^d \subset \mathbb{C}^d$.

Exercise 8.4. Suppose $-\infty < a < b < \infty$ and $f \in C\left([a,b],\mathbb{C}\right)$ satisfies

$$\int_a^b f\left(t\right) t^n dt = 0 \text{ for } n = 0,1,2\ldots.$$

Show $f \equiv 0$.

Exercise 8.5. Suppose $f \in C\left(\mathbb{R},\mathbb{C}\right)$ is a 2π – periodic function (i.e. $f\left(x+2\pi\right) = f\left(x\right)$ for all $x \in \mathbb{R}$) and

$$\int_0^{2\pi} f\left(x\right) e^{inx} dx = 0 \text{ for all } n \in \mathbb{Z},$$

show again that $f \equiv 0$. **Hint**: Use Example 8.34 to shows that any 2π – periodic continuous function g on $\mathbb{R}$ is the uniform limit of trigonometric polynomials of the form

$$p\left(x\right) = \sum_{k=-n}^{n} p_k e^{ikx} \text{ with } p_k \in \mathbb{C} \text{ for all } k.$$

[3] Note that it is easy to extend $f \in C(S^1)$ to a function $F \in C(\mathbb{C})$ by setting $F(z) = zf(\frac{z}{|z|})$ for $z \neq 0$ and $F(0) = 0$. So this special case does not require the Tietze extension theorem.

8.5 Iterated Integrals

Theorem 8.35 (Baby Fubini Theorem). *Let $a, b, c, d \in \mathbb{R}$ and $f(s,t) \in X$ be a continuous function of (s,t) for s between a and b and t between c and d. Then the maps $t \to \int_a^b f(s,t)ds \in X$ and $s \to \int_c^d f(s,t)dt$ are continuous and*

$$\int_c^d \left[\int_a^b f(s,t)ds\right] dt = \int_a^b \left[\int_c^d f(s,t)dt\right] ds. \tag{8.24}$$

Proof. See Exercise 8.7 for a sketch of another, more instructive, proof of this result. (BRUCE: Drop the following proof and leave it as an exercise.) With out loss of generality we may assume $a < b$ and $c < d$. By uniform continuity of f (Theorem 8.2),

$$\sup_{c \le t \le d} \|f(s,t) - f(s_0,t)\| \to 0 \text{ as } s \to s_0$$

and so by Lemma 8.7

$$\int_c^d f(s,t)dt \to \int_c^d f(s_0,t)dt \text{ as } s \to s_0$$

showing the continuity of $s \to \int_c^d f(s,t)dt$. The other continuity assertion is proved similarly.

Now let

$$\pi = \{a \le s_0 < s_1 < \cdots < s_m = b\} \text{ and } \pi' = \{c \le t_0 < t_1 < \cdots < t_n = d\}$$

be partitions of $[a,b]$ and $[c,d]$ respectively. For $s \in [a,b]$ let $s_\pi = s_i$ if $s \in (s_i, s_{i+1}]$ and $i \ge 1$ and $s_\pi = s_0 = a$ if $s \in [s_0, s_1]$. Define $t_{\pi'}$ for $t \in [c,d]$ analogously. Then

$$\begin{aligned}\int_a^b \left[\int_c^d f(s,t)dt\right] ds &= \int_a^b \left[\int_c^d f(s,t_{\pi'})dt\right] ds + \int_a^b \varepsilon_{\pi'}(s)ds \\ &= \int_a^b \left[\int_c^d f(s_\pi,t_{\pi'})dt\right] ds + \delta_{\pi,\pi'} + \int_a^b \varepsilon_{\pi'}(s)ds\end{aligned}$$

where

$$\varepsilon_{\pi'}(s) = \int_c^d f(s,t)dt - \int_c^d f(s,t_{\pi'})dt$$

and

$$\delta_{\pi,\pi'} = \int_a^b \left[\int_c^d \{f(s,t_{\pi'}) - f(s_\pi,t_{\pi'})\}\, dt\right] ds.$$

The uniform continuity of f and the estimates

$$\sup_{s\in[a,b]}\|\varepsilon_{\pi'}(s)\| \le \sup_{s\in[a,b]}\int_c^d \|f(s,t)-f(s,t_{\pi'})\|\,dt$$
$$\le (d-c)\sup\{\|f(s,t)-f(s,t_{\pi'})\| : (s,t)\in Q\}$$

and

$$\|\delta_{\pi,\pi'}\| \le \int_a^b \left[\int_c^d \|f(s,t_{\pi'})-f(s_\pi,t_{\pi'})\|\,dt\right] ds$$
$$\le (b-a)(d-c)\sup\{\|f(s,t)-f(s,t_{\pi'})\| : (s,t)\in Q\}$$

allow us to conclude that

$$\int_a^b \left[\int_c^d f(s,t)dt\right] ds - \int_a^b \left[\int_c^d f(s_\pi,t_{\pi'})dt\right] ds \to 0 \text{ as } |\pi|+|\pi'|\to 0.$$

By symmetry (or an analogous argument),

$$\int_c^d \left[\int_a^b f(s,t)ds\right] dt - \int_c^d \left[\int_a^b f(s_\pi,t_{\pi'})ds\right] dt \to 0 \text{ as } |\pi|+|\pi'|\to 0.$$

This completes the proof since

$$\int_a^b \left[\int_c^d f(s_\pi,t_{\pi'})dt\right] ds = \sum_{0\le i<m,0\le j<n} f(s_i,t_j)(s_{i+1}-s_i)(t_{j+1}-t_j)$$
$$= \int_c^d \left[\int_a^b f(s_\pi,t_{\pi'})ds\right] dt.$$

■

Proposition 8.36 (Equality of Mixed Partial Derivatives). *Let $Q = (a,b)\times(c,d)$ be an open rectangle in $\mathbb{R}^2$ and $f\in C(Q,X)$. Assume that $\frac{\partial}{\partial t}f(s,t)$, $\frac{\partial}{\partial s}f(s,t)$ and $\frac{\partial}{\partial t}\frac{\partial}{\partial s}f(s,t)$ exists and are continuous for $(s,t)\in Q$, then $\frac{\partial}{\partial s}\frac{\partial}{\partial t}f(s,t)$ exists for $(s,t)\in Q$ and*

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t}f(s,t) = \frac{\partial}{\partial t}\frac{\partial}{\partial s}f(s,t) \text{ for } (s,t)\in Q. \tag{8.25}$$

Proof. Fix $(s_0,t_0)\in Q$. By two applications of Theorem 8.13,

$$f(s,t) = f(s_{t_0},t) + \int_{s_0}^s \frac{\partial}{\partial\sigma}f(\sigma,t)d\sigma$$
$$= f(s_0,t) + \int_{s_0}^s \frac{\partial}{\partial\sigma}f(\sigma,t_0)d\sigma + \int_{s_0}^s d\sigma\int_{t_0}^t d\tau\frac{\partial}{\partial\tau}\frac{\partial}{\partial\sigma}f(\sigma,\tau) \tag{8.26}$$

and then by Fubini's Theorem 8.35 we learn

$$f(s,t) = f(s_0,t) + \int_{s_0}^s \frac{\partial}{\partial\sigma}f(\sigma,t_0)d\sigma + \int_{t_0}^t d\tau\int_{s_0}^s d\sigma\frac{\partial}{\partial\tau}\frac{\partial}{\partial\sigma}f(\sigma,\tau).$$

Differentiating this equation in t and then in s (again using two more applications of Theorem 8.13) shows Eq. (8.25) holds. ■

8.6 Exercises

Throughout these problems, $(X, \|\cdot\|)$ is a Banach space.

Exercise 8.6. Show $f = (f_1, \dots, f_n) \in \bar{\mathcal{S}}([a,b], \mathbb{R}^n)$ iff $f_i \in \bar{\mathcal{S}}([a,b], \mathbb{R})$ for $i = 1, 2, \dots, n$ and

$$\int_a^b f(t)dt = \left(\int_a^b f_1(t)dt, \dots, \int_a^b f_n(t)dt \right).$$

Here $\mathbb{R}^n$ is to be equipped with the usual Euclidean norm. **Hint:** Use Lemma 8.7 to prove the forward implication.

Exercise 8.7. Prove Theorem 8.35 using the following strategy.

1. Use the results from the proof in the text of Theorem 8.35 that

$$s \to \int_c^d f(s,t)dt \text{ and } t \to \int_a^b f(s,t)ds$$

are continuous maps.
2. For the moment take $X = \mathbb{R}$ and prove Eq. (8.24) holds by first proving it holds when $f(s,t) = s^m t^n$ with $m, n \in \mathbb{N}_0$. Then use this result along with Theorem 8.31 to show Eq. (8.24) holds for all $f \in C([a,b] \times [c,d], \mathbb{R})$.
3. For the general case, use the special case proved in item 2. along with Hahn - Banach theorem.

Exercise 8.8. Give another proof of Proposition 8.36 which does not use Fubini's Theorem 8.35 as follows.

1. By a simple translation argument we may assume $(0,0) \in Q$ and we are trying to prove Eq. (8.25) holds at $(s,t) = (0,0)$.
2. Let $h(s,t) := \frac{\partial}{\partial t}\frac{\partial}{\partial s} f(s,t)$ and

$$G(s,t) := \int_0^s d\sigma \int_0^t d\tau h(\sigma, \tau)$$

so that Eq. (8.26) states

$$f(s,t) = f(0,t) + \int_0^s \frac{\partial}{\partial \sigma} f(\sigma, t_0) d\sigma + G(s,t)$$

and differentiating this equation at $t = 0$ shows

$$\frac{\partial}{\partial t} f(s,0) = \frac{\partial}{\partial t} f(0,0) + \frac{\partial}{\partial t} G(s,0). \tag{8.27}$$

Now show using the definition of the derivative that

$$\frac{\partial}{\partial t} G(s, 0) = \int_0^s d\sigma h(\sigma, 0). \tag{8.28}$$

Hint: Consider

$$G(s, t) - t \int_0^s d\sigma h(\sigma, 0) = \int_0^s d\sigma \int_0^t d\tau \left[h(\sigma, \tau) - h(\sigma, 0)\right].$$

3. Now differentiate Eq. (8.27) in s using Theorem 8.13 to finish the proof.

Exercise 8.9. Give another proof of Eq. (8.24) in Theorem 8.35 based on Proposition 8.36. To do this let $t_0 \in (c, d)$ and $s_0 \in (a, b)$ and define

$$G(s, t) := \int_{t_0}^t d\tau \int_{s_0}^s d\sigma f(\sigma, \tau)$$

Show G satisfies the hypothesis of Proposition 8.36 which combined with two applications of the fundamental theorem of calculus implies

$$\frac{\partial}{\partial t} \frac{\partial}{\partial s} G(s, t) = \frac{\partial}{\partial s} \frac{\partial}{\partial t} G(s, t) = f(s, t).$$

Use two more applications of the fundamental theorem of calculus along with the observation that $G = 0$ if $t = t_0$ or $s = s_0$ to conclude

$$G(s, t) = \int_{s_0}^s d\sigma \int_{t_0}^t d\tau \frac{\partial}{\partial \tau} \frac{\partial}{\partial \sigma} G(\sigma, \tau) = \int_{s_0}^s d\sigma \int_{t_0}^t d\tau \frac{\partial}{\partial \tau} f(\sigma, \tau). \tag{8.29}$$

Finally let $s = b$ and $t = d$ in Eq. (8.29) and then let $s_0 \downarrow a$ and $t_0 \downarrow c$ to prove Eq. (8.24).

Exercise 8.10 (Product Rule). Prove items 1. and 2. of Lemma 8.9. This can be modeled on the standard proof for real valued functions.

Exercise 8.11 (Chain Rule). Prove the chain rule in Proposition 8.10. Again this may be modeled on the on the standard proof for real valued functions.

Exercise 8.12. To each $A \in L(X)$, we may define $L_A, R_A : L(X) \to L(X)$ by

$$L_A B = AB \text{ and } R_A B = BA \text{ for all } B \in L(X).$$

Show $L_A, R_A \in L(L(X))$ and that

$$\|L_A\|_{L(L(X))} = \|A\|_{L(X)} = \|R_A\|_{L(L(X))}.$$

Exercise 8.13. Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $U, V : \mathbb{R} \to L(X)$ are the unique solution to the linear differential equations

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I \tag{8.30}$$

and

$$\dot{U}(t) = -U(t)A(t) \text{ with } U(0) = I. \tag{8.31}$$

Prove that $V(t)$ is invertible and that $V^{-1}(t) = U(t)$[4], where by abuse of notation I am writing $V^{-1}(t)$ for $[V(t)]^{-1}$. **Hints:** 1) show $\frac{d}{dt}[U(t)V(t)] = 0$ (which is sufficient if $\dim(X) < \infty$) and 2) show compute $y(t) := V(t)U(t)$ solves a linear differential ordinary differential equation that has $y \equiv Id$ as an obvious solution. (The results of Exercise 8.12 may be useful here.) Then use the uniqueness of solutions to linear ODEs.

Exercise 8.14. Suppose that $(X, \|\cdot\|)$ is a Banach space, $J = (a, b)$ with $-\infty \le a < b \le \infty$ and $f_n : \mathbb{R} \to X$ are continuously differentiable functions such that there exists a summable sequence $\{a_n\}_{n=1}^{\infty}$ satisfying

$$\|f_n(t)\| + \left\|\dot{f}_n(t)\right\| \le a_n \text{ for all } t \in J \text{ and } n \in \mathbb{N}.$$

Show:

1. $\sup\left\{\left\|\frac{f_n(t+h)-f_n(t)}{h}\right\| : (t,h) \in J \times \mathbb{R} \ni t+h \in J \text{ and } h \neq 0\right\} \le a_n.$
2. The function $F : \mathbb{R} \to X$ defined by

$$F(t) := \sum_{n=1}^{\infty} f_n(t) \text{ for all } t \in J$$

   is differentiable and for $t \in J$,

$$\dot{F}(t) = \sum_{n=1}^{\infty} \dot{f}_n(t).$$

Exercise 8.15. Suppose that $A \in L(X)$. Show directly that:

1. e^{tA} define in Eq. (8.19) is convergent in $L(X)$ when equipped with the operator norm.
2. e^{tA} is differentiable in t and that $\frac{d}{dt}e^{tA} = Ae^{tA}$.

Exercise 8.16. Suppose that $A \in L(X)$ and $v \in X$ is an eigenvector of A with eigenvalue λ, i.e. that $Av = \lambda v$. Show $e^{tA}v = e^{t\lambda}v$. Also show that if $X = \mathbb{R}^n$ and A is a diagonalizable $n \times n$ matrix with

$$A = SDS^{-1} \text{ with } D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

then $e^{tA} = Se^{tD}S^{-1}$ where $e^{tD} = \operatorname{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})$. Here $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ denotes the diagonal matrix Λ such that $\Lambda_{ii} = \lambda_i$ for $i = 1, 2, \ldots, n$.

Exercise 8.17. Suppose that $A, B \in L(X)$ and $[A, B] := AB - BA = 0$. Show that $e^{(A+B)} = e^A e^B$.

[4] The fact that $U(t)$ must be defined as in Eq. (8.31) follows from Lemma 8.9.

Exercise 8.18. Suppose $A \in C(\mathbb{R}, L(X))$ satisfies $[A(t), A(s)] = 0$ for all $s, t \in \mathbb{R}$. Show

$$y(t) := e^{\left(\int_0^t A(\tau)d\tau\right)}x$$

is the unique solution to $\dot{y}(t) = A(t)y(t)$ with $y(0) = x$.

Exercise 8.19. Compute e^{tA} when

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and use the result to prove the formula

$$\cos(s+t) = \cos s \cos t - \sin s \sin t.$$

Hint: Sum the series and use $e^{tA}e^{sA} = e^{(t+s)A}$.

Exercise 8.20. Compute e^{tA} when

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

with $a, b, c \in \mathbb{R}$. Use your result to compute $e^{t(\lambda I + A)}$ where $\lambda \in \mathbb{R}$ and I is the 3×3 identity matrix. **Hint:** Sum the series.

Exercise 8.21. Prove Theorem 8.24 using the following outline.

1. Using the right continuity at 0 and the semi-group property for T_t, show there are constants M and C such that $\|T_t\|_{L(X)} \le MC^t$ for all $t > 0$.
2. Show $t \in [0, \infty) \to T_t \in L(X)$ is continuous.
3. For $\varepsilon > 0$, let $S_\varepsilon := \frac{1}{\varepsilon}\int_0^\varepsilon T_\tau d\tau \in L(X)$. Show $S_\varepsilon \to I$ as $\varepsilon \downarrow 0$ and conclude from this that S_ε is invertible when $\varepsilon > 0$ is sufficiently small. For the remainder of the proof fix such a small $\varepsilon > 0$.
4. Show
$$T_t S_\varepsilon = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} T_\tau d\tau$$
and conclude from this that
$$\lim_{t \downarrow 0}\left(\frac{T_t - I}{t}\right) S_\varepsilon = \frac{1}{\varepsilon}\left(T_\varepsilon - Id_X\right).$$
5. Using the fact that S_ε is invertible, conclude $A = \lim_{t\downarrow 0} t^{-1}\left(T_t - I\right)$ exists in $L(X)$ and that
$$A = \frac{1}{\varepsilon}\left(T_\varepsilon - I\right) S_\varepsilon^{-1}.$$
6. Now show, using the semigroup property and step 4., that $\frac{d}{dt}T_t = AT_t$ for all $t > 0$.

7. Using step 5, show $\frac{d}{dt}e^{-tA}T_t = 0$ for all $t > 0$ and therefore $e^{-tA}T_t = e^{-0A}T_0 = I$.

Exercise 8.22 (Duhamel' s Principle I). Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation in Eq. (8.30). Let $x \in X$ and $h \in C(\mathbb{R}, X)$ be given. Show that the unique solution to the differential equation:

$$\dot{y}(t) = A(t)y(t) + h(t) \text{ with } y(0) = x \tag{8.32}$$

is given by

$$y(t) = V(t)x + V(t)\int_0^t V(\tau)^{-1}h(\tau)\, d\tau. \tag{8.33}$$

Hint: compute $\frac{d}{dt}[V^{-1}(t)y(t)]$ (see Exercise 8.13) when y solves Eq. (8.32).

Exercise 8.23 (Duhamel' s Principle II). Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation in Eq. (8.30). Let $W_0 \in L(X)$ and $H \in C(\mathbb{R}, L(X))$ be given. Show that the unique solution to the differential equation:

$$\dot{W}(t) = A(t)W(t) + H(t) \text{ with } W(0) = W_0 \tag{8.34}$$

is given by

$$W(t) = V(t)W_0 + V(t)\int_0^t V(\tau)^{-1}H(\tau)\, d\tau. \tag{8.35}$$

Exercise 8.24. Give another proof Corollary 8.14 based on Remark 8.12. **Hint:** the Hahn Banach theorem implies

$$\|f(b) - f(a)\| = \sup_{\lambda \in X^*,\, \lambda \neq 0} \frac{|\lambda(f(b)) - \lambda(f(a))|}{\|\lambda\|}.$$