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 \((X, \|\|)\) (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 \( \Sigma \) in \( \mathbb{R}^3 \) is a metric space with the distance between two points on \( \Sigma \) 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 \( B_\delta(x) \). We also define the closed ball centered at \( x \in X \) with radius \( \delta > 0 \) as the set \( C_\delta(x) := \{ y \in X : d(x, y) \leq \delta \} \).

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 \( \lim_{n \to \infty} x_n = x \) or \( x_n \to x \) 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 if 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 if \( V^c \) is closed. We will write \( F \subset X \) to indicate the \( F \) is a closed subset of \( X \) and \( V \subset X \) to indicate the \( V \) is an open subset of \( X \). We also let \( \tau_d \) 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 \( V \subset X \) 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 \( F \) be a collection of closed subsets of \( X \), show \( \cap F := \cap_{F \in F} F \) 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, \( \tau_d \), 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

\[
|d_A(x) - d_A(y)| \leq d(x, y) \quad \forall x, y \in X
\]

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 \( F_\varepsilon := \{ x \in X | d_A(x) \geq \varepsilon \} \) is closed in \( X \).

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

\[
d(a, x) \leq d(x, y) + d(y, a).
\]

Take the inf over \( a \) in the above equation shows that

\[
d_A(x) \leq d(x, y) + d_A(y) \quad \forall x, y \in X.
\]

Therefore, \( d_A(x) - d_A(y) \leq d(x, y) \) and by interchanging \( x \) and \( y \) we also have that \( d_A(y) - d_A(x) \leq d(x, y) \) which implies

\[
|d_A(x) - d_A(y)| \leq d(x, y) \quad \text{as } n \to \infty
\]

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

\[
d_A(x) = \lim_{n \to \infty} d_A(x_n) \geq \varepsilon
\]

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

Corollary 6.7. The function \( d \) satisfies,

\[
|d(x, y) - d(x', y')| \leq d(y, y') + d(x, x')
\]

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.)
6.1 Continuity

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

\[ |d(x, y) - d(x', y')| \leq |d(x, y) - d(x, y')| + |d(x, y') - d(x', y')| \leq d(y, y') + d(x, x'). \]

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

Exercise 6.3. Show that \( V \subseteq X \) is open iff for every \( x \in V \) there is a \( \delta > 0 \) such that \( B_x(\delta) \subseteq V \). 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 \( A \subseteq X \) be a closed subset of \( X \) and \( F_\varepsilon := \{ x \in X : d_A(x) \geq \varepsilon \} \subseteq X \) 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 \( F_\varepsilon \subseteq A^c \) for each \( \varepsilon > 0 \) and hence \( \cup_{\varepsilon > 0} F_\varepsilon \subseteq A^c \). Now suppose that \( x \in A^c \subseteq X \). By Exercise 6.3 there exists an \( \varepsilon > 0 \) such that \( B_x(\varepsilon) \subseteq A^c \), i.e. \( d(x, y) \geq \varepsilon \) for all \( y \in A \). Hence \( x \in F_\varepsilon \) and we have shown that \( A^c \subseteq \cup_{\varepsilon > 0} F_\varepsilon \). Finally it is clear that \( F_\varepsilon \subseteq F_{\varepsilon'} \) whenever \( \varepsilon' \leq \varepsilon \).

Definition 6.10. Given a set \( A \) contained a metric space \( X \), let \( \bar{A} \subseteq X \) be the closure of \( A \) defined by

\[ A := \{ x \in X : \exists \{ x_n \}_{n=1}^{\infty} \subseteq A \ni x = \lim_{n \to \infty} x_n \}. \]

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

Exercise 6.4. Given \( A \subseteq X \), show \( \bar{A} \) is a closed set and in fact

\[ \bar{A} = \bigcap \{ F : A \subseteq F \subseteq X \text{ with } F \text{ closed} \}. \]  \hspace{1cm} (6.2)

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

6.1 Continuity

Suppose that \((X, \rho)\) 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

\[ d(f(x), f(x')) < \varepsilon \text{ provided that } \rho(x, x') < \delta. \]  \hspace{1cm} (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 \((X, \rho)\) 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 \( A \subseteq Y \) of \( f(x) \), \( f^{-1}(A) \) is a neighborhood of \( x \in X \).
3. For all sequences \( \{x_n\}_{n=1}^{\infty} \subseteq X \) 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. Suppose \( A \subseteq Y \) is a neighborhood of \( f(x) \), there exists \( \varepsilon > 0 \) such that \( B_{f(x)}(\varepsilon) \subseteq A \) and because \( f \) is continuous there exists a \( \delta > 0 \) such that Eq. (6.3) holds. Therefore

\[ B_x(\delta) \subseteq f^{-1} \left( B_{f(x)}(\varepsilon) \right) \subseteq f^{-1}(A) \]

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

2. \( \implies \) 3. Suppose that \( \{x_n\}_{n=1}^{\infty} \subseteq X \) 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 \( x_n \in B_x(\delta) \subseteq f^{-1} \left( B_{f(x)}(\varepsilon) \right) \) 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 \( \rho(x_n, x) < \frac{\varepsilon}{3} \), 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 \((X, \rho)\) 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. \( f^{-1}(V) \in \tau_d \) 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 \( V \subset X \) 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 \( x \in X \).


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

\[
f(x) = \frac{d_B(x)}{d_A(x) + d_B(x)} \quad \text{for } x \in X
\]

(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 used 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 \( d(x, y) = \|x - y\|_2^2 = \sum_{i=1}^{n} (x_i - y_i)^2 \).
3. \( X = C^p(\mu) \) 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 \((\bar{X}, \bar{d})\) and an isometric map \( i : X \to \bar{X} \) such that \( i(X) \) is dense in \( \bar{X} \), see Definition 6.10.

1. Let \( C \) denote the collection of Cauchy sequences \( \{a_n\}_{n=1}^\infty \subset X \). Given two elements \( a, b \in C \) show

\[
d_C(a, b) := \lim_{n \to \infty} d(a_n, b_n) \quad \text{exists},
\]

\[
d_C(a, b) \geq 0 \quad \text{for all } a, b \in C \quad \text{and} \quad d_C \quad \text{satisfies the triangle inequality,}
\]

\[
d_C(a, c) \leq d_C(a, b) + d_C(b, c) \quad \text{for all } a, b, c \in C.
\]

Thus \((C, d_C)\) would be a metric space if it were true that \( d_C(a, b) = 0 \) iff \( a = b \). This however is false, for example if \( a_n = b_n \) for all \( n \geq 100 \), then \( d_C(a, b) = 0 \) while \( a \neq b \).
2. Define two elements $a, b \in C$ to be equivalent (write $a \sim b$) whenever $d_C(a, b) = 0$. Show “$\sim$” is an equivalence relation on $C$ and that $d_C(a, b') = d_C(a, b)$ if $a \sim a'$ and $b \sim b'$. (Hint: see Corollary 6.7.)

3. Given $a \in C$ let $\bar{a} := \{b \in C : b \sim a\}$ denote the equivalence class containing $a$ and let $X := \{\bar{a} : a \in C\}$ denote the collection of such equivalence classes. Show that $d(\bar{a}, \bar{b}) := d_C(a, b)$ is well defined on $X \times X$ and verify $(X, d)$ 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 $n$.

5. Verify that $i : X \to X$ is an isometric map and that $i(\{x\})$ is dense in $\bar{X}$.

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 $B_x(1) = \{1\}$, $\overline{B_x(1)} = \{1\}$ while $C_x(1) = \bar{X}$. For another counter example, take

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

with the usually Euclidean metric coming from the plane. Then

$$B_{(0,0)}(1) = \{(0, y) \in \mathbb{R}^2 : |y| < 1\},$$

$$\overline{B_{(0,0)}(1)} = \{(0, y) \in \mathbb{R}^2 : |y| \leq 1\},$$

while

$$C_{(0,0)}(1) = \overline{B_{(0,0)}(1)} \cup \{(0, 1)\}.$$

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

$$\overline{B_x(\varepsilon)} = C_x(\varepsilon)$$

and

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

where the boundary operation, $\text{bd}(\cdot)$ 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

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

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 $y \in \bar{B}$. $\blacksquare$

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 $\sigma$. We define $\ell(\sigma) = \infty$ if $\sigma$ is not piecewise smooth.

Then

$$\overline{B_x(\varepsilon)} = C_x(\varepsilon)$$

and

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

where the boundary operation, $\text{bd}(\cdot)$ 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 $B^* \subset C^*$. Suppose that $y \in B^*$ 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 $y \in C^*$. This will be accomplished by showing: if $d(x, y) < \varepsilon + \delta$ then $B_y(\delta) \cap B_x(\varepsilon) = \emptyset$.

If $d(x, y) < \max(\varepsilon, \delta)$ then either $x \in B_y(\delta)$ or $y \in B_x(\varepsilon)$. In either case $B_y(\delta) \cap B_x(\varepsilon) = \emptyset$. 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 $\sigma$ 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 $z = \sigma(t_0) \in B_\eta(\delta)$ then

$$d(x, z) \leq \ell(\sigma|_{[t_0, t_1]}) = \ell(\sigma) - \ell(\sigma|_{[t_0, t_1]}) < \alpha - d(z, y) = \alpha - \delta' < \varepsilon$$

and therefore $z \in B_\delta(\varepsilon) \cap B_\eta(\delta) \neq \emptyset$.

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

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

So if $\sigma$ is within $\varepsilon$ of a length minimizing curve from $x$ to $y$ that $\sigma|_{[u, v]}$ is within $\varepsilon$ 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 $0 \leq u < v \leq 1$, i.e. if $\sigma$ 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

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

$$\geq d(x, \sigma(u)) + d(\sigma|_{[u, v]}),$$

and therefore

$$\ell(\sigma|_{[u, v]}) \leq d(x, y) + \varepsilon - d(x, \sigma(u)) - d(\sigma(v), y)$$

$$\leq d(\sigma(u), \sigma(v)) + \varepsilon.$$ 

### 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

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

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

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

### 6.4 Exercises

**Exercise 6.10.** Show that $(X, d)$ is a complete metric space if 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 $f'' \leq 0$. Show that $d(x, y) = f(d(x, y))$ is a metric on $X$. In particular show that

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

is a metric on $X$. (Hint: use calculus to verify that $f(a+b) \leq f(a) + f(b)$ 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

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

Show:

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

**Exercise 6.13.** Suppose $(X, \rho)$ 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

$$d(f(a), f(b)) < \varepsilon$$

for all $a, b \in A$ 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 $\mathbb{R}$. 
Banach Spaces

Let \((X, \|\cdot\|)\) be a normed vector space and \(d(x, y) := \|x - y\|\) be the associated metric on \(X\). We say \(\{x_n\}_{n=1}^{\infty} \subseteq X\) 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 \(\{x_n\}_{n=1}^{\infty} \subseteq X\) 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

\[
\|x\| - \|y\| \leq \|x - y\| \quad \text{for all } x, y \in X.
\]

It is also easily seen that the vector addition and scalar multiplication are continuous 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 \(BC(X) \subseteq \ell^\infty(X) = B(X)\) is closed subspace of \(\ell^\infty(X)\) and hence is also a Banach space.

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

\[
|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty
\]

which shows that \(\{f_n(x)\}_{n=1}^{\infty} \subseteq \mathbb{F}\) 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

\[
|f(x) - f_m(x)| \leq \lim \inf_{n \to \infty} \|f_n - f_m\|_\infty
\]

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

\[
\|f - f_m\|_\infty \leq \lim \inf_{n \to \infty} \|f_n - f_m\|_\infty \to 0 \quad \text{as } m \to \infty
\]

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

For the second assertion, suppose that \(\{f_n\}_{n=1}^{\infty} \subseteq BC(X) \subseteq \ell^\infty(X)\) 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

\[
|f(x) - f(y)| \leq |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)|
\]

\[
\leq \|f - f_n\|_\infty + |f_n(x) - f_n(y)|.
\]

Thus if \(\epsilon > 0\), we may choose \(n\) large so that \(2\|f - f_n\|_\infty < \epsilon/2\) and then for this \(n\) there exists an open neighborhood \(V_{\epsilon/2}^y\) of \(y\) in \(X\) such that \(|f_n(x) - f_n(y)| < \epsilon/2\) for \(y \in V_{\epsilon/2}^y\). Thus \(|f(x) - f(y)| < \epsilon\) for \(y \in V_{\epsilon/2}^y\), 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 \(a \in (0, 1)\) and \(f : D \to [0, a]\) is continuous function. Let \(A := f^{-1}(0, 1/2]\) and \(B := f^{-1}(1/2, a]\). By Lemma 6.15 there exists a function \(\tilde{g} \in C(A, [0, \alpha/3])\) such that \(\tilde{g} = 0\) on \(A\) and \(\tilde{g} = 1\) on \(B\). Letting \(g := 3 \tilde{g}^\alpha\), we have \(g \in C(X, [0, \alpha/3])\) such that \(g = 0\) on \(A\) and \(g = \alpha/3\) on \(B\).

Further notice that

\[
0 \leq f(x) - g(x) \leq \frac{2}{3} \alpha \quad \text{for all } x \in D.
\]

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 \(f_1 := f - g_1 \mid_D \in C(D, [0, 2/3])\). Apply step 2. with \(\alpha = 2/3\) and \(f = f_1\) to
find \( g_2 \in C(X, [0, \frac{1}{2\pi}]) \) such that \( f_2 := f - (g_1 + g_2) \mid_D \in C(D, [0, (\frac{1}{2})^3]) \).

Continue this way inductively to find \( g_n \in C(X, [0, \frac{1}{2\pi}) \) such that

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

(7.2)

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

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

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

\[\text{Fig. 7.1. Reducing } f \text{ by subtracting off a globally defined function } g_1 \in C \left( X, [0, \frac{1}{4} \right)) \]

\[\text{\textbullet \ Theorem 7.5 (Completeness of } \ell^p(\mu)) \text{. Let } X \text{ be a set and } \mu \colon X \to (0, \infty) \text{ be a given function. Then for any } p \in [1, \infty], \ell^p(\mu), \left\| \cdot \right\|_p \text{ is a Banach space.}
\]

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

\[
\left| f_n(x) - f_m(x) \right| \leq \frac{1}{\mu(x)} \left\| f_n - f_m \right\|_p \to 0 \text{ as } m, n \to \infty
\]

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,

\[\|f_n - f\|_p = \sum_{x \in X} \mu(x) \inf_{m} \|f_n - f_m\|_p \leq \sum_{x \in X} \mu(x) \inf_{m} \|f_n - f_m\|_p
\]

\[= \lim_{m \to \infty} \inf_{x \in X} \|f_n - f_m\|_p \to 0 \text{ as } n \to \infty.
\]

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

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

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

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 \( 1 \leq p < \infty \) 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 \( \ell^p(X, Y), \left\| \cdot \right\|_p \) is a Banach space.

\[\text{7.2 Bounded Linear Operators Basics}
\]

\[\text{Definition 7.7. Let } X \text{ and } Y \text{ be normed spaces and } T : X \to Y \text{ be a linear map. Then } T \text{ is said to be bounded provided there exists } C < \infty \text{ such that}
\]

\[\|T(x)\| \leq C \|x\| \text{ for all } x \in X \text{. We denote the best constant by } \|T\|, \text{ i.e.}
\]

\[\|T\| = \sup_{x \neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{x \neq 0} \left\{ ||T(x)|| : \|x\| = 1 \right\}.
\]

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

\[\text{Proposition 7.8. Suppose that } X \text{ and } Y \text{ are normed spaces and } T : X \to Y \text{ is a linear map. The following are equivalent:
}\]

\[\text{(a) } T \text{ is continuous.}
\]

\[\text{(b) } T \text{ is continuous at } 0.
\]

\[\text{(c) } T \text{ is bounded.}
\]

\[\text{Proof. (a) } \Rightarrow (b) \text{ trivial. (b) } \Rightarrow (c) \text{ If } T \text{ continuous at } 0 \text{ then there exist } \delta > 0 \text{ such that}
\]

\[\|T(x)\| \leq 1 \text{ if } \|x\| \leq \delta \text{. Therefore for any } x \in X, \|T(\delta x/\|x\|)\| \leq 1
\]

\[\text{and}
\]

\[\|T\| = \sup_{x \neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{x \neq 0} \left\{ ||T(x)|| : \|x\| = 1 \right\}.
\]

\[\text{The number } ||T|| \text{ is called the operator norm of } T.
\]
which implies that 
\[ \|T(x)\| \leq \frac{1}{2}\|x\| \text{ and hence } \|T\| \leq \frac{1}{2} < \infty. \]

For the second assertion, we have for \( q \in X \) and \( \varepsilon > 0 \) be given. Then

\[ \|Ty - Tx\| = \|T(y - x)\| \leq \|T\| \|y - x\| < \varepsilon \]

provided \( \|y - x\| < \varepsilon/\|T\| \). 

Exercise 7.1. Assume the norms on \( X \) and \( Y \) are the \( \ell^1 \) 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

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

Exercise 7.2. Suppose that norms on \( X \) and \( Y \) are the \( \ell^\infty \) 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

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

Exercise 7.3. Assume the norms on \( X \) and \( Y \) are the \( \ell^2 \) 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 \( T^*T : \mathbb{R}^n \rightarrow \mathbb{R}^n \), \( T^* : \mathbb{R}^n \rightarrow \mathbb{R}^m \). 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 \( T : X \rightarrow Y \) and \( S : Y \rightarrow Z \) 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:

\[ \|A + B\| = \sup_{x \neq 0} \frac{\|Ax + Bx\|}{\|x\|} \leq \sup_{x \neq 0} \frac{\|Ax\| + \|Bx\|}{\|x\|} \leq \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} + \sup_{x \neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\|. \]

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

\[ \|STx\| \leq \|S\|\|T\| \leq \|S\|\|T\|\|x\|. \]

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

\[ (S, x) \in L(X, Y) \times X \rightarrow Sx \in Y \]

and

\[ (S, T) \in L(X, Y) \times L(Y, Z) \rightarrow ST \in L(X, Z) \]

are continuous relative to the norms

\[ \|(S, x)\|_{L(X, Y) \times X} := \|S\|_{L(X, Y)} + \|x\|_X \text{ and} \]

\[ \|(S, T)\|_{L(X, Y) \times L(Y, Z)} := \|S\|_{L(X, Y)} + \|T\|_{L(Y, Z)} \]

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 \( L(X, Y) \), \( \|\cdot\|_\text{op} \) 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 \( x \in X \),

\[ \|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\| \rightarrow 0 \text{ as } m, n \rightarrow \infty \]

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

\[ \lim_{n \rightarrow \infty} \|T_n x - T x\| = 0. \]

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

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

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

\[ \|Tx - T_n x\| \leq \|T x - T_n x\| + \|T_n x - T_n x\| \leq \|T x - T_n x\| + \|T_n - T_n\| \|x\| \]

and therefore

\[ \|T x - T_n x\| \leq \lim_{m \rightarrow \infty} \inf_{m \in \mathbb{N}} (\|T x - T_m x\| + \|T_m - T_n\| \|x\|) = \|x\| \cdot \lim_{m \rightarrow \infty} \inf_{m \in \mathbb{N}} \|T_m - T_n\|. \]

Hence

\[ \|T - T_n\| \leq \lim_{m \rightarrow \infty} \inf_{m \in \mathbb{N}} \|T_m - T_n\| \rightarrow 0 \text{ as } n \rightarrow \infty. \]

Thus we have shown that \( T_n \rightarrow T \) 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 \(\sum_{n=1}^{\infty} \| x_n \| < \infty\) implies \(\lim_{N \to \infty} \sum_{n=1}^{N} x_n = s\) exists in \(X\) (that is to say every absolutely convergent series is a convergent series in \(X\).) As usual we will denote \(s\) by \(\sum_{n=1}^{\infty} x_n\).

Proof. This is very similar to Exercise 6.10.

\((\Rightarrow)\) If \(X\) is complete and \(\sum_{n=1}^{\infty} \| x_n \| < \infty\) then sequence \(s_N := \sum_{n=1}^{N} x_n\) 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 \(s = \lim_{n \to \infty} x_n := \lim_{n \to \infty} \sum_{n=1}^{N} x_n\) exists in \(X\).

\((\Leftarrow)\) 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 \(\sum_{n=1}^{\infty} \| y_{n+1} - y_n \| < \infty\). By assumption

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

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,

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

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 \(\sum_{n=1}^{\infty} \| T_n \| < \infty\). Then

\[\sum_{n=1}^{\infty} \| T_n x \| \leq \sum_{n=1}^{\infty} \| T_n \| \| x \| < \infty\]

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

\[\| Sx \| = \lim_{N \to \infty} \| S_N x \| \leq \lim_{N \to \infty} \sum_{n=1}^{N} \| T_n x \| \leq \sum_{n=1}^{\infty} \| T_n \| \| x \| ,\]

\(S\) is bounded and

\[\| S \| \leq \sum_{n=1}^{\infty} \| T_n \|.\]

Similarly,

\[\| Sx - S_M x \| = \lim_{N \to \infty} \| S_N x - S_M x \| \leq \lim_{N \to \infty} \sum_{n=1}^{N} \| T_n \| \| x \| = \sum_{n=M+1}^{\infty} \| T_n \| \| x \| \]

and therefore,

\[\| S - S_M \| \leq \sum_{n=M}^{\infty} \| T_n \| \to 0 \text{ as } M \to \infty.\]

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 \(\Gamma_\varepsilon \subseteq A\) such that \(\|s - \sum_{\alpha \in \Gamma_\varepsilon} x_\alpha\| < \varepsilon\) for any \(A \subseteq \mathbb{A}\) such that \(\Gamma_\varepsilon \subseteq A\).

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 \(\{x_\alpha : \alpha \in A\}\) 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

\[Ts = \sum_{\alpha \in A} T x_\alpha = \sum_{\alpha \in A} T x_\alpha.\]
2. If \( s = \sum_{\alpha \in A} x_\alpha \) exists in \( X \) then for every \( \varepsilon > 0 \) there exists \( \Gamma_\varepsilon \subset A \) such that
\[
\left\| \sum_{\alpha \in A} x_\alpha \right\| < \varepsilon \text{ for all } A \subset \subset A \setminus \Gamma_\varepsilon.
\]
3. If \( s = \sum_{\alpha \in A} x_\alpha \) exists in \( X \), the set \( \Gamma := \{ \alpha \in A : x_\alpha \neq 0 \} \) is at most countable. Moreover if \( \Gamma \) is infinite and \( \{ \alpha_n \}_{n=1}^\infty \) is an enumeration of \( \Gamma \), 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 \( \Gamma_\varepsilon \subset A \) such that \( \left\| \sum_{\alpha \in A} x_\alpha \right\| < \varepsilon \) whenever \( A \subset \subset A \setminus \Gamma_\varepsilon \), then \( \sum_{\alpha \in A} x_\alpha \) exists in \( X \).

Proof.
1. Let \( \Gamma_\varepsilon \) be as in Definition 7.15 and \( A \subset \subset A \) such that \( \Gamma_\varepsilon \subset A \). Then
\[
\left\| Ts - \sum_{\alpha \in A} T x_\alpha \right\| \leq \left\| T \right\| \left\| s - \sum_{\alpha \in A} x_\alpha \right\| < \|T\| \varepsilon
\]
which shows that \( \sum_{\alpha \in A} T x_\alpha \) 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 \( A \subset \subset A \setminus \Gamma_\varepsilon \),
\[
\left\| \sum_{\alpha \in A} x_\alpha \right\| = \left\| \sum_{\alpha \notin \Gamma_\varepsilon \cup A} x_\alpha - \sum_{\alpha \in \Gamma_\varepsilon} x_\alpha \right\|
\leq \left\| \sum_{\alpha \notin \Gamma_\varepsilon \cup A} x_\alpha - \sum_{\alpha \in \Gamma_\varepsilon} x_\alpha \right\| + \left\| \sum_{\alpha \in \Gamma_\varepsilon} x_\alpha - s \right\| < 2\varepsilon.
\]
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 \( \left\| \sum_{\alpha \in \Gamma_n} x_\alpha \right\| < \frac{1}{n} \) for all \( A \subset \subset A \setminus \Gamma_n \). Without loss of generality we may assume \( x_\alpha \neq 0 \) for all \( \alpha \in \Gamma_n \). Let \( \Gamma_\infty := \bigcup_{n=1}^\infty \Gamma_n \) be a countable subset of \( A \). Then for any \( \beta \notin \Gamma_\infty \), we have \( \{ \beta \} \cap \Gamma_n = \emptyset \) and therefore
\[
\left\| x_\beta \right\| = \left\| \sum_{\alpha \in \{ \beta \}} x_\alpha \right\| \leq \frac{1}{n} \to 0 \text{ as } n \to \infty.
\]
Let \( \{ \alpha_n \}_{n=1}^\infty \) be an enumeration of \( \Gamma \) and define \( \gamma_N := \{ \alpha_n : 1 \leq n \leq N \} \). Since for any \( M \in \mathbb{N} \), \( \gamma_N \) will eventually contain \( \Gamma_M \) for \( N \) sufficiently large, we have
\[
\lim_{N \to \infty} \sup_{n=1}^N \left\| s - \sum_{n=1}^N x_{\alpha_n} \right\| \leq \frac{1}{M} \to 0 \text{ as } M \to \infty.
\]
Therefore Eq. (7.4) holds.

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 \( GL(X) \) and \( GL(X, X) \) respectively.

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

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

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

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

$$(I - A) S = (I - A) \lim_{N \to \infty} S_N = \lim_{N \to \infty} (I - A) S_N$$

$$= \lim_{N \to \infty} (I - A) \sum_{n=0}^{N} A^n = \lim_{N \to \infty} (I - A^{N+1}) = I$$

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

$$\| (I - A)^{-1} \| = \| S \| \leq \sum_{n=0}^{\infty} \| A^n \|.$$

If we further assume $\| A \| < 1$, then $\| A^n \| \leq \| A \|^n$ and

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

\[ \square \]

**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

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

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

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

$$\| B^{-1} \| \leq \| A^{-1} \| \frac{1}{1 - \| A^{-1} \| \| A - B \|}$$

and

$$\| B^{-1} - A^{-1} \| \leq \| A^{-1} \|^2 \| A - B \| \frac{1}{1 - \| A^{-1} \| \| A - B \|}.$$  \hspace{3mm} (7.8)

In particular the map

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

is continuous.

---

70 7 Banach Spaces

**Proof.** Let $A$ and $B$ be as above, then

$$B = A - (A - B) = A [I_X - A^{-1}(A - B)] = A (I_X - A)$$

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

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

Now

$$\| A \| = \| A^{-1}(A - B) \| \leq \| A^{-1} \| \| A - B \| < \| A^{-1} \| \| A^{-1} \|^{-1} = 1.$$ 

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

$$B^{-1} = (I_X - A)^{-1} A^{-1} = [I_X - A^{-1}(A - B)]^{-1} A^{-1}.$$

Taking norms of the previous equation gives

$$\| B^{-1} \| \leq \| (I_X - A)^{-1} \| \| A^{-1} \| \leq \| A^{-1} \| \frac{1}{1 - \| A \|}$$

$$\leq \| A^{-1} \| \| A - B \|$$

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

$$\| B^{-1} - A^{-1} \| = \| B^{-1} (A - B) A^{-1} \| \leq \| B^{-1} \| \| A^{-1} \| \| A - B \|$$

$$\leq \| A^{-1} \|^2 \| A - B \| \frac{1}{1 - \| A^{-1} \| \| A - B \|}.$$  \hspace{3mm} (7.8)

\[ \square \]

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 $X^*_R$ denote $X$ thought of as a real vector space. If $f \in X^*$ and $u = \text{Re} f \in X^*_R$ then

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

\hspace{3mm} (7.10)

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

$$|f| \leq p \text{ iff } u \leq p.$$
7.5 Hahn Banach Theorem

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

\[ v(ix) = \text{Im} \, f(ix) = \text{Im} (if(x)) = \text{Ref}(x) = u(x). \]

Therefore

\[ f(x) = u(x) + iv(x) = u(x) + iu(-ix) = u(x) - iu(ix). \]

Conversely for \( u \in X_k^* \) let \( f(x) = u(x) - iu(ix) \). Then

\[ f((a + ib)x) = u(ax + ibx) = u(iax - bx) \]
\[ = au(x) + bu(ix) - i(au(ix) - bu(x)) \]

while

\[ (a + ib)f(x) = au(x) + bu(ix) + i(bu(x) - au(ix)). \]

So \( f \) is complex linear.

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

\[ |f(x)| = f(\lambda x) = u(\lambda x) \leq \|u\| |\lambda x| = \|u\| |x|. \]

Since \( x \in X \) is arbitrary, this shows that \( \|f\| \leq \|u\| \) so \( \|f\| = \|u\| \).

For the last assertion, it is clear that \( |f| \leq \|f\| \) implies that \( |u| \leq \|u\| \leq \|f\| \).

Conversely if \( |u| \leq \|f\| \) and \( x \in X \), choose \( \lambda \in S^1 \subseteq \mathbb{C} \) such that \( |f(x)| = \lambda f(x) \).

Then

\[ |f(x)| = \lambda f(x) = f(\lambda x) = u(\lambda x) \leq p(\lambda x) = p(x) \]

holds for all \( x \in X \). \( \blacksquare \)

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

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

Suppose that \( M = \sup_{x \in X \setminus M} |u(x)| \) 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 \( x_0 \),

\[ 0 = \frac{d}{dx} |u(x_0 + itx_0)| = \frac{d}{dx} \left\{ \frac{1}{|1 + it|} \left( u(x_0) + tu(ix_0) \right) \right\} = u(ix_0) \]

since \( \frac{d}{dx} |1 + it| = \frac{1}{\sqrt{1 + t^2}} = 0 \). This explains why \( \|f\| = \|u\| \). \( \blacksquare \)

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

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

Notice that if \( x \geq 0 \), then \( p(x) = x/2 \) and if \( x \leq 0 \) then \( p(x) = -x \), i.e.

\[ p(x) = \begin{cases} x/2 & \text{if } x \geq 0 \\ |x| & \text{if } x \leq 0. \end{cases} \]

From this formula it is clear that \( p(cx) = cp(x) \) for all \( c \geq 0 \) 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

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

which shows that \( p(x + y) \leq \lambda + \mu = p(x) + p(y) \). 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, \( M \subseteq X \) be a subspace \( f : M \to \mathbb{R} \) be a linear functional such that \( f \leq p \) on \( M \).
Then there exists a linear functional \( F : X \to \mathbb{R} \) such that \( F|_M = f \) and \( F \leq p \).

Proof. Step (1) We show for all \( x \in X \setminus M \) there exists and extension \( F \) to
\( M \oplus \mathbb{R} \) 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 \( f(y) + \lambda \alpha = F(y + \lambda x) \leq p(y + \lambda x) \) i.e.
\( \lambda \alpha \leq p(y + \lambda x) - f(y) \). Equivalently put we must find \( \alpha \in \mathbb{R} \) such that
\[ \alpha \leq \frac{p(y + \lambda x) - f(y)}{\lambda} \text{ for all } y \in M \text{ and } \lambda > 0 \]
\[ \alpha \geq \frac{p(z - \mu x) - f(z)}{\mu} \text{ for all } z \in M \text{ and } \mu > 0. \]

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
\[ f(z) - p(z - \mu x) \leq \frac{p(y + \lambda x) - f(y)}{\lambda} \]
or equivalently
But by assumption and the triangle inequality for \( p \),
\[
    f(\lambda z + \mu y) \leq \mu p(y + \lambda x) + \lambda p(z - \mu x) \\
    = p(\mu y + \mu \lambda x) + p(\lambda z - \lambda \mu x),
\]
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 \( D(F) \subseteq X \) and \( F : D(F) \to \mathbb{R} \) is linear. For \( F, G : X \to \mathbb{R} \) we will say \( F \prec G \) if \( D(F) \subseteq D(G) \) and \( F = G|_{D(F)} \), that is \( G \) is an extension of \( F \). Let
\[
    \mathcal{F} = \{ F : X \to \mathbb{R} : f \prec F \text{ and } F \preceq p \text{ on } D(F) \}.
\]
Then \( (\mathcal{F}, \prec) \) is a partially ordered set. If \( \Phi \subseteq \mathcal{F} \) is a chain (i.e. a linearly ordered subset of \( \mathcal{F} \)) then \( \Phi \) has an upper bound \( G \in \mathcal{F} \) defined by \( D(G) = \bigcup D(F) \) and \( G(x) = F(x) \) for \( x \in D(F) \). Then it is easily checked that \( F \in \Phi \)
\[
    G = D(G) \text{ is a linear subspace, } G \in \mathcal{F}, \text{ and } F \prec G \text{ for all } F \in \Phi.
\]
We may now apply Zorn’s Lemma\(^2\) (see Theorem B.7) to conclude there exists a maximal \( F \in \mathcal{F} \). 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, \( M \subseteq X \) 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 \( u = \text{Ref} \) then \( u \leq p \) on \( M \) and hence by Theorem 7.24, there exists \( U \in X'_0 \) 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 \( M \subseteq X \) 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 \( M \oplus \text{span}(\beta) \) 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 \( F : M \oplus \text{span}(\beta) \to \mathbb{R} \) with \( F(x) \leq p(x) \) for \( x \in M \oplus \text{span}(\beta) \). This function \( F \) then extends by continuity to \( X \) and gives the desired extension of \( f \).

**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 \( |\hat{x}(f)| = |f(x)| \leq \|f\| \|x\| \) for all \( f \in X^* \), it follows that \( \|\hat{x}\|_{X^*} \leq \|x\|_{X^*} \). 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 \( \|\hat{x}\|_{X^*} \geq \|x\|_{X^*} \). 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 \( M \subseteq X \) and \( N \subseteq X^* \) let
\[
    M^0 := \{ f \in X^* : f|_M = 0 \} \quad \text{and} \quad N^\perp := \{ x \in X : f(x) = 0 \text{ for all } f \in N \}.
\]

**Lemma 7.30.** Let \( M \subseteq X \) and \( N \subseteq X^* \), then
\[\]
1. \( M^0 \) and \( N^\perp \) are always closed subspace of \( X^* \) and \( X \) respectively.
2. \( M^0 \perp = \overline{M} \).

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

If \( x \in M \), then \( f(x) = 0 \) for all \( f \in M^0 \) so that \( x \in (M^0)^\perp \). Therefore \( \overline{M} \subseteq (M^0)^\perp \). If \( x \notin \overline{M} \), 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 \( x \notin (M^0)^\perp \) and we have shown \( (M^0)^\perp \subseteq \overline{M} \).

**Proposition 7.31.** Suppose \( X \) is a Banach space, then \( X^{**} = (X^*)^\perp \) where
\[
    (X^*)^\perp = \{ \lambda \in X^{**} : \lambda(\hat{x}) = 0 \text{ for all } x \in X \}.
\]
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 \( \psi' := \psi - f_\psi \). 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 \( X^{***} = \hat{X}^* + \hat{X}^0 \). If \( \psi \in \hat{X}^* \cap \hat{X}^0 \), then \( \psi = f \) for some \( f \in X^* \) and \( 0 = \hat{f}(\hat{e}) = \hat{x}(f) = f(x) \) for all \( x \in X \), i.e. \( f = 0 \) so \( \psi = 0 \). Therefore \( X^{***} = \hat{X}^* \oplus \hat{X}^0 \) as claimed.

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

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

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

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

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 \( \infty \),” i.e. for every \( \varepsilon > 0 \) there exists a finite subset \( A_\varepsilon \subset X \) such that \( |f(x)| < \varepsilon \) whenever \( x \notin A_\varepsilon \). 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 \( (\ell^p(\mu), \| \cdot \|_p) \) for \( 1 \leq p < \infty \), 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 \( f \in \ell^p(\mu) \) define \( \phi_f : \ell^q(\mu) \to \mathbb{F} \) by

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

Then

\[
\| \phi_f \|_{\ell^q(\mu)} \leq \| f \|_p \|.g\|_q,
\]

which shows that \( \phi_f \) is well defined. The \( \phi_f : \ell^p(\mu) \to \mathbb{F} \) is linear by the linearity of sums and since

\[
|\phi_f(g)| = \sum_{x \in X} f(x) g(x) \mu(x) \leq \sum_{x \in X} \| f(x) \|_p \| g(x) \|_q \mu(x) \| f \|_p \| g \|_q,
\]

we learn that

\[
\| \phi_f \|_{\ell^q(\mu)} \leq \| f \|_p \|.g\|_q,
\]

Therefore \( \phi_f \in \ell^q(\mu)^* \).

2. The map \( \phi \) in Eq. (7.12) is linear in \( f \) by the linearity properties of infinite sums.

For \( p \in (1, \infty) \), define \( g(x) = \text{sgn}(f(x)) |f(x)|^{q-1} \) where

\[
\text{sgn}(z) := \begin{cases} \frac{1}{|z|} & \text{if } z \neq 0 \\ 0 & \text{if } z = 0. \end{cases}
\]

Then

\[
\| g \|_p^p = \sum_{x \in X} |f(x)|^{(q-1)p} \mu(x) = \sum_{x \in X} |f(x)|^{(q-1)p} \mu(x) = \sum_{x \in X} |f(x)|^q \| \mu \|_q = \| f \|_q^q
\]

and

\[
1. \phi_f(g) \text{ is well defined and } \phi_f \in \ell^q(\mu)^*.
\]

2. The map

\[
f \in \ell^p(\mu) \mapsto \phi_f \in \ell^q(\mu)^*
\]

is an isometric linear map of Banach spaces.

3. If \( p \in [1, \infty) \), then the map in Eq. (7.12) is also surjective and hence \( \ell^q(\mu)^* \) is isometrically isomorphic to \( \ell^p(\mu) \). When \( p = \infty \), the map

\[
f \in \ell^1(\mu) \mapsto \phi_f \in c_0^*
\]

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.
\[ \phi_f (g) = \sum_{x \in X} f (x) \text{sgn}(f(x)) |f(x)|^{q-1} \mu (x) = \sum_{x \in X} |f(x)| |f(x)|^{q-1} \mu (x) \]

\[ = \| f \|_{q}^{q} \| f \|_{q}^{q-1} \| f \|_{q}^{q-1} \mu (x) \]

Hence \[ \| \phi_f \|_{p}^{p} \geq \| f \|_{q} \] which combined with Eq. (7.13) shows \[ \| \phi_f \|_{p} = \| f \|_{q} \] for \( p = \infty \), let \( g (x) = \text{sgn}(f(x)) \), then \( \| g \|_{\infty} = 1 \) and

\[ \| \phi_f (g) \| = \| f \|_{q} \text{sgn}(f(x)) \mu (x) \]

\[ = \| f \|_{q} \| f \|_{q}^{q-1} \mu (x) \]

which shows \( \| \phi_f \|_{q}^{q} \geq \| f \|_{q}^{q} \). Combining this with Eq. (7.13) shows \( \| \phi_f \|_{q} = \| f \|_{q} \).

For \( p = 1 \),

\[ \| \phi_f (g) \| = \| f \|_{1} \| f \|_{1} \]

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

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 \( f \in \ell^p (\mu) \) 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,

\[ |f (x)| = \left| \frac{\lambda (\delta_x)}{\mu (x)} \right| \leq \frac{|| \lambda ||_{\ell^p (\mu)^*} \| \delta_x \|_{\ell^p (\mu)}}{\mu (x)} \]

\[ = \frac{\| \lambda \|_{\ell^p (\mu)^*}}{\mu (x)^{1/p}} \| \mu (x) \|_{1/p} \]

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

\[ \| f \|_{p}^{p} := \sum_{x \in X} |f (x)|^{p} \mu (x) = \sum_{x \in X} f (x) \text{sgn}(f(x)) |f(x)|^{q-1} \mu (x) \]

\[ = \sum_{x \in X} \| f (x) \|_{q}^{q} \| f (x) \|_{q}^{q-1} \mu (x) \]

\[ = \sum_{x \in X} \| f (x) \|_{q}^{q} \| f (x) \|_{q}^{q-1} \mu (x) \]

\[ = \lambda \left( \sum_{x \in X} \| f (x) \|_{q}^{q} \| f (x) \|_{q}^{q-1} \delta_x \right) \]

\[ \leq \| \lambda \|_{\ell^p (\mu)^*} \left( \sum_{x \in X} \| f (x) \|_{q}^{q} \| f (x) \|_{q}^{q-1} \delta_x \right) \]
7.6 Exercises

Exercise 7.7. Let $(X, \| \cdot \|)$ be a normed space over $F$ (R or C). Show the map $(\lambda, x, y) \in F \times X \times X \rightarrow x + \lambda y \in X$ is continuous relative to the norm on $F \times X \times X$ defined by
\[
\|(\lambda, x, y)\|_{F \times X \times X} := |\lambda| + \|x\| + \|y\|.
\]
(See Exercise 10.21 for more on the metric associated to this norm.) Also show that $\| \cdot \| : X \rightarrow [0, \infty)$ is continuous.

Exercise 7.8. Let $X = \mathbb{N}$ and for $p, q \in [1, \infty)$ let $\| \cdot \|_p, \| \cdot \|_q$ 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 $\overline{S}$ 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 $\overline{S}$ 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 $Q : X \times Y \rightarrow Z$ is a bilinear form, i.e. we are assuming $x \in X \rightarrow Q(x, y) \in Z$ is linear for each $y \in Y$ and $y \in Y \rightarrow Q(x, y) \in Z$ 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 $X \times Y$ if there is a constant $M < \infty$ such that
\[
\|Q(x, y)\|_Z \leq M \|x\|_X \cdot \|y\|_Y \text{ for all } (x, y) \in X \times Y.
\]
Then apply this result to prove Lemma 7.11.

Exercise 7.13. Let $d : C(\mathbb{R}) \times C(\mathbb{R}) \rightarrow [0, \infty)$ 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 $n \rightarrow \infty$ if $f_n$ 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 a 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. $A = \left\{f \in X : \sup_{t \in [0, 1]} f(t) \leq 5 \right\}$.
3. $A = \left\{f \in X : \int_0^1 f(t) \, dt = 0 \right\}$.

Exercise 7.16. Suppose $\{x_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in a Banach space $X$. Show the map
\[
\sum_{\alpha \in A} x_\alpha \text{ exists in } X \text{ and} \quad \|\sum_{\alpha \in A} x_\alpha\| \leq \sum_{\alpha \in A} \|x_\alpha\|
\]
if $\sum_{\alpha \in A} \|x_\alpha\| < \infty$. That is to say “absolute convergence” implies convergence in a Banach space.

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

7.6.1 Hahn — Banach Theorem Problems


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 $|\ell(x_\ell)| \geq \frac{1}{2} \|\ell\|$. 

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 $\varepsilon$ 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 $J_1 \supset J_2 \supset J_3 \supset \ldots$ 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 $t_{n_k} = t_{n_k} \in J_k$ for all $k$. The sequence $\{n_k\}_{k=1}^\infty$ 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 $t_n$ are convergent. But this implies $\lim_{n \to \infty} u_n = u$ 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 exist an $\varepsilon > 0$ and sequences $\{u_n\}_{n=1}^\infty$ and $\{v_n\}_{n=1}^\infty$ in $K$ such that

$$\|f(u_n) - f(v_n)\| \geq \varepsilon$$

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

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

which is a contradiction.

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

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

(8.1)

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

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

(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 $\pi$ 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 $\tilde{\mathcal{S}}$ 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 $S \subset Z$ is a dense linear subspace of $Z$. If $T : S \to X$ is a bounded linear transformation (i.e., there exists $C < \infty$ such that $\|Tz\| \leq C \|z\|$ for all $z \in S$), then $T$ has a unique extension to an element $\tilde{T} \in L(S, X)$ and this extension still satisfies

$$\|Tz\| \leq C \|z\| \quad \text{for all } z \in S.$$  

Exercise 8.2. Prove Theorem 8.4.

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

$$\|\tilde{I}(f)\| \leq (b - a) \|f\|_{\infty} \quad \text{for all } f \in S. \quad (8.3)$$

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

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

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 $\pi$ and $c_i^\pi$ 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.](image)

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

$$\|I(f)\| \leq \sum_{i=0}^{n-1} (t_{i+1} - t_i) \|x_i\| \leq \sum_{i=0}^{n-1} (t_{i+1} - t_i) \|f\|_{\infty} \leq (b - a) \|f\|_{\infty}. \quad (8.5)$$
6. If \( f, g \in \mathcal{S}(\mathbb{R}) \) and \( f \leq g \), then
\[
\int_a^b f(t) \, dt \leq \int_a^b g(t) \, dt.
\]

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
\[
L := \lim_{h \to 0} \frac{f(t + h) - f(t)}{h} \quad \text{exists in } X.
\]
The limit \( L \), if it exists, will be denoted by \( f'(t) \) 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 \( f \in C((a, b) \to X) \).

As for the case of real valued functions, the derivative operator \( \frac{df}{dt} \) is easily seen to be linear. The next two results have proofs 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. \( \frac{d}{dt}vert_{t_0} [U(t) x(t)] \in X \) exists and
\[
\frac{d}{dt}vert_{t_0} [U(t) x(t)] = \left[ U(t_0) x(t_0) + U(t_0) x'(t_0) \right].
\]

2. \( \frac{d}{dt}vert_{t_0} [U(t) V(t)] \in L(X) \) exists and
\[
\frac{d}{dt}vert_{t_0} [U(t) V(t)] = \left[ U(t_0) V(t_0) + U(t_0) V'(t_0) \right].
\]

3. If \( U(t_0) \) is invertible, then \( t \to U(t)^{-1} \) is differentiable at \( t = t_0 \) and
\[
\frac{d}{dt}vert_{t_0} U(t)^{-1} = -U(t_0)^{-1} U'(t_0) U(t_0)^{-1}.
\]

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
\[
\frac{d}{dt}vert_{t_0} x(T(t)) = x'(T(t_0)) T'(t_0).
\]

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 \( f(t) \) 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
\[
\| f(t) - f(\tau) \| = \| f(t) - f(\tau) - f'(\tau)(t - \tau) \| \leq \varepsilon |t - \tau| \quad \text{if } |t - \tau| < \delta_\tau.
\]
Let
\[
A = \{ t \in [a, b]: \| f(t) - f(\alpha) \| \leq \varepsilon (t - \alpha) \}
\]
and \( t_0 \) be the least upper bound for \( A \). We will now use a standard argument which is referred to as \textit{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),
\[
\|f(t) - f(\alpha)\| \leq \|f(t) - f(t_0)\| + \|f(t_0) - f(\alpha)\| \\
\leq \varepsilon(t_0 - \alpha) + \varepsilon(t - t_0) = \varepsilon(t - \alpha)
\]
for \( 0 \leq t - t_0 < \delta_0 \), 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. ■

\textbf{Remark 8.12.} The usual real variable proof of Proposition 8.11 makes use of 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.

\textbf{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. \( F \) exists and is continuous for \( t \in (a, b) \)) and \( F \) extends to a continuous function on \( [a, b] \) which is still denoted by \( F \). Then
\[
\int_a^b F'(t)\,dt = F(b) - F(a).
\]

\textbf{Proof.} Let \( h > 0 \) be a small number and consider
\[
\left\| \int_a^{t+h} f(\tau)\,d\tau - \int_a^t f(\tau)\,d\tau - f(t)h \right\| \\
\leq \int_t^{t+h} \|f(\tau) - f(t)\|\,d\tau \leq h\varepsilon(h),
\]
where \( \varepsilon(h) := \max_{t \in [t, t+h]} \|f(\tau) - f(t)\| \). Combining this with a similar computation when \( h < 0 \) shows, for all \( h \in \mathbb{R} \) sufficiently small, that
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 $U \subseteq \mathbb{R}^d$, let $C_c(U)$ denote the collection of real valued continuous functions $f$ on $U$ such that

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

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

$$|Tf(x) - Tf(z)| \leq \int_0^1 |K(x,y) - K(z,y)||f(y)|dy \leq \|f\|_{\infty} \max_y |K(x,y) - K(z,y)|$$

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,

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

where

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

(8.12)

This shows $\|T\| \leq A < \infty$ 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

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

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

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

Therefore,

$$\|T\| \geq \lim_{\varepsilon \to 0} \frac{1}{\|f\|_{\infty}} \int_0^1 \frac{|K(x_0,y)|^2}{\varepsilon + |K(x_0,y)|^2}dy$$

$$= \lim_{\varepsilon \to 0} \int_0^1 \frac{|K(x_0,y)|^2}{\varepsilon + |K(x_0,y)|^2}dy = A$$

since

$$0 \leq |K(x_0,y)| - \frac{|K(x_0,y)|^2}{\varepsilon + |K(x_0,y)|^2}$$

$$= \frac{|K(x_0,y)|}{\varepsilon + |K(x_0,y)|^2} \left[ \sqrt{\varepsilon + |K(x_0,y)|^2} - |K(x_0,y)| \right]$$

$$\leq \frac{|K(x_0,y)|}{\varepsilon + |K(x_0,y)|^2} \left[ \sqrt{\varepsilon + |K(x_0,y)|^2} - |K(x_0,y)| \right]$$

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 $\int_I^1 f(x)dx = 1$. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $\alpha K(x_0, y_0) = M$, then

$$\int_0^1 \frac{|K(x_0,y)|^2}{\varepsilon + |K(x_0,y)|^2}dy$$

$$\int_0^1 \frac{|K(x_0,y)|^2}{\varepsilon + |K(x_0,y)|^2}dy$$

and

$$\int_0^1 \frac{|K(x_0,y)|^2}{\varepsilon + |K(x_0,y)|^2}dy$$
\[ |(T\alpha f)(x_0)| = \left| \int_0^1 K(x_0, y)\alpha f(y) dy \right| = \left| \int_1^M (M - \varepsilon) f(y) dy \right| \]

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,
\[ y'(t) = A(t)y(t) + h(t) \quad \text{where} \quad y(0) = x \in X, \quad (8.13) \]

where \( y \) is an unknown function in \( C^1(J \to X) \). This equation can be written in its integral form (as the reader should verify) namely we are looking for \( y \in C(J, X) \) such that
\[ y(t) = x + \int_0^t h(\tau) \, d\tau + \int_0^t A(\tau)y(\tau) \, d\tau. \quad (8.14) \]

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

**Remark 8.19.** Suppose that \( \psi \in C(\mathbb{R}, \mathbb{R}) \), then
\[ (-1)^{n-1} \int_{\Delta_n(t)} \psi(\tau_1) \cdots \psi(\tau_n) \, d\tau = \frac{1}{n!} \left( \int_0^t \psi(\tau) \, d\tau \right)^n, \quad (8.15) \]

**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
\[ (-1)^{1-1} \int_{\Delta_1(t)} \psi(\tau_1) \, d\tau_1 = \int_0^t \psi(\tau) \, d\tau = \Psi(t). \]

Now assume the truth of Eq. (8.15) for \( n - 1 \) for some \( n \geq 2 \), then
\[ (-1)^{n-1} \int_{\Delta_n(t)} \psi(\tau_1) \cdots \psi(\tau_n) \, d\tau = \int_0^t \psi(\tau_1) \cdots \psi(\tau_n) \, d\tau. \]

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
\[ \int_{\Delta_n(t)} \psi(\tau_1) \cdots \psi(\tau_n) \, d\tau = \frac{1}{n!} \left( \int_{\Delta_1(t)} \psi(\tau) \, d\tau \right)^n \]

and another way to understand this equality is to view \( \int_{\Delta_n(t)} \psi(\tau_1) \cdots \psi(\tau_n) \, d\tau \) 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 \( \{1, 2, \ldots, n\} \) we have
\[ [0, t]^n = \bigcup_{\sigma \in S_n} \{ (\tau_1, \ldots, \tau_n) \in \mathbb{R}^n : 0 \leq \tau_{\sigma 1} \leq \cdots \leq \tau_{\sigma n} \leq t \} \]

with the union being “essentially” disjoint. Therefore, making a change of variables and using the fact that \( \psi(\tau_1) \cdots \psi(\tau_n) \) is invariant under permutations, we find
\[
\left( \int_0^t \psi(\tau) d\tau \right)^n = \int_0^t \psi(\tau_1) \ldots \psi(\tau_n) d\tau \\
= \sum_{\sigma \in \mathcal{S}_n} \int_{\{(\tau_1, \ldots, \tau_n) \in \mathbb{R}^n : 0 \leq \tau_1 \leq \ldots \leq \tau_n \leq t\}} \psi(\tau_1) \ldots \psi(\tau_n) d\tau \\
= \sum_{\sigma \in \mathcal{S}_n} \int_{\{(s_1, \ldots, s_n) \in \mathbb{R}^n : 0 \leq s_1 \leq \ldots \leq s_n \leq t\}} \psi(s_{\sigma^{-1}1}) \ldots \psi(s_{\sigma^{-1}n}) ds \\
= \sum_{\sigma \in \mathcal{S}_n} \int_{\{(s_1, \ldots, s_n) \in \mathbb{R}^n : 0 \leq s_1 \leq \ldots \leq s_n \leq t\}} \psi(s_1) \ldots \psi(s_n) ds \\
= n! \int_{\Delta_n(t)} \psi(\tau_1) \ldots \psi(\tau_n) d\tau.
\]

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

\[
y(t) = \phi(t) + \int_0^t A(\tau)y(\tau) d\tau \quad (8.16)
\]

has a unique solution given by

\[
y(t) = \phi(t) + \sum_{n=1}^{\infty} (-1)^{n-1} \int_0^t A(\tau_n) \ldots A(\tau_1) \phi(\tau_1) d\tau \quad (8.17)
\]

and this solution satisfies the bound

\[
\|y\|_\infty \leq \|\phi\|_\infty e^{f_J} \|A(\tau)\| d\tau.
\]

**Proof.** Define \( A : BC(J, X) \to BC(J, X) \) by

\[
(Ay)(t) = \int_0^t A(\tau)y(\tau) d\tau.
\]

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

An induction argument shows

\[
(A^n \phi)(t) = \int_0^t d\tau_n A(\tau_n)(A^{n-1} \phi)(\tau_n) \\
= \int_0^t \int_0^{\tau_n} d\tau_{n-1} A(\tau_n)A(\tau_{n-1})(A^{n-2} \phi)(\tau_{n-1}) \\
\vdots \\
= \int_0^t \int_0^{\tau_n} \ldots \int_0^{\tau_2} d\tau_1 A(\tau_n) \ldots A(\tau_1) \phi(\tau_1) \\
= (-1)^{n-1} \int_{\Delta_n(t)} A(\tau_n) \ldots A(\tau_1) \phi(\tau_1) d\tau.
\]

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

\[
\|A^n \phi\| \leq \|\phi\|_{\infty} \int_{\Delta_n(t)} \|A(\tau_n)\| \ldots \|A(\tau_1)\| d\tau \\
\leq \|\phi\|_{\infty} \cdot \frac{1}{n!} \left( \int_{\Delta_n(t)} \|A(\tau)\| d\tau \right)^n \\
\leq \|\phi\|_{\infty} \cdot \frac{1}{n!} \left( \int_{J} \|A(\tau)\| d\tau \right)^n.
\]

Therefore,

\[
\|A^n\|_{op} \leq \frac{1}{n!} \left( \int_{J} \|A(\tau)\| d\tau \right)^n \quad (8.18)
\]

and

\[
\sum_{n=0}^{\infty} \|A^n\|_{op} \leq e^{f_J} \|A(\tau)\| |d\tau| < \infty
\]

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

\[
\|(I - A)^{-1}\|_{op} \leq e^{f_J} \|A(\tau)\| |d\tau|.
\]

It is now only a matter of working through the notation to see that these assertions prove the theorem. \( \blacksquare \)

**Corollary 8.22.** Suppose \( h \in C(J \to X) \) and \( x \in X \), then there exists 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 \( J_0 \) 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 \( J_1 \) 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 \( t \in J \). \( \blacksquare \)

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

\[
y(t) = Ay(t) \quad \text{with} \quad y(0) = x
\]

\(^1\) We do this so that \( \phi_{J_0} \) will be bounded.
is given by \( y(t) = e^{tA}x \) where

\[
e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!}A^n.
\] (8.19)

Moreover,

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

\[
\frac{d}{dt}y(t) = \frac{d}{dt}e^{tA}e^{sA}x = Ae^{sA}x = Ae^{t+s}A \text{ with } y(0) = e^{sA}x.
\]

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.) \( T_0 = I \) is a semi-group with \( T_tT_s = T_{t+s} \) for all \( s, t \geq 0 \).
2. (Norm Continuity) \( t \rightarrow T_t \) is continuous at \( 0 \), i.e. \( \|T_t - I\|_{L(X)} \rightarrow 0 \) 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 \( f : X \rightarrow Y \) be a function from a topological space \( (X, \tau_X) \) to a vector space \( Y \). Then we define the support of \( f \) by

\[
supp(f) := \{x \in X : f(x) \neq 0\}.
\]

a closed subset of \( X \).

**Example 8.26.** For example if \( f : \mathbb{R} \rightarrow \mathbb{R} \) 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 \( \text{supp}(f) = [0, 4\pi] \).

**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 \( \delta \) – 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

\[
\int_{\mathbb{R}} q_n(x) \, dx = 1 \text{ and } \lim_{n \to \infty} \int_{|x| \geq \varepsilon} q_n(x)dx = 0 \text{ for all } \varepsilon > 0.
\] (8.21)

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

\[
q_n \ast f(x,w) := \int_{\mathbb{R}} q_n(y)f(x-y,w)dy
\]

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),

\[
\left| q_n \ast f(x,w) - f(x,w) \right| = \left| \int_{\mathbb{R}} q_n(y)(f(x-y,w) - f(x,w))dy \right|
\]

\[
\leq \int_{\mathbb{R}} q_n(y)\left| f(x-y,w) - f(x,w) \right|dy.
\]

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

\[
\left| q_n \ast f(x,w) - f(x,w) \right| \leq \int_{|y| \leq \varepsilon} q_n(y)\left| f(x-y,w) - f(x,w) \right|dy
\]

\[
+ \int_{|y| > \varepsilon} q_n(y)\left| f(x-y,w) - f(x,w) \right|dy
\]

\[
\leq \sup_{|y| \leq \varepsilon} \left| f(x+z,w) - f(x,w) \right| + 2M \int_{|y| > \varepsilon} q_n(y)dy.
\]

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

\[
\left| q_n \ast f(x,w) - f(x,w) \right| \leq \sup_{|z| \leq \varepsilon, (x,w) \in K \times W} \left| f(x+z,w) - f(x,w) \right| + 2M \int_{|y| > \varepsilon} q_n(y)dy.
\]
and hence by Eq. (8.22),
\[
\lim_{n \to \infty} \sup_{(x,w) \in K \times W} |q_n \ast f(x,w) - f(x,w)| \\
\leq \sup_{|z| \leq \varepsilon, (x,w) \in K \times W} |f(x+z, w) - f(x,w)|.
\]
This finishes the proof since the right member of this equation tends to 0 as \( \varepsilon \to 0 \) by uniform continuity of \( f \) on compact subsets of \( \mathbb{R} \times W \).

Let \( q_n : \mathbb{R} \to [0, \infty) \) be defined by
\[
q_n(x) := \frac{1}{c_n}(1 - x^2)^n 1_{|x| \leq 1} \quad \text{where } c_n := \int_{-1}^1 (1 - x^2)^n dx. \tag{8.23}
\]
Figure 8.2 displays the key features of the functions \( q_n \).

![Graph of q1, q90, and q100](image)

**Fig. 8.2.** A plot of \( q_1, q_{90}, \text{and } q_{100}. \) 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 \( \delta \) – 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

\[
\int_{|x| \geq \varepsilon} q_n(x) dx = \frac{2}{\varepsilon} \int_0^1 (1 - x^2)^n dx \\
= \frac{2}{\varepsilon} \int_0^1 (1 - x^2)^{n-1} dx + 2 \int_0^1 (1 - x^2)^n dx \\
\leq \frac{1}{\varepsilon} \frac{1}{n!} (1 - \varepsilon^2)^{n+1} \\
\leq \frac{1}{\varepsilon} \frac{1}{n!} \left( 1 - (1 - \varepsilon^2)^{n+1} \right) \\
\to 0 \quad \text{as } n \to \infty,
\]
the proof is complete.

**Notation 8.30** Let \( Z_n := \mathbb{N} \cup \{0\} \) and for \( x \in \mathbb{R}^d \) and \( \alpha \in Z_n^d \) 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 \( \alpha \) such that \( |\alpha| = N \), then we define \( \text{deg}(p) := N \) 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 \( f \in C(K,\mathbb{C}) \). 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
\[
K_0 := \lambda K - b := \{\lambda x - b \mid x \in K\} \subset B_d
\]
where \( B_d := (0,1)^d \). The function \( F(y) := f(\lambda^{-1}(y+b)) \) for \( y \in K_0 \) is in \( C(K_0, \mathbb{C}) \) and if \( p_n(y) \) are polynomials on \( \mathbb{R}^d \) such that \( p_n \to F \) 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
\[
g(x) = \begin{cases} 
  f(x) & \text{if } x \in K \\
  0 & \text{if } x \in B_d^c 
\end{cases}
\]
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 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 $f \in C(\mathbb{R}^d, \mathbb{C})$, $K = B_d = [0,1]^d$, and $f \equiv 0$ on $B_d$.

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

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

$$= \frac{1}{\epsilon_n} \int_{[0,1]} f(y,w) \left[ (1 - (x-y)^2)^n \right]_{|x-y| \leq 1} dy$$

$$= \frac{1}{\epsilon_n} \int_{[0,1]} f(y,w)(1 - (x-y)^2)^n dy = \sum_{k=0}^{2n} A^n_k(w) x^k$$

where

$$A^n_k(w) = \int_{[0,1]} f(y,w) \rho^n_k(y) dy$$

and $\rho_k$ is a polynomial function in $y$ for each $k$. Then $A^n_k(w) = 0$ if $w \notin (0,1)^{d-1}$ and using the uniform continuity of $f$ on $[0,1]^d$, one easily shows $A^n_k \in C(\mathbb{R}^{d-1}, \mathbb{C})$. 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 $A^n_k$ 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 $a^n_k$ on $\mathbb{R}^{d-1}$ such that

$$\sup_{w \in [0,1]^{d-1}} |A^n_k(w) - a^n_k(w)| \leq \frac{1}{2(n+1)}$$

Then

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

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

$$|f(x,w) - p_n(x,w)| \leq |f(x,w) - f_n(x,w)| + |f_n(x,w) - p_n(x,w)|$$

$$\leq \frac{\epsilon_n}{2(n+1)} + \sum_{k=0}^{2n} |A^n_k(w) - a^n_k(w)| x^k$$

$$\leq \frac{\epsilon_n}{n} + \frac{1}{n}$$

where

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

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

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

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 $C(S^1)^3$. This example generalizes in an obvious way to $K = (S^1)^d \subset \mathbb{C}^d$.

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

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

Show $f \equiv 0$.

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

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

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

$$p(x) = \sum_{k=-n}^{n} p_k e^{ikx} \text{ with } p_k \in \mathbb{C} \text{ for all } k.$$
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 $t \to \int_d^c f(s, t)dt$ are continuous and

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

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

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

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

be partitions of $[a, b]$ and $[c, d]$ respectively. For $s \in [a, b]$ let $s_i = s$ if $s \in (s_i, s_{i+1}]$ and $i \geq 1$ and $s_n = s = a$ if $s \in [s_0, s_1]$. Define $t_i$, for $t \in [c, d]$ analogously. Then

$$\int_a^b \left[ \int_c^d f(s, t)dt \right] ds = \int_a^b \left[ \int_c^d f(s, t_i)dt \right] ds + \int_a^b \varepsilon_{\pi'}(s)ds$$

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

where

$$\varepsilon_{\pi'}(s) = \int_a^b f(s, t)dt - \int_a^b f(s, t_{i'})dt$$

and

$$\delta_{\pi, \pi'} = \int_a^b \int_c^d \{f(s, t_{i'}) - f(s, t_{i'})\} dt ds.$$ (8.25)

The uniform continuity of $f$ and the estimates

$$\sup_{s \in [a, b]} \|f(s, t)\| \leq \sup_{s \in [a, b]} \int_c^d \|f(s, t) - f(s, t_{i'})\| dt$$

$$\leq (d - c) \sup_{s \in [a, b]} \|f(s, t) - f(s, t_{i'})\| : (s, t) \in Q$$

and

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

allow us to conclude that

$$\int_a^b \left[ \int_c^d f(s, t)dt \right] ds \to 0 \text{ as } |\pi| + |\pi'| \to 0.$$ (8.26)

This completes the proof since

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

$$= \int_c^d \left[ \int_a^b f(s, t)dt \right] ds.$$ (8.27)

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

$$\frac{\partial}{\partial s} \frac{\partial}{\partial t} f(s, t) = \frac{\partial}{\partial s} \frac{\partial}{\partial t} f(s, t) \text{ for } (s, t) \in Q.$$ (8.28)

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

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

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

and then by Fubini’s Theorem 8.35 we learn

$$f(s, t) = f(s_0, t) + \int_{s_0}^s \frac{\partial}{\partial \sigma} f(\sigma, t_0)d\sigma + \int_{t_0}^t dt \int_{s_0}^s d\sigma \frac{\partial}{\partial \sigma} f(\sigma, \tau).$$ (8.30)

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 \( f = (f_1, \ldots, f_n) \in \mathcal{S}([a, b], \mathbb{R}^n) \) iff \( f_i \in \mathcal{S}([a, b], \mathbb{R}) \) for \( i = 1, 2, \ldots, n \) and

\[
\int_a^b f(t) dt = \left( \int_a^b f_1(t) dt, \ldots, \int_a^b f_n(t) dt \right).
\]

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 \( h(s, t) := \frac{\partial}{\partial s} f(s, t) \) and

\[
G(s, t) := \int_0^s ds \int_0^t d\tau h(\sigma, \tau)
\]

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

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

**Hint:** Consider

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

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

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

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

\[
\frac{\partial}{\partial t} \frac{\partial}{\partial s} G(s, t) = \frac{\partial}{\partial s} \frac{\partial}{\partial \tau} G(s, t) = f(s, t).
\]

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

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

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

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

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

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

**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

\[
\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I \tag{8.30}
\]
8.6 Exercises

and

\[ \dot{U}(t) = -U(t)A(t) \] with \( U(0) = I \). \hspace{1cm} (8.31)

Prove that \( V(t) \) is invertible and that \( V^{-1}(t) = U(t)^t \), where by abuse of notation I am writing \( V^{-1}(t) \) for \( [V(t)]^{-1} \). \textbf{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 \( y \equiv I \)d 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 \((X, \|\|)\) is a Banach space, \( J = (a, b) \) with \(-\infty \leq a < b \leq \infty \) and \( f_n : \mathbb{R} \rightarrow X \) are continuously differentiable functions such that there exists a summable sequence \( \{a_n\}^\infty_{n=1} \) satisfying

\[ \|f_n(t)\| + \|f_n'(t)\| \leq a_n \] for all \( t \in J \) and \( n \in \mathbb{N} \).

Show:

1. \( \sup \{ \|f_n(t+h) - f_n(t)\| : (t, h) \in J \times \mathbb{R}, t \in J \) and \( h \neq 0 \} \leq a_n \).

2. The function \( F : \mathbb{R} \rightarrow X \) defined by

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

is differentiable and for \( t \in J \),

\[ \dot{F}(t) = \sum_{n=1}^{\infty} f_n'(t) \].

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

1. \( e^{tA} \) as defined 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 \( \lambda \), i.e. that \( Av = \lambda v \). Show \( e^{tA}v = e^{\lambda t}v \). Also show that if \( X = \mathbb{R}^n \) and \( A \) is a diagonalizable \( n \times n \) matrix with

\[ A = SDS^{-1} \]

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

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

\(^4\) The fact that \( U(t) \) must be defined as in Eq. (8.31) follows from Lemma 8.9.
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:

\[
\dot{y}(t) = A(t)y(t) + h(t) \quad \text{with} \quad y(0) = x
\]

is given by

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

**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:

\[
\dot{W}(t) = A(t)W(t) + H(t) \quad \text{with} \quad W(0) = W_0
\]

is given by

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

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

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