Functional Analysis

Foundations and Advanced Topics

Answer Book

Martín Argerami

Martín Argerami

Functional Analysis

Foundations and Advanced Topics
Answer Book

First Edition
Version A.1.07– February 1, 2026

Martín Argerami
Department of Mathematics and Statistics
University of Regina
Regina, SK, Canada

ISBN 978-1-0698019-1-3 (eBook)

Mathematics Subject Classification: 46-01, 28-01, 47-01, 46A03, 46A22, 46B03, 46C05, 46E30, 46L05, 46L06, 46L07, 46L10

First Edition. Version A.1.07– February 1, 2026
This digital version may be updated periodically; the most recent version is available at https://book.argerami.ca/

To Romina, who has been giving me happiness for decades. And to Jose, Euge, Vicky, Tere, Tom, Santi, Lucas, and Emilia, who have also been doing a pretty good job at it.

Preface

This is the companion book to *Functional Analysis*; it consists of my answers to all exercises. There are compelling reasons both to publish an answer book, and to not publish an answer book. A strong reason from the "not publish" camp is that the only way to really learn mathematics is by trying hard on your own, and getting stuck often. The same way one needs lots of hours and repetition to excel at sports, arts, or other human activities. There is also an elation that needs to be experienced, when one sees the light after being trying and trying on a problem for hours; or, sometimes, after apparently fruitless hours on a problem, the solution will come while on the shower, or on a walk, or another activity very far from mathematics. All those efforts train our minds, and prepare us better to appreciate a certain trick that makes things work, and the emotions involved will make it easier to remember the idea or at least part of it. The "for publish" reasons are varied. There is a risk that the student will give up early on a problem due to the availability of a full answer to the problem. This is unavoidable these days since for common problems it is simple to find a solution online (a number of them will likely be mine, if the problem is related to the topics in this book). This means that these days the student has a stronger responsibility, compared to days past, to be a shepherd of their own mathematical path. The temptation to quickly go read an answer should be fought if progress is to be made. We have all experienced reading someone else's ideas and saying "I could have done that!" but if our knowledge is put to the test, we might not be able to recreate the idea we have just read.

The headlines of the few sections from the book without exercises have been also included here so that the chapter/section numbering stays coherent with the book. The equations in the answers have a different numbering scheme than that in the book, so that both are recognizable and coherent. Namely, equations in the book are of the form (12.3) (meaning equation 3 in chapter 12), while equations in the answers are numbered in the form (AB.2.3) (meaning the third equation in the answers to chapter 2). While care has been put in checking the answers for correctness and typos, most certainly some mistakes are still there. This, or any other feedback, is very welcome! I can be reached at my email address below.

Martín Argerami
Regina, SK, Canada
argerami@uregina.ca
November 2025

Contents

CHAPTER

1

Prerequisites

1.1. Set Theory

(1.1.1) Let $\{A_j\}_{j\in J}$ be a collection of subsets of a set A. Show that
$$\Big[\bigcup_{j\in J} A_j\Big]^c = \bigcap_{j\in J} A_j^c, \qquad \Big[\bigcap_{j\in J} A_j\Big]^c = \bigcup_{j\in J} A_j^c$$

Answer. If $a \notin \bigcup_{j\in J} A_j$, then $a \notin A_j$ for all j; this means that $a \in \bigcap_{j\in J} A_j^c$. Conversely, if $a \in \bigcap_{j\in J} A_j^c$, then $a \notin A_j$ for all j, so $a \notin \bigcup_{j\in J} A_j$.

The second equality is obtained from the first one by taking complements, since $(B^c)^c = B$ for any set B.

(1.1.2) For sets A, B, C, show that
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

and

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Answer. If $a \in A$ and $a \in B \cap C$, we have that $a \in A \cap B$ and $a \in A \cap C$, so $a \in (A \cup B) \cap (A \cup C)$; and the converse also holds: if $a \in A \cap B$ and $a \in A \cap C$, we have that either $a \in A$, or otherwise $a \in B$ and $a \in C$, so $a \in A \cup (B \cap C)$.

The second equality can be proven in a similar manner, or we can use Exercise 1.1.1 to get

$$\begin{aligned} A \cap (B \cup C) &= \left[A^c \cup (B^c \cap C^c)\right]^c = \left[(A^c \cup B^c) \cap (A^c \cup C^c)\right]^c \\ &= (A^c \cup B^c)^c \cup (A^c \cup C^c)^c = (A \cap B) \cup (A \cap C). \end{aligned}$$

(1.1.3) Let $\{A_j\}_{j\in J}, \{B_j\}_{j\in J}$ be collections of subsets of a set A. Are the equalities

$$\Big[\bigcup_{j\in J} A_j\Big] \cap \Big[\bigcup_{j\in J} B_j\Big] = \bigcup_{j\in J} (A_j \cap B_j)$$

and

$$\Big[\bigcap_{j\in J} A_j\Big] \cup \Big[\bigcap_{j\in J} B_j\Big] = \bigcap_{j\in J} (A_j \cup B_j)$$

true? Prove them, or find a counterexample.

Answer. Let $A_1 = \{1\}$, $A_2 = \{2\}$, $B_1 = \{2\}$, $B_2 = \{1\}$. Then

$$(A_1 \cup A_2) \cap (B_1 \cup B_2) = \{1, 2\},$$

while $A_1 \cap B_1 = A_2 \cap B_2 = \varnothing$. What is true is the inclusion

$$\bigcup_{j\in J} (A_j \cap B_j) \subset \Big[\bigcup_{j\in J} A_j\Big] \cap \Big[\bigcup_{j\in J} B_j\Big],$$

for if $a \in A_k \cap B_k$ for some k, then $a \in \bigcup_j A_j$ and $a \in \bigcup_j B_j$.

The second equality is the complement of the first one, so it cannot be true either. We have

$$\Big[\bigcap_{j\in J} A_j\Big] \cup \Big[\bigcap_{j\in J} B_j\Big] \subset \bigcap_{j\in J} (A_j \cup B_j),$$

because if a is in every A_j or a is in every B_j, then $a \in A_j \cup B_j$ for all j. The inclusion is proper in general; consider the same sets $A_1 = \{1\}$, $A_2 = \{2\}$, $B_1 = \{2\}$, $B_2 = \{1\}$ from before. Then $A_1 \cap A_2 = B_1 \cap B_2 = \varnothing$, but $A_1 \cup B_1 = A_2 \cup B_2 = \{1, 2\}$ and the intersection is nonempty.

(1.1.4) Prove Proposition 1.1.1. Show that the inclusion $f\Big(\bigcap_j B_j\Big) \subset \bigcap_j f(B_j)$ can be strict.

Answer. We have

$$a \in f^{-1}\Big(\bigcup_k B_k\Big) \iff f(a) \in \bigcup_k B_k \iff \exists k:\ f(a) \in B_k$$
$$\iff \exists k:\ a \in f^{-1}(B_k) \iff a \in \bigcup_k f^{-1}(B_k).$$

Similarly,

$$a \in f^{-1}\Big(\bigcap_k B_k\Big) \iff f(a) \in \bigcap_k B_k \iff \forall k:\ f(a) \in B_k$$
$$\iff \forall k:\ a \in f^{-1}(B_k) \iff a \in \bigcap_k f^{-1}(B_k).$$

For the complement,

$a \in f^{-1}(B\backslash B_0) \iff f(a) \in B\backslash B_0 \iff a \notin f^{-1}(B_0) \iff a \in A\backslash f^{-1}(B_0).$

As for the images,

$$f\Big(\bigcup_j A_j\Big) = \Big\{f(a) : a \in \bigcup_j A_j\Big\} = \bigcup_j f(A_j).$$

And if $b \in f\Big(\bigcap_j A_j\Big)$, then $b = f(a)$ with $a \in A_j$ for all j. Hence $a \in \bigcap_j f(A_j)$.

Regarding counterexamples, we can have $f : \{1, 2\} \to \{1\}$ be the only possible function, $f(x) = 1$. If $A_1 = \{1\}$ and $A_2 = \{2\}$, then $f(A_1 \cap A_2) = f(\varnothing) = \varnothing$, while $f(A_1) \cap f(A_2) = \{1\} \cap \{1\} = \{1\}$.

If the empty set makes the example above look unconvincing, we can tweak it slightly. Let $f : \{1, 2, 3\} \to \{1, 2\}$ be given by $f(2) = 2$, $f(1) = f(3) = 1$. Put

$$A_1 = \{1, 2\}, \qquad A_2 = \{2, 3\}.$$

Then $f(A_1 \cap A_2) = \{2\}$, while $f(A_1) \cap f(A_2) = \{1, 2\}$.

(1.1.5) Let $f : A \to B$. For any $B_0 \subset B$ and $A_0 \subset A$, show that

$$f(f^{-1}(B_0)) \subset B_0, \qquad A_0 \subset f^{-1}(f(A_0)).$$

Show that equality does not always hold, but that

$$f(f^{-1}(B_0)) = B_0$$

whenever f is surjective or, more generally, if $B_0 \subset f(A)$.

Answer. If $a \in f^{-1}(B_0)$, it means that $f(a) \in B_0$. Hence

$$f(f^{-1}(B_0)) \subset B_0.$$

If $f : \{1, 2\} \to \{1, 2\}$ is given by $f(x) = 1$, then $f(f^{-1}(\{1, 2\})) = f(\{1, 2\}) = \{1\} \subsetneq \{1, 2\}$. When $B_0 \subset f(A)$, given any $b \in B_0$ there exists $a \in A$ with $f(a) = b$. Then $a \in f^{-1}(B_0)$ and $b = f(a) \in f(f^{-1}(B_0))$, so $B_0 \subset f(f^{-1}(B_0))$.

As for the second inclusion, if $a \in A_0$, then $f(a) \in f(A_0)$, so $a \in f^{-1}(f(A_0))$. Thus $A_0 \subset f^{-1}(f(A_0))$. To see that the inclusion can be strict, let $A = B = \{1, 2\}$, $f : A \to B$ given by $f(x) = 1$. Put $A_0 = \{1\}$. Then $f^{-1}(f(A_0)) = f^{-1}(\{1\}) = \{1, 2\} \supsetneq A_0$.

(1.1.6) Let $f : A \to B$ be a function. Show that

(i) f is injective if and only if there exists $g : B \to A$ with $g \circ f = \mathrm{id}_A$;

(ii) f is surjective if and only if there exists $h : B \to A$ with $f \circ h = \mathrm{id}_B$;

(iii) f is bijective if and only if it is invertible.

Answer.

(i) Suppose that f is injective. Fix $a_0 \in A$. Define $g : B \to A$ by $g(f(a)) = a$ on $f(A)$, and $g(b) = a_0$ for $b \in B \setminus f(A)$; the injectivity of f guarantees that g is well-defined. Then $g(f(a)) = a$ for all $a \in A$ by construction. Conversely, if g exists with $g \circ f = \mathrm{id}_A$ and $f(a_1) = f(a_2)$, then

$$a_1 = g(f(a_1)) = g(f(a_2)) = a_2,$$

and f is injective.

(ii) Suppose that f is surjective. Given $b \in B$, choose one element $a_b \in f^{-1}(\{b\})$; these always exist because f is surjective. Let $h(b) = a_b$. Then $f(h(b)) = f(a_b) = b$. Conversely, if h exists with $f \circ h = \mathrm{id}_B$, given $b \in B$ we have $b = f(h(b))$, and so f is surjective.

(iii) If f is bijective, by the previous part there exist $g : B \to A$ and $h : B \to A$ with $g \circ f = \mathrm{id}_A$ and $f \circ h = \mathrm{id}_B$. Then

$$g = g \circ \mathrm{id}_B = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_A \circ h = h,$$

so $g = h$ and hence f is invertible. Conversely, if f is invertible then we can take $g = h = f^{-1}$ and then the arguments above show that f is bijective.

(1.1.7) Let A be a set with an associative operation $(a, b) \longmapsto ab$ and with a unit $e \in A$ (that is, $ae = ea = a$ for all $a \in A$). Show that the unit is unique. Show also that if $a \in A$ is invertible (that is there exists $b \in A$ with $ab = ba = e$) then b is unique with that property. More generally, show that if a has a left inverse b and a right inverse c, then $b = c$.

Answer. If e and f are units, then $e = ef = f$.

Now suppose that $ba = ac = e$. Then

$$b = be = b(ac) = (ba)c = ec = c.$$

As for inverses, if there exist b, c such that $ba = ac = e$, then

$$b = be = b(ac) = (ba)c = ec = c.$$

(1.1.8) Let R be a relation on $\mathbb{Z}$ defined by:

$$a\,R\,b \iff 3 \text{ divides } (a - b).$$

Show that R is an equivalence relation. Determine its equivalence classes. The quotient $\mathbb{Z}/R$ is often denoted by $\mathbb{Z}_3$.

Answer. The relation is reflexive, because 3 divides 0. It is symmetric, for if $a - b$ is a multiple of 3, so is $b - a = -(a - b)$. And it is transitive, because if $a - b = 3n$ and $b - c = 3m$, then

$$a - c = (a - b) + (b - c) = 3n + 3m = 3(n + m).$$

If $m \in \mathbb{Z}$ then $m = 3q + r$ (via the Division Algorithm) for unique $q, r \in \mathbb{Z}$ and $0 \leq r < 3$. As $m - r = 3q \in 3\mathbb{Z}$, we have that $m \sim r$. So $\{0, 1, 2\}$ forms a set of representatives. The classes are $3\mathbb{Z}$, $3\mathbb{Z} + 1$ and $3\mathbb{Z} + 2$.

(1.1.9) Let $f : X \to Y$ be a function. Define a relation $\sim$ on X by:

$$x_1 \sim x_2 \iff f(x_1) = f(x_2).$$

(i) Prove that $\sim$ is an equivalence relation.

(ii) Show that the equivalence classes are the "fibers" of f (i.e., sets of the form $f^{-1}(\{y\})$ for $y \in Y$).

Answer. We have $f(x) = f(x)$ so $x \sim x$. If $x \sim y$ then $f(x) = f(y)$, so $y \sim x$. And if $f(x) = f(y)$ and $f(y) = f(z)$, then $f(x) = f(z)$ and so $x \sim z$.

For the equivalence classes, given $x \in X$

$$[x] = \{z \in X : f(z) = f(x)\} = f^{-1}(\{f(x)\}).$$

(1.1.10) Consider the relation $\sim$ on $\mathbb{R}^2$ defined by:

$$(x_1, y_1) \sim (x_2, y_2) \iff x_1^2 + y_1^2 = x_2^2 + y_2^2.$$

(i) Prove that $\sim$ is an equivalence relation.

(ii) Describe the equivalence classes geometrically.

Answer.

(i) This is a relation of the form considered in Exercise 1.1.9, so it is an equivalence relation.

(ii) Two points in $\mathbb{R}^2$ are equivalent if their distance to the origin is the same. So the classes are the distinct circles centered at the origin.

(1.1.11) Let $S = \{1, 2, 3, 4, 5\}$ and define R as:

$$R = \{(1,1), (2,2), (3,3), (4,4), (5,5), (1,2), (2,1), (3,4), (4,3)\}.$$

(i) Show that R is an equivalence relation.

(ii) Determine the quotient S/R.

Answer.

(i) The relation is reflexive because $(x, x) \in R$ for all $x \in S$. It is symmetric because for every pair $(x, y) \in R$ the corresponding pair (y, x) is in R. And it is transitive: the only way to have pairs (x, y) and (y, z) in R is the constant pairs (x, x) and $(1, 2)$, $(2, 1)$ and $(3, 4)$, $(4, 3)$, all of the form (x, y) and (y, x), where in all cases we also have $(x, x) \in R$.

(ii) The classes are $\{1,2\}$, $\{3,4\}$, and $\{5\}$.

(1.1.12) Prove that any partition $\mathcal{P}$ of a set S induces an equivalence relation $\sim$ on S where:

$$a \sim b \iff a \text{ and } b \text{ belong to the same subset in } \mathcal{P}.$$

Conversely, show that any equivalence relation on S induces a partition of S.

Answer. Suppose that $\mathcal{P} = \{S_j : j \in J\}$ is a partition of S. Let $\sim$ be given by $a \sim b$ if there exists $j \in J$ with $a, b \in S_j$. This is reflexive and symmetric by definition, and transitivity is also automatic: if $a, b \in S_j$ and $b, c \in S_k$, then $b \in S_j \cap S_k$ which implies that $k = j$ since the sets in the partition are disjoint; so $a \sim c$.

Conversely, if $\sim$ is an equivalence relation, let $\mathcal{P}$ be the sets of classes for $\sim$. For any $a \in S$ there exists $P \in \mathcal{P}$ with $a \in P = [a]$ since every element belongs to its own class. Therefore the union of all the classes is equal to S. It remains to show that they are pairwise disjoint. Suppose that $a \in [b] \cap [c]$. Then $a \sim b$ and $a \sim c$, so $b \sim c$ by the transitivity. If $d \sim b$ then $d \sim c$ by the transitivity, so $[b] \subset [c]$; exchanging roles we get that $[b] = [c]$. We have shown that if $[b] \cap [c] \neq \varnothing$ then $[b] = [c]$, so the classes are pairwise disjoint.

(1.1.13) Let $\sim$ be an equivalence relation on $\mathbb{N} \times \mathbb{N}$ defined by:

$$(a,b) \sim (c,d) \iff a + d = b + c.$$

Attention: the set R here consists of pairs of ordered pairs!

(i) Prove that $\sim$ is an equivalence relation.

(ii) Show that addition defined as:

$$[(a,b)] + [(c,d)] = [(a+c, b+d)]$$

is **well-defined** (i.e., independent of the choice of representatives).

(iii) Show that multiplication defined as

$$[(a,b)][(c,d)] = [ac + bd, ad + bc]$$

is well-defined, and it is distributive with respect to addition.

Answer.

(i) From $a+b=b+a$ we get that the relation is reflexive. If $a+d=b+c$ then $c+b=d+a$, so the relation is symmetric. And if $a+d=b+c$ and $c+f=d+e$, then

$$(a+f)+d=(a+d)+f=b+c+f=b+d+e=(b+e)+d.$$

As we can cancel d, we get that $a+f=b+e$ and so $(a,b)\sim(e,f)$; therefore the relation is transitive.

(ii) If $(a',b')\in[(a,b)]$ and $(c',d')\in[(c,d)]$, then using that $a'+b=a+b'$ and $c'+d=c+d'$,

$$a'+c'+b+d=(a'+b)+(c'+d)=(a+b')+(c+d')=a+c+b'+d',$$

showing that $[(a+c,b+d)]=[(a'+c',b'+d')]$.

(iii) If $(a',b')\in[(a,b)]$ and $(c',d')\in[(c,d)]$, then $a'+b=a+b'$ and $c'+d=c+d'$. We need to show that

$$(ac+bd,ad+bc)\sim(a'c'+b'd',a'd'+b'c').$$

That is, we need to show that

$$ac+bd+a'd'+b'c'=ad+bc+a'c'+b'd'.$$

From $a+b'=b+a'$ and $c+d'=d+c'$, multiplying by c',

$$a'c'+bc'=ac'+b'c'. \tag{AB.1.1}$$

Multiplying $c'+d=c+d'$ by a:

$$ac'+ad=ac+ad'. \tag{AB.1.2}$$

Multiplying $a+b'=a'+b$ by d':

$$ad'+b'd'=a'd'+bd'. \tag{AB.1.3}$$

Multiplying $c+d'=d+c'$ by b:

$$bc+bd'=bc'+bd. \tag{AB.1.4}$$

Adding the four equalities and cancelling ac', ad', bd', and bc' from both sides,

$$ac+bd+a'd'+b'c'=ad+bc+a'c'+b'd'.$$

Which is exactly what we needed to show; thus, multiplication is well-defined.

It remains to check the distributivity. Since the operations are well-defined, we can work with representatives without concern. We have

$$\begin{aligned}[(a,b)]\big([(c,d)]+[(e,f)]\big) &= [(a,b)][(c+e,d+f)]\\ &= [(a(c+e)+b(d+f),a(d+f)+b(c+e)]\\ &= [(ac+bd+ae+bf,ad+bc+af+be)]\\ &= [(ac+bd,ad+bc)]+[(ae+bf,af+be)]\\ &= [(a,b)][(c,d)]+[(a,b)][(e,f)].\end{aligned}$$

(1.1.14) Let $\sim$ be an equivalence relation on $\mathbb{Z}\times\mathbb{Z}^*$ (where $\mathbb{Z}^* = \mathbb{Z}\setminus\{0\}$) defined by:

$$(a,b)\sim(c,d) \iff ad = bc.$$

(i) Prove that $\sim$ is an equivalence relation.

(ii) Show that addition defined as:

$$[(a,b)]+[(c,d)] = [(ad+bc,bd)]$$

is **well-defined** (i.e., independent of the choice of representatives).

(iii) Show that multiplication defined as

$$[(a,b)][(c,d)] = [ab,cd]$$

is well-defined, and it is distributive with respect to addition.

Answer.

(i) We have $ab = ba$, so $(a,b)\sim(a,b)$ and the relation is reflexive. If $(a,b)\sim(c,d)$ then $ad = bc$; written as $cb = da$ this says that $(c,d)\sim(a,b)$ and the relation is symmetric. Finally, if $(a,b)\sim(c,d)$ and $(c,d)\sim(e,f)$, then $ad = bc$ and $cf = de$. If $c = 0$, then from $b,d\neq 0$ we get $a = e = 0$ and $af = be$ holds. Otherwise, if $c\neq 0$, then

$$afcd = (ad)(cf) = (bc)(de) = becd.$$

As $cd\neq 0$ we can cancel and get $af = be$; that is, $(a,b)\sim(e,f)$.

(ii) If $(a',b')\sim(a,b)$ and $(c',d')\sim(c,d)$, then $a'b = ab'$ and $c'd = cd'$. Then

$$\begin{aligned}(ad+bc)b'd' &= adb'd' + bcb'd' = (ab')d'd + (cd')b'b\\ &= a'd'bd + b'c'bd = (a'd'+b'c')bd.\end{aligned}$$

So $[(a',b')]+[(c',d')]=[(a,b)]+[(c,d)]$.

(iii) We have
$$acb'd'=a'c'bd,$$
so $[(a',b')][(c',d')]=[(a,b)][(c,d)]$ and the multiplication is well-defined.

For the distributivity,
$$\begin{aligned}[(a,b)]\big([(c,d)]+[(e,f)]\big)&=[(a,b)][(cf+de,df)]=[(acf+ade,bdf)]\\&=[(a,b)][(c,d)]+[(a,b)][(e,f)].\end{aligned}$$

1.2. The Axiom of Choice

(1.2.1) Let R be a nonzero unital commutative ring and $x\in R$ non-invertible. Show that there exists a proper maximal ideal J of R with $x\in J$.

Answer. The ideal xR generated by x cannot be all of R because if it were then x would be invertible. Let
$$\mathcal{J}=\{J\subset R:\quad \text{proper ideal with } x\in J\},$$
ordered by inclusion. If $\{J_k\}$ is a chain in $\mathcal{J}$, let $J_\infty=\bigcup_k J_k$. As the union is monotone, J_∞ is an ideal. And it is proper, for if $1\in J_\infty$ then there exists k with $1\in J_k$, a contradiction. So J_∞ is an upper bound for the chain in $\mathcal{J}$, and by Zorn's Lemma there exists $J\in\mathcal{J}$, maximal proper ideal with $x\in J$.

(1.2.2) Let X be a set. Prove that Zorn's Lemma implies the Well Ordering Principle by applying Zorn's Lemma to the collection
$$\mathcal{P}=\{(A,\le_A): A\subset X,\ \le_A\quad \text{is a well-ordering on } A\},$$
where the order is given by saying that $(A,\le_A)\preceq(B,\le_B)$ if $A\subset B$, $\le_B$ extends $\le_A$, and every element of A is less (in the $\le_B$ order) than every element of $B\setminus A$.

Answer. We have that $\mathcal{P}$ is nonempty because it contains singletons with the only possible order on each. Suppose that $\{(A_j, \leq_{A_j}\}$ is a chain in $\mathcal{P}$. Let $U = \bigcup_j A_j$ with the order $\leq_U$ defined as follows: if $u, v \in U$ there exists j such that $u, v \in A_j$, and we say that $u \leq_U v$ if $u \leq_{A_j} v$. This is well-defined because if $u, v \in A_k$ with $k > j$, the order in A_j is the restriction of the order in A_k and so $u \leq_{A_j} v \iff u \leq_{A_k} v$. We claim that U is well-ordered. Let $S \subset U$ be nonempty. Then there exists j with $S \cap A_j \neq \varnothing$. As A_j is well-ordered by $\leq_{A_j}$, there exists a least element $s \in S \cap A_j$. Because $\leq_U$ restricts to $\leq_{A_j}$, the element s is also least for $\leq_U$. Thus $U \in \mathcal{P}$ and it is an upper bound for the chain. By Zorn's Lemma, $\mathcal{P}$ admits a maximal element $(M, \leq_M)$.

If we had $M \subsetneq X$, pick $x \in X \setminus M$ and let $M' = M \cup \{x\}$ with the order $\leq_{M'}$ defined to be $\leq_M$ for all elements of M, and $m \leq_{M'} x$ for all $m \in M$. Then $(M', \leq_{M'}) \in \mathcal{P}$ and $(M, \leq_M) \preceq (M', \leq_{M'})$, contradicting the maximality.

1.3. Real Numbers and Calculus

(1.3.1) Let $E \subset \mathbb{R}$. Show that $\inf E = -\sup(-E)$ and
$$\liminf_n a_n = -\limsup_n (-a_n).$$

Answer. Let c be a lower bound for E. Then $-c$ is an upper bound for $-E$, which gives us $\sup(-E) \leq -c$, and so $c \leq -\sup(-E)$. As this occurs for every lower bound of E, we get $\inf E \leq -\sup(-E)$. Conversely, if c is an upper bound for $-E$, then $-c$ is a lower bound for E, which means that $-c \leq \inf E$, which we can write as $-\inf E \leq c$. So $-\inf E$ is below every upper bound for $-E$, and so $-\inf E \leq \sup(-E)$, which is $-\sup(-E) \leq \inf E$.

Now, using the above twice,
$$-\limsup_n (-a_n) = -\inf_m \sup_{n \geq m} (-a_n) = \sup_m (-\sup_{n \geq m} (-a_n))$$
$$= \sup_m \inf_{n \geq m} a_n = \liminf_n a_n.$$

(1.3.2) Let $\{a_n\} \subset \mathbb{R}$ be a sequence. Allowing $\pm\infty$ to be cluster points for unbounded sequences, show that

$$\limsup_n a_n = \max\{\text{cluster points of } \{a_n\}\},$$

and

$$\liminf_n a_n = \min\{\text{cluster points of } \{a_n\}\}.$$

Answer. Assume first that the set of cluster points is bounded (this is equivalent to $\{a_n\}$ being bounded).

Let $A = \limsup_n a_n$ and B the maximum cluster point of $\{a_n\}$. Since B is a cluster point, there exists a subsequence $\{a_{n_j}\}$ with $a_{n_j} \to B$. Then for any m there exists j with $n_j \geq m$, and so

$$B = \lim_j a_{n_j} \leq \sup_{n \geq m} a_n.$$

Then B is a lower bound for $\sup_{n \geq m}\{a_n\}$ for all m, and thus $B \leq A$. Conversely, since $A = \inf_m \sup_{n \geq m} a_n$, there exists a subsequence $\{m_j\}$ such that $\sup_{n \geq m_j} a_n \searrow A$. For each j we can find n_j such that $|a_{n_j} - \sup_{n \geq m_j}| < \frac{1}{j}$. Then $a_{n_j} \to A$, that is A is a cluster point for $\{a_n\}$; this immediately gives us $A \leq B$ and so $A = B$. We also have

$$\liminf_n a_n = -\limsup_n(-a_n). \tag{AB.1.5}$$

By the first part of the answer, the right hand side is $-C$, where C is the largest cluster point of $\{-a_n\}$. Then $-C$ is the smallest cluster point of $-\{-a_n\} = \{a_n\}$.

If $\{a_n\}$ is unbounded above, then $B = \infty$. If $\{a_{n_j}\}$ is a subsequence of $\{a_n\}$ such that $a_{n_j} \geq j$ for all j, then $\sup_{n \geq m} a_n = \infty$ for all m, and so $A = \infty = B$. This also gives the result when $\{a_n\}$ is unbounded below, via (AB.1.5).

(1.3.3) Let $\{a_n\} \subset \mathbb{R}$ be a sequence. Show that $\lim_n a_n$ exists and is equal to $L \in \mathbb{R}$ if and only if $\limsup_n a_n = \liminf_n a_n = L$.

Answer. If $\lim_n a_n = L$, then L is the only cluster point of the sequence. By Exercise 1.3.2 we have that $\liminf_n a_n = L = \limsup_n a_n$. Exercise 1.3.2 also provides the converse, for if $\liminf_n a_n = L = \limsup_n a_n$ then L is the only cluster point of $\{a_n\}$ and hence $\lim_n a_n = L$.

(1.3.4) Prove (1.3).

Answer. We start, from the Fundamental Theorem of Calculus, with

$$f(x) = f(0) + \int_0^x f'(s)\,ds = f(0) + x\int_0^1 f'(tx)\,dt.$$

Now we proceed by induction. If (1.3) holds for n and $f^{(n+1)}$ exists,

$$\begin{aligned}\frac{x^n}{(n-1)!}\int_0^1 (1-t)^{n-1} f^{(n)}(tx)\,dt &= \frac{x^n}{(n-1)!}\left[-\frac{(1-t)^n}{n} f^{(n)}(tx)\Big|_0^1\right.\\ &\qquad \left.+\frac{x}{n}\int_0^1 (1-t)^n f^{(n+1)}(tx)\,dt\right]\\ &= \frac{f^{(n)}(0)}{n!}\,x^n + \frac{x^{n+1}}{n!}\int_0^1 (1-t)^n f^{(n+1)}(tx)\,dt.\end{aligned}$$

(1.3.5) Let $f : [0,\infty) \to \mathbb{R}$ be continuously differentiable, with $f'(x) > 0$ for all x, and such that $f(x) \le c$ for all x.

(i) Show that $\lim_{x\to\infty} f(x)$ exists and it is equal to $\sup\{f(x) : x \ge 0\}$;

(ii) show that there exists f as above and such that $\lim\limits_{x\to\infty} f'(x)$ does not necessarily exist;

(iii) show that if in addition f' is differentiable and $f''(x) < 0$ for all sufficiently large x, then $\lim\limits_{x\to\infty} f'(x) = 0$.

Answer.

(i) Let $s = \sup\{f(x) :\ x \ge 0\}$. By hypothesis, $s \le c < \infty$. Fix $\varepsilon > 0$. Then there exists x_0 such that $s - f(x_0) < \varepsilon$. Suppose that $x \ge x_0$. Then

$$f(x) = f(x_0) + \int_{x_0}^x f'(t)\,dt > f(x_0)$$

since $f' > 0$. We have

$$|s - f(x)| = s - f(x) < s - f(x_0) < \varepsilon,$$

showing that $\lim_{x\to\infty} f(x) = s$.

(ii) Let

$$g(x) = \sum_{n=0}^{\infty} \frac{1}{n^2} 1_{[n,n+1)}(x) + \left(1 - n^2 \left| x - n - \frac{1}{n^2} \right| \right) 1_{\left[n, n+\frac{2}{n^2}\right]}$$

and

$$f(x) = \int_0^x g(t)\, dt.$$

The idea is that g has bumps so it has no limit, but the bumps shrink in width so that the integral is bounded. The first term of the sum guarantees that $g(x) > 0$ for all x, and the second provides the bumps. The function g is continuous and positive by construction, so $f'(x) = g(x) > 0$ for all $x > 0$. We have

$$f(x) \leq \int_0^\infty g(t)\, dt = \sum_{n=0}^{\infty} \frac{1}{n^2} + \frac{1}{n^2} = \frac{\pi^2}{3},$$

so f is bounded. Since $g(n + \frac{1}{n^2}) = 1$ and $g(n) = \frac{1}{n^2}$ for all n, the limit of $f' = g$ at infinity does not exist.

(iii) We have that $f'(x) > 0$ for all x, and since $f''(x) < 0$ for all x we have that f' is decreasing. Let $c_0 = \inf\{f'(x) : x > 0\}$; this exists for the set is bounded below by 0. Now the function $-f'$ satisfies the condition of the original question, so by part (i) we have that $c_0 = \lim_{x\to\infty} f'(x)$. Write $s_0 = \lim_{x\to\infty} f(x)$. If $c_0 > 0$, we can choose $x_0 > 0$ such that for all $x \geq x_0$ we have $f(x) > s_0 - \frac{c_0}{2}$. Then we would have

$$f(x_0 + 1) = f(x_0) + \int_{x_0}^{x_0+1} f'(t)\, dt \geq s_0 - \frac{c_0}{2} + \int_{x_0}^{x_0+1} c_0\, dt = s_0 + \frac{c_0}{2},$$

a contradiction since $s_0 \geq f(x)$ for all x. It follows that $c_0 = 0$, as desired.

(1.3.6) Let $[a, b] \subset \mathbb{R}$ and P, P' partitions with $P \subset P'$. Show that $L(f, P) \leq L(f, P') \leq U(f, P') \leq U(f, P)$.

Answer. That $L(f,P') \leq U(f,P')$ follows from $m_j(f) \leq M_j(f)$. Consider the partition $P \cup \{y\}$. Suppose that $x_j < y < x_{j+1}$. Then

$$\begin{aligned} U(f, P\cup\{y\}) &= \sum_{k=1}^{j} M_j(f)\,(x_k - x_{k-1}) + M_{[x_j,y]}(f)\,(y - x_j) \\ &\qquad + M_{[y,x_{j+1}]}(f)\,(x_{j+1} - y) + \sum_{k=j+2}^{m} M_j(f)\,(x_k - x_{k-1}) \\ &\leq \sum_{k=1}^{j} M_j(f)\,(x_k - x_{k-1}) + M_j(f)\,(y - x_j) \\ &\qquad + M_j(f)\,(x_{j+1} - y) + \sum_{k=j+2}^{m} M_j(f)\,(x_k - x_{k-1}) \\ &= U(f,P). \end{aligned}$$

Iterating this we get that $U(f,P') \leq U(f,P)$. The corresponding inequalities for the lower sums follow from $L(f,P) = -U(-f,P)$.

(1.3.7) Let $f : [a,b] \to \mathbb{R}$ be bounded. Show that f is Riemann integrable if and only if for each $\varepsilon > 0$ there exists a partition P of $[a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$.

Answer. Suppose first that f is Riemann integrable and fix $\varepsilon > 0$. By definition of supremum and infimum there exist partitions P, Q of $[a,b]$ such that $|U(f,P) - L(f,Q)| < \varepsilon$. Using Exercise 1.3.6,

$$U(f, P\cup Q) - L(f, P\cup Q) \leq U(f,P) - L(f,Q) < \varepsilon.$$

Conversely, assuming that for each $\varepsilon > 0$ there exists a partition P with $U(f,P) - L(f,P) < \varepsilon$, we have $U(f,P) < L(f,P) + \varepsilon$. This gives

$$U(f,P) < \varepsilon + \sup\{L(f,Q) : \ Q\}.$$

As ε was arbitrary, this shows that $\sup\{L(f,Q) : \ Q\}$ is an upper bound for $\{U(f,P) : \ P\}$. In particular $\inf\{U(f,P) : \ P\} \leq \sup\{L(f,P) : \ P\}$. The reverse inequality is trivial by Exercise 1.3.6, and therefore f is integrable.

(1.3.8) Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. Show that

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(a + \frac{k(b-a)}{n}\right)\frac{1}{n}. \tag{1.4}$$

Answer. Let

$$P_n = \left\{a + \frac{k(b-a)}{n}\right\}_{k=1}^{n}.$$

This is a partition of $[a, b]$, so we have

$$L(f, P_n) \leq \sum_{k=1}^{n} f\left(a + \frac{k(b-a)}{n}\right)\frac{1}{n} \leq U(f, P_n).$$

Now fix $\varepsilon > 0$. By definition of integrability there exists a partition Q such that $U(f, Q) - L(f, Q) < \varepsilon$. Let m be the size of Q, and choose n big enough so that each element of Q lies in a single interval of P_n. Then

$$L(f, Q) \leq L(f, P_n) \leq \sum_{k=1}^{n} f\left(a + \frac{k(b-a)}{n}\right)\frac{1}{n} \leq U(f, P_n) \leq U(f, Q).$$

If follows that for $m \geq n$,

$$\left|\int_a^b f(t)\,dt - \sum_{k=1}^{n} f\left(a + \frac{k(b-a)}{n}\right)\frac{1}{n}\right| < \varepsilon.$$

Hence (1.4) holds.

(1.3.9) Show that a bounded function f is continuous at x if and only if $o(f, x) = 0$.

Answer. By definition of sup and inf, the numbers $m(f, x, \delta)$ increase with δ and the numbers $M(f, x, \delta)$ decrease with δ.

Suppose that f is continuous at x. Given $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $y \in (x - \delta, x + \delta)$. Then $M(f, x, \delta) \leq f(x) + \varepsilon$. As $f(y) > f(x) - \varepsilon$ we also have $m(f, x, \delta) > f(x) - \varepsilon$. Then

$$M(f, x, \delta) - m(f, x, \delta) < 2\varepsilon.$$

As ε was arbitrary, $o(f, x) = 0$.

Conversely, suppose that $o(f, x) = 0$ and fix $\varepsilon > 0$. Then there exists δ with $M(f, x, \delta) - m(f, x, \delta) < \varepsilon$. This says that $|f(x) - f(y)| < \varepsilon$ on $(x - \delta, x + \delta)$, so f is continuous.

1.4. Trigonometric Functions

(1.4.1) Find exact formulas for $\sin\frac{\pi}{5}$ and $\cos\frac{\pi}{5}$.

Answer. Write $s = \sin\frac{\pi}{5}$, $c = \cos\frac{\pi}{5}$, $t = \frac{\pi}{5}$. We have that $\sin 3t = \sin(\pi - 3t) = \sin 2t$. This we can rewrite as

$$\sin 2t \cos t + \cos 2t \sin t = \sin 2t.$$

Using that

$$\sin 2t = 2\sin t\cos t, \qquad \cos 2t = \cos^2 t - \sin^2 t = 1 - 2\sin^2 t,$$

the first equality becomes

$$2sc^2 + (1 - 2s^2)s = 2sc.$$

After dividing both sides by s (which is nonzero since $0 < \frac{\pi}{5} < \frac{\pi}{2}$), using that $s^2 = 1 - c^2$, and simplifying, we get

$$4c^2 - 2c - 1 = 0.$$

Knowing that $c > 0$, from the quadratic equation we get

$$\cos\frac{\pi}{5} = c = \frac{1+\sqrt{5}}{4}.$$

And then

$$\sin\frac{\pi}{5} = \sqrt{1 - c^2} = \frac{\sqrt{10 - 2\sqrt{5}}}{4}.$$

(1.4.2) Write formulas to express the sine and the cosine in terms of the tangent for $x \in \left[0, \frac{\pi}{2}\right)$. Explain how to adapt the formulas for arbitrary x.

Answer. When $0 \le x < \frac{\pi}{2}$, both the sine and the cosine are non-negative. Then

$$\tan^2 x = \frac{\sin^2 x}{\cos^2 x} = \frac{\sin^2 x}{1 - \sin^2 x},$$

and solving (we can happily take square roots because everything is non-negative)

$$\sin x = \frac{\tan x}{\sqrt{1+\tan^2 x}}.$$

In an analogous way we get

$$\cos x = \frac{1}{\sqrt{1+\tan^2 x}}.$$

For $x \in \left(\frac{\pi}{2}, \pi\right]$ we have $\sin x = \sin(\pi - x)$ and $\cos x = -\cos(\pi - x)$. Then

$$\sin x = \sin(\pi - x) = \frac{\tan(\pi - x)}{\sqrt{1+\tan^2(\pi - x)}} = -\frac{\tan x}{\sqrt{1+\tan^2 x}},$$

and

$$\cos x = -\cos(\pi - x) = -\frac{1}{\sqrt{1+\tan^2(\pi - x)}} = -\frac{\tan x}{\sqrt{1+\tan^2 x}}.$$

In the third quadrant the tangent is again non-negative, so we get

$$\sin x = \frac{\tan x}{\sqrt{1+\tan^2 x}}$$

and

$$\cos x = \frac{1}{\sqrt{1+\tan^2 x}}.$$

And in the fourth quadrant the sine and cosine are both negative; this makes tan positive, and so the formulas will be the same as in the second quadrant.

(1.4.3) Find an addition formula for the tangent; that is, express $\tan(x+y)$ as a formula on $\tan x$ and $\tan y$.

Answer. We have, factoring $\cos x \cos y$ out from numerator and denominator when they are nonzero,

$$\begin{aligned} \tan(x+y) &= \frac{\sin(x+y)}{\cos(x+y)} = \frac{\sin x \cos y + \cos x \sin y}{\cos x \cos y - \sin x \sin y} \\ &= \frac{\tan x + \tan y}{1 - \tan x \tan y} \end{aligned}$$

If $\cos x = 0$ and $\cos y \neq 0$ we have

$$\tan(x+y) = -\frac{\sin x \cos y}{\sin x \sin y} = -\frac{1}{\tan y},$$

which actually agrees with the limit as $x \to \frac{\pi}{2}$ of the full expression for the sum. When $\cos y = 0$ and $\cos x \neq 0$ we get a similar expression. And when $\cos x = \cos y = 0$, then $x = \frac{2k+1}{2}\pi$, $y = \frac{2j+1}{2}\pi$, and $x + y = (2k+2j+1)\pi$, so $\tan(x+y) = 0$.

(1.4.4) For each x, let

$$T(x) = \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix}.$$

Show that $T(x+y) = T(x)T(y)$.

Answer. We have

$$\begin{aligned} T(x+y) &= \begin{bmatrix} \cos x \cos y - \sin x \sin y & -\sin x \cos y - \cos x \sin y \\ \sin x \cos y + \cos x \sin y & \cos x \cos y - \sin x \sin y \end{bmatrix} \\ &= \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix} \begin{bmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{bmatrix} = T(x)T(y). \end{aligned}$$

(1.4.5) Find formulas for $\sin(2x)$ and $\cos(2x)$ in terms of $\tan x$.

Answer. We have, from Exercise 1.4.2

$$\sin 2x = 2 \sin x \cos x = \frac{2 \tan x}{1 + \tan^2 x}.$$

And

$$\cos 2x = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x = 1 - \frac{2\tan^2 x}{1+\tan^2 x} = \frac{1 - \tan^2 x}{1 + \tan^2 x}.$$

(1.4.6) Show that

$$\sin x + \sin y = 2 \sin\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right),$$

$$\cos x + \cos y = 2 \cos\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right),$$

and

$$\cos x - \cos y = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right).$$

Answer. With $r = \frac{x+y}{2}$ and $s = \frac{x-y}{2}$, we have

$$\begin{aligned} \sin x + \sin y &= \sin(r+s) + \sin(r-s) \\ &= \sin r \cos s + \cos r \sin s + \sin r \cos s - \cos r \sin s \\ &= 2 \sin r \cos s \\ &= 2 \sin\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right). \end{aligned}$$

Also,

$$\begin{aligned}\cos x + \cos y &= \cos(r+s) + \cos(r-s)\\ &= \cos r \cos s - \sin r \sin s + \cos r \cos s + \sin r \sin s\\ &= 2\cos r \cos s = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)\end{aligned}$$

Similarly,

$$\begin{aligned}\cos x - \cos y &= \cos(r+s) - \cos(r-s)\\ &= \cos r \cos s - \sin r \sin s - \cos r \cos s - \sin r \sin s\\ &= -2\sin r \sin s = -2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right).\end{aligned}$$

(1.4.7) Show that

$$\tan\frac{x+y}{2} = \frac{\sin x + \sin y}{\cos x + \cos y}$$

whenever the denominator is nonzero.

Answer. If $r = (x+y)/2$ and $s = (x-y)/2$, then

$$\begin{aligned}\frac{\sin x + \sin y}{\cos x + \cos y} &= \frac{\sin(r+s) + \sin(r-s)}{\cos(r+s) + \cos(r-s)}\\ &= \frac{\sin r \cos s + \cos r \sin s + \sin r \cos s - \cos r \sin s}{\cos r \cos s - \sin r \sin s + \cos r \cos s + \sin r \sin s}\\ &= \frac{2\sin r \cos s}{2\cos r \cos s} = \frac{\sin r}{\cos r} = \tan\frac{x+y}{2}.\end{aligned}$$

(1.4.8) Show that

$$\frac{\sin^2 x - \sin^2 y}{\cos^2 x - \cos^2 y} = -1$$

whenever the denominator is nonzero.

Answer. The denominator is

$$\cos^2 x - \cos^2 y = 1 - \sin^2 x - (1 - \sin^2 y) = -(\sin^2 x - \sin^2 y).$$

(1.4.9) Show that
$$\frac{\sin x + \sin y}{\cos x + \cos y} = -\frac{\cos x - \cos y}{\sin x - \sin y}$$
whenever the denominators are nonzero.

Answer. Cross multiplying the denominators,
$$\begin{aligned}(\sin x + \sin y)(\sin x - \sin y) &= \sin^2 x - \sin^2 y = -(\cos^2 x - \cos^2 y) \\ &= (\cos x - \cos y)(\cos x + \cos y).\end{aligned}$$

(1.4.10) Show that
$$\{\lambda \sin(x+r) : \ \lambda, r \in \mathbb{R}\} = \{\alpha \cos x + \beta \sin x : \ \alpha, \beta \in \mathbb{R}\}.$$

Answer. We have
$$\lambda \sin(x+r) = \lambda \sin r \, \cos x + \lambda \, \cos r \, \sin x.$$
Conversely, given $\alpha, \beta \in \mathbb{R}$ let
$$\alpha' = \alpha/\sqrt{\alpha^2+\beta^2}, \qquad \beta' = \beta/\sqrt{\alpha^2+\beta^2}.$$
Then we have $\alpha'^2 + \beta'^2 = 1$ and so there exists $r \in \mathbb{R}$ with $\alpha' = \sin r$, $\beta' = \cos r$. Then
$$\alpha \cos x + \beta \sin x = \sqrt{\alpha^2+\beta^2}\,\left[\sin r \cos x + \cos r \sin x\right] = \sqrt{\alpha^2+\beta^2}\, \sin(x+r).$$

(1.4.11) Use the idea in the proof of Lemma 1.4.1 to show that the initial problem $y' = y$, $y(0) = 1$ has at most one solution.

Answer. Having two solutions y_1, y_2 means that $y_1 - y_2$ is a solution of the initial problem $y' = y$, $y(0) = 0$. So it is enough to show that this latter problem has the only solution $y = 0$. We have, since $y = y$, $y^{(n)}(0) = \cdots = y'(0) = y(0) = 0$. Also, working on the interval $[-a, a]$, say, by continuity we have that there exists $c > 0$ with $|y(x)| \le c$. Then the Taylor polynomial of y is just the error term, that is for each $n \in \mathbb{N}$
$$y(x) = \frac{y^{(n)}(\xi(x))\, x^{n+1}}{(n+1)!}, \qquad x \in [-a, a],$$

with $\xi(x)$ between 0 and x. As $y^{(n)} = y$ and $|y| \leq c$, we have

$$|y(x)| \leq \frac{c\,a^{n+1}}{(n+1)!}.$$

As this is true for all $n \in \mathbb{N}$ and the limit on the right is zero, we get that $y(x) = 0$ on $[-a, a]$. But this can be done for all $a > 0$, so $y = 0$.

1.5. Complex Numbers

(1.5.1) Without using series, show that a nonzero complex number z can be written in a unique way as $z = r(\cos\theta + i\sin\theta)$ with $r > 0$ and $\theta \in [0, 2\pi)$.

Answer. If $z \in \mathbb{C}$ is nonzero, then $z = a+ib$ with at least one of a, b nonzero. Then

$$\frac{a}{\sqrt{a^2+b^2}} \qquad \text{and} \qquad \frac{b}{\sqrt{a^2+b^2}}$$

are two real numbers such that their squares add to 1. That is, they form the coordinates of a point in the unit circle. As shown on page 17 of the Book, there exists $\theta \in [0, 2\pi)$ such that

$$\frac{a}{\sqrt{a^2+b^2}} = \cos\theta \qquad \text{and} \qquad \frac{b}{\sqrt{a^2+b^2}} = \sin\theta.$$

If $r = \sqrt{a^2+b^2}$, then $a = r\cos\theta$ and $b = r\sin\theta$.

(1.5.2) Still without using series, follow up from Exercise 1.5.1 by showing that if we formally use the notation $e^{i\theta}$ for the complex number $\cos\theta + i\sin\theta$, then $e^{i(\theta_1+\theta_2)} = e^{i\theta_1}e^{i\theta_2}$.

Answer. Using the addition formulas and the multiplication of complex numbers,

$$\begin{aligned}\big(\cos\theta_1 + i\sin\theta_1\big)\big(\cos\theta_2 + i\sin\theta_2\big) &= \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 \\ &\qquad\qquad + i\big(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2\big) \\ &= \cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2).\end{aligned}$$

(1.5.3) Let $z \in \mathbb{C}$ be nonzero. Show that if $z = r(\cos\theta + i\sin\theta)$, then

$$z^n = r^n(\cos n\theta + i\sin n\theta).$$

Answer. We have $z = re^{i\theta}$. Then $z^n = r^n(e^{i\theta})^n = r^n e^{in\theta}$. The last equality is obtained inductively from (1.5.2).

(1.5.4) Show that the equation $z^n = 1$ has precisely n distinct solutions, that can be expressed as

$$\omega_k = e^{2i\pi k/n}, \qquad k = 0, \ldots, n-1.$$

Answer. Suppose that $z^n = 1$. From Exercise 1.5.3 we know that if $z = r(\cos\theta + i\theta)$ the equality $z^n = 1$ can be written as

$$r^n(\cos n\theta + i\sin n\theta) = 1.$$

This means, by the uniqueness of the polar form, that $r^{1/n} = 1$ (so $r = 1$) and $\cos n\theta = 1$, $\sin n\theta = 0$. Thus $n\theta = 2k\pi$ for $k \in \mathbb{Z}$. As $e^{2i\pi k/n} = e^{2i\pi(k-mn)/n}$ for all $m \in \mathbb{Z}$, the unique solutions can be parametrized by

$$\omega_k = e^{2i\pi k/n}, \qquad k = 0, \ldots, n-1.$$

(1.5.5) Solve the equation $(z+1)^5 = z^5$.

Answer. Dividing both sides by z^5 (note that $z \neq 0$, since for $z = 0$ we get $0 = 1$ so $z = 0$ cannot satisfy the equation) we get

$$\left(1 + \frac{1}{z}\right)^5 = 1.$$

The expression in brackets cannot be 1, so we are left with the four non-trivial fifth roots of unity:

$$1 + \frac{1}{z} = e^{2\pi ik/5}, \quad k = 1, 2, 3, 4.$$

Hence we get four solutions, namely

$$z = \frac{1}{e^{2\pi ik/5} - 1}, \quad k = 1, 2, 3, 4.$$

(1.5.6) Show that $e^{z+w} = e^z e^w$ by using the series expansion.

Answer.

$$\begin{aligned} e^{z+w} &= \sum_{k=0}^{\infty} \frac{(z+w)^k}{k!} = \sum_{k=0}^{\infty}\sum_{j=0}^{k} \binom{k}{j} \frac{1}{k!} z^j w^{k-j} \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{k} \frac{1}{j!}\frac{1}{(k-j)!} z^j w^{k-j} \\ &= \sum_{j=0}^{\infty}\sum_{k=j}^{\infty} \frac{1}{j!}\frac{1}{(k-j)!} z^j w^{k-j} \\ &= \sum_{j=0}^{\infty}\sum_{r=0}^{\infty} \frac{1}{j!}\frac{1}{r!} z^j w^r \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} z^j\, e^w = e^z e^w. \end{aligned}$$

Formally, we are using Dominated Convergence (or the fact that the series converges uniformly) to justify the exchange of summation order. In light of Exercise 1.5.7 below, we will not worry about it for now.

(1.5.7) Show that $e^{z+w} = e^z e^w$ by showing that e^{z+w} is the unique solution of the initial value problem $y'(z) = y(z)$, $y(0) = e^w$.

Answer. Let $g(z) = e^{z+w}$. We have $g(0) = e^w$, $g'(z) = g(z)$. Then $g(z) = ce^z$ with $c = g(0) = e^w$ (guaranteed by Exercise 1.4.11, which works with complex scalars too). Hence

$$e^{w+z} = g(z) = e^w e^z.$$

1.6. Cardinality

(1.6.1) Show that equipotency is an equivalence relation.

Answer. We define the relation $A \sim B$ if there exists a bijection $f : A \to B$. Given a set A, the identity $\mathrm{id}_A : A \to A$ is a bijection, so the relation is reflexive. If $A \sim B$ and $f : A \to B$ is a bijection, then $f^{-1} : B \to A$ is a bijection, and so $B \sim A$, making the relation symmetric. Finally, if $f : A \to B$ and $g : B \to C$ are bijections, then $g \circ f : A \to C$ is also a bijection by Exercise 1.1.6.

(1.6.2) Let $n < m$ be positive integers. Show by induction that there is no bijection between $\{1, \dots, n\}$ and $\{1, \dots, m\}$.

Answer. We proceed by induction on m. When $m = 2$, any function $\gamma : \{1\} \to \{1, 2\}$ will clearly not be surjective. Assume as inductive hypothesis that there is no bijection between $\{1, \dots, n\}$ and $\{1, \dots, m\}$ for all $n < m$. Suppose that $\gamma : \{1, \dots, n+1\} \to \{1, \dots, m+1\}$ is bijective, where $n < m$. Because reordering is a bijection, we may assume without loss of generality that $\gamma(n+1) = m+1$. This means that γ restricts to a bijection $\{1, \dots, n\} \to \{1, \dots, m\}$ a contradiction.

(1.6.3) Write an explicit bijection $\gamma : (0, 1) \to (0, 1) \cup (1, 2)$.

Answer. We can map $(0, 1)$ to $(0, 2)$ and then use the countable shift idea to hide the middle point 1. So $\gamma : (0, 1) \to (0, 1) \cup (0, 2)$ given by

$$\gamma(t) = \begin{cases} \frac{1}{n+1}, & t = \frac{1}{2n},\ n \in \mathbb{N} \\ 2t, & t \notin \{1, \frac{1}{2}, \frac{1}{3}, \dots\} \end{cases}$$

is a bijection as desired.

(1.6.4) Write γ^{-1} explicitly for Example 1.6.5.

Answer. We need to identify whether the positive rational q is in an interval of the form $(2n, 2n+1]$ or $(2n+1, 2n+2]$. Let $m(q) = \lfloor q \rfloor$. Then we put

$$\eta(q) = \begin{cases} q - \frac{m(q)}{2}, & m(q) \text{ even} \\ \frac{m(q)-1}{2} + 1 - q, & m(q) \text{ odd} \end{cases}$$

(1.6.5) Write in detail the proof of Proposition 1.6.27.

Answer. Let $\gamma : \mathcal{P}(A) \to \{0,1\}^A$ be given by $\gamma(B) = 1_B$. If $\gamma(B) = \gamma(C)$, then $1_B = 1_C$. So if $b \in B$ then $1_C(b) = 1_B(b) = 1$, so $b \in C$; and, similarly, $C \subset B$. So $B = C$ and γ is injective. Let $g : A \to \{0,1\}$ be a function, and let $B = \{a \in A : \ g(a) = 1\}$. Then $\gamma(B) = 1_B = g$ and hence γ is surjective. Thus γ is bijective and $|\{0,1\}^A| = |\mathcal{P}(A)|$.

(1.6.6) Let A be a set and $A = \bigcup_{j\in J} A_j$, with $\{A_j\}$ nonempty and pairwise disjoint. Fix sets $\{B_a\}_{a\in A}$. Show that

$$|\prod_{a\in A} B_a| = |\prod_{j\in J}\prod_{a\in A_j} B_a|.$$

Answer. Let $g \in \prod_{j\in J}\prod_{a\in A_j} B_a$. Then, for each $j \in J$, $g(j) : A_j \to \bigcup_{a\in A_j} B_a$ with $g(j)(a) \in B_a$. Define $\tilde{g} \in \prod_{a\in A} B_a$ by $\tilde{g}(a) = g(j)(a)$ where j is the unique index such that $a \in A_j$. If $\tilde{g} = \tilde{h}$, then $g(j)(a) = h(j)(a)$ for $a \in A_j$, and so $g(j) = h(j)$, and so $g = h$. That is, the assignment $g \longmapsto \tilde{g}$ is injective. And given $h \in \prod_{a\in A} B_a$, define $h_0 \in \prod_{j\in J}\prod_{a\in A_j} B_a$ by $h_0(j)(a) = h(a)$. Then $h = \tilde{h_0}$ and the assignment is surjective, thus $|\prod_{a\in A} B_a| = |\prod_{j\in J}\prod_{a\in A_j} B_a|$.

(1.6.7) Let $A_1, A_2, \ldots$ be countable (finite or not), with $|A_n| \geq 2$ for all n. Show that

$$\left|\prod_{\mathbb{N}} A_n\right| = \left|\prod_{\mathbb{N}} \{0,1\}\right|.$$

Answer. Since each A_n has at least two elements, we have injections $\gamma_n : \{0,1\} \to A_n$. Then we can map each $g \in \prod_{\mathbb{N}}\{0,1\}$ to $\tilde{g} \in \prod_{\mathbb{N}} A_n$ by $\tilde{g}(n) = \gamma_n(g(n))$. The assignment $g \longmapsto \tilde{g}$ is clearly injective, and so $|\prod_{\mathbb{N}}\{0,1\}| \leq |\prod_{\mathbb{N}} A_n|$. Conversely, let $\{B_n\}_{n\in\mathbb{N}}$ be a pairwise disjoint family of subsets of $\mathbb{N}$ with $|B_n| = |\mathbb{N}|$ for each n; for instance we can fix the sequence of prime numbers $\{p_n\}$ and define $B_n = \{p_n^k : k \in \mathbb{N}\}$. Then $|A_n| = |\mathbb{N}| < |\prod_{\mathbb{N}}\{0,1\}| = |\prod_{k\in B_n}\{0,1\}$. And

$$\left|\prod_{n\in\mathbb{N}} A_n\right| \leq \left|\prod_{n\in\mathbb{N}}\prod_{k\in B_n}\{0,1\}\right| = \left|\prod_{n}\{0,1\}\right|.$$

Having shown both injections, Schröder–Bernstein gives us the equality

$$\left|\prod_{\mathbb{N}} A_n\right| = \left|\prod_{\mathbb{N}}\{0,1\}\right|.$$

(1.6.8) Show that, in the proof of Proposition 1.6.32, $g : X \times \{0,1\} \to X$ is a bijection.

Answer. Let $(x,t),(y,s) \in X \times \{0,1\}$ with $g(x,t) = g(y,s)$. Because $X = \bigcup_j X_j$, there exists j such that $x, y \in X_j$ (find a j for each of x and y and then choose the largest of both). Then $g_j(x,t) = g(x,t) = g(y,s) = g_j(y,s)$ and then $x = y$ and $t = s$ by the injectivity of g_j. Given $z \in X$, there exists j with $z \in X_j$. As g_j is surjective, there exist $x \in X$ and $t \in \{0,1\}$ with $g_j(x,t) = z$. Then $g(x,t) = g_j(x,t) = z$ and g is surjective. Being both injective and surjective, g is bijective.

(1.6.9) Show that, in the proof of Proposition 1.6.33, $g : X \times X \to X$ is a bijection.

Answer. If $g(x_1,x_2) = g(z_1,z_2)$ for $x_1,x_2,z_1,z_2 \in X$, by construction of X there exists j such that $x_1,x_2,z_1,z_2 \in X_j$ (choose a j for each, and then keep the largest). Then $g_j(x_1,x_2) = g(x_1,x_2) = g(z_1,z_2) = g_j(z_1,z_2)$ and $(x_1,x_2) = (z_1,z_2)$ by the injectivity of g_j. Hence g is injective. Given $z \in X$, there exists j with $z \in X_j$. By the surjectivity of g_j there exists $(x_1,x_2) \in X\times X$ with $g_j(x_1,x_2) = z$. Then $g(x_1,x_2) = g_j(x_1,x_2) = z$ and g is surjective. Being both injective and surjective, g is bijective.

1.7. Linear Algebra

(1.7.1) Let $\mathbb{F}$ be a field. Show that $0\alpha = 0$ for all $\alpha \in \mathbb{F}$.

Answer. We have
$$0\alpha = (0+0)\alpha = 0\alpha + 0\alpha.$$
If we now add $-(0\alpha)$ to both sides, we get $0 = 0\alpha$.

(1.7.2) Let $\mathbb{F}$ be a field, $\alpha, \beta \in \mathbb{F}$. Show that

(i) $(-\alpha)\beta = \alpha(-\beta) = -(\alpha\beta)$;

(ii) $-(-\alpha) = \alpha$;

(iii) $(-1)\alpha = -\alpha$.

Answer.

(i) We have $(-\alpha)\beta + \alpha\beta = (-\alpha+\alpha)\beta = 0\beta = 0$. As additive inverses are unique, $(-\alpha)\beta = -(\alpha\beta)$. Now $\alpha(-\beta) = (-\beta)\alpha = -(\beta\alpha) = -(\alpha\beta)$.

(ii) This is just the definition of $-\alpha$. We can see the equality $\alpha + (-\alpha) = 0$ as saying that $-\alpha$ is the additive inverse of α, but also as saying that α is the additive inverse of $-\alpha$.

(iii) This follows from the above: $(-1)\alpha = -(1\alpha) = -\alpha$.

(1.7.3) Prove Proposition 1.7.5.

Answer.

(i) If W is a subspace, then $\alpha v + w \in W$ for all $\alpha \in \mathbb{F}$ and $v, w \in W$. Conversely, taking $\alpha = 1$ we get that $v + w \in W$ for all $v, w \in W$, so the operation of addition is defined on W. We have $0 = (-1)v + v \in W$, and $-v = (-1)v + 0 \in W$; as associativity and commutativity are inherited from V, $(W, +)$ is an abelian group, hence a subspace.

(ii) If $W_j \subset V$ is a subspace for all $j \in J$, then given $\alpha \in \mathbb{F}$ and $v, w \in \bigcap_j W_j$ by the first part of the exercise we have that $\alpha v + w \in W_j$ for each j, and so $\alpha v + w \in \bigcap_j W_j$; hence $\bigcap_j W_j$ is a subspace of V. In particular when $W \subset V$ is a subset we can take the family of all subspaces of V that contain W (this family is non-empty, because it contains V) and so the intersection is the smallest subspace that contains W.

(iii) Let

$$W_0 = \left\{ \sum_{j=1}^{m} \alpha_j w_j : \ m \in \mathbb{N}, \ \alpha_1, \dots, \alpha_m \in \mathbb{F}, \ w_1, \dots, w_m \in W \right\}.$$

A linear combination of linear combinations is a linear combination, so W_0 is a subspace; hence $\operatorname{span} W \subset W_0$ by definition. At the same time, the elements of W_0 belong to any subspace that contains W, so $W_0 \subset \operatorname{span} W$ and hence $\operatorname{span} W = W_0$.

(1.7.4) Let $\{p_n\}_{n=0}^{\infty} \subset \mathbb{F}[x]$ such that $\deg p_n = n$ for all n. Show that $\{p_n\}$ is a basis for $\mathbb{F}[x]$.

Answer. Since $\deg p_0 = 0$, we have $p_0 = \alpha \in \mathbb{F} \setminus \{0\}$. Then $\mathbb{F}p_0 = \mathbb{F}$. Assume for induction that $\operatorname{span}\{p_0, \dots, p_k\} = \operatorname{span}\{1, x, \dots, x^k\}$. We have $p_{k+1} = \alpha x^{k+1} + q(x)$, where $\alpha \neq 0$ and $\deg q \leq k$. By the inductive hypothesis $q \in \operatorname{span}\{p_0, \dots, p_k\}$; so $x^{k+1} = \alpha^{-1}(p_{k+1} - q) \in \operatorname{span}\{p_0, \dots, p_{k+1}\}$. By induction, $\operatorname{span}\{p_0, p_1, \dots\} = \mathbb{F}[x]$.

It remains to show that $\{p_n\}$ is linearly independent. Suppose that $\alpha_1 p_{n_1} + \cdots + \alpha_k p_{n_k} = 0$, where $n_1 > n_2 > \cdots > n_k$. The monomial of highest degree in the expression is $\alpha_1 x^{n_1}$; so $\alpha_1 = 0$. But then the monomial of highest degree is $\alpha_2 x^{n_2}$, forcing $\alpha_2 = 0$. This can be repeated until obtaining that $\alpha_j = 0$ for all j, and so $p_{n_1}, \dots, p_{n_k}$ are linearly independent. As they were arbitrary elements in $\{p_n\}$, we have shown that $\{p_n\}$ is linearly independent.

(1.7.5) Show that if V, W are vector spaces and $\phi : V \to W$ is bijective and linear, then ϕ^{-1} is linear.

Answer. Let $\alpha \in \mathbb{F}$, $w_1, w_2 \in W$. Since ϕ is surjective, there exist $v_1, v_2 \in V$ such that $\phi(v_1) = w_1$, $\phi(v_2) = w_2$. As $\phi(\alpha v_1 + v_2) = \alpha\phi(v_1) + \phi(v_2) =$

$\alpha w_1 + w_2$, we have that

$$\phi^{-1}(\alpha w_1 + w_2) = \alpha v_1 + v_2 = \alpha\phi^{-1}(w_1) + \phi^{-1}(w_2)$$

and so ϕ^{-1} is linear.

(1.7.6) Prove Proposition 1.7.10.

Answer. If $\phi : V \to W$ is an isomorphism and X is a basis for V, then $\phi(X)$ is a basis for W. Indeed, if $v = \sum_j \alpha_j x_j$ then $\phi(v) = \sum_j \alpha_j \phi(x_j)$, so $W = \operatorname{span}\phi(X)$. And if $0 = \sum_j \alpha_j \varphi(x_j)$, then $0 = \varphi(\sum_j \alpha_j x_j)$; the injectivity of φ gives $0 = \sum_j \alpha_j x_j$, and the linear independence of X gives $\alpha_j = 0$ for all j. So $\phi(X)$ spans W and is linearly independent: a basis. And ϕ is a bijection, so $\dim W = |\phi(X)| = |X| = \dim V$.

Conversely, if $\dim V = \dim W$, fix bases $X = \{x_j\}_{j\in J}$ of V and $\{y_j\}_{j\in J}$ of W. Define $\phi : V \to W$ by

$$\phi\Big(\sum_j \alpha_j x_j\Big) = \sum_j \alpha_j y_j.$$

This is well-defined because the α_j are uniquely determined for each element of V. The fact that $\{y_j\}$ is spans W makes ϕ surjective. And if $\phi(\sum_j \alpha_j x_j) = 0$, then $\sum_j \alpha_j y_j = 0$ and $\alpha_j = 0$ for all j since $\{y_j\}$ is a basis. So ϕ is bijective.

It remains to show that ϕ is linear. Given $v_1, v_2 \in V$, by using zeros if necessary we may write $v_1 = \sum_{j=1}^n \alpha_j x_j$, $v_2 = \sum_{j=1}^n \beta_j x_j$. Then

$$\begin{aligned}\phi(\alpha v_1 + v_2) &= \phi\Big(\sum_{j=1}^n (\alpha\alpha_j + \beta_j)x_j\Big) = \sum_{j=1}^n (\alpha\alpha_j + \beta_j)y_j \\ &= \alpha\sum_{j=1}^n \alpha_j y_j + \sum_{j=1}^n \beta_j y_j = \alpha\phi(v_1) + \phi(v_2)\end{aligned}$$

and ϕ is linear.

(1.7.7) Prove Proposition 1.7.11.

Answer. If $\varphi_1, \varphi_2 : V \to \mathbb{F}$ are linear and $a \in \mathbb{F}$, then $(a\varphi_1 + \varphi_2)(v) = a\varphi_1(v) + \varphi_2(v)$, so V^* is a vector space. Given a basis $\{e_1, \ldots, e_n\}$ of V, define

$$e_k^*\Big(\sum_{j=1}^n c_j e_j\Big) = c_k.$$

Then $e_k^* \in V^*$. Given any $\varphi \in V^*$ and $x \in V$, we have

$$\varphi(x) = \varphi\Big(\sum_{j=1}^n c_j e_j\Big) = \sum_{j=1}^n \varphi(e_j)\, c_j = \sum_{j=1}^n \varphi(e_j)\, e_j^*(x).$$

So $\varphi = \sum_{j=1}^n \varphi(e_j)\, e_j^*$, showing that $V^* = \operatorname{span}\{e_1^*, \ldots, e_n^*\}$. And if $\sum_j a_j e_j^* = 0$, evaluating at e_k we get $a_k = 0$. So $\{e_1^*, \ldots, e_n^*\}$ are linearly independent and hence a basis for V^*. Which also shows that $\dim V^* = \dim V = n$.

(1.7.8) Let V, W be a finite-dimensional vector spaces with $\dim W = \dim V$ and $\phi : V \to V$ linear. Show that ϕ is injective if and only if it is surjective.

Answer. Suppose that ϕ is injective, and let $\{e_1, \ldots, e_n\}$ be a basis for V. If $\alpha_1\phi(e_1) + \cdots + \alpha_n\phi(e_n) = 0$, then

$$0 = \alpha_1\phi(e_1) + \cdots + \alpha_n\phi(e_n) = \phi(\alpha_1 e_1 + \cdots + \alpha_n e_n).$$

As ϕ is injective, we get $\alpha_1 e_1 + \cdots + \alpha_n e_n = 0$ and then by the linear independence we get $\alpha_1 = \cdots = \alpha_n = 0$. Thus $\phi(e_1), \ldots, \phi(e_n)$ are linearly independent. Being a linearly independent set with the same cardinality as a basis, it is a basis for W. Then for any $w \in W$ there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ with $w = \alpha_1\phi(e_1) + \cdots + \alpha_n\phi(e_n) = \phi(\alpha_1 e_1 + \cdots + \alpha_n e_n)$ and ϕ is surjective.

Conversely, suppose that ϕ is surjective. Then $\phi(e_1), \ldots, \phi(e_n)$ span W; as the dimension of W is n, necessarily $\phi(e_1), \ldots, \phi(e_n)$ are linearly independent (otherwise we could choose a proper linearly independent subset and then $\dim W < \dim V$, a contradiction). If $0 = \phi(\alpha_1 e_1 + \cdots + \alpha_n e_n)$ then $0 = \alpha_1\phi(e_1) + \cdots + \alpha_n\phi(e_n)$ and so $\alpha_1 = \cdots \alpha_n = 0$, showing that ϕ is injective.

(1.7.9) Prove Proposition 1.7.13.

Answer. If $\phi(v) = \lambda v$ for nonzero v, then $(\phi - \lambda I)(v) = \phi(v) - \lambda v = 0$, so $v \in \ker(\phi - \lambda I)$. These implications also work the other way: if $v \in \ker(\phi - \lambda I)$ is nonzero, then $\phi(v) = \lambda v$.

When $\ker(\phi - \lambda I) \neq \{0\}$, this means that $\phi - \lambda I$ is not injective, hence not invertible. Conversely, if $\ker(\phi - \lambda I) = \{0\}$, then $\phi - \lambda I$ is injective. By Exercise 1.7.8, $\phi - \lambda I$ is invertible.

That $\phi - \lambda I$ is invertible if and only if $\det(\phi - \lambda I) \neq 0$ is proven in Theorem A.3.3.

(1.7.10) Let V, W be vector spaces over $\mathbb{F}$ and $\phi_1, \phi_2 : V \to W$ linear. Fix bases E, F for V and W respectively, and show that

$$[\phi_1 + \phi_2]_{E,F} = [\phi_1]_{E,F} + [\phi_2]_{E,F}.$$

Answer. We have that $[\phi_1]_{E,F}$ is the matrix $\{\alpha_{kj}\}$ with

$$\phi_1(e_j) = \sum_{k=1}^{m} \alpha_{kj} f_k.$$

Similarly, $[\phi_2]_{E,F}$ is the matrix $\{\beta_{kj}\}$ with

$$\phi_2(e_j) = \sum_{k=1}^{m} \beta_{kj} f_k.$$

Then

$$(\phi_1 + \phi_2)(e_j) = \sum_{k=1}^{m} (\alpha_{kj} + \beta_{kj}) f_k.$$

Thus

$$[\phi_1 + \phi_2]_{E,F} = [\alpha_{kj} + \beta_{kj}] = \alpha + \beta = [\phi_1]_{E,F} + [\phi_2]_{E,F}.$$

(1.7.11) Let V, W, Z be vectors spaces over $\mathbb{F}$ and $\phi : V \to W$, $\psi : W \to Z$ linear maps. Fix bases $\{e_1, \dots, e_n\}$ for V, $\{f_1, \dots, f_m\}$ for W, and $\{g_1, \dots, g_p\}$ for Z. Show that

$$[\psi \circ \phi]_{E,G} = [\psi]_{F,G} [\phi]_{E,F}.$$

Answer. We have that $[\phi]_{E,F}$ is the matrix $\{\alpha_{kj}\}$ with

$$\phi(e_j) = \sum_{k=1}^{m} \alpha_{kj} f_k.$$

Similarly, $[\psi]_{F,G}$ is the matrix $\{\beta_{rs}\}$ with

$$\psi(f_s) = \sum_{r=1}^{m} \beta_{rs} g_r.$$

Then

$$(\psi \circ \phi)(e_j) = \sum_{k=1}^{m} \alpha_{kj}\, \psi(f_k) = \sum_{k=1}^{m} \alpha_{kj} \sum_{r=1}^{p} \beta_{rk} g_r = \sum_{r=1}^{p} \left(\sum_{k=1}^{m} \beta_{rk} \alpha_{kj} \right) g_r.$$

Thus if $\gamma = [\psi \circ \phi]_{E,G}$, then $\gamma = \beta\alpha$, where

$$\gamma_{rj} = \sum_{k=1}^{m} \beta_{rk}\alpha_{kj}.$$

(1.7.12) Let V be a finite-dimensional vector space and $\phi : V \to V$ linear. Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues for ϕ. Show that if $v_1, \ldots, v_n$ are eigenvectors for $\lambda_1, \ldots, \lambda_n$ respectively, then $v_1, \ldots, v_n$ are linearly independent.

Answer. We proceed by induction. A single eigenvector is linearly independent, so this is our base case. Suppose as inductive hypothesis that $n-1$ eigenvectors corresponding to distinct eigenvalues are linearly independent. If

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0, \tag{AB.1.6}$$

applying ϕ we get

$$\alpha_1\lambda_1 v_1 + \cdots + \alpha_n\lambda_n v_n = 0 \tag{AB.1.7}$$

Multiplying (AB.1.6) by λ_n and subtracting from (AB.1.7),

$$\alpha_1(\lambda_1 - \lambda_n)v_1 + \cdots + \alpha_{n-1}(\lambda_1 - \lambda_n)v_{n-1} = 0.$$

As $v_1, \ldots, v_{n-1}$ are linearly independent we get $\alpha_j(\lambda_j - \lambda_n) = 0$ for $j = 1, \ldots, n-1$. And $\lambda_j \neq \lambda_n$, so $\alpha_1 = \cdots = \alpha_{n-1} = 0$. Going back to (AB.1.6) we get $\alpha_n = 0$ and so $v_1, \ldots, v_n$ are linearly independent.

1.8. Basic Point Set Topology

(1.8.1) Let X be a metric space, $x \in X$ and $r > 0$. Show that $B_r(x)$ is open.

Answer. Fix $y \in B_r(x)$. Let $s = \frac{r - d(y,x)}{2}$. If $z \in B_s(y)$, then

$$d(x,z) \leq d(x,y) + d(y,z) = d(x,y) + \frac{r - d(x,y)}{2} = \frac{d(x,y) + r}{2} < r,$$

so $z \in B_r(x)$, showing that $B_s(y) \subset B_r(x)$; hence $B_r(x)$ is open.

(1.8.2) Let X be a metric space. Let $\{A_j\}$ be a collection of open sets, and let $\{B_k\}$ be a collection of closed sets. Show that $\bigcup_j A_j$ is open, and that $\bigcap_k B_k$ is closed.

Answer. Let $a \in \bigcup_j A_j$. Then there exists j such that $a \in A_j$. Since A_j is open, there exists $r > 0$ with $B_r(a) \subset A_j$. Then $B_r(a) \subset \bigcup_j A_j$ and, as this can be done for any $a \in \bigcup_j A_j$, we conclude that $\bigcup_j A_j$ is open.

We have

$$X \setminus \bigcap_k B_k = \Big(\bigcap_k B_k\Big)^c = \bigcup_k B_k^c,$$

a union of open sets. That is, the complement of $\bigcap_k B_k$ is open, which proves that $\bigcap_k B_k$ is closed.

(1.8.3) Find an example of a metric space X and a collection $\{A_j\}$ of open sets such that $\bigcap_j A_j$ is not open. Find also an example of a collection $\{B_k\}$ of closed sets such that $\bigcup_k B_k$ is not closed.

Answer. This can be easily done in the real line. Let $X = \mathbb{R}$ and $A_n = \big(-\frac{1}{n}, \frac{1}{n}\big)$. Then each A_n is open, but $\bigcap_n A_n = \{0\}$, which is not open. With a similar idea, let $B_n = \big[\frac{1}{n}, 1\big]$. Each B_n is closed, but $\bigcup_n B_n = (0, 1]$, which is not closed.

(1.8.4) Let X be a topological space and $V, W \subset X$ open and disjoint. Show that $\overline{V} \cap W = \varnothing$.

Answer. If $x \in W$, as W is open and $V \cap W = \varnothing$, this means that $x \notin \overline{V}$ (because x has a neighbourhood that does not touch V).

(1.8.5) Show that a metric space is normal.

Answer. Let X be a metric space and $C_1, C_2 \subset X$ be closed and disjoint. Since $C_1 \subset X \setminus C_2$ and this latter set is open, for each $x \in C_1$ there exists $r_x > 0$ such that $B_{r_x}(x) \subset X \setminus C_2$. Similarly, for each $y \in C_2$ there exists $r_y > 0$ such that $B_{r_y}(y) \subset X \setminus C_1$. Consider the open sets

$$V_1 = \bigcup_{x \in C_1} B_{r_x/2}(x), \qquad V_2 = \bigcup_{y \in C_2} B_{r_y/2}(y).$$

Then V_1, V_2 are open, and $C_1 \subset V_1$, $C_2 \subset V_2$. We will be done if we show that $V_1 \cap V_2 = \varnothing$. Let $z \in V_1 \cap V_2$. Then there exist $x \in C_1$ and $y \in C_2$ such that $z \in B_{r_x/2}(x) \cap B_{r_y/2}(y)$. Suppose that $r_y \leq r_x$ (otherwise, we exchange roles). We have

$$d(x, y) \leq d(x, z) + d(z, y) < \frac{r_x}{2} + \frac{r_y}{2} \leq r_x.$$

Thus $y \in B_{r_x}(x) \subset X \setminus C_2$, contradicting the fact that $y \in C_2$.

(1.8.6) Let X be a topological space. Show that the following statements are equivalent:

(i) X is normal;

(ii) given $K \subset V$ with K closed and V open, there exists $W \subset X$, open, with $K \subset W \subset \overline{W} \subset V$.

Answer. If X is normal and $K \subset V$ with K closed and V open, consider the disjoint closed sets K and $X \setminus V$. By hypothesis there exist disjoint open sets W and W' with $K \subset W$ and $X \setminus V \subset W'$. By Exercise 1.8.4, $\overline{W} \cap (X \setminus V) = \varnothing$, so $\overline{W} \subset V$. Thus $K \subset W \subset \overline{W} \subset V$.

Conversely, suppose that for all $K \subset V$ with K closed and V open, there exists $W \subset X$, open, with $K \subset W \subset \overline{W} \subset V$. Given C_1, C_2 closed and disjoint, we have $C_1 \subset X \setminus C_2$. Then there exists W open with $C_1 \subset W \subset \overline{W} \subset X \setminus C_2$. Hence $C_1 \subset W$ and $C_2 \subset X \setminus \overline{W}$, which are disjoint open sets.

(1.8.7) Prove Proposition 1.8.6.

Answer.

(i) A is closed if and only if $C(A) \subset A$. Assume first that A is closed. Then A^c is open. For any $b \in A^c$, there exists a neighbourhood N such that $b \in N \subset A^c$; then $(N \setminus \{b\}) \cap A = \varnothing$. Thus $C(A) \subset (A^c)^c = A$.

Conversely, if $C(A) \subset A$, then for any $b \in A^c$ we have that $b \notin C(A)$, so there exists a neighbourhood N with $b \in N \subset A^c$. So A^c is open, which shows that A is closed.

(ii) A is closed if and only if $\partial A \subset A$. Suppose first that $\partial A \not\subset A$. This means that there exists $x \in \partial A$ with $x \notin A$. So $x \in A^c$ and every neighbourhood of x touches A; this means that A^c is not open, and no neighbourhood of x is entirely contained in A^c; therefore A is not closed. Conversely, suppose that A is not closed; then A^c is not open. This means that there exists $x \in A^c$ and such that no neighbourhood of x is entirely contained in A^c. Then $x \in \partial A$ and so $\partial A \not\subset A$.

(iii) $\overline{A} = A \cup C(A)$. Let $x \in \overline{A}$. If $x \in A$, there is nothing to prove. If $x \notin A$, as $x \in \overline{A}$ every neighbourhood V of x satisfies $V \cap A \neq \varnothing$ (otherwise $x \in V \subset A^c$ and $A \subset (A^c \setminus V)^c$, contradicting that $x \in \overline{A}$). As $x \notin A$, we get that $(V \setminus \{x\}) \cap A \neq \varnothing$, which means precisely that $x \in C(A)$. This shows that $\overline{A} \subset A \cup C(A)$.

Conversely, if $x \in C(A)$ then by definition for every neighbourhood V of x, $(V \setminus \{x\}) \cap A \neq \varnothing$, and hence $V \cap A \neq \varnothing$. This allows us to construct a net $\{x_V\} \subset A$ with $x_V \to x$, and so $x \in \overline{A}$. This shows that $A \cup C(A) \subset \overline{A}$.

(iv) $\overline{A} = A \cup \partial A$. Since $C(A) = (\int A) \cup \partial A$, we have

$$A \cup \partial A = A \cup (\int A) \cup C(A) = A \cup C(A) = \overline{A}.$$

(v) $\overline{A \cup B} = \overline{A} \cup \overline{B}$. Using that $C(A \cup B) = C(A) \cup C(B)$ and the previous item, $\overline{A \cup B} = A \cup B \cup C(A) \cup C(B) = \overline{A} \cup \overline{B}$.

(vi) $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$. Using that $C(A \cap B) \subset C(A) \cap C(B)$, we have $\overline{A \cap B} \subset (A \cap B) \cup (C(A) \cap C(B)) = (A \cup C(A)) \cup (B \cup C(B)) = \overline{A} \cap \overline{B}$. The inclusion can be strict; this is easy to see if $A \cap B = \varnothing$. For example, in the real line, let $A \subset \mathbb{R}$ be the rationals, and $B \subset \mathbb{R}$ the irrationals. Then $A \cap B = \varnothing$, while $\overline{A} \cap \overline{B} = \mathbb{R}$.

(1.8.8) Let M be a separable metric space and $X \subset M$ be uncountable. Show that X has infinitely many accumulation points.

Answer. Let D be a countable dense subset of M, and let $\mathcal{B} = \{B_q(s) : s \in D,\ q \in \mathbb{Q}^+\}$ be a countable base for the topology. Form the set

$$E = \{x \in X : \exists B \in \mathcal{B},\ \mathcal{B} \cap X \text{ countable}\}.$$

For each $x \in E$, denote by B_x the corresponding ball with $B_x \cap X$ countable. As there are only countable many balls available in $\mathcal{B}$, the set $E_0 = \{B_x : x \in E\}$ is countable. Thus

$$X_0 = \bigcup_{B_x \in E_0} (B_x \cap X)$$

is countable. Since $X \setminus X_0$ is uncountable, for every $x \in X \setminus X_0$ every ball around it contains uncountably many points in X, so it is an accumulation point.

(1.8.9) Let $X = \mathbb{R}$ and $\mathcal{T} = \text{Top}\{[a,b) : a, b \in \mathbb{R},\ a < b\}$. The topological space $(\mathbb{R}, \mathcal{T})$ is called the **Sorgenfrey Line**, and $\mathcal{T}$ is called the **lower limit topology**.

(i) Show that $\mathcal{T}$ is finer than the usual topology on $\mathbb{R}$.

(ii) Show that $[a,b)$ is both open and closed for all $a < b$.

(iii) Show that $(\mathbb{R}, \mathcal{T})$ is normal.

(iv) Show that $x_n \to x$ if and only if there exists n_0 such that $x_n \geq x$ for all $n \geq n_0$, and $x_n \to x$ in the usual topology.

(v) Show that if $K \subset \mathbb{R}$ is compact then K is countable.

Answer.

(i) Any interval (a,b) can be written $(a,b) = \bigcup_n [a + \frac{1}{n}, b) \in \mathcal{T}$.

(ii) The interval $[a,b)$ is open by definition of $\mathcal{T}$. To see that it is closed, its complement

$$\mathbb{R} \setminus [a,b) = (-\infty, a) \cup [b, \infty) = \bigcup_n \left[a - n, a - \frac{1}{n}\right) \cup \bigcup_n [b, b+n)$$

is open.

(iii) Let $A, B \subset \mathbb{R}$ be disjoint closed sets. For each $a \in A$, since it is in the complement of B there exists $\tilde{a}$ such that $[a, \tilde{a}) \cap B = \varnothing$. Let $V = \bigcup_{a \in A} [a, \tilde{a})$. Then V is open, $A \subset V$, and $V \cap B = \varnothing$. Similarly, for each

$b \in B$ there exists $\tilde{b}$ with $[b, \tilde{b}) \cap A = \varnothing$ and $W = \bigcup_{b \in B}[b, \tilde{b})$ is an open set with $B \subset W$, and $W \cap A = \varnothing$. For any $a \in A$, $b \in B$, if $a < b$ then $\tilde{a} < b$ and so $[a, \tilde{a}) \cap [b, \tilde{b}) = \varnothing$, and the same happens if $b < a$. It follows that $V \cap W = \varnothing$.

(iv) Suppose that $x_n \to x$. Then for each $y > x$ we have that eventually $x_n \in [x, y)$, so $x_n \geq x$ for all big enough n. Conversely, if $x_n \to x$ in the usual topology and $x_n \geq x$ for all n (which we may assume after discarding finitely many elements in the sequence if necessary) given $V \in \mathcal{T}$ open with $x \in V$, there exists y such that $[x, y) \subset V$. Then $x_n \in [x, y) \subset V$ for all big enough n, and so $x_n \to x$.

(v) Suppose that K is compact. Fix $k \in K$. Then

$$\Big\{\big(-\infty, k - \frac{1}{n}\big) : \ n \in \mathbb{N}\Big\} \cup [k, \infty)$$

is an open cover for K; so it admits a finite subcover. This means that there exists n_0 such that $[k - \frac{1}{n}, k) \cap K = \varnothing$ for all $n \geq n_0$. In particular we can fix a rational $q_k \in [k - \frac{1}{n}, k)$. We claim that the intervals $\{(q_k, k)\}_{k \in K}$ are pairwise disjoint. Indeed, suppose that $z \in (q_x, x) \cap (q_y, y)$ for $x, y \in K$. Assume without loss of generality that $x < y$ Then $q_y < z < x < y$; this means that $x \in (q_y, y)$. But we established above that $(q_y, y) \cap K = \varnothing$. As the real line admits at most countably many disjoint intervals, K is countable.

(1.8.10) Let $(\mathbb{R}, \mathcal{T})$ be the Sorgenfrey Line and consider the topological space $(\mathbb{R}^2, \mathcal{T} \times \mathcal{T})$. This is called the **Sorgenfrey Plane**.

(i) Show that the Sorgenfrey Plane is separable.

(ii) Consider the set $Y = \{(x, -x) : \ x \in \mathbb{R}\} \subset \mathbb{R}^2$. Show that Y is not separable.

(iii) Conclude that the Sorgenfrey Plane is not metric.

(iv) Show that $Y_0 = \{(x, -x) : \ x \in \mathbb{Q}\}$ and $Y \setminus Y_0$ are closed, and use this information to show that $(\mathbb{R}, \mathcal{T})$ is not normal.

Answer.

(i) Since $(x_n, y_n) \to (x, y)$ if and only if $x_n \to x$ and $y_n \to y$ in the usual topology together with $x_n \geq x$ and $y_n \geq y$ eventually, the countable subset $\mathbb{Q}^2$ is dense.

(ii) Given $(x, -x)$ let $V_x = [x, x+1) \times [-x, -x+1)$. Then V_x is open, and $V_x \cap Y = \{x\}$; indeed, if $(y, -y) \in V_x \cap Y$, then $y \geq x$ and $-y \geq -x$, so

$y = x$. If $\{q_j\}_{j\in J}$ is dense in Y, then for each x there exists j_x such that $q_{j_x} \in V_x$; it follows that $q_{j_x} \neq q_{j_y}$ for all $y \neq x$, and so $|J| \geq |\mathbb{R}|$. That is, Y admits no countable dense subset.

(iii) In a separable metric space its subsets are separable Proposition 1.8.5, so $(\mathbb{R}, \mathcal{T})$ cannot be metric as Y is not separable.

(iv) Let $x, y \in \mathbb{R}$ such that $(x, y) \notin Y_0$. Suppose first that $y \neq -x$. This means that the Euclidean distance from (x, y) to Y_0 is positive. That is, there exists $\delta > 0$ such that $\sqrt{(x-q)^2 + (y+q)^2} \geq \delta$ for all $q \in \mathbb{Q}$. Let $V = [x, x+\frac{\delta}{2}) \times [y, y+\frac{\delta}{2})$. Then $(x, y) \in V$ and $V \cap Y_0 = \varnothing$; for if $(q, -q) \in V$, then $|x-q| < \frac{\delta}{2}$ and $|y+q| < \frac{\delta}{2}$ and so $(x-q)^2+(y+q)^2 < \frac{\delta^2}{2}$. The second possibility is that $y = -x$. In such case we put $V = [x, x+1) \times [-x, -x+1)$ and V is open, $(x, y) \in V$ and $V \cap Y_0 = \varnothing$. We have shown that $\mathbb{R}^2 \setminus Y_0$ is open, so Y_0 is closed. The same proof shows that $Y \setminus Y_0$ is closed.

The closed subsets Y_0 and $Y \setminus Y_0$ cannot be separated. Fix $(q, -q) \in Y_0$ and V open with $(q, -q) \in V$. Then there exists a sequence $\{r_n\}$ such that $r_n \in \mathbb{R} \setminus \mathbb{Q}$, $r_n \geq q$ for all n, and $r_n \to q$. Then $(r_n, -r_n) \to (q, -q)$, so eventually $(r_n, -r_n) \in V$. That is, $V \cap Y \setminus Y_0 \neq \varnothing$. So Y_0 and $Y \setminus Y_0$ cannot be separated, and $(\mathbb{R}, \mathcal{T})$ is not normal.

(1.8.11) Show that a discrete compact topological space is finite.

Answer. Suppose that T is discrete and compact. Since T is discrete, the family $\big\{\{t\}\big\}_{t\in T}$ is an open cover for T. By compactness, there exist $t_1, \ldots, t_n \in T$ with $T \subset \{t_1, \ldots, t_n\}$; that is, T is finite.

(1.8.12) Show that completeness is actually required in Lemma 1.8.25.

Answer. Let $X = (0, 1)$ with the usual topology. Then X is not compact, but given any $\varepsilon > 0$ it is possible to cover X with finitely many balls of radius ε. Namely, let $x_k = k\varepsilon/2$, $k = 1, \ldots, m$ with m the smallest integer greater than $2/\varepsilon$. Then $X \subset \displaystyle\bigcup_{k=1}^{m} B_\varepsilon(x_k)$.

(1.8.13) Using the ε-δ definition of continuity in a metric space, show that everywhere continuity of $f : X \to Y$ is equivalent to saying that $f^{-1}(E)$ is open in X for every open set $E \subset Y$.

Answer. Assume that f satisfies the ε-δ definition of continuity at every point. Let $E \subset Y$ be open, and consider $x \in f^{-1}(E)$. Since $f(x) \in E$ and E is open, there exists a ball surrounding $f(x)$ and inside E; that is, there exists $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subset E$. The continuity of f gives us a δ such that $|y - x| < \delta$ implies $|f(y) - f(x)| < \varepsilon$. This means that if $y \in B_\delta(x)$, then $f(y) \in B_\varepsilon(f(x)) \subset E$; that is, $y \in f^{-1}(E)$ and so $B_\delta(x) \subset E$, showing that $f^{-1}(E)$ is open since x was arbitrary.

Conversely, suppose that $f^{-1}(E)$ is open for all E open. Fix $x \in X$ and $\varepsilon > 0$. Consider the open ball $B_\varepsilon(f(x)) \subset Y$; by hypothesis, $f^{-1}(B_\varepsilon(f(x)))$ is open. Since x is a point in this open set, this means that there exists $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$. So, if $|y - x| < \delta$, then $y \in B_\delta(x)$ and so $y \in f^{-1}(B_\varepsilon(f(x)))$, which is to say that $f(y) \in B_\varepsilon(f(x))$. This is precisely $|f(y) - f(x)| < \varepsilon$.

(1.8.14) Show that $f : X \to Y$ is continuous if and only if $f^{-1}(E)$ is open for every E in a subbase for X.

Answer. If f is continuous, then $f^{-1}(E)$ is open for every open set $E \subset Y$, in particular for those in a subbase. Conversely, if B is a subbase for Y, then the set B' of finite intersections of sets in B is a base for X. As $f^{-1}(E_1 \cap \cdots \cap E_m) = f^{-1}(E_1) \cap \cdots \cap f^{-1}(E_m)$ is open, we may assume without loss of generality that B is a base. Then any $V \subset Y$ open can be written as $V = \bigcup_j E_j$, with $E_j \in B$, and

$$f^{-1}(V) = f^{-1}\Big(\bigcup_j E_j\Big) = \bigcup_j f^{-1}(E_j),$$

which is open in X.

(1.8.15) Let X be a topological space and $H \subset X$ a subset. Show that 1_H is continuous if and only if H is clopen.

Answer. Assume first that 1_H is continuous. Then $H = (1_H)^{-1}(\{1\})$ is closed. We can also write

$$H = (1_H)^{-1}\Big(1 - \frac{1}{2}, 1 + \frac{1}{2}\Big)$$

so H is open.

Conversely, suppose that H is clopen. We have, for $B \subset \mathbb{C}$ open,

$$(1_H)^{-1}(B) = \begin{cases} X, & \{0,1\} \subset B \\ H, & 1 \in B,\ 0 \notin B \\ X \setminus H, & 1 \notin B,\ 0 \in B \\ \varnothing, & \{0,1\} \cap B = \varnothing \end{cases}$$

In all four cases the preimage is open (even if B is not open, though we don't need that), so 1_H is continuous.

(1.8.16) Let (X, d) be a metric space and $\{f_n\}$ a sequence of continuous functions such that $f_n \to f$ uniformly. Show that f is continuous.

Answer. Let $\varepsilon > 0$. By definition of uniform convergence, there exists n_0 such that $d(f_n(x), f(x)) < \varepsilon$ for all $n \geq n_0$ and all $x \in X$. Fix $x \in X$ and $n \geq n_0$. Then

$$\begin{aligned} d(f(x), f(y)) &\leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) \\ &< 2\varepsilon + d(f_n(x), f_n(y)). \end{aligned}$$

As f_n is continuous, there exists $\delta > 0$ such that $d(x, y) < \delta$ implies that $d(f_n(x), f_n(y)) < \varepsilon$. Then, for all y such that $d(x, y) < \delta$, we have

$$d(f(x), f(y)) < 3\varepsilon,$$

and so f is continuous.

(1.8.17) Let (M, d) and (N, d') be metric spaces and $f : M \to N$ a function. As mentioned, f is continuous at $x \in M$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d(y, x) < \delta \implies d'(f(y), f(x)) < \varepsilon$. When δ does not depend on $\mathcal{X}$, we say that f is **uniformly continuous**. Show that if M is compact then f is uniformly continuous.

Answer. Fix $\varepsilon > 0$. By the continuity of f, for each $x \in M$ there exists $\delta_x > 0$ such that $d(y, x) < \delta_x \implies d'(f(y), f(x)) < \varepsilon/2$. The balls $B_{\delta_x/2}(x)$ form an open cover for M; so there is a finite subcover, given by say $x_1, \ldots, x_m$. Let $\delta = \frac{1}{2}\min\{\delta_{x_1}, \ldots, \delta_{x_m}\}$. If $d(y,x) < \delta$, choose j so that $d(x, x_j) < \delta_{x_j}/2$. Then

$$d(y, x_j) \le d(y,x) + d(x, x_j) < \frac{1}{2}\,\delta_{x_j} + \frac{1}{2}\,\delta_{x_j} = \delta_{x_j},$$

and so $d'(f(y), f(x_j)) < \varepsilon/2$ and $d'(f(x), f(x_j)) < \varepsilon/2$. The triangle inequality then gives $d'(f(y), f(x)) < \varepsilon$; indeed,

$$d'(f(x), f(y)) \le d'(f(x), f(x_j)) + d'(f(y), f(x_j)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(1.8.18) Let $X = \{1,2,3\}$ with the topology $\{\varnothing, X, \{1\}, \{2,3\}\}$. Show that $f : X \to \mathbb{R}$ is continuous if and only if $f(2) = f(3)$.

Answer. Suppose first that $f(2) \ne f(3)$. Then $f^{-1}(f(2) - 1/2, f(2) + 1/2)$ is either $\{1,2\}$ or $\{2\}$, neither of which is open; so f is not continuous.

Now assume that $f(2) = f(3)$; name this number r. Let $V \subset \mathbb{R}$ be open. If $r \in V$, then $f^{-1}(V)$ is either $\{2,3\}$ (when $f(1) \notin V$) or X (when $f(1) \in V$); in either case, $f^{-1}(V)$ is open, and so f is continuous. And if $r \notin V$, then $f^{-1}(V)$ is either $\varnothing$ or $\{1\}$, both open.

(1.8.19) Let $a, b \in \mathbb{R}$ and $f : (a,b) \to \mathbb{R}$ be continuously differentiable. Show that f is piecewise monotonic, i.e., there exists disjoint intervals (a_k, b_k) such that $[a,b] = \bigcup_k [a_k, b_k]$, f is monotone on each $[a_k, b_k]$, and there exist intervals $(a'_k, b'_k) \subset (a_k, b_k)$ such that f is strictly monotone on (a'_k, b'_k) and constant on (a_k, a'_k) and on (b'_k, b_k).

Answer. The set $\{t : \ f'(t) \ne 0\} = (f')^{-1}(\mathbb{R} \setminus \{0\})$ is open by hypothesis. By Proposition 1.8.1, there exist disjoint intervals such that

$$\{t : \ f'(t) \ne 0\} = \bigcup_k (a'_k, b'_k).$$

Since f' is continuous and nonzero on each of (a'_k, b'_k), we conclude that f' does not change sign there and so f is monotone on each (a'_k, b'_k). On any interval contained in the complement of $V = \bigcup_k (a_k, b_k)$ we have $f' = 0$ and so f is constant on such intervals. Assuming that the endpoints are ordered

$$a'_1 < b'_1 < a'_2 < b'_2 < \cdots$$

Then we choose $a_1 = a$, $b_1 = a_2$ in between b'_1 and a'_2, and so on.

(1.8.20) Let $X = \{1, 2, 3\}$. Show that $\mathscr{T} = \{\varnothing, X, \{1, 2\}, \{2\}, \{2, 3\}\}$ is a topology, and that it is not Hausdorff.

Answer. $\mathscr{T}$ is closed under taking unions and under taking intersections, and it has X and $\varnothing$, so it is a topology, It is not Hausdorff, because the topology cannot separate 1 and 2 (nor 2 and 3).

(1.8.21) Let $X = \{1, 2, 3\}$ with the topology $\{\varnothing, X, \{1, 2\}, \{2\}, \{2, 3\}\}$. Show that $f : X \to \mathbb{R}$ is continuous if and only if f is constant.

Answer. If f is constant, then $f^{-1}(V)$ is either X or $\varnothing$, so open. Conversely, if f is continuous, fix any $\delta > 0$; then $V = f^{-1}(f(1) - \delta, f(1) + \delta)$ is open and it contains 1, so either $V = X$ or $V = \{1, 2\}$. The case $V = X$ forces $f(2) = f(3) = f(1)$, for otherwise we get a contradiction by taking δ small enough. And when $V = \{1, 2\}$ we get that $f(2) = f(1)$. A similar argument then shows that $f(2) = f(3)$. In either case, f is constant.

(1.8.22) Prove Proposition 1.8.10.

Answer. (i) $\Longrightarrow$ (ii) Trivial.

(ii) $\Longrightarrow$ (iii) Let $a \in \overline{A}$. Then there exists a net $\{a_j\} \subset A$ such that $a_j \to a$. As f is continuous at a we have $f(a) = \lim_j f(a_j) \in \overline{f(A)}$.

(iii) $\Longrightarrow$ (iv) Let $C \subset Y$ be closed, with $f^{-1}(C)$ not closed. Then $XA \setminus f^{-1}(C)$ is not open. So there exists $z \in X \setminus f^{-1}(C)$ that is not interior, meaning that there is a net $\{a_j\} \subset f^{-1}(C)$ with $a_j \to z$. Then

$$f(z) \in f(\overline{f^{-1}(C)}) \subset \overline{f(f^{-1}(C))} \subset \overline{C} = C,$$

giving us $z \in f^{-1}(C)$, a contradiction. It follows that $f^{-1}(C)$ is closed.

(iv) $\Longrightarrow$ (i) If $B \subset Y$ is open, then $f^{-1}(B)^c = f^{-1}(B^c)$ is closed by hypothesis, so $f^{-1}(B)$ is open. That is, f is continuous.

(v) $\Longrightarrow$ (i) Suppose that f is not continuous. Then there exists $V \subset Y$ open with $f^{-1}(V)$ not open. This implies that there exists $a \in f^{-1}(V)$ and a net $\{x_j\} \subset X \setminus f^{-1}(V)$ and $x_j \to a$. As $x_j \notin f^{-1}(V)$, we have that $f(x_j) \notin V$ for all j. Then $f(x_j)$ cannot converge to $f(a)$, for V is a neighbourhood of $f(a)$ with not points from the net.

(i) $\Longrightarrow$ (v) Suppose that $x_j \to x$. Let $V \subset Y$ be an open neighbourhood of $f(x)$. As $f^{-1}(V)$ is an open neighbourhood of x, there exists j_0 such that

$x_j \in f^{-1}(V)$ for all $j \geq j_0$. Then $f(x_j) \in V$ for all $j \geq j_0$. As this can be done for any open neighbourhood of $f(x)$, this shows that $f(x_k) \to f(x)$.

(1.8.23) Show that an interval $(a,b) \subset \mathbb{R}$ is connected.

Answer. Suppose that $(a,b) = V \cup W$, with V, W open and disjoint. Define $f : (a,b) \to \mathbb{R}$ by $f(x) = 1$ if $x \in V$, $f(x) = 0$ if $x \in W$. It is easy to see that f is continuous. Indeed, given any $Z \subset \mathbb{R}$ open, we have

$$f^{-1}(Z) = \begin{cases} \varnothing, & 0,1 \not\in Z \\ W, & 0 \in Z,\ 1 \not\in Z \\ V, & 0 \not\in Z,\ 1 \in Z \\ (a,b), & 0,1 \in Z \end{cases}$$

In all cases the preimage of Z is open, so f is continuous. But this contradicts the Intermediate Value Theorem. Thus necessarily one of V and W is empty, and (a,b) is connected.

Next is a different argument. Again suppose that $(a,b) = V \cup W$, with V, W open and disjoint. Fix $v \in V$ and $w \in W$. Assume without loss of generality that $v < w$ (otherwise, exchange roles). Let

$$c = \sup\{t \in (a,b) : \ (a,b) \cap [v,t) \subset V;$$

it exists because $v \in V$. Note that $c \leq w$; for otherwise we would have $w \in V$, a contradiction. Then $a < v \leq c \leq w < b$, which means that $c \in (a,b)$. We cannot have $c \in V$, because if it were there would exist $\delta > 0$ with $c + \delta \in V$ since V is open. But we cannot have $c \in W$ either; we would have $\delta > 0$ with $c - 2\delta \in W$, contradicting the definition of c.

(1.8.24) Let (X,d) be a metric space. Show the **reverse triangle inequality**

$$|d(x,y) - d(y,w)| \leq d(x,w), \qquad x,y,w \in X.$$

Answer. Using the triangle inequality,

$$d(x,y) \leq d(x,w) + d(y,w), \qquad d(y,w) \leq d(x,y) + d(x,w).$$

We can rewrite these as

$$-d(x,w) \leq d(x,y) - d(y,w) \leq d(x,w),$$

which in turn is

$$|d(x,y) - d(y,w)| \leq d(x,w).$$

(1.8.25) Let X be a complete topological space and $C \subset X$ a closed subset. Show that C is complete.

Answer. Let $\{c_j\} \subset C$ be a Cauchy net. Because X is complete, there exists $x \in X$ such that $c_j \to x$. Then $x \in \partial C$, and so by Proposition 1.8.6 $x \in C \cup \partial C = \overline{C} = C$.

(1.8.26) Let (X,d) be a metric space. Construct a completion for X in the following way. Let $\tilde{X}$ be the set of Cauchy sequences in X, and R the equivalence relation

$$(x_n)\, R\, (y_n) \iff d(x_n, y_n) \to 0.$$

On $\bar{X} = \tilde{X}/R$ define

$$d'(\,(x_n), (y_n)\,) = \lim_n d(x_n, y_n).$$

(i) Show that d' is well-defined.

(ii) Show that the map $\rho : X \to \bar{X}$ that maps x to the constant sequence (x) is isometric.

(iii) Show that $\rho(X)$ is dense in $\bar{X}$.

(iv) Show that $\bar{X}$ is complete.

Answer.

(i) First, if (x_n) and (y_n) are Cauchy, then

$$\begin{aligned}|d(x_m, y_m) - d(x_n, y_n)| &\leq |d(x_m, y_m) - d(y_m, x_n)| + |d(y_m, x_n) - d(x_n, y_n)| \\ &\leq d(x_m, x_n) + d(y_m, y_n) \to 0\end{aligned}$$

since each sequence is Cauchy. So the number sequence $(d(x_n, y_n))$ is Cauchy and hence its limit $d' = \lim_n d(x_n, y_n)$ exists.

If $d(x_n, z_n) \to 0$ and $d(y_n, w_n) \to 0$, then

$$\begin{aligned}|d(x_n, y_n) - d(z_n, w_n)| &\leq |d(x_n, y_n) - d(y_n, z_n)| + |d(y_n, z_n) - d(z_n, w_n)| \\ &\leq d(x_n, z_n) + d(y_n, w_n) \to 0.\end{aligned}$$

Thus $d'(\,(x_n), (y_n)\,) = d'(\,(z_n), (w_n)\,)$.

(ii) This is

$$d'(\rho(x),\rho(y)) = d'(\,(x),(y)\,) = \lim_n d(x,y) = d(x,y).$$

(iii) Let (x_n) be a representative in $\bar{X}$. By definition, the sequence is Cauchy. Let A_n be the constant sequence $(x_n, x_n, \ldots)$. Then

$$d'(A_m,(x_n)\,) = \lim_n d(x_m,x_n)$$

(note that the limit exists, as we proved above that d' always exists). Because (x_n) is Cauchy, for m big enough the limit can be made as small as we want. Thus $\lim_m A_m = (x_n)$, showing that $\rho(X)$ is dense in $\bar{X}$.

(iv) This one is a bit cumbersome to write because we need to deal with sequences of sequences. If (A_m) is a Cauchy sequence in $\overline{X}$, for every $r \in \mathbb{N}$, there exists n_r such that

$$d'(A_m, A_\ell) < \frac{1}{r}, \quad \text{for all } m, \ell \geq n_r.$$

Each A_m is the class of a Cauchy sequence $(A_{mn})_n \subset X$. In turn, using the definition of d', this means that there exists $\{m_r\} \subset \mathbb{N}$, with $m_r \geq m_{r-1}$ for all r, such that

$$d(A_{n_r,k}, A_{n_r+\ell,k}) < \frac{1}{r}, \quad \text{for all } k \geq m_r, \quad \text{and for all } \ell, \tag{AB.1.8}$$

and such that

$$d(A_{n_r,h}, A_{n_r,j}) < \frac{1}{r}, \qquad \text{for all } h, j \geq m_r. \tag{AB.1.9}$$

(this, because the sequence A_{n_r} is Cauchy).

Now consider the sequence $(A_{n_r,m_r})_r \subset X$. This sequence is Cauchy, since for any $s \geq r$

$$\begin{aligned} d(A_{n_r,m_r}, A_{n_s,m_s}) &\leq d(A_{n_r,m_r}, A_{n_r,m_s}) + d(A_{n_r,m_s}, A_{n_s,m_s}) \\ &\leq d(A_{n_r,m_r}, A_{n_r,m_s}) + \frac{1}{r} \\ &\leq \frac{1}{r} + \frac{1}{r} = \frac{2}{r} \end{aligned}$$

(this first estimate by (AB.1.8), and the second one by (AB.1.9).

So the sequence $(A_{n_r,m_r})_r$ is Cauchy, and $\lim_m A_m = (A_{n_r,m_r})_r$.

(1.8.27) Let (X,d) be a metric space and $\{x_n\}$ a Cauchy sequence. Show that $\{x_n\}$ is bounded; that is, there exists $x \in X$ and a ball B centered at x such that $\{x_n\} \subset B$.

Answer. Let $\varepsilon = 1$. Then there exists n_0 such that $d(x_n, x_m) < 1$ whenever $n, m \geq n_0$. Put $x = x_{n_0}$ and

$$r = 1 + \max\{1, d(x_1, x), \ldots, d(x_{n_0-1}, x)\}.$$

Then $d(x_n, x) < r$ for all n.

(1.8.28) Let X, Y be complete metric spaces with dense subsets X_0, Y_0 respectively. Let $\gamma : X_0 \to Y_0$ be an isometric surjection. Show that there exists a unique $\tilde{\gamma} : X \to Y$, bijective and isometric.

Answer. If $x \in X$, there exists $\{x_n\} \subset X_0$ with $x_n \to x$. The sequence (x_n) is Cauchy. As $d_Y(\gamma(x_n), \gamma(x_m)) = d_X(x_n, x_m)$, the sequence $(\gamma(x_n))$ is also Cauchy. We want to define $\tilde{\gamma}(x) = \lim \gamma(x_n)$. To see that this makes sense, if $x_n' \to x$, then $d_Y(\gamma(x_n), \gamma(x_n')) = d_X(x_n, x_n')$, so $\gamma(x_n') \to \tilde{\gamma}(x)$.

Next we see that $\tilde{\gamma}$ is isometric. If $x_n \to x$, $z_n \to z$, then

$$d_Y(\tilde{\gamma}(x), \tilde{\gamma}(z)) = \lim_n d_Y(\gamma(x_n), \gamma(z_n)) = \lim_n d_X(x_n, z_n) = d_X(x, z).$$

Thus $\tilde{\gamma}$ is isometric, and in particular it is injective. Finally, if $y \in Y$, there exists $(y_n) \subset Y_0$ with $y_n \to y$. Put $x_n = \gamma^{-1}(y_n)$. Since γ is isometric, (x_n) is Cauchy and so there exists $x \in X$ with $x_n \to x$. Then

$$d_Y(\tilde{\gamma}(x), y) = \lim_n d_Y(\gamma(x_n), y) = \lim_n d_Y(y_n, y) = 0,$$

so $\tilde{\gamma}(x) = y$, and $\tilde{\gamma}$ is surjective.

(1.8.29) Let $X = \mathbb{Q}$ and define

$$d(x, y) = |\arctan x - \arctan y|.$$

Show that d is a distance, and find the completion of (X, d).

Answer. We have $d(x, x) = 0$ for all x, and $d(x, y) \geq 0$ by definition. The absolute value and the difference also give us that $d(x, y) = d(y, x)$. And

$$\begin{aligned} d(x, z) &= |\arctan x - \arctan z| \\ &\leq |\arctan x - \arctan y| + |\arctan y - \arctan z| \\ &= d(x, y) + d(y, z). \end{aligned}$$

As for the completion, let $\tilde{X} = \mathbb{R} \cup \{\pm\infty\}$, where

$$\tilde{d}(x, +\infty) = \left|\frac{\pi}{2} - \arctan x\right|, \qquad \tilde{d}(x, -\infty) = \left|\arctan x + \frac{\pi}{2}\right|.$$

The arctan is bicontinuous, so $|q_n - x| \to 0$ if and only if $d(q_n, x) \to 0$. This, together with the fact that $\lim_{x\to\pm\infty} \arctan x = \pm\frac{\pi}{2}$ guarantees that $\tilde{d}$ is still a distance and that X is dense in $\tilde{X}$. So it remains to show that $\tilde{X}$ is complete. Let $\{q_n\} \subset X$ be Cauchy. If there exists $c > 0$ with $|q_n| \le c$ by the Mean Value Theorem there exists $\xi(x, y) \in [-c, c]$ with

$$|\arctan x - \arctan y| = |\arctan' \xi(x,y)|\,|x-y| = \frac{1}{1+\xi(x,y)^2}\,|x-y| \le |x-y|.$$

Thus

$$|q_n - q_m| \le d(q_n, q_m)$$

and so $\{q_n\}$ is Cauchy in the usual sense and converges to some $x \in \mathbb{R}$. The same estimate as above shows that $d(q_n, x) \to 0$.

When $\{q_n\}$ is Cauchy for the metric d but unbounded for the usual metric, the above does not work. If $\{q_n\}$ has a limit point $x \in \mathbb{R}$, we could apply the above to said subsequence and also to a subsequence that increases to infinity (or decreases to minus infinity). This would have x and $\pm\infty$ as accumulation points for the sequence on (X, d), a contradiction. If follows that $q_n \to \infty$ or $q_n \to -\infty$. In both cases, the fact that arctan converges to $\pm\frac{\pi}{2}$ at infinity implies that $q_n \to \pm\infty$ in (X, d). Hence $(\tilde{X}, \tilde{d})$ is complete.

(1.8.30) Let $X = \mathbb{R} \setminus \mathbb{Q}$, with

$$d(x,y) = \begin{cases} |x-1| + |y-1|, & x \ne y \\ 0, & x = y \end{cases}$$

Show that d is a metric, and find the completion of (X, d).

Answer. The expression of d is symmetric on x and y. And

$$|x-1| + |z-1| \le |x-1| + |y-1| + |z-1| + |y-1| = d(x,y) + d(y,z).$$

We claim that 1 is the only accumulation point in (X, d). Let $\{x_n\}$ be a Cauchy sequence. Then

$$|x_n - 1| + |x_m - 1| = d(x_n, x_m) \to 0,$$

so $x_n \to 1$ in the usual topology and also in the d topology, and $d(x_n, 1) \to 0$. So $(\tilde{X}, \tilde{d}) = (X \cup \{1\}, \tilde{d})$, where $\tilde{d}$ is defined with the same formula as d.

(1.8.31) Let X be a complete metric space, and $E_1 \supset E_2 \supset \cdots$ a decreasing sequence of closed sets, such that $\lim_n \operatorname{diam}(E_n) =$

0. Show that $\bigcap_n E_n$ is nonempty and it consists of a single point. Can the "closed" condition be removed?

Answer. Fix $x_n \in E_n$ for each $n \in \mathbb{N}$. Given $\varepsilon > 0$ there exists m such that $\operatorname{diam}(E_m) < \varepsilon/2$. Then for $k, n \geq m$ we have $x_n, x_k \in E_m$, so $d(x_n, x_k) < \varepsilon/2 < \varepsilon$; which shows that the sequence $\{x_n\}$ is Cauchy. By the completeness, there exists $x = \lim_n x_n$. For any m, since $x_n \in E_m$ for all $n \geq m$ (from $E_n \subset E_m$), we get by the closedness of E_m that $x \in E_m$. Thus $x \in \bigcap_n E_n$. If y is another element in the intersection, then $d(x, y) \leq \operatorname{diam}(E_n)$ for all n, so $d(x, y) = 0$ and $x = y$.

The closedness of the E_n is necessary. For instance consider $X = \mathbb{R}$ with the usual topology and let $E_n = \left(0, \frac{1}{n}\right)$. Then $E_1 \supset E_2 \supset \cdots$ but $\bigcap_n E_n = \varnothing$.

(1.8.32) Let X be a topological space, and let $V, W \subset X$ be disjoint open subsets. Show that $\overline{U} \cap W = \varnothing$.

Answer. Since $V \subset W^c$ and W^c is closed, we have $\overline{V} \subset W^c$. This is $\overline{V} \cap W = \varnothing$.

(1.8.33) Let X be a topological space, $E \subset X$. Show that if E is connected, then $\overline{E}$ is also connected.

Answer. Write $\overline{E} = (\overline{E} \cap V) \cup (\overline{E} \cap W)$, with V, W open and disjoint. Taking intersection with E, we get that $E = (E \cap V) \cup (E \cap W)$. As E is connected, one of the two sets is empty, say $E \cap W = \varnothing$. That is, $E \subset W^c$, which is closed. Thus $\overline{E} \subset W^c$, which we may write as $\overline{E} \cap W = \varnothing$. Hence $\overline{E}$ is connected.

(1.8.34) Let X be a topological space and $E, F \subset X$ connected. Show that if $E \cap F \neq \varnothing$, then $E \cup F$ is connected.

Answer. Suppose that $E \cup F = A \cup B$, with $A = (E \cup F) \cap V$, $B = (E \cup F) \cap W$ disjoint, and V, W open. Fix $x \in E \cap F$. As $A \cap B = \varnothing$, either $x \in A$ or $x \in B$. Without loss of generality, assume that $x \in A$. So $x \in V$. We may write $E = (E \cap V) \cup (E \cap W)$. These two sets are relatively open; as E is connected, one of them is empty; and as $x \in V$, we get that $E = E \cap V$, $E \cap W = \varnothing$. We may do the same for F, so $F \cap W = \varnothing$. Thus $B = (E \cap W) \cup (F \cap W) = \varnothing$, and $E \cup F$ is connected.

(1.8.35) Let X be a topological space. For each x, denote by E_x a maximal connected subset with $x \in E_x$. Define a relation by $x \sim y$ if $y \in E_x$. Show that $\sim$ is an equivalence relation.

Answer. **Reflexive:** $x \in E_x$ by definition, so $x \sim x$.

Symmetric: Suppose that $x \sim y$. Then $y \in E_x$. We also have $y \in E_y$. By Exercise 1.8.34, the set $E_x \cup E_y$ is connected; the maximality of E_y then shows that $E_x \cup E_y = E_y$, so $E_x \subset E_y$. Now the maximality of E_x gives us that $E_x = E_y$. So $x \in E_x = E_y$, showing that $y \sim x$.

Transitive: If $x \sim y$ and $y \sim z$, then by the above $E_x = E_y = E_z$; and so $x \sim z$.

(1.8.36) Show that a path-connected space is connected.

Answer. Suppose that X is not connected. Then $X = V \cup W$, with V, W open, nonempty, and $V \cap W = \varnothing$. Let $v \in V$, $w \in W$. Let $f : [0,1] \to X$ be continuous, with $f(0) = v$ and $f(1) = w$. We get

$$[0,1] = f^{-1}(X) = f^{-1}(V \cup W) = f^{-1}(V) \cup f^{-1}(W).$$

As $[0,1]$ is connected, we get that either $f^{-1}(V)$ or $f^{-1}(W)$ is empty; but this contradicts the fact that $f(0) = v$ (so $v \in f^{-1}(V)$) and $f(1) = w$ (so $w \in f^{-1}(W)$).

(1.8.37) Let X, Y be topological spaces, $f : X \to Y$ continuous, and $K \subset X$ compact. Show that $f(K)$ is compact.

Answer. Let $\{V_j\}$ be an open cover of $f(K)$. Then

$$K \subset f^{-1}(f(K)) \subset f^{-1}\left(\bigcup_j V_j\right) = \bigcup_j f^{-1}(V_j).$$

As f is continuous, each $f^{-1}(V_j)$ is open, so we have an open cover of K. By the compactness of K there exist $j_1, \ldots, j_m$ such that $K \subset f^{-1}(V_{j_1}) \cup \cdots \cup f^{-1}(V_m)$. Then, as images preserve unions (Proposition 1.1.1),

$$f(K) \subset \bigcup_{k=1}^m V_{j_k}.$$

Hence $f(K)$ admits a finite subcover and so it is compact.

(1.8.38) Let X be compact Hausdorff, Y a Hausdorff topological space, and $\psi : X \to Y$ continuous. Show that if ψ is injective, then ψ is a homeomorphism onto $\psi(X)$.

Answer. By hypothesis $\psi : X \to \psi(X)$ is a continuous bijection. So all we need to address is the continuity of ψ^{-1}. Let $X_0 \subset X$ be closed. As X is compact, X_0 is compact (Lemma 1.8.16). By Exercise 1.8.37, $f(X_0)$ is compact, and by Lemma 1.8.16 $f(X_0)$ is closed. We have shown that f maps closed sets to closed sets, which means that the pre-images of closed sets by f^{-1} are closed. Then f^{-1} is continuous by Proposition 1.8.10.

(1.8.39) Use Exercise 1.8.38 to show that if X is compact Hausdorff, any weaker topology on X is not Hausdorff, and any stronger topology on X is Hausdorff but not compact.

Answer. Let $\mathcal{T}_1$ be the topology on X, and $\mathcal{T}_2 \subset \mathcal{T}_1$. If $\mathcal{T}_2$ is Hausdorff, then the identity function $\mathrm{id} : (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ is continuous and Exercise 1.8.38 implies that $\mathcal{T}_2 = \mathcal{T}_1$. Similarly, if $\mathcal{T}_3 \supset \mathcal{T}_1$ it is necessarily Hausdorff by the fact that it contains $\mathcal{T}_1$; and if it is compact, then by Exercise 1.8.38, applied to $\mathrm{id} : (X, \mathcal{T}_3) \to (X, \mathcal{T}_1)$ we get $\mathcal{T}_3 = \mathcal{T}_1$.

(1.8.40) Let X, Y be topological spaces and $f : X \to Y$ continuous. If $X_0 \subset X$ is such that $\overline{X_0}$ is compact, show that $f(\overline{X_0}) = \overline{f(X_0)}$.

Answer. We have $f(\overline{X_0}) \subset \overline{f(X_0)}$ by Proposition 1.8.10. Conversely, let $y \in \overline{f(X_0)}$. Then there exists a net $\{x_j\} \subset X_0$ with $f(x_j) \to y$. By the compactness of $\overline{X_0}$ and Proposition 1.8.19, there exists a convergent subnet $\{x_{j_k}\}$, say $x_{j_k} \to x \in \overline{X_0}$. Then

$$y = \lim_k f(x_{j_k}) = f(\lim_k x_{j_k}) = f(x),$$

so $\overline{f(X_0)} \subset f(\overline{X_0})$.

(1.8.41) Show that the set $\mathbb{R}\cup\{-\infty,\infty\}$ can be given a topology such that it is a compactification of $\mathbb{R}$.

Answer. We mimic the proof of Proposition 1.8.27. On $T_{\pm\infty}=\mathbb{R}\cup\{-\infty,\infty\}$ we consider the topology

$$\mathcal{T}=\operatorname{Top}\Big\{\{V\subset\mathbb{R},\ \ \text{open}\}\cup\{(\mathbb{R}\setminus K)\cup N: \\ K\subset\mathbb{R}\ \ \text{compact},\ N\subset\{\infty,-\infty\}\setminus\{\varnothing\}\}\Big\}.$$

The open sets that do not contain $\pm\infty$ are precisely the open sets in $\mathbb{R}$, so this topology restricts to the usual topology on $\mathbb{R}$. Given an open cover $\{V_j\}$ of $\mathbb{R}_{\pm\infty}$, since ∞ is covered there has to exist an open set of the form $(\mathbb{R}\setminus K_1)\cup\{\infty\}$ on the cover; and similarly there exists K_2 compact with $(\mathbb{R}\setminus K_2)\cup\{-\infty\}$ in the cover. If we now let $\{V_j\}$ consist of all open sets in the cover with the two points $\pm\infty$ removed, we have that $\{V_j\}$ is an open cover for $K_1\cap K_2$. Hence there exist $j_1,\ldots,j_m$ such that $K_1\cap K_2\subset\bigcup_{k=1}^m V_{j_k}$. This implies that $\mathbb{R}_{\pm\infty}=V_{j_1}\cup\cdots\cup V_{j_m}\cup\big((\mathbb{R}\setminus K_1)\cup\{\infty\}\big((\mathbb{R}\setminus K_1)\cup\{-\infty\}$, and this gives us a finite subcover. Thus $\mathbb{R}_{\pm\infty}$ is compact.

(1.8.42) Let S,T be homeomorphic topological spaces such that both are locally compact Hausdorff . Show that S_∞ is homeomorphic to T_∞.

Answer. We have a homeomorphism $\gamma:S\to T$. We extend it as $\tilde\gamma(\infty)=\infty$, and we need to show that the extension is still a homeomorphism. We already have that it is bijective, and as the domain S_∞ is compact, by Exercise 1.8.38 we just need to show that $\tilde\gamma$ is continuous. Exercise 1.8.14 tells us that we can test continuity of the elements of a subbase, so we can use open subsets of T or complements of compacts together with infinity. The first possibility is that $V\subset T$ is open. In that case, $\tilde\gamma^{-1}(V)=\gamma^{-1}(V)$ is open by the continuity of γ. The second possibility is that $V=(T\setminus K)\cup\{\infty\}$ for some $K\subset T$ compact. Then, since preimages preserve all set operations and $\tilde\gamma(\infty)=\infty$,

$$\tilde\gamma^{-1}(V)=(S\setminus\gamma^{-1}(K))\cup\{\infty\}.$$

Since γ is a homeomorphism, $\gamma^{-1}(K)$ is compact, and then $\tilde\gamma^{-1}(V)$ is open in S_∞. Thus $\tilde\gamma$ is continuous and hence a homeomorphism.

(1.8.43) Let T be a locally compact Hausdorff space. Let R and S be one-point compactifications of T, that is $R = T \cup \{\infty_R\}$ is a compact Hausdorff space such that the restriction to T agrees with the topology of T, and similarly for S. Show that R and S are homeomorphic; that is, the one-point compactification is unique.

Answer. Since $R = T \cup \{\infty_R\}$ and $S = T \cup \{\infty_S\}$, we have an obvious bijection γ between the two sets, that is the identity on T. Since the domain is compact, it is enough to show that γ is continuous to conclude that it is a homeomorphism. Let $V \subset S$ be open. If $V \subset T$, then $\gamma^{-1}(V) = V$ is open in R. If $V \not\subset T$, then $\infty_S \in V$. Because points are closed (due to S being Hausdorff), $T = R \setminus \{\infty_S\}$ is open on S. Let $V_0 = V \cap T$, which is open in S; we have $K = T \setminus V_0 = S \setminus V$ is closed, and hence compact in S; but K is entirely inside T, where both topologies agree, so K is compact in T. Then $V = \big(S \setminus K\big) \cup \{\infty_S\}$, and as $\gamma^{-1}(K) = K$,

$$\gamma^{-1}(V) = (T \setminus K) \cup \{\infty_R\},$$

open. Hence γ is continuous, and thus a homeomorphism.

(1.8.44) Show that the one-point compactification $\mathbb{R}_\infty$ of $\mathbb{R}$ is homeomorphic to the unit circle $\mathbb{T}$.

Answer. By Exercise 1.8.42 it is enough to show that $\mathbb{R}$ is homeomorphic to $\mathbb{T} \setminus \{1\}$. The set $\mathbb{T} \setminus \{1\}$ is homeomorphic to $\mathbb{T} \setminus \{-1\}$. We have

$$\mathbb{T} \setminus \{-1\} = \{e^{it} :\ t \in (-\pi, \pi)\}.$$

So the exponential provides a homeomorphism between $\mathbb{T}\setminus\{-1\}$ and $(-\pi, \pi)$. We showed in Example 1.6.8 that $\mathbb{R}$ is homeomorphic (as the inverse tangent function is bicontinuous) to the interval $\big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$. Then, with $\simeq$ denoting homeomorphism, we have

$$\mathbb{R} \simeq \Big(-\frac{\pi}{2}, \frac{\pi}{2}\Big) \simeq (-\pi, \pi) \simeq \mathbb{T} \setminus \{-1\} \simeq \mathbb{T} \setminus \{1\}.$$

By Exercise 1.8.43 the one-point compactification of $\mathbb{T} \setminus \{1\}$ is $\mathbb{T}$, and hence $\mathbb{R}_\infty \simeq \mathbb{T}$ by Exercise 1.8.42.

(1.8.45) Let S, T be topological spaces and $f : S \to T$ be continuous and surjective and such that f is not surjective when restricted to any proper closed subset of S. Let $U \subset S$ be open. Show that $f(U) \subset \overline{T \setminus f(S \setminus U)}$.

Answer. Fix $t \in f(U)$ and let $V \subset T$ be an open neighbourhood of t. Since $W = U \cap f^{-1}(V)$ is open and nonempty, its complement $S \setminus W$ is a proper closed subset of S; by hypothesis there exists $z \in T \setminus f(S \setminus W)$. As f is surjective we have $T = f(W) \cup f(S \setminus W)$, so $z = f(w)$ for some $w \in W$. Then $z = f(w) \in f(U) \cap V$. Hence $z \in V \cap (T \setminus f(S \setminus U))$ (since $T \setminus f(S \setminus W) \subset T \setminus f(S \setminus U)$). We have shown that any neighbourhood of t touches $T \setminus f(S \setminus U)$, and so $f(U) \subset \overline{T \setminus f(S \setminus U)}$.

(1.8.46) Let X be a separable metric space, and $V \subset X$ open. Show that V is a countable union of balls.

Answer. By hypothesis there exists $D \subset X$, dense. Since V is open, for each $x \in V$ there exists an open ball $B_{r_x}(x) \subset V$. Hence

$$V = \bigcup_{x \in V} B_{r_x}(x).$$

Because D is dense, for each $x \in V$ there exists $z_x \in D$ with $d(x, z_x) < r_x/2$. Fix $q_x \in \mathbb{Q} \cap (0, 1)$ with $q_x < r_x/2$. This guarantees that

$$x \in B_{q_x}(z_x) \subset B_{r_x}(x).$$

Therefore

$$V = \bigcup_{x \in V} B_{q_x}(z_x).$$

As there are countably many q_x (being rational) and countably many z_x (being elements of D), only countably many balls appear in the union.

CHAPTER 2

Measure and Integration

2.1. Motivation

(2.1.1) Show that the integral in (2.1) exists if f is continuous with the possible exception of finitely many jump discontinuities.

Answer. Because f has finitely many jump discontinuities, it is bounded. Let $t_1, \ldots, t_m$ be the points where f has discontinuities. For each n, let

$$k(n,j) = \left\lfloor \frac{n(t_j - a)}{b-a} \right\rfloor.$$

Then

$$a + \frac{k(n,j)(b-a)}{n} \leq t_j < a + \frac{(k(n,j)+1)(b-a)}{n}.$$

As f is continuous on (t_{j-1}, t_j), we get from (2.1) and Exercise 1.3.8

$$\int_{t_{j-1}}^{t_j} f(t)\,dt = \lim_{n\to\infty} \sum_{k=k(n,j-1)+1}^{k(n,j)-1} f\Big(a + \frac{k(b-a)}{n}\Big)\Delta_k.$$

So

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{k=1}^{n} f\Big(a + \frac{k(b-a)}{n}\Big)\,\Delta_k - \frac{1}{n}\sum_{j=1}^{m} f\Big(a + \frac{k(n,j)(b-a)}{n}\Big)\,\Delta_k$$
$$= \lim_{n\to\infty} \sum_{k=1}^{n} f\Big(a + \frac{k(b-a)}{n}\Big)\,\Delta_k.$$

2.2. The Cantor Set

(2.2.1) Show that for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, 2^{n-1}\}$, the interval $C_{n,k}$ is of the form $\left(\frac{r}{3^n}, \frac{r+1}{3^n}\right)$, with neither r nor $r+1$ multiples of 3.

Answer. Each interval $C_{n+1,k}$ is of the form $\left(\frac{r(n,k)}{3^n}, \frac{r(n,k)+1}{3^n}\right)$. It is obtained as the middle third an interval of the form $\left(\frac{a}{3^n}, \frac{a+1}{3^n}\right)$, with $a \in \mathbb{N}$. Thus

$$\frac{r(n+1,k)}{3^{n+1}} = \frac{a}{3^n} + \frac{1}{3^{n+1}}$$

and so

$$r(n+1,k) = 3a+1,$$

not a multiple of 3. And neither is $r(n+1,k)+1 = 3a+2$.

(2.2.2) Let $t \in [0,2]$.

(i) Show that there exist $a, b \in \mathcal{C}$ such that $t = a + b$.

(ii) Find $a, b \in \mathcal{C}$, expressed as fractions, such that $a + b = 1$.

(iii) Are such a, b unique? If they are not, find another suitable pair a, b.

Answer.

(i) Since t is 2 times an element of $[0,1]$, we may write

$$t = \sum_{k=1}^{\infty} \frac{2a_k}{3^k},$$

with $a_k \in \{0,1,2\}$ for all k. Let $U = \{k : a_k = 1\}$, and put

$$a = \sum_{k\in U} \frac{2a_k}{3^k} + \sum_{k\notin U} \frac{a_k}{3^k}, \qquad b = \sum_{k\notin U} \frac{a_k}{3^k}.$$

Then $a + b = t$, and by Proposition 2.2.1 we have that $a, b \in \mathcal{C}$.

(ii) In ternary, $1 = 0.2\cdots_3$. So we may take for instance

$$a = 0.20202020\cdots_3\,, \qquad b = 0.020202\cdots_3\,,$$

which are in $\mathcal{C}$ by Proposition 2.2.1. Noting that $a = 3b$ (since in base 3 it is multiplication by 3 that "moves the period to the right") and $a+b = 1$, we immediately determine that $a = \frac{3}{4}$, $b = \frac{1}{4}$. Or we can go the hard way and calculate

$$b = \sum_{k=1}^{\infty} \frac{2}{9^n} = 2\,\frac{\frac{1}{9}}{1-\frac{1}{9}} = \frac{2}{8} = \frac{1}{4}.$$

(iii) There are infinitely many suitable a, b. We can "pass" any part of the expansion to the other. For instance we can take $\frac{2}{3}$ off a (that would be the first 2 in the expansion) and put it in b: we get

$$a' = 0.00202020\cdots_3\,, \qquad b = 0.22020202\cdots_3\,.$$

That is,

$$a' = a - \frac{2}{3} = \frac{3}{4} - \frac{2}{3} = \frac{1}{12}, \qquad b' = b + \frac{2}{3} = \frac{1}{4} + \frac{2}{3} = \frac{11}{12}.$$

An even simpler observation is that in this particular case $\frac{1}{3} + \frac{2}{3} = 1$, and both numbers are in $\mathcal{C}$; so that's another possible choice. There are infinitely many choices, as there are infinitely many ways to shift some ternary digit from a to b. One particular example is

$$a_n = \frac{3}{4} - \frac{2}{3^n}, \qquad b_n = \frac{1}{4} + \frac{2}{3^n};$$

these satisfy $a_n, b_n \in \mathcal{C}$, $a_n + b_n = 1$.

(2.2.3) Complete the details the proof of Proposition 2.2.2. That is, justify why if

$$|s - t| = \left|\sum_n \frac{a_n}{3^n} - \sum_n \frac{b_n}{3^n}\right| < \frac{1}{3^m}$$

with $a_n, b_n \in \{0,2\}$, then $a_j = b_j$ for $j = 1, \ldots, m-1$.

Answer. By hypothesis we have that $|a_j - b_j| \leq 2$ for all j. Suppose that $a_j = b_j$ for $j = 1, \ldots, k-1$, $|a_k - b_k| = 2$, and $k < m$. Then

$$\begin{aligned}\Big|\sum_n \frac{a_n}{3^n} - \sum_n \frac{b_n}{3^n}\Big| &= \Big|\frac{a_k - b_k}{3^k} + \sum_{n>k} \frac{a_n - b_n}{3^n}\Big| \geq \frac{|a_k - b_k|}{3^k} - \Big|\sum_{n>k} \frac{a_n - b_n}{3^n}\Big| \\ &\geq \frac{2}{3^k} - \sum_{n>k} \frac{|a_n - b_n|}{3^n} \geq \frac{2}{3^k} - \sum_{n>k} \frac{2}{3^n} \\ &= \frac{2}{3^k} - 2\,\frac{\frac{1}{3^{k+1}}}{1-\frac{1}{3}} = \frac{1}{3^k} \\ &> \frac{1}{3^m},\end{aligned}$$

a contradiction.

(2.2.4) Show that the dyadic numbers in $[0,1]$ are dense.

Answer. Let $t \in [0,1]$. If we write t in binary, we have $t = \sum_{k=1}^\infty \frac{t_k}{2^k}$. Since the series converges, given $\varepsilon > 0$ there exists k_0 such that $\sum_{k=k_0+1}^\infty \frac{t_k}{2^k} < \varepsilon$. If $s = \sum_{k=1}^{k_0} \frac{t_k}{2^k}$, then

$$s = \sum_{k=1}^{k_0} \frac{t_k}{2^k} = \frac{\sum_{k=1}^{k_0} 2^{k_0-k} t_k}{2^{k_0}}$$

is dyadic and $|t - s| = |\sum_{k=k_0+1}^\infty \frac{t_k}{2^k}| < \varepsilon$.

(2.2.5) Consider the function β from Proposition 2.2.2, and recall the notation $C_{n,k}$ for the removed intervals in the construction of $\mathcal{C}$.

(i) Show that the right endpoints of all the removed intervals $C_{n,k}$ are those numbers in $[0,1]$ such that their ternary expansion is finite, without any 1, and ends in 2.

(ii) Show that for any two endpoints of a removed interval $C_{n,k} = (a,b)$, we have $\beta(a) = \beta(b)$ and that this is a dyadic number.

(iii) Show that if $a, b \in \mathcal{C}$ are distinct and $\beta(a) = \beta(b)$, then there exist n, k such that $C_{n,k} = (a,b)$.

(iv) Conclude that if E is the set of endpoints of the removed intervals, then β is injective on $\mathcal{C} \setminus E$.

(v) Conclude that $\beta^{-1}(\{t\})$ is a singleton if t is not dyadic, and that it consists of two points when t is dyadic.

Answer.

(i) We proceed by induction on n. When $n = 1$, $C_{1,1} = (\frac{1}{3}, \frac{2}{3})$, and $\frac{2}{3} = 0.2_3$. Now assume that the right endpoint of $C_{n,k}$ is of the form $b_{n,k} = 0.b_1 \cdots b_r 2_3$, with $r < n$. When we remove the middle third $C_{n+1,j} = (s,t)$ immediately to the right of $C_{n,k}$, the length of this middle third is 3^{-n-1}, and it will be situated $2 \times 3^{-n-1}$ to the right of the endpoint $0.b_1 \cdots b_r 2_3$, with $r \leq n-1$. Thus

$$t = 0.b_1 \cdots b_r 2_3 + \frac{2}{3^{n+1}} = 0.b_1 \cdots b_r 2_3 + 0.\overbrace{0 \cdots 0}^{n} 2_3 = 0.b_1 \cdots b_r 2 \overbrace{0 \cdots 0}^{n-r-1} 2_3,$$

which completes the induction. For each n there are precisely 2^{n-1} intervals $C_{n,k}$. And that's also the precise amount of numbers in $\mathcal{C}$ that finish with a 2 in the n^{th} position. So every such number has to be an endpoint.

(ii) We know that the right endpoint of $C_{n,k} = (a,b)$ is

$$b = 0.b_1 \cdots b_r 2 \overbrace{0 \cdots 0}^{n-r-1} 2_3.$$

The left endpoint is 3^{-n} units to the left, that is

$$a = 0.b_1 \cdots b_r 2 \overbrace{0 \cdots 0}^{n-r-1} 2_3 - 3^{-n} = 0.b_1 \cdots b_r 2 \overbrace{0 \cdots 0}^{n-r-1} 2_3 - 0.\overbrace{0 \cdots 0}^{n-1} 1_3$$
$$= 0.b_1 \cdots b_r 2 \overbrace{0 \cdots 0}^{n-r-1} 1_3 = 0.b_1 \cdots b_r 2 \overbrace{0 \cdots 0}^{n-r-1} 022 \cdots_3 .$$

Now, using b_j' to denote $b_j/2$,

$$\beta(a) = 0.b_1' \cdots b_r' 1 \overbrace{0 \cdots 0}^{n-r-1} 011 \cdots_2 = 0.b_1' \cdots b_r' 1 \overbrace{0 \cdots 0}^{n-r-1} 1_2,$$

and

$$\beta(b) = 0.b_1' \cdots b_r' 1 \overbrace{0 \cdots 0}^{n-r-1} 1_2 = \beta(a).$$

Since dyadic numbers are those with a finite expansion in base 2, $\beta(a) = \beta(b)$ is dyadic.

(iii) Write $a = \sum_{k=1}^{\infty} \frac{a_k}{3^k}$, $b = \sum_{k=1}^{\infty} \frac{b_k}{3^k}$, with $a_k, b_k \in \{0,2\}$ for all k. Assume $a < b$. By hypothesis we have that

$$\sum_{k=r}^{\infty} \frac{a_k'}{2^k} = \sum_{k=r}^{\infty} \frac{b_k'}{2^k}$$

(still using the notation $a'_j = a_j/2$), where r is the smallest index such that $a_r \neq b_r$. Necessarily, from $a < b$, $a'_r = 0$, $b'_r = 1$. Then

$$\frac{1}{2^r} = \frac{b'_r - a'_r}{2^r} = \sum_{k=r+1}^{\infty} \frac{a'_k - b'_k}{2^k}.$$

The right-hand-side is at most

$$\sum_{k=r+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^r},$$

which forces $a'_k - b'_k = 1$ for all $k > r$. That is, $a'_k = 1$, $b'_k = 0$ for all $k > r$. Thus

$$a = 0.a_1 \cdots a_{r-1}0222 \cdots_3\,, \qquad b = 0.a_1 \cdots a_{r-1}2_3.$$

By part (i), the interval (a, b) is one of the $C_{n,k}$.

(iv) By part (iii), β has to be injective on $\mathcal{C} \setminus E$, for the equality $\beta(a) = \beta(b)$ implies that $a, b \in E$.

(v) In base 2, dyadic numbers are those with a finite binary expansion, which in our convention translates to those that finish with $0111 \cdots$. So if $t = 0.t'_1 \cdots t'_r 0111 \cdots_2$, then $t = \beta(a)$, where $a = 0.t_1 \cdots t_r 0222 \cdots_3$ (and still denoting $t_k = 2t'_j$). By (i), this means that $a \in E$. Combined with (ii), this gives us that t is dyadic if and only if $\beta^{-1}(\{t\})$ consists of the two endpoints of a $C_{n,k}$. By part (iii), β is injective on $\mathcal{C} \setminus E$, so $\beta^{-1}(\{t\})$ is a singleton when t is not dyadic.

(2.2.6) Show that $\mathcal{C}$ has no isolated points.

Answer. Fix $t \in \mathcal{C}$. If we denote the removed middle thirds by $C_{n,k}$, with $k = 1, \ldots, 2^{n-1}$ and $m(C_{n,k}) = 3^{-n}$, then for each n there exists $k(n)$ such that t is in between $C_{n,k(n)}$ and $C_{n,k(n)+1}$. As the endpoints of each $C_{n,k}$ are in $\mathcal{C}$, this guarantees that there exists $t_n \in \mathcal{C}$ with $|t - t_n| < 3^{-n}$. It follows that there exists $\{t_n\} \subset \mathcal{C}$ with $t_n \to t$.

(2.2.7) Let $s, t \in \mathcal{C}$ with $s < t$. Show that there exists $b \in [0, 1] \setminus \mathcal{C}$ with $s < b < t$. This shows that $\mathcal{C}$ is **totally disconnected**.

Answer. We may write

$$s = \sum_{k=1}^{\infty} \frac{a_n}{3^n}, \qquad t = \sum_{k=1}^{\infty} \frac{b_n}{3^n},$$

with $a_n, b_n \in \{0, 2\}$ for all n. Since $s < t$, there exists a minimum index r such that $a_r = 0$, $b_r = 2$ and $a_k = b_k$ for $k = 1, \dots, r-1$. Let

$$b = \sum_{k=1}^{r-1} \frac{a_k}{3^n} + \frac{1}{3^r} + \frac{1}{3^{r+1}}.$$

Then $b \notin \mathcal{C}$ because its ternary expansion has a non-terminating 1, and $s < t < b$ by construction.

(2.2.8) Suppose that you create a set with a similar idea as the Cantor set, you start with the unit interval $[0, 1]$ but instead of removing in each step middle intervals of measure 3^{-n}, you remove middle intervals of measure 4^{-n}. Discuss the set $\mathcal{D}$ you obtained. Does it contain any intervals? What properties are the same as in the Cantor set, and what properties are different?

Answer. The intervals we remove are

$$\left(\frac{3}{8}, \frac{5}{8}\right), \left(\frac{5}{32}, \frac{7}{32}\right), \left(\frac{25}{32}, \frac{27}{32}\right),$$

etc. In each step we are removing 2^{n-1} intervals, each of length 4^{-n}. Thus

$$\mathcal{D} = [0, 1] \setminus \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{2^{n-1}} D_{n,k},$$

where each $D_{n,k}$ is an interval of length 4^{-n}. Thus

$$\begin{aligned} m(\mathcal{D}) &= 1 - \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} m(D_{n,k}) = 1 - \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \frac{1}{4^n} \\ &= 1 - \sum_{n=1}^{\infty} \frac{2^{n-1}}{4^n} = 1 - \frac{1}{2} \sum_{n=1}^{\infty} 2^{-n} = \frac{1}{2}. \end{aligned}$$

This is the main difference between $\mathcal{D}$ and $\mathcal{C}$, that $\mathcal{D}$ has some "mass". Other than that,

- the set $\mathcal{D}$ is uncountable for the same reasons as $\mathcal{C}$, just working with expansions in base 4;
- it is also compact, being the complement of an open set inside $[0, 1]$;

- it has no isolated points, since in each step we are removing a middle interval, so the length of the remaining closed intervals after each step decreases by a factor of more than 2.
- $\mathcal{D}$ cannot contain an interval since we can do an analog of the argument from Exercise 2.2.6.

(2.2.9) Define a function $f : [0,1] \to [0,1]$ in the following way. Write the complement of $\mathcal{D}$ (as in Exercise 2.2.8) as $\bigcup_n \bigcup_{k=1}^{2^n} D_{n,k}$ and the complement of $\mathcal{C}$ as $\bigcup_n \bigcup_{k=1}^{2^n} C_{n,k}$. For each n, k let $f_{n,k}$ be the natural increasing bijection $D_{n,k} \to C_{n,k}$. That is, if $D_{n,k} = (a,b)$ and $C_{n,k} = (c,d)$, then $f_{n,k}(x) = c + \frac{(x-a)(d-c)}{b-a}$. Patch them together so that

$$f(x) = f_{n,k}(x), \qquad x \in D_{n,k}.$$

As the complement of the union of the $D_{n,k}$ is nowhere dense, we can extend f by continuity to get $f : [0,1] \to [0,1]$. Show that f is continuous, monotone non-decreasing, and fails the property that the preimage of a nullset is a nullset.

Answer. The function f is continuous by construction. It is monotone because it respects the order of the intervals $D_{k,n}$. And $f^{-1}(\mathcal{C}) = \mathcal{D}$, so there is a nullset whose preimage has positive measure.

(2.2.10) Show that the sequence $\{f_n\}$ defined on page 88 of the Book converges uniformly to α. This gives an alternative proof that α is continuous (and other properties, too).

Answer. We first show the uniform convergence by induction. We have $|f_0(x) - f_1(x)| \leq 1$ for all x, since $0 \leq f_0(x), f_1(x) \leq 1$. Now assume for induction that $|f_n(x) - f_{n-1}(x)| \leq 2^{-n+1}$ for all x. The trivial case is $x \in \left[\frac{1}{3}, \frac{2}{3}\right]$, since then $f_{n+1}(x) = f_n(x)$. For $x \in \left[0, \frac{1}{3}\right]$,

$$|f_{n+1}(x) - f_n(x)| = \left|\frac{1}{2}\, f_n(3x) - \frac{1}{2}\, f_{n-1}(3x)\right| \leq \frac{1}{2}\, 2^{-n+1} = 2^n.$$

And for $x \in \left[\frac{2}{3}, 1\right]$,

$$\begin{aligned} |f_{n+1}(x) - f_n(x)| &= \left|\frac{1}{2} + \frac{1}{2} f_n(3x-2) - \frac{1}{2} - \frac{1}{2} f_{n-1}(3x-2)\right| \\ &= \left|\frac{1}{2} f_n(3x-2) - \frac{1}{2} f_{n-1}(3x-2)\right| \le \frac{1}{2} 2^{-n+1} = 2^n. \end{aligned}$$

So by induction we have shown that $|f_{n+1}(x) - f_n(x)| \le 2^{-n}$ for all x and all n. This implies that the sequence is (uniformly) Cauchy. Indeed, by telescoping we get

$$|f_{n+k}(x) - f_n(x)| \le \sum_{j=0}^{k-1} |f_{n+j+1}(x) - f_{n+j}(x)| \le \sum_{j=0}^{k-1} 2^{-n-j} \le 2^{-n}.$$

Let $f(x) = \lim_n f_n(x)$. Since the convergence is uniform, $f(x)$ is continuous. The function f satisfies

$$f(x) = \begin{cases} \frac{1}{2} f(3x), & x \in \left[0, \frac{1}{3}\right] \\ \frac{1}{2}, & x \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \frac{1}{2} + \frac{1}{2} f(3x-2), & x \in \left[\frac{2}{3}, 1\right] \end{cases}$$

And α also satisfies the relations

$$\alpha(x) = \begin{cases} \frac{1}{2} \alpha(3x), & x \in \left[0, \frac{1}{3}\right] \\ \frac{1}{2}, & x \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \frac{1}{2} + \frac{1}{2} \alpha(3x-2), & x \in \left[\frac{2}{3}, 1\right] \end{cases}$$

Indeed, when $x \in \left[0, \frac{1}{3}\right)$, we have $x = 0.0\mathrm{R}_3$, where R denotes the rest of the expansion. Then $3x = 0.\mathrm{R}_3$. Now $\alpha(3x) = 0.\mathrm{R'}_2$, where R′ is obtained from R by truncating at the first 1 and replacing all remaining 2 with 1. And then $\frac{1}{2}\alpha(3x) = 0.0\mathrm{R'}_2$, which is precisely $\alpha(x)$. When $x \in \left[\frac{1}{3}, \frac{2}{3}\right]$, we have that $x = 0.1\mathrm{R}_3$, and then $\alpha(x) = 0.1_2 = \frac{1}{2}$. And when $x \in \left(\frac{2}{3}, 1\right]$, now $x = 0.2\mathrm{R}_3$. Then $3x - 2 = 2.\mathrm{R}_3 - 2 = 0.\mathrm{R}_3$; so $\alpha(3x-2) = 0.\mathrm{R'}_2$ and

$$\frac{1}{2} + \frac{1}{2}\alpha(3x-2) = 0.1_2 + 0.0\mathrm{R'}_2 = 0.1\mathrm{R'}_2 = \alpha(x).$$

Let us now show that $\alpha = f$. Since $0 \le \alpha(x), f(x) \le 1$ for all $x \in [0,1]$, we have $|\alpha(x) - f(x)| \le 1$ for all x. If $x \in \left[0, \frac{1}{3}\right]$,

$$|\alpha(x) - f(x)| = \frac{1}{2} |\alpha(3x) - f(3x)|. \tag{AB.2.1}$$

If $x \in \left[\frac{1}{3}, \frac{2}{3}\right]$, then $\alpha(x) = f(x) = \frac{1}{2}$. And if $x \in \left(\frac{2}{3}, 1\right]$,

$$\begin{aligned} |\alpha(x) - f(x)| &= \left|\frac{1}{2} + \frac{1}{2}\alpha(3x-2) - \frac{1}{2} - \frac{1}{2} f(3x-2)\right| \\ &= \frac{1}{2} |\alpha(3x-2) - f(3x-2)|. \end{aligned} \tag{AB.2.2}$$

Iterating the inequalities (AB.2.1) and (AB.2.2) we obtain

$$|\alpha(x) - f(x)| \leq \frac{1}{2^n}, \qquad x \in [0,1]$$

for arbitrary n, and hence $\alpha = f$.

(2.2.11) Consider the metric space

$$\mathcal{X} = \{f : [0,1] \to [0,1], \ \text{ continuous}, \ f(0) = 0, \ f(1) = 1\},$$

with the metric

$$d(f,g) = \max\{|f(x) - g(x)| : \ x \in [0,1]\}.$$

Define $\Phi : \mathcal{X} \to \mathcal{X}$ by

$$(\Phi f)(x) = \begin{cases} \frac{1}{2} f(3x), & x \in \left[0, \frac{1}{3}\right] \\ \frac{1}{2}, & x \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \frac{1}{2} + \frac{1}{2} f(3x-2), & x \in \left[\frac{2}{3}, 1\right] \end{cases}$$

We use Φ^n to denote composition of Φ with itself n times.

(i) Show that $d(\Phi f, \Phi g) \leq \frac{1}{2} d(f,g)$ for all $f, g \in \mathcal{X}$.

(ii) Show that, for any $f \in \mathcal{X}$, the sequence $\{\Phi^n f\}$ converges.

(iii) Show that, for any $f, g \in \mathcal{X}$, $\lim_n \Phi^n f = \lim_n \Phi^n g$.

(iv) Deduce that, for any $f \in \mathcal{X}$, $\lim_n \Phi^n f = \alpha$.

Answer.

(i) When $x \in \left[\frac{1}{3}, \frac{2}{3}\right]$, we have $\Phi f(x) = \Phi g(x)$. For the other two cases,

$$|\Phi f(x) - \Phi g(x)| = \frac{1}{2} |f(3x) - g(3x)| \leq \frac{1}{2} d(f,g), \qquad x \in \left[0, \frac{1}{3}\right),$$

and

$$\begin{aligned} |\Phi f(x) - \Phi g(x)| &= \frac{1}{2} \left|1 + f(3x-2) - 1 - g(3x-2)\right| \\ &= \frac{1}{2} |f(3x-2) - g(3x-2)| \\ &\leq \frac{1}{2} d(f,g) \end{aligned}$$

for $x \in \left[\frac{2}{3}, 1\right)$. Hence $d(\Phi f, \Phi g) \leq \frac{1}{2} d(f,g)$.

(ii) The proof is the typical proof of the fixed point theorem. Note that $d(f,g) \leq 1$ for all $f, g \in \mathcal{X}$. We have

$$|\Phi^n f(x) - \Phi^{n+1} f(x)| \leq \frac{1}{2} |\{\Phi^{n-1} f(x) - \Phi^n f(x)| \leq \cdots \leq \frac{1}{2^n}.$$

So $d(\Phi^n f, \Phi^{n+1} f) \le \frac{1}{2^n}$. Then by the triangle inequality

$$d(\Phi^{n+k} f, \Phi^n f) \le \sum_{j=1}^{k-1} d(\Phi^{n+j}, \Phi^{n+j-1}) \le \sum_{j=1}^{k-1} \frac{1}{2^{n+j-1}} \le \frac{1}{2^n}.$$

Thus the sequence $\{\Phi^n f\}$ is Cauchy on $\mathcal{X}$. The space $\mathcal{X}$ is complete because a uniform limit of continuous functions is continuous, and the values at the endpoints will be unchanged. It follows that $\lim_n \Phi^n f \in \mathcal{X}$ exists.

(iii) This is the uniqueness in the fixed point theorem. Note that we proved above that

$$d(\Phi f, \Phi g) \le d(f, g).$$

This means that Φ is continuous. Let $\tilde{f} = \lim_n \Phi^n f$, $\tilde{g} = \lim_n \Phi^n g$. Then $\Phi\tilde{f} = \tilde{f}$, $\Phi\tilde{g} = \tilde{g}$. Iterating,

$$d(\tilde{f}, \tilde{g}) = d(\Phi\tilde{f}, \Phi\tilde{g}) \le \frac{1}{2} d(\tilde{f}, \tilde{g}) = \frac{1}{2} d(\Phi\tilde{f}, \Phi\tilde{g}) \le \frac{1}{4} d(\tilde{f}, \tilde{g}) \le \cdots \le \frac{1}{2^n}.$$

As this can be done for any $n \in \mathbb{N}$, it follows that $\tilde{f} = \tilde{g}$.

(iv) Given $f \in \mathcal{X}$, let $\gamma = \lim_n \Phi^n f$. Then $\Phi\gamma = \gamma$. This guarantees that γ satisfies the relations

$$\gamma(x) = \begin{cases} \frac{1}{2}\gamma(3x), & x \in \left[0, \frac{1}{3}\right] \\ \frac{1}{2}, & x \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \frac{1}{2} + \frac{1}{2}\gamma(3x-2), & x \in \left[\frac{2}{3}, 1\right] \end{cases}$$

The computation in the answer to Exercise 2.2.10 shows that any $f \in \mathcal{X}$ satisfying the recursive relation above equals α. Thus $\gamma = \alpha$.

(2.2.12) Consider α as the fixed point in Exercise 2.2.11; that is do not use the other equivalent definitions.

(i) Let $x \in [0, 1]$, written as $\sum_{k=1}^{\infty} \frac{a_k}{3^k}$. Put $m = \min\{k : a_k = 1\}$ when $x \notin \mathcal{C}$, and $m = \infty$ when $x \in \mathcal{C}$. Show that

$$\alpha(x) = \sum_{k=1}^{m-1} \frac{a_k}{2^{k+1}} + \frac{1}{2^m} \qquad x \in [0, 1]. \tag{2.7}$$

This works even when $m = \infty$, if we interpret $\frac{1}{2^\infty} = 0$.

(ii) Show that $\alpha(1 - x) = 1 - \alpha(x)$ for all $x \in [0, 1]$.

Answer.

(i) We proceed by induction on m. When $m = 1$, we have $x \in [1/3, 2/3]$, $\alpha(x) = \frac{1}{2}$, and the (2.7) holds. Assume as inductive hypothesis that (2.7) holds for m.

Suppose first that $x \in [0, 1/3)$, and a_{m+1} is the leftmost 1. Then $a_1 = 0$ and we have

$$\begin{aligned}\alpha(x) &= \frac{1}{2}\,\alpha(3x) = \frac{1}{2}\,\alpha\Big(\sum_{k=2}^{\infty}\frac{3a_k}{3^k}\Big)\\ &= \frac{1}{2}\,\alpha\Big(\sum_{k=1}^{\infty}\frac{a_{k+1}}{3^k}\Big) = \frac{1}{2}\,\alpha\Big(\sum_{k=1}^{m-1}\frac{a_{k+1}}{2^{k+1}} + \frac{1}{2^m}\Big)\\ &= \frac{1}{2}\,\alpha\Big(\sum_{k=0}^{m-1}\frac{a_{k+1}}{2^{k+1}} + \frac{1}{2^m}\Big) = \sum_{k=1}^{m}\frac{a_k}{2^{k+1}} + \frac{1}{2^{m+1}}.\end{aligned}$$

Similarly, when $x \in (2/3, 1]$ we have $a_1 = 2$ and

$$\begin{aligned}\alpha(x) &= \frac{1}{2} + \frac{1}{2}\,\alpha(3x-2) = \frac{1}{2} + \frac{1}{2}\,\alpha\Big(2 + \sum_{k=2}^{\infty}\frac{3a_k}{3^{k-1}} - 2\Big)\\ &= \frac{1}{2} + \frac{1}{2}\,\alpha\Big(\sum_{k=1}^{\infty}\frac{a_{k+1}}{3^k}\Big) = \frac{1}{2} + \frac{1}{2}\,\alpha\Big(\sum_{k=1}^{m-1}\frac{a_{k+1}}{2^{k+1}} + \frac{1}{2^m}\Big)\\ &= \frac{1}{2} + \frac{1}{2}\,\alpha\Big(\sum_{k=2}^{m-1}\frac{a_k}{2^k} + \frac{1}{2^m}\Big) = \frac{1}{2} + \sum_{k=2}^{m}\frac{a_k}{2^{k+1}} + \frac{1}{2^{m+1}}\\ &= \sum_{k=1}^{m}\frac{a_k}{2^{k+1}} + \frac{1}{2^{m+1}}.\end{aligned}$$

This completes the induction. When $m = \infty$, that is when $x \in \mathcal{C}$, we may write

$$x = \lim_{m\to\infty}\sum_{k=1}^{m}\frac{a_k}{3^k} + \frac{1}{3^{m+1}}.$$

As α is continuous,

$$\alpha(x) = \lim_{m\to\infty}\sum_{k=1}^{m}\frac{a_k}{2^{k+1}} + \frac{1}{2^{m+1}} = \sum_{k=1}^{\infty}\frac{a_k}{2^{k+1}}.$$

(ii) We write $x = \sum_{k=1}^{\infty}\frac{a_k}{3^k}$. Since $1 = \sum_{k=1}^{\infty}\frac{2}{3^k}$, we have that

$$1 - x = \sum_{k=1}^{\infty}\frac{2-a_k}{3^k}.$$

If m is the least index such that $a_k = 1$ when $x \notin \mathcal{C}$, it is clear that m is also the least index such that $2 - a_k = 1$.

Then

$$\begin{aligned}\alpha(x) + \alpha(1-x) &= \sum_{k=1}^{m-1} \frac{a_k}{2^{k+1}} + \frac{1}{2^m} + \sum_{k=1}^{m-1} \frac{2-a_k}{2^{k+1}} + \frac{1}{2^m} \\ &= \sum_{k=1}^{m-1} \frac{1}{2^k} + \frac{1}{2^{m-1}} = \frac{\frac{1}{2} - \frac{1}{2^m}}{1 - \frac{1}{2}} + \frac{1}{2^{m-1}} \\ &= 1 - \frac{1}{2^{m-1}} + \frac{1}{2^{m-1}} = 1.\end{aligned}$$

(2.2.13) Let $\alpha : [0,1] \to [0,1]$ be the Cantor function. We saw that α is surjective and continuous. By Exercise 1.1.6 it admits a right-inverse. Show that such right-inverse cannot possibly be continuous.

Answer. Let $h : [0,1] \to [0,1]$ such that $\alpha \circ h = \mathrm{id}$. By Exercise 1.1.6, h is injective. Also, from α non-decreasing we get that h is non-decreasing, for if $h(s) > h(t)$ then

$$s = \alpha(h(s)) \geq \alpha(h(t)) = t;$$

hence if $s < t$ then $h(s) \leq h(t)$. As h is bounded and monotone, its side limits exist. Let $\ell = \sup\{h(t) : t < \frac{1}{2}\}$ and $r = \inf\{h(t) : t > \frac{1}{2}\}$ be the left and right limits at $\frac{1}{2}$. If $t < \frac{1}{2}$, then $h(t) < \frac{1}{3}$, for if $h(t) \geq \frac{1}{3}$ then $t = \alpha(h(t)) \geq \alpha(\frac{1}{3}) = \frac{1}{2}$. It follows that $\ell \leq \frac{1}{3}$. Similarly, if $t > \frac{1}{2}$ we have $h(t) > \frac{2}{3}$, for if $h(t) \leq \frac{2}{3}$, then $t = \alpha(h(t)) \leq \alpha(\frac{2}{3}) = \frac{1}{2}$; hence $r \geq \frac{2}{3}$. As the side limits do not agree, h is not continuous at $t = \frac{1}{2}$. Although not needed here, the same argument shows that h fails to be continuous at every dyadic number in $[0,1]$.

2.3. Measures and Lebesgue Measure

(2.3.1) Let X be a set. Show that $(X, \mathcal{P}(X), \mu)$ is a measure space, where μ is the counting measure, given by

$$\mu(S) = \begin{cases} |S|, & \text{if } S \text{ is finite} \\ \infty, & \text{if } S \text{ is infinite} \end{cases}$$

and $|S|$ denotes the number of elements in S.

Answer. Since $\mathcal{P}(X)$ contains all subsets of X, it is a σ-algebra. We have $\mu(\varnothing) = 0$ since $|\varnothing| = 0$. And if $\{A_n\}$ are pairwise disjoint subsets of X, we first remove empty sets from the list—so possibly the list becomes finite, and we write $A_1, \ldots, A_s$ with $s \in \mathbb{N}$ or $s = \infty$—and we consider two cases:

- if $|A_m| = \infty$ for some m, then $|\bigcup_n A_n| = \infty$ and
$$\mu\Big(\bigcup_n A_n\Big) = \infty = \mu(A_m) = \sum_n \mu(A_n);$$
- When all A_n are finite, since we have countably many finite disjoint sets, we may consider them as disjoint subsets of $\mathbb{N}$. Establish bijections $\mu_n : A_n \to \{1, \ldots, |A_n|\}$, and put $k_1 = 0$, $k_n = \sum_{j=1}^{n-1} |A_n|$. Write $A_n = \{a_{n,1}, \ldots, a_{n,r_n}\}$. So $r_n = |A_n|$. Put
$$\gamma(a_{n,k}) = k_n + k.$$
Then $\gamma : \bigcup_n A_n \to \{1, \ldots, \sum_n r_n\}$ is a bijection, since
$$k_n \leq \gamma(a_{n,k}) \leq k_n + r_n < k_{n+1}, \qquad n \in \{1, \ldots, s\},\ k \in \{1, \ldots, r_n\}.$$
This shows that
$$\mu\Big(\bigcup_n A_n\Big) = \sum_n r_n = \sum_n \mu(A_n).$$

(2.3.2) Let X be a set and let

$$\mathcal{A} = \{A \subset X : \quad A \text{ is countable or } A^c \text{ is countable}\}.$$

(i) Show that $\mathcal{A}$ is a σ-algebra.

(ii) Show that $\mathcal{A}$ is the σ-algebra generated by the singletons (the family of all subsets of A consisting of a single element).

(iii) Show that $\mathcal{A} = \mathcal{P}(X)$ if and only if X is countable.

Answer. The empty set is countable, so $\varnothing \in \mathcal{A}$. The definition of $\mathcal{A}$ is symmetric on A and A^c, so $\mathcal{A}$ contains complements. If $\{A_k\} \subset \mathcal{A}$ is countable, we want to show that $\bigcup_k A_k$ is either countable or has countable complement. If A_k is countable for all k, then $\bigcup_k A_k$ is countable and thus in $\mathcal{A}$. If at least one of the A_k is not countable, say A_j, then A_j^c is countable and

$$\Big(\bigcup_k A_k\Big)^c = \bigcap_k A_k^c \subset A_j^c$$

is countable.

Let $\mathcal{S}$ be the σ-algebra generated by the singletons. As $\{x\}$ is obviously countable for all $x \in X$, we have $\mathcal{S} \subset \mathcal{A}$. Conversely, if $A \in \mathcal{A}$ is countable, then $A = \bigcup_{a\in A}\{a\} \in \mathcal{S}$; and similarly, if A^c is countable then $A^c \in \mathcal{S}$ and so $A \in \mathcal{S}$. This shows that $\mathcal{A} \subset \mathcal{S}$, and so $\mathcal{S} = \mathcal{A}$.

If X is countable, then every $A \in \mathcal{P}(X)$ is countable, and so it is in $\mathcal{A}$. Conversely, if X is uncountable then it can be partitioned into two uncountable disjoint subsets, $X = X_0 \cup X_1$, $X_0 \cap X_1 = \varnothing$. Then $X_0 \notin \mathcal{A}$, and so $\mathcal{A} \subsetneq \mathcal{P}(X)$.

(2.3.3) Consider a set X and $\mathcal{A}$ as in Exercise 2.3.2. Let

$$\mu(A) = \begin{cases} 0, & A \text{ countable} \\ 1, & \text{otherwise} \end{cases}$$

Show that $(X, \mathcal{A}, \mu)$ is a measure space.

Answer. Since $\varnothing$ is countable, $\mu(\varnothing) = 0$. Now suppose that $\{A_k\} \subset \mathcal{A}$ are pairwise disjoint. If all A_k are countable, then $\bigcup_k A_k$ is countable and

$$\mu\Big(\bigcup_k A_j) = 0 = \sum_k \mu(A_k).$$

If at least one A_j is uncountable with A_j^c countable, then $\bigcup_k A_k$ is uncountable, and $\bigcup_{k\neq j} A_k \subset A_j^c$ is countable. So $\mu(A_k) = 0$ if $k \neq j$. Thus

$$\mu\Big(\bigcup_k A_j\Big) = 1 = \sum_k \mu(A_k).$$

(2.3.4) Let $\mathcal{A}$ be a σ-algebra. Show $|\mathcal{A}| \neq |\mathbb{N}|$. *(Hint: if $\mathcal{A}$ is countably infinite, consider for each $x \in X$ the smallest set in $\mathcal{A}$ that contains x)*

Answer. Assume that $\mathcal{A}$ is countable. For each $x \in X$, let

$$S_x = \bigcap \{A \in \mathcal{A} : \ x \in A\}.$$

Because $\mathcal{A}$ is countable, each intersection uses at most countably many sets, and so $S_x \in \mathcal{A}$.

Given $x, y \in \mathcal{A}$, suppose that $x \notin S_y$. Then $x \in S_x \setminus S_y \in \mathcal{A}$. Because S_x is the smallest set in $\mathcal{A}$ that contains x, this implies that $S_x \setminus S_y = S_x$; this in turn is equivalent to $S_x \cap S_y = \varnothing$. And if $x \in S_y$, then $x \in S_x \cap S_y$, which by the minimality of S_x implies that $S_x \cap S_y = S_x$; this means that $S_x = S_y$. We have shown that either $S_x = S_y$ or $S_x \cap S_y = \varnothing$.

As $\mathcal{A}$ is infinite, the family $\{S_x : \ x \in X\}$ has to be infinite (otherwise, $\mathcal{A}$ would be finite, as the S_x are minimal in $\mathcal{A}$ and so every element of $\mathcal{A}$ is a union of some S_x). Let $\{x_n\} \subset X$ be chosen so that $\{S_{x_n} : \ n\}$ is infinite. Now consider the map $\Phi : \mathcal{P}(\mathbb{N}) \to \mathcal{A}$, given by

$$\Phi(N) = \bigcup_{n \in N} S_{x_n}.$$

Because the sets $\{S_{x_n}\}_n$ are pairwise disjoint, the function Φ is injective. Thus $|\mathcal{A}| \geq |\mathcal{P}(\mathbb{N})| > |\mathbb{N}|$, a contradiction.

(2.3.5) Let $(X, \mathcal{A})$ be a measurable space. Define

$$\mu(A) = \begin{cases} 0, & A \ \text{finite} \\ \infty, & A \ \text{infinite} \end{cases}$$

Show that μ is always additive, and discuss when it is σ-additive.

Answer. Let $A_1, \ldots, A_m \in \mathcal{A}$. If all m sets are finite, then so is $\bigcup_j A_j$ and

$$\mu\Big(\bigcup_{j=1}^m A_j\Big) = 0 = \sum_{j=1}^m \mu(A_j).$$

If at least one of the sets is infinite, then so is the union and we have

$$\mu\Big(\bigcup_{j=1}^m A_j\Big) = \infty = \sum_{j=1}^m \mu(A_j).$$

The problem with σ-additivity is this: if $\mathcal{A}$ has infinitely many finite sets then we can, as in Exercise 2.3.4, obtain countably many pairwise disjoint finite sets $\{A_n\} \subset \mathcal{A}$. Then $\bigcup_n A_n$ is infinite, and if we had σ-additivity then

$$\infty = \mu\Big(\bigcup_n A_n\Big) = \sum_n \mu(A_n) = 0,$$

a contradiction. So $\mathcal{A}$ has to have finitely many finite sets (this is possible even when $\mathcal{A}$ is infinite, for instance write $\mathbb{N} \setminus \{1\} = \bigcup_n A_n$ with $\{A_n\}$ all infinite and pairwise disjoint, and put $\mathcal{A} = \big\{\varnothing, \mathbb{N}, \{1\}\big\} \cup \Sigma(A_1, A_2, \ldots)$).

(2.3.6) Complete the details in Example 2.3.7. That is, show that δ is a measure, and that x_0 is an atom.

Answer. Since $x_0 \notin \varnothing$, $\delta(\varnothing) = 0$.

Let $\{E_k\} \subset X$ be a disjoint countable family. If $x_0 \notin \bigcup_k E_k$, then $\delta(\bigcup_k E_k) = 0 = \sum_k \delta(E_k)$, since $x_0 \notin E_k$ for all k. If $x_0 \in \bigcup_k E_k$, then there exists a single k_0 with $x_0 \in E_{k_0}$. Then $\delta(\bigcup_k E_k) = 1 = \delta(E_{k_0}) = \sum_k \delta(E_k)$.

Finally, $\delta(\{x_0\}) = 1$ by definition, so x_0 is an atom.

(2.3.7) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $E \in \mathcal{A}$. Show that

$$\mathcal{A}_E = \{A \cap E : \ A \in \mathcal{A}\}$$

is a σ-algebra on E, and that $\mu_E(A) = \mu(A)$ defines a measure on $\mathcal{A}_E$.

Answer. Since $\varnothing = \varnothing \cap E$ and $E = X \cap E$, we have $\varnothing, E \in \mathcal{A}_E$. If $A_1, A_2, \ldots$ are sets in $\mathcal{A}_E$, then $A_n = A_n \cap E$ for all n; hence

$$\bigcup_n A_n = \bigcup_n A_n \cap E = \Big(\bigcup_n A_n\Big) \cap E \in \mathcal{A}_E.$$

And if $A \in \mathcal{A}_E$, then $E \setminus A = E \cap A^c \in \mathcal{A}_E$. As for μ_E, we clearly have $\mu_E(\varnothing) = \mu(\varnothing \cap E) = 0$, and if $\{A_n\} \subset \mathcal{A}_E$ are pairwise disjoint, then

$$\begin{aligned}\mu_E\Big(\bigcup_n A_n\Big) &= \mu\Big(E \cap \bigcup_n A_n\Big) = \mu\Big(\bigcup_n E \cap A_n\Big) = \mu\Big(\bigcup_n A_n\Big) \\ &= \sum_n \mu(A_n) = \sum_n \mu(A_n \cap E) = \sum_n \mu_E(A_n).\end{aligned}$$

(2.3.8) Let $(X, \mathcal{A}, \mu)$ be a measure space and Y a set with $X \subset Y$. Show that there exists a measure space $(Y, \mathcal{A}', \mu')$ such that $\mu'(Y \setminus X) = 0$, $\mathcal{A} = \mathcal{A}'_X$, and $\mu = \mu'_X$.

Answer. Let $\mathcal{A}' = \mathcal{A} \cup \{A \cup (Y \setminus X) : A \in \mathcal{A}\}$. Then $\mathcal{A}'$ is a σ-algebra:

- we have $\varnothing, Y = X \cup (Y \setminus X) \in \mathcal{A}'$ by construction.
- If $B \in \mathcal{A}'$, then either $B \in \mathcal{A}$, in which case $Y \setminus B = B \cup (Y \setminus X) \in \mathcal{A}'$, or $B = B_0 \cup (Y \setminus X)$, in which case $Y \setminus B = X \setminus B_0 \in \mathcal{A}'$.
- If $\{B_n\} \subset \mathcal{A}'$, suppose first that $B_n \in \mathcal{A}$ for all n. Then $B = \bigcup_n B_n \in \mathcal{A}$, so $B \in \mathcal{A}'$. Otherwise, there exists m with $B_m = B'_m \cup (Y \setminus X)$ and $B'_m \in \mathcal{A}$. Then

$$B = \bigcup_n B_n = \bigcup_n (B_n \cap X) \cup (B_n \cap (Y \setminus X))$$
$$= \left(\bigcup_n (B_n \cap X)\right) \cup (Y \setminus X) \in \mathcal{A}'.$$

Now we define $\mu'(A \cup (Y \setminus X)) = \mu(A)$. Then $\mu(\varnothing) = 0$. Suppose that $\{B_n\} \subset \mathcal{A}'$ are pairwise disjoint. Then

$$\mu'\Big(\bigcup_n B_n\Big) = \mu\Big(X \cap \bigcup_n B_n\Big) = \mu\Big(\bigcup_n (B_n \cap X)\Big) = \sum_n \mu(B_n \cap X) = \sum_n \mu'(B_n).$$

Hence μ' is a measure on $\mathcal{A}'$ and $\mu'_X = \mu$.

(2.3.9) Let $\mathcal{M}$ be an infinite σ-algebra. Show that there exists nonempty $E \in \mathcal{M}$ such that $\mathcal{M}_{E^c}$ is infinite.

Answer. Suppose that such E does not exist. This means that for any nonempty $E \in \mathcal{M}$ the σ-algebra $\mathcal{M}_{E^c}$ is finite. If we fix any nonempty $E \in \mathcal{M}$, then $\mathcal{M}_{E^c}$ and $\mathcal{M}_E$ are finite (because E^c is also an element of $\mathcal{M}$ and so the negation of the statement does apply to it). This gives us, from $A = (A \cap E) \cup (A \cap E^c)$, that

$$\mathcal{M} = \{A \cup B : A \in \mathcal{M}_E, \ B \in \mathcal{M}_{E^c}\},$$

finite. The contradiction implies that the desired E exists.

(2.3.10) Let $\mathcal{M}$ be an infinite σ-algebra. Show that $\mathcal{M}$ contains a pairwise disjoint sequence of sets. *(Hint: the naive approach does not work; instead, use Exercise 2.3.9)*

Answer. Let E_1 as in Exercise 2.3.9. As $\mathcal{M}_{E_1^c}$ is infinite, we can apply Exercise 2.3.9 again to obtain $E_2 \in \mathcal{M}$, nonempty, disjoint with E_1, and such that $(\mathcal{M}_{E_1^c})_{E_2^c} = \mathcal{M}_{E_1^c \cap E_2^c}$ is infinite. Continuing inductively we produce a pairwise disjoint sequence of nonempty sets $\{E_n\} \subset \mathcal{M}$.

(2.3.11) Show that the equality

$$\mu\Big(\bigcap_k E_k\Big) = \lim_{k\to\infty} \mu(E_k)$$

can fail if $\{E_k\}$ is non-increasing but $\mu(E_k) = \infty$ for all k. Examples can be found in the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ with μ the counting measure.

Answer. Let

$$E_n = \{m \in \mathbb{N} : \ m \geq n\}.$$

Then $E_{n+1} \subset E_n$, $\mu(E_n) = \infty$ for all n, and $\bigcap_n E_n = \varnothing$, so $\mu\Big(\bigcap_n E_n\Big) = 0$.

Here is another example, using Lebesgue measure. For each n, let $E_n = \bigcup_m \big(m - \frac{1}{n}, m + \frac{1}{n}\big)$. Then $E_n \supset E_{n+1}$ for all n, and $m(E_n) = \infty$. But $\bigcap_n E_n = \mathbb{N}$, and $m(\mathbb{N}) = 0$ with the same proof we used for $\mathbb{Q}$ in Section 2.1. Thus $m(E_n) = \infty$ for all n, while $m\Big(\bigcap_n E_n\Big) = 0$.

(2.3.12) Let X be a set and $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be given by

$$\mu^*(E) = \begin{cases} 0, & E = \varnothing \\ 1, & E \neq \varnothing \end{cases}$$

Show that μ^* is an outer measure and find $\mathcal{M}(X)$.

Answer. We have $\mu^*(\varnothing) = 0$ by definition. For any $A \in \mathcal{P}(X)$, $A^c = \varnothing$ if and only if $A = \varnothing$, which shows that $\mu^*(A^c) = \mu^*(A)$.

If $E \in \mathcal{M}(X)$, then $\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c)$ for all $S \in \mathcal{P}(X)$. If $E \subsetneq X$ and $E \neq \varnothing$, then $E^c \neq \varnothing$. Let $S = X$. Then $\mu^*(S) = 1$, and

$$\mu^*(S \cap E) = \mu^*(E) = 1 = \mu^*(E^c) = \mu^*(S \cap E^c).$$

As $1 \neq 2$, $E \notin \mathcal{M}(X)$. Hence $\mathcal{M}(X) = \{\varnothing, X\}$.

(2.3.13) Let $\nu^* : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ be given by

$$\nu^*(E) = \begin{cases} 0, & E = \varnothing \\ \frac{|E|}{1+|E|}, & E \text{ finite} \\ 1, & E \ \text{ infinite} \end{cases}$$

Show that ν^* is an outer measure and find $\mathcal{M}(X)$.

Answer. The function $f(t) = \frac{t}{1+t}$ is increasing and subadditive (proof at the end) on $[0, \infty)$.

We have $\nu^*(\varnothing) = 0$ by definition. If $A \subset B$ and $|A| = \infty$, then $|B| = \infty$ and $\nu^*(A) = 1 = \nu^*(B)$. If $|A| < \infty$ and $|B| = \infty$, then

$$\nu^*(A) = \frac{|A|}{1 + |A|} < 1 = \nu^*(B).$$

When both A, B are finite, $\nu^*(A) = f(|A|) \leq f(|B|) = \nu^*(B)$. So $\nu^*(A) \leq \nu^*(B)$ every time we have $A \subset B$.

Let $E_1, \ldots, E_n \subset \mathbb{N}$. If any of these sets is infinite, then their union is infinite and we have

$$\nu^*(E_1 \cup \cdots \cup E_n) = \infty \leq \infty = \sum_k \nu^*(E_k)$$

since at least one term on the right is infinite. If instead all of $E_1, \ldots, E_n$ are finite, then their union is finite. We have, since

$$\Big| \bigcup_{k=1}^n E_k \Big| \leq \sum_{k=1}^n |E_k|$$

and the function $f(t) = \frac{t}{1+t}$ is increasing and subadditive,

$$\nu^*\Big(\bigcup_{k=1}^n E_k \Big) = \frac{\left| \bigcup_{k=1}^n E_k \right|}{1 + \left| \bigcup_{k=1}^n E_k \right|} \leq \frac{\sum_{k=1}^n |E_k|}{1 + \sum_{k=1}^n |E_k|} \leq \sum_{k=1}^n \frac{|E_k|}{1 + |E_k|} = \sum_{k=1}^n \nu^*(E_k).$$

Now consider infinitely many $E_1, E_2, \ldots \subset \mathbb{N}$ with infinitely many of them nonempty. Then

$$\sum_{k=1}^\infty \nu^*(E_k) \geq \sum_{k=1}^\infty \frac{1}{2} = \infty.$$

So we have

$$\nu^*\Big(\bigcup_{k=1}^\infty E_k\Big) \leq \sum_{k=1}^\infty \nu^*(E_k)$$

regardless. Hence ν^* is an outer measure.

As for $\mathcal{M}(X)$, if $E \subset \mathbb{N}$ is a proper subset then E^c is nonempty and at least one of them is infinite. Then $\nu^*(E) + \nu^*(E^c) > 1$. Therefore

$$\nu^*(\mathbb{N}) = 1 < \nu^*(E) + \nu^*(E^c),$$

showing that $E \notin \mathcal{M}(X)$. Hence $\mathcal{M}(X) = \{\varnothing, X\}$.

Let us now finish the answer by proving that f is increasing and subadditive. We have

$$f'(t) = \Big(\frac{t}{1+t}\Big)' = \Big(1 - \frac{1}{1+t}\Big)' = \frac{1}{(1+t)^2} > 0$$

for all $t \in \mathbb{R}$. So f is increasing. As for the subadditivity, if $t, s \geq 0$

$$\frac{1}{1+t} + \frac{1}{1+s} \leq \frac{2}{1+t+s} \leq \frac{2+t+s}{1+t+s} = 1 + \frac{1}{1+t+s}.$$

Then

$$f(t+s) = 1 - \frac{1}{1+t+s} \leq 2 - \frac{1}{1+t} - \frac{1}{1+s} = f(t) + f(s).$$

(2.3.14) Suppose that in Definition 2.3.11 we replace "countable cover" with "finite cover". Show that this would give $m^*(\mathbb{Q} \cap [0,1]) = 1$, and that this would imply that m^* is not an outer measure.

Answer. Let $I_1, \ldots, I_m$ be open intervals such that $\mathbb{Q} \cap [0,1] \subset I_1 \cup \cdots \cup I_m$. By removing any interval entirely contained in another and reordering if needed, we may assume that $I_k = (a_k, b_k)$, where $b_k \geq a_{k+1} > a_k$ for all k (the second inequality by prescription, and the first one because there can be no gaps between the intervals), $a_1 < 0$, $b_m > 1$. Then

$$\sum_k \ell(I_k) = \sum_k b_k - a_k = b_m - a_1 + \sum_{k=1}^{m-1} b_k - a_{k+1} > b_m - a_1 > 1.$$

Therefore $m^*(\mathbb{Q} \cap [0,1]) \geq 1$. As $(-\varepsilon, 1+\varepsilon)$ is also a cover for each $\varepsilon > 0$, $m^*(\mathbb{Q} \cap [0,1]) = 1$.

We would then have

$$\sum_{q \in \mathbb{Q} \cap [0,1]} m^*(\{q\}) = 0 < 1 = m^*\Big(\bigcup_{q \in \mathbb{Q} \cap [0,1]} \{q\}\Big),$$

contradicting the definition of outer measure.

(2.3.15) Let $E \subset \mathbb{R}$ such that $m^*(E) > 0$. Show that there exist $a, b \in E$ such that $a - b$ is irrational.

Answer. Fix $a \in E$. Suppose that $a - b \in \mathbb{Q}$ for all $b \in E$. If we enumerate $\mathbb{Q} = \{q_k\}$, this means that $b = a - q_k$ for some k. Thus $E \subset \{a - q_k : k \in \mathbb{N}\} \subset a - \mathbb{Q}$. As m^* is translation invariant,

$$m^*(E) \leq m^*(a - \mathbb{Q}) = m^*(\mathbb{Q}) = 0.$$

Therefore if $m^*(E) > 0$ there has to exist $b \in E$ with $a - b \notin \mathbb{Q}$.

(2.3.16) Let $E \subset \mathbb{R}$ be the set of all numbers in $[0, 1]$ that do not have a 1 anywhere in their decimal expansion. Is E measurable? Find $m^*(E)$.

Answer. We assume, in base 10, the same convention as we did when dealing with the Cantor set, which is that we consider $0.199\cdots$ instead of 0.2. This is important because it means that 0.2 has a 1 in its expansion! For each $k \in \mathbb{N}$, let

$$E_k = \Big\{ \sum_{n=1}^{\infty} \frac{a_n}{10^n} \in [0, 1] : \ a_k = 1, \ a_1, \dots, a_{k-1} \neq 1 \Big\}.$$

Each E_k is a finite union of intervals, so measurable. To see this, note that

$$E_1 = [0.1, 0.2], \qquad E_2 = [0.01, 0.02] \cup [0.21, 0.22] \cup [0.31, 0.32] \cup \cdots \cup [0.91, 0.92],$$

and

$$E_k = \bigcup_{a_1, \dots, a_{k-1} \neq 1} [0.a_1 \cdots a_{k-1}1, 0.a_1 \cdots a_{k-1}2].$$

That is, the E_k are pairwise disjoint and each is made of 9^{k-1} (because we have 9 choices for each of the $a_1, \dots, a_{k-1}$) intervals of length 10^{-k}. Then $E = [0, 1] \setminus \bigcup_n E_n$ is measurable, and

$$m(E) = 1 - \sum_k m(E_k) = 1 - \sum_k \frac{9^{k-1}}{10^k} = 1 - \frac{1}{9} \sum_k \frac{9^k}{10^k}$$

$$= 1 - \frac{1}{9} \frac{\frac{9}{10}}{1 - \frac{9}{10}} = 1 - \frac{1}{9} \frac{9}{1} = 0.$$

(2.3.17) Let $A, B \subset \mathbb{R}$ be Lebesgue measurable with $m(A) < \infty$. Show that the function $f : \mathbb{R} \to [0, \infty)$ given by $f(x) = m((A + x) \cap B)$ is continuous. *(Hint: the assertion is easier to prove for intervals)*

Answer. Let $\varepsilon > 0$. Then there exists V open with $V \supset A$ and $m(V \setminus A) < \varepsilon/4$. This implies that $m(V) < \infty$ since $m(A) < \infty$. Now

$$\begin{aligned} m((V + x) \cap B) - m((A + x) \cap B) &= m(((V \setminus A) + x) \cap B) \\ &\leq m(V \setminus A) < \frac{\varepsilon}{4}. \end{aligned}$$

Then, for any x, y,

$$|f(y) - f(x)| \leq \varepsilon/2 + m((V + y) \cap B) - m((V + x) \cap B).$$

Now since V is an open subset of $\mathbb{R}$, we may write it as a disjoint union of open intervals, $V = \bigcup_n (a_n, b_n)$. The finite measure of V gives $\sum_n (b_m - a_n) = m(V) < \infty$. For any $E, F \subset \mathbb{R}$, from $E \cup F = E \cup (F \setminus E) = (E \setminus F) \cup F$ we obtain

$$m(E) - m(F) = m(E \setminus F) - m(F \setminus E).$$

Then, for $x < y$, with $I_n = (a_n, b_n)$,

$$\begin{aligned} |m((I_n + y) \cap B) - m((I_n + x) \cap B)| &\leq m((a_n + x, a_n + y)) \\ &\quad + m((b_n + x, b_n + y)) \\ &= 2(y - x). \end{aligned}$$

Choose n_0 such that $\sum_{n > n_0} (b_n - a_n) < \varepsilon/8$. Then, if $|y - x| < \varepsilon/(8n_0)$,

$$\begin{aligned} |f(y) - f(x)| &\leq \frac{\varepsilon}{2} + m(V + y) \cap B) - m(V + x) \cap B) \\ &= \frac{\varepsilon}{2} + \sum_n m((a_n, b_n) + y) \cap B) - m((a_n, b_n) + x) \cap B) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \sum_{n \leq n_0} m((a_n, b_n) + y) \cap B) - m((a_n, b_n) + x) \cap B) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{2n_0 \varepsilon |y - x|}{8n_0} = \varepsilon. \end{aligned}$$

(2.3.18) Show that the relation $x \sim y$ if $x - y \in \mathbb{Q}$ is an equivalence relation in $\mathbb{R}$.

Answer. Reflexive: $x - x = 0 \in \mathbb{Q}$. Symmetric: if $x - y \in \mathbb{Q}$, then $y - x = (-(x-y) \in \mathbb{Q}$. Transitive: if $x - y \in \mathbb{Q}$ and $y - z \in \mathbb{Q}$, then $x - z = (x-y) + (y-z) \in \mathbb{Q}$.

(2.3.19) Fix $c \in (0,1)$. Construct an open set $V \subset [0,1]$, dense in $[0,1]$ and with $m(V) = c$ *(Hint: in Exercise 2.2.8 this was done for $c = \frac{1}{2}$; an entirely different approach is possible with an idea similar—but not equal—to Exercise 2.3.17)* .

Answer. When we consider middle thirds (length of the interval is a power of 3) for the Cantor set, all the intervals together form a dense open set of measure 1. So we need to remove smaller intervals; this guarantees that there will be no overlaps, as the intervals we will consider are subintervals of those middle thirds removed for the Cantor set.

Fix $a \in \left[0, \frac{1}{3}\right]$. On the n^{th} step we remove (in the end, in this case, we want to keep them) 2^{n-1} middle open intervals $V_{n,1}, \ldots, V_{n,2^{n-1}}$ each of length a^n. Then we put

$$V = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{2^{n-1}} V_{n,k}.$$

This set is open, being a union of open intervals, and it is dense because in each step, when we consider the middle interval of length a^n inside (c,d), this latter interval gets divided into two intervals each of length less than $\frac{d-c}{2}$. This means that for $\varepsilon > 0$ and $t \in [0,1]$ either $t \in V$ or there exist n, k such that $\operatorname{dist}(t, V_{n,k}) < \varepsilon$.

Finally,

$$m(V) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} a^n = \frac{1}{2} \sum_{n=1}^{\infty} (2a)^n = \frac{1}{2} \frac{2a}{1-2a} = \frac{a}{1-2a}.$$

Solving for a in $\frac{a}{1-2a} = c$, we get $a = \frac{c}{1+2c}$. This works as expected: when $c = 0$ we get $a = 0$, and when $c = 1$ we get $a = \frac{1}{3}$. So by continuity any value $c \in [0,1]$ can be achieved by an appropriate $a \in \left[0, \frac{1}{3}\right]$.

For the second approach, let $\{q_k\}$ be an enumeration of $\mathbb{Q} \cap [0,1]$, and put

$$A = (0,1) \cap \bigcup_k \left(q_k - \frac{c}{2^{k+2}}, q_k + \frac{c}{2^{k+2}}\right).$$

Then $A \subset [0,1]$ is open, dense, and

$$m(A) \leq \sum_{k=1}^{\infty} \frac{c}{2^{k+1}} = \frac{c}{2}.$$

Let $f(x) = m(A \cup (0, x))$. Then $f(0) = m(A) < c$, $f(1) = m((0, 1)) = 1$. The function f is continuous by an argument similar to Exercise 2.3.17. By the Intermediate Value Theorem, there exists x such that $f(x) = c$. So $A \cup (0, x)$ is a dense open set of measure c.

Here is yet another approach, related to the first one. Write c in base 3, $c = \sum_{n=1}^{\infty} a_n 3^{-n}$, with $a_n \in \{0, 1, 2\}$ for all n. Let $\{r(n)\}$ be the increasing sequence of indices such that $a_{r(n)} \neq 0$. So

$$c = \sum_{n=1}^{\infty} \frac{a_{r(n)}}{3^{r(n)}}$$

Define numbers $b_n = \frac{a_{r(n)}}{2^{n-1}}$. Now we remove middle thirds as with the Cantor set, but in each step the length of the removed intervals will be $b_n 3^{-r(n)}$ instead of 3^{-n}. Explicitly, if $\{(a_{n,k}, b_{n,k})\}_{n \in \mathbb{N},\ 1 \leq k \leq 2^{n-1}}$ are the intervals removed from the usual Cantor ternary set, we can define

$$\delta_n = \frac{1}{2} \left(\frac{1}{3^{r(n)}} - \frac{2a_{r(n)}}{2^n 3^{r(n)}} \right),$$

and let

$$V = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{2^{n-1}} \big(a_{r(n),k} + \delta_n, b_{r(n),k} - \delta_n \big).$$

By construction, V is open. Also,

$$m(V) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \left(\frac{1}{3^{r(n)}} - 2\delta_n \right) = \sum_{n=1}^{\infty} \frac{2^n a_{r(n)}}{2^n 3^{r(n)}} = c.$$

Finally, if $t \in V^c$, then $t \notin \bigcup_{k=1}^{2^{n-1}} (a_{n,k} + \delta_n, b_{n,k} - \delta_n)$. These are 2^{n-1} disjoint intervals inside $[0, 1]$. Thus $\operatorname{dist}(t, V) < \frac{1}{2^{n-1}}$. This can be done for any n, so $\operatorname{dist}(t, V) = 0$, which shows that $\overline{V} = [0, 1]$.

(2.3.20) Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$. Show that μ is outer regular if and only if it is inner regular by closed sets.

Answer. Fix $\varepsilon > 0$. If μ is outer regular and $E \in \mathcal{A}$, then $X \setminus E \in \mathcal{A}$ and by hypothesis there exists V open such that $X \setminus E \subset V$ and

$$\mu(V \setminus (X \setminus E)) = \mu(V) - \mu(X \setminus E) < \varepsilon.$$

Let $K = X \setminus V$. Then K is closed, $K = X \setminus V \subset X \setminus (X \setminus E) = E$, and

$$\mu(E) - \mu(K) = \mu(E \setminus K) = \mu(E \cap (X \setminus K)) = \mu(E \cap V$$
$$= \mu((X \setminus (X \setminus E)) \cap V) = \mu(V \setminus (X \setminus E)) < \varepsilon.$$

As this can be done for any $\varepsilon > 0$,

$$\mu(E) = \sup\{\mu(K) : \ K \subset E, \ K \ \text{ closed}\}.$$

The converse is proven in an entirely similar way.

The condition $\mu(X) < \infty$ was used to guarantee the equalities $\mu(V \setminus (X \setminus E)) = \mu(V) - \mu(X \setminus E)$ and $\mu(E) - \mu(K) = \mu(E \setminus K) = \mu(E \cap (X \setminus K))$.

(2.3.21) Show that Lebesgue measure is not outer regular by closed sets, nor inner regular by open sets.

Answer. Let A be the open set from Exercise 2.3.19, where $c = 1/2$. If C is closed and $C \supset A$ then $C \supset [0, 1]$, forcing $m(C) \geq 1$. This shows that outer regularity by closed sets fails.

If we consider the complement of A in $[0, 1]$ we get a closed set that has no nontrivial open subsets; for any open $V \subset [0, 1] \setminus A$ would give us $A \subset V^c$ and so $m(V^c) = 1$ and hence $m(V) = 0$. So m is not inner regular by open sets either.

(2.3.22) Show that a Borel measure on a locally compact Hausdorff space is locally finite if and only if it is finite on compact sets.

Answer. Suppose that μ is locally finite. Fix E compact. For each $x \in E$, there exists V_x open with $x \in V_x$ and $\mu(V_x) < \infty$. By compactness, there exist $x_1, \ldots, x_n \in E$ such that $E \subset V_{x_1} \cup \cdots \cup V_{x_n}$. Then $\mu(E) \leq \sum_{k=1}^{n} \mu(V_{x_k}) < \infty$.

Conversely, if μ is not locally finite, then there exists $x \in X$ such that every open set V with $x \in V$ satisfies $\mu(V) = \infty$. Since X is locally compact, there exists an open set W with $x \in W$ and $\overline{W}$ compact. Then $\mu(\overline{W}) \geq \mu(W) = \infty$. That is, there exists a compact set of infinite measure.

(2.3.23) Show that if X is a Hausdorff topological space, a Dirac delta is a Radon measure on $\mathcal{B}(X)$. Show by example that the assertion can fail if X is not Hausdorff.

Answer. Denote by $0 \in X$ the distinguished element such that $\delta(E) = 1$ if and only if $0 \in E$. For any $E \in \mathcal{B}(X)$, if $0 \in E$ and $V \supset E$ with V open, then $\delta(V) = \delta(E) = 1$; and if $0 \notin E$, by X being Hausdorff $X \setminus \{0\}$ is open, and $E \subset X \setminus \{0\}$ with $\delta(X \setminus \{0\}) = 0$. We have shown that δ is outer regular. For inner regularity, if $0 \in E$ then $K = \{0\}$ is compact with $\delta(K) = 1 = \delta(E)$; and if $0 \notin E$, then any $K \subset E$ does not have 0, and hence $\delta(K) = \delta(E) = 0$.

Let $X = \{0, 1, 2\}$ with the topology $\{\varnothing, X, \{0\}, \{0, 2\}, \{0, 1\}\}$. Then 0 and 1 cannot be separated, and X is not Hausdorff. The set $E = \{1\}$ is Borel, as it is closed: $\{1\} = X \setminus \{0, 2\}$. We have $\delta(E) = 0$, while $\delta(V) = 1$ for all $V \supset E$ open. Thus δ is not outer regular.

(2.3.24) Let X be a metric and compact. Show that any finite measure on $\mathcal{B}(X)$ is Radon.

Answer. Fix $E \subset X$ closed. As the function $x \longmapsto d(x, F)$ is continuous, the sets $V_n = \{x : \ d(x, F) < 1/n\}$ are open. We have $F \subset V_{n+1} \subset V_n$, and (since F is closed) $F = \bigcap_n V_n$. By continuity of the measure (and the finiteness), $\mu(F) = \lim_m \mu(V_n)$.

Now consider the family $\mathcal{E}$ of Borel sets $E \subset X$ that satisfy

$$\mu(E) = \sup\{\mu(F) : \ F \subset E, \ \ \text{closed}\} = \inf\{\mu(V) : \ E \subset V, \ \ \text{open}\}.$$

We proved in the first paragraph that $\mathcal{E}$ contains the closed sets. If $E \in \mathcal{E}$, from $\mu(X) < \infty$ if $E \subset V$ with $\mu(V) < \mu(E) + \varepsilon$ we get $X \setminus V \subset X \setminus E$ with $\mu(X \setminus V) + \varepsilon > \mu(X \setminus E)$. It follows that $X \setminus E \in \mathcal{E}$. If $\{E_n\} \subset \mathcal{E}$ are pairwise disjoint, fix $\varepsilon > 0$ and for each n choose V_n open, F_n closed with $F_n \subset E_n \subset V_n$ and $\mu(V_n \setminus F_n) < \varepsilon/2^n$. Then $V = \bigcup_n V_n$ and $F = \bigcap_n F_n$ are open and closed respectively, with $F \subset E \subset V$, and

$$\begin{aligned}\mu(V \setminus F) &= \mu\Big(\bigcup_n V_n \setminus \bigcap_n F_n\Big)\\ &\leq \mu\Big(\bigcup_n (V_n \setminus F_n)\Big)\\ &\leq \sum_n \mu(V_n \setminus F_n) < \varepsilon.\end{aligned}$$

Hence $\bigcup_n E_n \in \mathcal{E}$. We have established that $\mathcal{E}$ is a σ-subalgebra of $\mathcal{B}(X)$ that contains the closed sets, so $\mathcal{E} = \mathcal{B}(X)$. This proves the outer regularity and the inner regularity by closed sets. As the measure is finite and every closed subset of X is compact, μ is Radon.

(2.3.25) Show that the counting measure on $\mathbb{R}^n$ is not Radon.

Answer. Any open set on $\mathbb{R}^n$ is uncountable, so $\mu(V) = \infty$ for all open V. Hence no finite set can be approximated in measure from above by open sets. The counting measure is inner regular, though: if E is finite, then it is compact and can be approximated; and if E is infinite, it can be approximated by finite sets of arbitrary size, and hence of arbitrary large measure. The counting measure also fails to be locally finite, and infinite compact sets will have infinite measure.

(2.3.26) Let $(X, \mathcal{A}, \mu)$ be a measure space. The measure μ is **semifinite** if whenever $\mu(E) = \infty$ for some $E \in \mathcal{A}$, there exists $F \in \mathcal{A}$ with $F \subset E$ and $0 < \mu(F) < \infty$. Show that for such μ any $E \in \mathcal{A}$ satisfies

$$\mu(E) = \sup\{\mu(F) : \ F \in \mathcal{A}, \ F \subset E, \ \mu(F) < \infty\}. \qquad (2.12)$$

Show also that the equality above fails for a measure that is not semifinite.

Answer. Let $\alpha = \sup\{\mu(F) : \ F \in \mathcal{A}, \ F \subset E, \ \mu(F) < \infty\}$. It is immediate that $\alpha \leq \mu(E)$ since $\mu(E)$ is an upper bound. So if $\alpha = \infty$ we are done. When $\alpha < \infty$, choose $\{F_n\} \subset \mathcal{A}$ with $F_n \subset E$, of finite measure, with $F_n \subset F_{n+1}$ for all n, and such that $\alpha = \lim_n \mu(F_n)$. Let $F = \bigcup_n F_n \in \mathcal{A}$. For each n, $\bigcup_{k<n} F_k \subset E$ and has finite measure, so by continuity of the measure

$$\mu(F) = \lim_n \mu\Big(\bigcup_{k<n} F_n\Big) \leq \alpha = \lim_n \mu(F_n) \leq \mu(F).$$

Thus $\mu(F) = \alpha$. If $G \subset E \setminus F$ is measurable and $\mu(G) < \infty$, then $F \cup G \subset E$ and $\alpha = \mu(F) \leq \mu(F) + \mu(G) = \mu(F \cup G) = \alpha$, so $\mu(G) = 0$. This prevents $\mu(E \setminus F) = \infty$, for in such case it would have subsets of positive measure. Then $\mu(E \setminus F) < \infty$ but now the argument we just did implies that $\mu(E \setminus F) = 0$. Therefore $\mu(E) = \mu(F) = \alpha$.

When μ is not semifinite, there exists $E \in \mathcal{A}$ with $\mu(E) = \infty$ and $\mu(F) = 0$ for every measurable subset of E with finite measure. Thus the supremum in (2.12) is zero.

(2.3.27) Let K be a topological space and $\{\mu_j\}$ a collection of Borel measures on K. Let $X = \bigcup_j K \times \{j\}$. We write π_1 for the coordinate function $\pi_1(a, b) = a$.

(i) Show that

$$\Sigma = \big\{B \subset X : \ \pi_1(B \cap (K \times \{j\})) \in \mathcal{B}(K) \ \text{ for all } j\big\}$$

is a σ-algebra.

(ii) Show that μ given by $\mu(B) = \sum_j \mu_j\big(\pi_1(B \cap (K \times \{j\})\big)$ is a measure on Σ.

Answer.

(i) We have $X \in \Sigma$, since K is a Borel subset of itself. Given $B \in \Sigma$, we have $B = \bigcup_j B_j \times \{j\}$ with $B_j \subset K$ Borel. Then

$$\pi_1\big((X \setminus B) \cap (K \times \{j\})\big) = \pi_1\big((K \setminus B_j) \times \{j\}\big) = K \setminus B$$

is Borel for each j. So $X \setminus B \in \Sigma$. If now $\{B_n\}_{n\in\mathbb{N}} \subset \Sigma$, we have $B_n = \bigcup_j B_{n,j} \times \{j\}$ with $B_{n,j} \subset K$ Borel for all j. Then

$$\begin{aligned}\pi_1\Big(\bigcup_n B_n \cap (K \cap \{j\})\Big) &= \pi_1\Big(\bigcup_n \bigcup_j (B_{n,k} \times \{k\}) \cap (K \cap \{j\})\Big)\\ &= \pi_1\Big(\bigcup_n B_{n,j} \times \{j\}\Big) = \bigcup_n B_{n,j}\end{aligned}$$

which is Borel in K. So Σ is a σ-algebra.

(ii) All we need to check is the σ-additivity. If $\{B_n\}_{n\in\mathbb{N}} \subset \Sigma$ are pairwise disjoint, we have

$$\begin{aligned}\varnothing = B_n \cap B_m &= \Big(\bigcup_j B_{n,j} \times \{j\}\Big) \cap \Big(\bigcup_k B_{n,k} \times \{k\}\Big)\\ &= \bigcup_j (B_{n,j} \cap B_{n,j}) \times \{j\}.\end{aligned}$$

So $\{B_{n,j}\}_n$ are pairwise disjoint for each j. Then (using Tonelli)

$$\begin{aligned}\mu\Big(\bigcup_n B_n\Big) &= \sum_j \mu_j\Big(\pi_1\Big(\bigcup_n B_n \cap (K \times \{j\})\Big)\Big) = \sum_j \mu_j\Big(\bigcup_n B_{n,j}\Big)\\ &= \sum_j \sum_n \mu_j(B_{n,j}) = \sum_n \sum_j \mu_j(B_{n,j}) = \sum_n \mu(B_n).\end{aligned}$$

2.4. Measurable Functions

(2.4.1) Let $(X, \mathcal{A})$ be a measurable spaces and $f : X \to Y$ a function. Show that

$$\mathcal{F} = \{E \subset Y : \ f^{-1}(E) \in \mathcal{A}\}$$

is a σ-algebra.

Answer. We have $\varnothing \in \mathcal{F}$, since $f^{-1}(\varnothing) = \varnothing \in \mathcal{A}$. If $E \in \mathcal{F}$, then $f^{-1}(E^c) = (f^{-1}(E))^c \in \mathcal{A}$. And if $\{E_n\}$ is a countable family with $E_n \in \mathcal{F}$ for all n, then

$$f^{-1}\Big(\bigcup_n E_n\Big) = \bigcup_n f^{-1}(E_n) \in \mathcal{A}$$

since $\mathcal{A}$ is a σ-algebra and $f^{-1}(E_n) \in \mathcal{A}$ for all n.

(2.4.2) Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces and $f : X \to Y$ a function. Show that

$$\mathcal{A}_0 = \{f^{-1}(B) : \ B \in \mathcal{B}\}$$

is a σ-algebra. What is the relation between $\mathcal{A}_0$ and $\mathcal{A}$?

Answer. We have $\varnothing = f^{-1}(\varnothing) \in \mathcal{A}_0$. If $A \in \mathcal{A}_0$, then $A = f^{-1}(B)$ for some $B \in \mathcal{B}$; hence, since preimages preserve all set operations, $A^c = f^{-1}(B^c) \in \mathcal{A}_0$ since $B^c \in \mathcal{B}$. If $\{E_n\} \subset \mathcal{A}_0$ is a countable family with $E_n \in \mathcal{A}_0$ for all n, there exist $\{B_n\} \subset \mathcal{B}$ with $E_n = f^{-1}(B_n)$ for each n. Then

$$\bigcup_n E_n = \bigcup_n f^{-1}(B_n) = f^{-1}\Big(\bigcup_n B_n\Big) \in \mathcal{A}_0$$

since the countable unions of sets in $\mathcal{B}$ stays in $\mathcal{B}$. Therefore $\mathcal{A}_0$ is a σ-algebra.

In general there is not much relation between $\mathcal{A}_0$ and $\mathcal{A}$. For instance take $X = \mathbb{R}$, $\mathcal{A} = \mathcal{M}(\mathbb{R})$, $Y = \mathbb{R}$, $\mathcal{B} = \mathcal{P}(\mathbb{R})$, $f = \mathrm{id}$. Then $\mathcal{A}_0 = \mathcal{P}(\mathbb{R}) \supsetneq \mathcal{A}$. If we reverse the roles of $\mathcal{A}$ and $\mathcal{B}$ we get $\mathcal{A} \supsetneq \mathcal{A}_0$. There need be no inclusion either. For instance fix some nonempty $A, B \subset \mathbb{R}$ such that $A \cap B \neq \varnothing$; then $A \subsetneq A \cup B$, $B \subsetneq A \cup B$. Let

$$\mathcal{A} = \{\varnothing, A, A^c, \mathbb{R}\}, \qquad \mathcal{B} = \{\varnothing, B, B^c, \mathbb{R}\}.$$

Then, with $f = \mathrm{id}$ we have $\mathcal{A}_0 = \mathcal{B}$ and $\mathcal{A} \cap \mathcal{B} = \{\varnothing, \mathbb{R}\}$.

(2.4.3) Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing. Show that f is Borel-measurable (that is, pre-image of open is Borel).

Answer. We need to show that $A = f^{-1}(a, \infty)$ is Borel for all $a \in \mathbb{R}$. So fix $a \in \mathbb{R}$. If $A = \varnothing$, then it is Borel. Now we assume that $A \neq \varnothing$.

We have $A = \{x : f(x) > a\}$. If $z \in A$, we have $f(z) > a$; for any $y > z$, $f(y) \geq f(z) > a$. So $[z, \infty) \subset A$.

Let $b = \inf A$. If $b = -\infty$, then for any $z \in \mathbb{R}$ there exists $z' \in A$ with $z' < z$; by the above $[z, \infty) \subset A$ for all $z \in \mathbb{R}$. So $A \supset \bigcup_{z \in \mathbb{R}} [z, \infty) = \mathbb{R}$, giving us $A = \mathbb{R}$. Otherwise, if $b > -\infty$, for any $n \in \mathbb{N}$ there exists $z_n \in A$ with $b < z_n < b + \frac{1}{n}$. By the above, $[b + \frac{1}{n}, \infty) \subset A$. If $b \notin A$, then

$$A \supset \bigcup_{n \in \mathbb{N}} [b + \frac{1}{n}, \infty) = (b, \infty);$$

as $b = \inf A$ it follows that $A = (b, \infty)$, so Borel. The other possibility is that $b \in A$; in that case we can repeat what we did and obtain $A = [b, \infty)$. In all cases, A is Borel.

(2.4.4) Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing. Show that f is continuous almost everywhere, by showing that the set of discontinuity points of f is at most countable.

Answer. Fix an interval $[-m, m]$, $m \in \mathbb{N}$ and $x_0 \in [-m, m]$. By the monotonicity we have that $f([-m, m]) \subset [f(-m), f(m)]$.

Since $f(x) \leq f(x_0)$ for all $x < x_0$, $l(x_0) = \lim_{x \to x_0^-} f(x)$ exists. Similarly, $r(x_0) = \lim_{x \to x_0^+} f(x)$ also exists. Since a function is continuous at a point if and only if the lateral limits exist and are equal, we have that f is continuous at x_0 if and only if $l(x_0) = r(x_0)$.

If we let X be the set of discontinuity points of f and we write $X_n = \{x : r(x) - l(x) > \frac{1}{n}\}$, then $X = \bigcup_n X_n$. Suppose that $x_1 < \cdots < x_w \in X_n$. Then

$$r(x_1) - l(x_1) + r(x_2) - l(x_2) + \cdots + r(x_w) - l(x_w) > \frac{w}{n}$$

We also have

$$\sum_{j=1}^{w} r(x_j) - l(x_j) = -l(x_1) + r(x_w) + \sum_{j=2}^{w-1} [r(x_j) - l(x_{j+1})] \leq f(m) - f(-m).$$

From the two inequalities we obtain $w \leq n(f(m) - f(-m))$; so X_n is finite. Then X is countable.

For the general case, $\mathbb{R} = \bigcup_m [-m, m]$ and each $[-m, m]$ contains at most countably many discontinuity points of f, so $\mathbb{R}$ contains at most countably many discontinuity points of f.

(2.4.5) Let $f : \mathbb{R} \to \mathbb{C}$ be Lebesgue-measurable. Show that there exists $g : \mathbb{R} \to \mathbb{C}$, Borel-measurable, with $f = g$ a.e. *(Hint: do it first for $f \geq 0$)*

Answer. By the usual trick of writing a complex valued function as a linear combination of four nonnegative functions, we may assume without loss of generality that $f \geq 0$. Let $\{s_n\}$ be an nondecreasing sequence of Lebesgue-measurable simple functions with $0 \leq s_n \nearrow f$. As in Theorem 2.4.13 we may write

$$s_n = n\, 1_{f^{-1}([n,\infty])} + \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\, 1_{E_{k,n}},$$

where $E_{k,n} = f^{-1}\big(\big[\frac{k-1}{2^n}, \frac{k}{2^n}\big)\big)$. By Proposition 2.3.25 there exist sets $F_{k,n} \in \mathcal{B}(\mathbb{R})$ with $F_{k,n} \subset E_{k,n}$ and $m(F_{k,n}) = m(E_{k,n})$ for all k, n. Define

$$g_n = n\, 1_{f^{-1}([n,\infty])} + \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\, 1_{F_{k,n}},$$

Then each g_n is Borel-measurable and so is $g = \lim_n g_n$ (this converges for every point because $g_n = s_n$ on each $F_{k,n}$). The set $\{f \neq g\}$ is contained in the countable union $\bigcup_{k,n} \big[E_{k,n} \setminus F_{n,k}\big]$, a nullset.

(2.4.6) Let $(X, \mathcal{A})$ be a measurable space. If $f : E \to \mathbb{R}$ is a measurable function relative to the measurable space $(E, \mathcal{A}_E)$ (cfr. Exercise 2.3.7), show that the extension $\tilde{f} : X \to \mathbb{R}$ of f given by $\tilde{f} = f\, 1_E$ is measurable with respect to $\mathcal{A}$.

Answer. Given $a \in \mathbb{R}$, we consider

$$\tilde{f}^{-1}(a, \infty) = \{x \in X : \ f(x)\, 1_E(x) > a\}.$$

We have

$$\begin{aligned}\tilde{f}^{-1}(a,\infty) &= [\tilde{f}^{-1}(a,\infty)\cap E]\cup[\tilde{f}^{-1}(a,\infty)\cap E^c]\\ &= \{x\in E:\ f(x)\,1_E(x)>a\}\cup\{x\in E^c:\ f(x)\,1_E(x)>a\}\\ &= \{x\in E:\ f(x)>a\}\cup\{x\in E^c:\ 0>a\}\\ &= f^{-1}(a,\infty)\cup\{x\in E^c:\ 0>a\}.\end{aligned}$$

The first set is in $\mathcal{A}_E\subset\mathcal{A}$ by hypothesis, while the second is either E^c or $\varnothing$, and in both cases it is in $\mathcal{A}$. The union is then in $\mathcal{A}$.

(2.4.7) If $\tilde{f}:X\to\mathbb{R}$ is a measurable function with respect to $\mathcal{A}$, and if $f=\tilde{f}|_E$ (the restriction of $\tilde{f}$ to E), show that $f:E\to\mathbb{R}$ is measurable with respect $\mathcal{A}_E$.

Answer. We have, for any $a\in\mathbb{R}$, and using that $\tilde{f}$ is measurable,

$$f^{-1}(a,\infty)=\{x\in E:\ f(x)>a\}=E\cap\tilde{f}^{-1}(a,\infty)\in\mathcal{A}_E.$$

(2.4.8) Let $(X,\mathcal{A})$ be a measurable space and $f:X\to[0,\infty]$ measurable. Show that if s_n is as in Theorem 2.4.13, then for $x\in f^{-1}[0,n)$ we have

$$s_n(x)=\frac{\lfloor 2^n f(x)\rfloor}{2^n}.$$

Answer. By definition, $s_n(x)=\frac{k-1}{2^n}$ when $\frac{k-1}{2^n}\le f(x)<\frac{k}{2^n}$. These inequalities can be written as $k-1\le 2^n f(x)<k$, which in turn are precisely $\lfloor 2^n f(x)\rfloor=k-1$. Thus

$$s_n(x)=\frac{k-1}{2^n}=\frac{\lfloor 2^n f(x)\rfloor}{2^n}.$$

(2.4.9) Let $(X,\mathcal{A},\mu)$ be a measure space and $f:X\to\mathbb{C}$ measurable. Show that $\operatorname{ess\,ran} f$ is closed.

Answer. Let $\alpha\in\overline{\operatorname{ess\,ran} f}$. Fix $\varepsilon>0$. Then there exists $\beta\in\operatorname{ess\,ran} f$ such that $|\alpha-\beta|<\varepsilon/2$. If $|\gamma-\beta|<\varepsilon/2$, then

$$|\alpha-\gamma|\le|\alpha-\beta|+|\beta-\gamma|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

This means that $B_{\varepsilon/2}(\beta) \subset B_\varepsilon(\alpha)$. Thus

$$\mu(f^{-1}(B_\varepsilon(\alpha))) \geq \mu(f^{-1}(B_{\varepsilon/2}(\beta))) > 0$$

since $\beta \in \operatorname{ess\,ran} f$. So $\alpha \in \operatorname{ess\,ran} f$ and $\operatorname{ess\,ran} f$ is closed.

(2.4.10) Show an example of a measure space, g measurable, and $f = g$ a.e. with f not measurable.

Answer. Consider $([0,1], \mathcal{B}(\mathbb{R}), m)$, and let $g = 1 - 1_{\mathcal{C}}$, the characteristic of the complement of the Cantor set. We know that there exists $V \subset \mathcal{C}$ that is not in $\mathcal{B}(\mathbb{R})$, as in the proof of Proposition 2.3.24. Let $f = g + 1_V$. Then $f = g$ outside of $\mathcal{C}$, but f is not measurable, since $V = f^{-1}(\{1\}) \cap \mathcal{C} \notin \mathcal{B}(\mathbb{R})$.

(2.4.11) Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) = 1$. Show that the following statements are equivalent:

(i) $\mu(E) \in \{0,1\}$ for all $E \in \mathcal{A}$;

(ii) if $f : X \to \mathbb{R}$ is measurable, there exists $c \in \mathbb{R}$ with $f = c$ a.e.

Answer. Suppose that $\mu(\mathcal{A}) = \{0,1\}$ and let f be measurable. We may assume without loss of generality that f maps into the interval $[0,1]$, for we might replace it by composition with a continuous bijection, say $g \circ f$ where $g(t) = \frac{1}{2} + \frac{1}{\pi} \arctan t$. Write $I_{0,1} = [0,1]$ and consider disjoint dyadic partitions

$$[0,1] = \bigcup_{j=1}^{2^n} I_{n,j}.$$

Then, for each n, the 2^n disjoint sets $X_{n,j} = f^{-1}(I_{n,j})$, $j = 1, \ldots, 2^n$, form a partition of X. Since $1 = \mu(X) = \sum_j \mu(X_{n,j})$, for each n there exists a single j_n with $\mu(X_{n,j_n}) = 1$. Then $f = f\,1_{X_{n,j_n}}$ a.e. Necessarily (by the fact that they are either disjoint or one inside the other, and looking at their measures) we have $X_{n+1,j_{n+1}} \subset X_{n,j_n}$ for all n. Let $X_0 = \bigcap_n X_{n,j_n}$; by continuity of the measure, $\mu(X_0) = 1$. We also have $I_{n+1,j+1} \subset I_{n,j}$ for all n. Then $\bigcap_n I_{n,j_n} = \{s\}$ for some $s \in [a,b]$ by Proposition 1.8.19 (or by forming s in base 2). We have

$$\mu(f^{-1}(\{s\})) = \mu\Big(f^{-1}\Big(\bigcap_n I_{n,j_n}\Big)\Big) = \mu\Big(\bigcap_n f^{-1}(I_{n,j_n})\Big) = \mu(X_0) = 1.$$

Thus $f = s$ a.e.

Conversely, if every measurable function is constant a.e., let $E \in \mathcal{A}$. If $1_E = 0$ a.e., then $\mu(E) = 0$; otherwise $1_E = 1$ a.e., which means that $\mu(E) = 1$.

2.5. The Lebesgue Integral

(2.5.1) Show, without using the convergence theorems, that if μ is the counting measure on $\mathbb{N}$, for a non-negative sequence $\{a_n\}$ we have

$$\int_{\mathbb{N}} a\,d\mu = \sum_{n=1}^{\infty} a_n.$$

Answer.

We can do it by working just with the definition. For any $k \in \mathbb{N}$,

$$\sum_{n=1}^{k} a_n = \int_{\mathbb{N}} \sum_{n=1}^{k} a_n\,1_{\{n\}}\,d\mu \leq \int_{\mathbb{N}} a\,d\mu,$$

since $s_n = \sum_{n=1}^{k} a_n\,1_{\{n\}}$ is a simple function that satisfies $0 \leq s_n \leq a$. If $\sum_{n=1}^{\infty} a_n = \infty$, then the equality holds. So we may assume that $\sum_{n=1}^{\infty} a_n < \infty$.

Fix $\varepsilon > 0$ and let s be a non-negative simple function with $0 \leq s \leq a$ and

$$\int_{\mathbb{N}} a\,d\mu < \varepsilon + \int_{\mathbb{N}} s\,d\mu.$$

We have $s = \sum_{j=1}^{r} \lambda_j\,1_{E_j}$, with $\lambda_j > 0$ for all j. If $|E_j| = \infty$ for some j, this would imply (from $0 \leq s \leq a$) that $\sum_{n=1}^{\infty} a_n = \infty$, contradicting our assumption. So each E_j is finite; by writing each E_j as a finite union of points, we can write $s = \displaystyle\sum_{n=1}^{m} \alpha_n\,1_{\{n\}}$. From $s \leq a$ we immediately get $\alpha_n \leq a_n$ for $n = 1, \ldots, m$. Then

$$\int_{\mathbb{N}} a\,d\mu < \varepsilon + \int_{\mathbb{N}} s\,d\mu = \varepsilon + \sum_{n=1}^{m} \alpha_n \leq \varepsilon + \sum_{n=1}^{\infty} a_n.$$

As this can be done for all $\varepsilon > 0$, we obtain $\displaystyle\int_{\mathbb{N}} a\,d\mu \le \sum_{n=1}^{\infty} a_n$, which shows the equality.

(2.5.2) Show, using the convergence theorems, that if μ is the counting measure on $\mathbb{N}$, for a non-negative sequence $\{a_n\}$ we have

$$\int_{\mathbb{N}} a\,d\mu = \sum_{n=1}^{\infty} a_n.$$

Answer. We can write

$$a = \sum_{n=1}^{\infty} a_n\,1_{\{n\}}.$$

Note that for each $k \in \mathbb{N}$ the series $a(k)$ has only one nonzero term, so there is no issue with convergence. Then Corollary 2.5.8 gives directly

$$\int_{\mathbb{N}} a\,d\mu = \int_{\mathbb{N}} \sum_{n=1}^{\infty} a_n\,1_{\{n\}}\,d\mu = \sum_{n=1}^{\infty} \int_{\mathbb{N}} a_n\,1_{\{n\}}\,d\mu = \sum_{n=1}^{\infty} a_n.$$

(2.5.3) Show that if $f \ge 0$ a.e. and $\displaystyle\int_X f\,d\mu = 0$, then $f = 0$ a.e. Conclude that if f is measurable and $\displaystyle\int_X |f|\,d\mu = 0$, then $f = 0$ a.e.

Answer. Let $E = \{f > 0\}$, $E_n = \{f > \frac{1}{n}\}$, $n \in \mathbb{N}$. Then $E = \bigcup_n E_n$. Suppose $\mu(E) > 0$. Since the union is increasing, by continuity of the measure (Proposition 2.3.8) there exists m with $\mu(E_m) > 0$. Then

$$\int_X f\,d\mu \ge \int_{E_m} f\,d\mu \ge \frac{1}{m}\int_{E_m} 1\,d\mu = \frac{\mu(E_m)}{m} > 0,$$

a contradiction. So $\mu(E) = 0$, that is $f = 0$ a.e.

This can also be proven by definition, in the following way. Let $s = \sum_j s_j\,1_{E_j}$ be simple, with $0 \le s \le f$. Then

$$0 \le \sum_j s_j\,\mu(E_j) = \int_X s\,d\mu \le \int_X f\,d\mu = 0.$$

So, for each j, either $s_j = 0$ or $\mu(E_j) = 0$. It follows that $s = 0$ a.e. Now, as f is measurable, there exists a monotone sequence $\{s_n\}$ of nonnegative simple functions with $0 \le s_n \le f$ and $s_n \to f$. By the above, $s_n = 0$ a.e. for all n, so $f = \lim s_n = 0$ a.e.

For arbitrary measurable f, if $\int_X |f|\,d\mu = 0$ we apply the above to $|f|$ to conclude that $|f| = 0$ a.e., and so $f = 0$ a.e.

(2.5.4) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $f : X \to \overline{\mathbb{R}}$. Let $\{E_n\} \subset \mathcal{A}$ be a pairwise disjoint sequence. Let $E = \bigcup_n E_n$. Show that, if the integral on the left exists,

$$\int_E f\,d\mu = \sum_n \int_{E_n} f\,d\mu.$$

Answer. We have that $1_E = \sum_n 1_{E_n}$ and the sequence of partial sums is monotone. Suppose $f \ge 0$. By Monotone Convergence and linearity,

$$\int_E f\,d\mu = \int_X 1_E f\,d\mu = \sum_n \int_X 1_{E_n} f\,d\mu = \sum_n \int_{E_n} f\,d\mu.$$

For general f, if the integral exists, the equality follows by linearity.

(2.5.5) Let X be any set, $\Sigma = \mathcal{P}(X)$, $x_0 \in X$. Consider the Dirac measure (cfr. Exercise 2.3.6) $\delta : \Sigma \to [0, \infty]$ by

$$\delta(A) = \begin{cases} 1, & x_0 \in A \\ 0, & x_0 \notin A \end{cases}$$

Show that for any $f : X \to R$,

$$\int_X f\,d\delta = f(x_0).$$

Answer.

If $x_0 \in E$ we have

$$\int_X f\,d\delta = \int_E f\,d\delta$$

since $\delta(X \setminus E) = 0$. In particular

$$\int_X f\,d\delta = \int_{\{x_0\}} f\,d\delta = \int_X f\,1_{\{x_0\}}\,d\delta = \int_X f(x_0)\,1_{\{x_0\}}\,d\delta$$
$$= f(x_0)\delta(\{x_0\}) = f(x_0).$$

(2.5.6) Let $f : [0,\infty) \to \mathbb{C}$ be integrable with respect to Lebesgue measure, and such that

$$\int_{[0,x]} f\,dm = 0, \qquad x > 0.$$

Show that $f = 0$ a.e. *(This can be certainly done with the techniques from this chapter, but the argument might not be all that direct)*

Answer. For any $0 < a < b$,

$$\int_{(a,b]} f\,dm = \int_{[0,b]} f\,dm - \int_{[0,a]} f\,dm = 0 - 0 = 0.$$

By Dominated convergence

$$\int_{\{b\}} f\,dm = \lim_n \int_{(b-1/n,b]} f\,dm = 0.$$

So we also have $\int_{(a,b)} f\,dm = 0$ for all $0 < a < b$. Given any open set $V \subset [0,\infty)$, by Proposition 1.8.1 we can write $V = \bigcup_n (a_n, b_n)$ as a disjoint union. Then, by Exercise 2.5.4 (or, by Dominated Convergence),

$$\int_V f\,dm = \sum_n \int_{(a_n,b_n)} f\,dm = 0.$$

For any $E \in \mathcal{M}([0,\infty))$, by Proposition 2.3.25 we can find a decreasing sequence $\{V_n\}$ of open sets with $E = \bigcap_n V_n$ a.e. Then $1_E = \lim_n 1_{V_n}$ a.e. and by Dominated Convergence

$$\int_E f\,dm = \int f\,1_E\,dm = \lim_n \int f\,1_{V_n}\,dm = \lim_n \int_{V_n} f\,dm = 0.$$

Now we provide two different arguments for the rest of the proof.

(i) As any simple function is a linear combination of characteristic functions of measurable set, linearity of the integral implies that

$$\int_{[0,\infty)} f\,g\,dm = 0$$

for all g simple. Let $K_n = \{|f| \leq n\}$. On K_n the function $\overline{f}$ is bounded, so by Theorem 2.4.13 we have a uniform limit $\overline{f} = \lim_n g_n$ with g_n simple. Then

$$\int_{K_n} |f|^2\,dm = \left|\int_{K_n} |f|^2\,dm - \int_{K_n} fg_n\,dm\right| \leq \int_{K_n} |f|\,|\overline{f} - g_n|\,dm$$
$$\leq \sup_{K_n}\{|\overline{f}(x) - g_n(x)|\} \int_{[0,\infty)} |f|\,dm.$$

As $\|\bar{f} - g_n\|_\infty$ can be made arbitrarily small, we get that

$$\int_{K_n} |f|^2 \, dm = 0,$$

and hence $f = 0$ a.e. on K_n by Proposition 2.5.2. As this can be done for any n, $f = 0$ a.e.

(ii) Let $E = \{\operatorname{Re} f > 0\}$. As f is measurable, so is E. Then

$$0 = \operatorname{Re} \int_E f \, dm = \int_E \operatorname{Re} f \, dm.$$

By Proposition 2.5.2, $m(E) = 0$. In a similar way we can prove that $m(\{\operatorname{Re} f < 0\}) = 0$, $m(\{\operatorname{Im} f > 0\}) = 0$, and $m(\{\operatorname{Im} f < 0\}) = 0$. This implies that $m(\{f \neq 0\}) = 0$, so $f = 0$ a.e.

With more tools available, a very straightforward proof is available, see Exercise 2.11.2.

(2.5.7) *(Change of variable)* Let $(X, \mathcal{A}, \mu)$ be a measure space and Y a topological space. Let $g : X \to Y$ measurable. Let $\nu(E) = \mu(g^{-1}(E))$. Show that ν is a measure on $\mathcal{B}(Y)$ and that

$$\int_Y f \, d\nu = \int_X f \circ g \, d\mu \tag{2.23}$$

for all $f : Y \to \overline{\mathbb{R}}$ which are Borel measurable and such that either integral exists.

Answer. Because g is measurable and E is Borel, $g^{-1}(E) \in \mathcal{A}$ (Proposition 2.4.3); as preimages preserve all set operations and μ is a measure, ν is a measure.

If $E \in \mathcal{B}(Y)$ and $f = 1_E$,

$$\int_Y 1_E \, d\nu = \nu(E) = \mu(g^{-1}(E)) = \int_X 1_{g^{-1}(E)} \, d\mu = \int_X 1_E \circ g \, d\mu.$$

By linearity of the integral the equality holds for all simple functions. Theorem 2.4.13 and Monotone Convergence then give us the equality (2.23) for all non-negative measurable f.

Finally, for arbitrary f we write $f = f^+ - f^-$ and, because by hypothesis either the integral for f^+ or f^- is finite (on both sides!) the equality holds.

(2.5.8) Consider the interval $[0,1]$ with Lebesgue measure. Let $g : [0,1] \to \mathbb{C}$ measurable and such that $\int_{[0,1]} |g| < \infty$. Prove that the function

$$\gamma : s \longmapsto \int_0^s |g|$$

is continuous.

Answer. Since $|g| \geq 0$, the function γ is monotone non-decreasing. Let $s \in [0,1]$ and $\{s_n\}$ a sequence with $s_n \to s$. We denote by (s_n, s) the interval between these two numbers which could be (s, s_n) if $s_n > s$. The functions $1_{(s_n,s)} |g|$ converge pointwise to 0. By Dominated Convergence (since $|g|$ is integrable and $1_{(s_n,s)} |g| \leq |g|$) we have

$$\lim_{n\to\infty} |\gamma(s_n) - \gamma(s)| = \lim_{n\to\infty} \int_0^1 1_{(s_n,s)} |g| = \int_0^1 0 = 0.$$

So $\gamma(s_n) \to \gamma(s)$ for any such sequence, and thus γ is continuous at s.

(2.5.9) *(Change of variable, II)* Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to [0, \infty]$ measurable. Define $\nu : \mathcal{A} \to [0, \infty]$ by

$$\nu(E) = \int_E f \, d\mu.$$

(i) Show that ν is a measure on $(X, \mathcal{A})$ and that $\nu(E) = 0$ whenever $E \in \mathcal{A}$ with $\mu(E) = 0$.

(ii) Show that for any measurable function $g : X \to \overline{\mathbb{R}}$,

$$\int_X g \, d\nu = \int_X g f \, d\mu \tag{2.24}$$

if the left integral exists.

Answer.

(i) If $\displaystyle\int_X f \, 1_E \, d\mu = \nu(E) > 0$, then by Exercise 2.5.3 we have the fact that $\mu(\{f \, 1_E > 0\}) > 0$. As $\{f \, 1_E > 0\} \subset E$, we get that $\mu(E) > 0$. Thus $\mu(E) = 0$ forces $\nu(E) = 0$, showing the last part. Also, $\nu(\varnothing) = 0$.

Now let $\{E_k\} \subset \mathcal{A}$ be a countable family of pairwise disjoint sets. We have

$$1_{\bigcup_k E_k} = \sum_{k=1}^{\infty} 1_{E_k}.$$

The equality is easily checked by evaluating at each $x \in X$. Note that, for all x, terms in the series are all zero with the exception of at most one term. Then, by Monotone Convergence (in the second to last equality),

$$\nu\Big(\bigcup_k E_k\Big) = \int_{\bigcup_k E_k} f\,d\mu = \int_X f\,1_{\bigcup_k E_k}\,d\mu = \int_X \sum_k f\,1_{E_k}\,d\mu$$
$$= \sum_k \int_X f\,1_{E_k}\,d\mu = \sum_k \nu(E_k).$$

(ii) When $g = 1_E$,

$$\int_X g\,d\nu = \nu(E) = \int_E f\,d\mu = \int_X 1_E\,f\,d\mu = \int_X gf\,d\mu.$$

Since integrals are linear and simple functions are linear combinations of characteristic functions, this implies that

$$\int_X s\,d\nu = \int_X sf\,d\mu$$

for all s simple. Now if $g \geq 0$, by Theorem 2.4.13 there exists a sequence $\{s_n\}$ of simple functions such that $0 \leq s_n \nearrow g$. By Monotone Convergence (twice, and note that since $f \geq 0$ we also have $0 \leq s_n f \nearrow gf$),

$$\int_X g\,d\nu = \lim_n \int_X s_n\,d\nu = \lim_n \int_X s_n f\,d\mu = \int_X gf\,d\mu.$$

For arbitrary g, the equality (2.24) holds for g^+ and g^-, and so it holds for g if the integral exists; note that since $f \geq 0$, $(gf)^+ = g^+ f$, $(gf)^- = g^- f$.

(2.5.10) *(Change of variable, III)* Use Exercises 2.5.7, 2.5.9 and 1.8.19 to show that if $g : [a, b] \to [c, d]$ is continuously differentiable then (2.23) gives the usual calculus change of variable (substitution) formula

$$\int_{g(a)}^{g(b)} f(t)\,dt = \int_a^b f(g(t))\,g'(t)\,dt.$$

Answer. Take (a_k, b_k) and (a'_k, b'_k) as in Exercise 1.8.19 applied to the function g. Since $g' = 0$ on each (a_k, a'_k) and each (b'_k, b_k) we have

$$\int_a^b f(g(t))\,g'(t)\,dt = \sum_k \int_{a'_k}^{b'_k} f(g(t))\,g'(t)\,dt,$$

and (since $g(a'_k) = g(a_k)$ and $g(b'_k) = g(b_k)$)

$$\int_{g(a)}^{g(b)} f(t)\,dt = \sum_k \int_{g(a_k)}^{g(b_k)} f(t)\,dt = \sum_k \int_{g(a'_k)}^{g(b'_k)} f(t)\,dt.$$

The above equalities show that we can assume, without loss of generality, that g is strictly monotone on $[a,b]$, which makes it invertible by Exercise 1.8.38.

Assume first that $g'>0$. Take $X=[a,b]$, $Y=[c,d]$, and μ the measure given by

$$\mu(E)=\int_E g'\,dm.$$

Let ν be the measure $\nu(E)=\mu(g^{-1}(E))$. For any $[c',d']\subset[c,d]$,

$$\begin{aligned}\nu([c',d'])=\mu(g^{-1}([c',d']))&=\int_{g^{-1}(c')}^{g^{-1}(d')}g'(t)\,dt\\&=g(g^{-1}(d'))-g(g^{-1}(c'))=d'-c'=m([c',d']).\end{aligned}$$

As they agree on intervals, $\nu=m$ on Borel sets. Then, using first (2.23) and then (2.24),

$$\int_c^d f(t)\,dt=\int_{[c,d]}f\,d\nu=\int_{[a,b]}f\circ g\,d\mu=\int_{g^{-1}(c)}^{g^{-1}(d)}f(g(t))\,g'(t)\,dt.$$

As $c=g(a)$ and $d=g(b)$, we are done.

When $g'<0$, we can replace g' with $-g'$ in the definition of μ, and all that this does is to account for the reverse the order of $g^{-1}(c')$ and $g^{-1}(d')$.

(2.5.11) Use Dominated Convergence and Exercise 1.3.8 to show that if $f:\mathbb{R}\to[0,\infty)$ is piecewise continuous with jump discontinuities, then with respect to Lebesgue measure

$$\int_{[a,b]}f\,dm=\int_a^b f(t)\,dt.$$

Answer. Because all finitely many discontinuities are jump discontinuities, f is bounded. We also get that f is measurable, as it can be written as a linear combination of products of characteristic functions and continuous functions.

If $\{a_k\}$ are the discontinuity points of f, then f is continuous on each interval (a_k,a_{k+1}) and (with $a_0=a$, $a_n=b$)

$$\int_{[a,b]}f\,dm=\sum_{k=1}^n\int_{a_{k-1}}^{a_k}f\,dm,$$

$$\int_a^b f(t)\,dt=\sum_{k=1}^n\int_{a_{k-1}}^{a_k}f(t)\,dt.$$

So we may assume without loss of generality that f is continuous. From (1.4) in Exercise 1.3.8 we get that the Riemann integral is a limit of Lebesgue integrals of simple functions. As f is continuous on $[a,b]$ is it uniformly continuous, and so the simple functions

$$\sum_{k=1}^{n} f\Big(a + \frac{k(b-a)}{n}\Big)\, 1_{\left[\frac{(k-1)(b-a)}{n}, \frac{k(b-a)}{n}\right]}$$

converge to f. By Dominated Convergence, their integrals converge to the Lebesgue integral of f.

(2.5.12) Show that a linear combination of integrable functions is integrable.

Answer. If f, g are integrable and $\lambda \in \mathbb{C}$, then

$$\int_X |f + \lambda g|\, d\mu \leq \int_X \big(|f| + |\lambda|\,|g|\big)\, d\mu = \int_X |f| + |\lambda| \int_X |g| < \infty.$$

(2.5.13) Compute $\displaystyle\int_{[0,1]} \alpha\, dm$, where α is Cantor's ternary function.

Answer. Since the Cantor set is a nullset, we can simply calculate the integral on its complement. So we use that α is equal a.e. to

$$\sum_{k=1}^{\infty} \sum_{j=1}^{2^{k-1}} \frac{2j-1}{2^k}\, 1_{(b_{k,j}/3^k, (b_{k,j}+1)/3^k)},$$

where the numbers $b_{k,j}$ are the left endpoints of the ternary intervals $C_{k,j}$ of length 3^{-k} that are removed in the k^{th} step. Note that we don't need to know the $b_{k,j}$ since, because α is constant on the interval, all that matters for the integral is the length of the interval times the value of the function. On the interval $(b_{k,j}/3^k, (b_{k,j}+1)/3^k)$, the function α takes the value $(2j-1)/2^k$.

Then

$$\begin{aligned}
\int_{[0,1]} \alpha\, dm &= \sum_{k=1}^{\infty} \sum_{j=1}^{2^{k-1}} \frac{2j-1}{2^k} \frac{1}{3^k} = \sum_{k=1}^{\infty} \frac{1}{6^k} \sum_{j=1}^{2^{k-1}} 2j - 1 \\
&= \sum_{k=1}^{\infty} \frac{1}{6^k} \left(2 \frac{2^{k-1}(2^{k-1}+1)}{2} - 2^{k-1} \right) \\
&= \sum_{k=1}^{\infty} \frac{2^k}{6^k} = \sum_{k=1}^{\infty} \frac{1}{3^k} = \frac{1/3}{1-1/3} \\
&= \frac{1}{2}.
\end{aligned}$$

A simpler possibility is to notice that $\alpha(1-x) = 1-\alpha(x)$ (this is obvious visually, and not too hard to show by looking at the intervals symmetric with respect to $x = 1/2$; see Exercise 2.2.12). Then

$$\int_0^1 \alpha(x)\, dx = \int_0^1 \alpha(1-x)\, dx = \int_0^1 (1-\alpha(x))\, dx = 1 - \int_0^1 \alpha(x)\, dx,$$

and the result follows.

Yet another way, if we know that the sequence $\{f_n\}$ as in Section 2.2 converges to α, is to notice that

$$\int_0^1 f_{n+1} = \int_0^{1/3} \frac{1}{2} f_n(3x)\, dx + \frac{1}{6} + \int_{2/3}^1 \left[\frac{1}{2} + \frac{1}{2} f_n(3x-2)\right] dx = \frac{1}{3} + \frac{1}{3} \int_0^1 f_n.$$

So Dominated Convergence gives us the relation $L = \frac{1}{3} + \frac{L}{3}$, from where $L = \frac{1}{2}$.

(2.5.14) Consider the function $f : [0,1] \to \mathbb{R}$ given by the series

$$f(x) = \sum_{n=0}^{\infty} \frac{x}{(1+x)^n}.$$

(i) Show that $f(x) = x + 1$ if $x \in (0,1]$, and that $f(0) = 0$.

(ii) Does the series converge uniformly on $[0,1]$?

(iii) Is is true that

$$\int_0^1 \left(\sum_{n=0}^{\infty} \frac{x}{(1+x)^n} \right) dx = \sum_{n=0}^{\infty} \int_0^1 \frac{x}{(1+x)^n}\, dx\,?$$

Answer.

(i) It is clear that $f(0) = 0$. When $x > 0$,

$$\sum_{n=0}^{\infty} \frac{x}{(1+x)^n} = \frac{x}{1 - \frac{1}{1+x}} = x + 1.$$

(ii) The series does not converge uniformly on $[0, 1]$. If it did, its limit would be continuous.

One can also check this explicitly:

$$\left| f(x) - \sum_{n=0}^{N} \frac{x}{(1+x)^n} \right| = \left| \sum_{n=N+1}^{\infty} \frac{x}{(1+x)^n} \right| = x \frac{\frac{1}{(1+x)^{N+1}}}{1 - \frac{1}{1+x}} = \frac{1}{(1+x)^N}.$$

For fixed N, taking x very close to 0 allows us to have the difference as close to 1 as we want.

(iii) Yes. The series has positive terms, so Monotone Convergence (more specifically, Corollary 2.5.8) applies. Let us verify:

$$\int_0^1 \left(\sum_{n=0}^{\infty} \frac{x}{(1+x)^n} \right) dx = \int_0^1 (x+1)\, dx = \frac{3}{2},$$

while (applying dominated convergence to exchange the integral with the series, and noting that the series is telescopic)

$$\begin{aligned}
\sum_{n=0}^{\infty} \int_0^1 \frac{x}{(1+x)^n}\, dx &= \int_0^1 x\, dx + \int_0^1 \frac{x}{1+x}\, dx + \int_0^1 \frac{x}{(1+x)^2} \\
&\quad + \sum_{n=3}^{\infty} \int_0^1 \frac{x}{(1+x)^n}\, dx \\
&= \frac{1}{2} + (1 - \log 2) + (\log 2 - \frac{1}{2}) + \sum_{n=3}^{\infty} \int_0^1 \frac{x}{(1+x)^n}\, dx \\
&= 1 + \sum_{n=3}^{\infty} \int_0^1 \left(\frac{1}{(1+x)^{n-1}} - \frac{1}{(1+x)^n} \right) dx \\
&= 1 + \int_0^1 \sum_{n=3}^{\infty} \left(\frac{1}{(1+x)^{n-1}} - \frac{1}{(1+x)^n} \right) dx \\
&= 1 + \int_0^1 \frac{1}{(1+x)^2}\, dx \\
&= 1 + \frac{1}{2} = \frac{3}{2}.
\end{aligned}$$

(2.5.15) Prove that

$$s = \lim_{k\to\infty} \sum_{n=1}^{\infty} \exp\left(-n + \frac{k}{n}\, e^{-k/n}\right)$$

exists and find its value.

Answer. We have

$$\exp\left(-n + \frac{k}{n}\, e^{-k/n}\right) = \exp(-n) \exp\left(\frac{k}{n}\, e^{-k/n}\right).$$

This function is always positive, and

$$\exp\left(\frac{k}{n}\, e^{-k/n}\right) \le \max\{xe^{-x} :\ x \in [0,\infty)\} = e^{-1}.$$

Then

$$\exp\left(-n + \frac{k}{n}\, e^{-k/n}\right) \le \exp(e^{-1})\, e^{-n} \le 2e^{-n},$$

which is integrable, and so by Dominated Convergence

$$\begin{aligned}\lim_{k\to\infty} \sum_{n=1}^{\infty} \exp\left(-n + \frac{k}{n}\, e^{-k/n}\right) &= \sum_{n=1}^{\infty} \exp\left(-n + \lim_{k\to\infty} \frac{k}{n}\, e^{-k/n}\right) \\ &= \sum_{n=1}^{\infty} e^{-n} = \frac{e^{-1}}{1-e^{-1}} = \frac{1}{e-1}.\end{aligned}$$

(2.5.16) Find

$$\lim_{n\to\infty} \int_0^n x^{1/n}\, e^{-x}\, dx.$$

Answer. We have

$$\int_0^n x^{1/n}\, e^{-x}\, dx = \int_0^\infty x^{1/n}\, e^{-x}\, 1_{[0,n]}\, dx$$

We also have

$$0 \le x^{1/n}\, e^{-x}\, 1_{[0,n]} \le 1_{[0,1]} + xe^{-x},$$

which is integrable. By Dominated Convergence we get

$$\lim_{n\to\infty} \int_0^n x^{1/n}\, e^{-x}\, dx = \lim_{n\to\infty} \int_0^\infty x^{1/n}\, e^{-x}\, 1_{[0,n]}\, dx = \int_0^\infty e^{-x}\, dx = 1.$$

(2.5.17) Fix $\alpha, \beta > 0$.

(i) Show that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\alpha + n\beta} = \int_0^1 \frac{x^{\alpha-1}}{1+x^\beta}\,dx.$$

(ii) Show that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log 2.$$

(iii) Show that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}.$$

Answer.

(i) We have

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{(-1)^n}{\alpha + n\beta} &= \sum_{n=0}^{\infty} (-1)^n \int_0^1 x^{\alpha-1+n\beta}\,dx \\
&= \lim_{M\to\infty} \sum_{n=0}^{M} (-1)^n \int_0^1 x^{\alpha-1+n\beta}\,dx \\
&= \lim_{M\to\infty} \int_0^1 x^{\alpha-1} \sum_{n=0}^{M} (-1)^n x^{n\beta}\,dx \\
&= \lim_{M\to\infty} \int_0^1 x^{\alpha-1} \frac{1-(-1)^{M+1}x^{n\beta(M+1)}}{1+x^\beta}\,dx \\
&= \int_0^1 \frac{x^{\alpha-1}}{1+x^\beta}\,dx.
\end{aligned}$$

The last equality is justified using Dominated Convergence, using as upper bound the function $g(x) = \frac{2x^{\alpha-1}}{1+x^\beta}$.

An alternative proof that does not use measure theory is as follows. Consider the power series

$$f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n\, x^{\alpha+n\beta}}{\alpha+n\beta}.$$

One can check that the radius of convergence is 1, so for $0 \le x < 1$ we can differentiate term by term. As $f(0) = 0$ (since $\alpha, \beta > 0$) and f' is continuous,

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\alpha + n\beta} = f(1) - f(0) = \int_0^1 f'(x)\,dx = \int_0^1 \sum_{n=0}^{\infty} (-1)^n\, x^{\alpha+n\beta-1}\,dx$$
$$= \int_0^1 x^{\alpha-1} \sum_{n=0}^{\infty} (-x^{\beta})^n\,dx$$
$$= \int_0^1 \frac{x^{\alpha-1}}{1+x^{\beta}}\,dx.$$

(ii) Taking $\alpha = \beta = 1$ above, we get

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \sum_{n=0}^{\infty} \frac{(-1)^n}{1+n} = \int_0^1 \frac{1}{(1+x)}\,dx = \log 2.$$

(iii) Taking $\alpha = 1$ and $\beta = 2$, we get

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \int_0^1 \frac{1}{1+x^2}\,dx = \frac{\pi}{4}.$$

(2.5.18) If $a_{n,m} \geq 0$ for all $n, m \in \mathbb{N}$, use Monotone Convergence to show that

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m}.$$

Show by example that the equality can fail in general.

Answer. Because each $a_{n,m} \geq 0$, the series $\sum_{m=1}^{\infty} a_{n,m}$ always exists (even if it is infinite) as a non-negative function of n. By Monotone Convergence and seeing the sum over n as an integral with respect to the counting measure,

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \sum_{n=1}^{\infty} \lim_M \sum_{m=1}^{M} a_{n,m} = \lim_M \sum_{n=1}^{\infty}\sum_{m=1}^{M} a_{n,m}$$
$$= \lim_M \sum_{m=1}^{M}\sum_{n=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m}.$$

As for the example, consider

$$a_{n,m} = \frac{(-1)^n}{m + \left\lfloor \frac{n+1}{2} \right\rfloor}.$$

Since $a_{2n-1,m} + a_{2n,m} = 0$ and $a_{n,m} \to 0$ with n, we have that $\sum_n a_{n,m} = 0$. Then

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m} = 0.$$

Meanwhile,

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \sum_{n=1}^{\infty}(-1)^n \sum_{m=1}^{\infty} \frac{1}{m + \left\lfloor \frac{n+1}{2} \right\rfloor}$$

does not exist, as the inner series is infinite and so the alternating series on n is not defined.

Another more or less canonical example is to take

$$a_{n,m} = \begin{cases} 1, & n = m \\ -1, & n = m+1 \\ 0, & \text{otherwise} \end{cases}$$

For fixed m, we have $a_{n,m} = 0, 0, \cdots, 0, 1, -1, 0, \cdots$ with the 1 and -1 in the m and $m+1$ positions. Thus

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m} = 0.$$

Meanwhile, when n is fixed the same happens, but there is an exception for $n = 1$. In that case, we have $a_{1,m} = 1, 0, \ldots$. Therefore

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = 1.$$

(2.5.19) Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{X_n\} \subset \mathcal{A}$ a non-decreasing sequence such that $X = \bigcup_n X_n$. Show that, for any $f \geq 0$ measurable,

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_{X_n} f\, d\mu.$$

Answer. The sequence $\{f\, 1_{X_n}\}$ is increasing, since $f \geq 0$ and $1_{X_{n+1}} \geq 1_{X_n}$. Then, by Monotone Convergence,

$$\lim_{n\to\infty} \int_{X_n} f\, d\mu = \lim_{n\to\infty} \int_X f\, 1_{X_n}\, d\mu = \int_X \lim_{n\to\infty} f\, 1_{X_n}\, d\mu = \int_X f\, d\mu.$$

(2.5.20) Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{C}$ integrable and such that

$$\Big|\int_X f\,d\mu\Big| = \int_X |f|\,d\mu. \tag{2.25}$$

Show that there exists $\alpha \in \mathbb{C}$, with $|\alpha| = 1$ and $g : X \to [0, \infty)$ integrable such that $f = \alpha g$ a.e.

Answer. If $\int_X f\,d\mu = 0$, then by (2.25) $\int_X |f|\,d\mu = 0$ and so $f = 0$ a.e. Otherwise, let

$$\beta = \frac{\Big|\int_X f\,d\mu\Big|}{\int_X f\,d\mu}.$$

Then $|\beta| = 1$ and

$$\int_X |f|\,d\mu = \Big|\int_X f\,d\mu\Big| = \beta\int_X f\,d\mu = \int_X \beta f\,d\mu.$$

So

$$\int_X (|f| - \beta f)\,d\mu = 0.$$

Looking at the real part,

$$0 = \int_X (|f| - \operatorname{Re}\beta f)\,d\mu.$$

As $|\operatorname{Re}\beta f| \leq |\beta f| = |f|$, we get from Exercise 2.5.3 that $|f| = \operatorname{Re}\beta f$ a.e. As $|\beta f| = |f|$, this forces $\beta f = |f|$ a.e. Because if $\operatorname{Im}\beta f \neq 0$, then $|f| = \operatorname{Re}\beta f < |\beta f| = |f|$.

Now $|f| = \beta f$, so we can define $g = |f|$ and take the scalar to be $\alpha = \beta^{-1}$.

(2.5.21) Let $(X, \mathcal{A}, \mu)$ be a measure space. If $f : X \to \mathbb{C}$ is measurable and $E \in \mathcal{A}$ with $0 < \mu(E) < \infty$, show that

$$\frac{1}{\mu(E)}\int_E f\,d\mu \in \overline{\operatorname{conv}}\,f(E),$$

the closed convex hull of $f(E)$; that is, the average of f is a limit of convex combinations of values of f.

Answer. Assume first that $f \geq 0$. We will construct a sequence of simple functions that increase to f. The only difference with the simple functions from Theorem 2.4.13 will be that instead of taking the coefficients to be $\frac{k-1}{2^n}$, we will take some value of f inside the interval $\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]$. So we can write

$$s_n = \sum_{j=1}^{k(n)} \lambda_{n,j}\, E_{n,j},$$

where $\{E_{n,j}\}_n$ is a partition of E and $\lambda_{n,j} \in f(E)$. Since $E = \bigcup_{j=1}^{k(n)} E_j$ is a disjoint union, we have

$$\sum_{j=1}^{k(n)} \frac{\mu(E_j)}{\mu(E)} = 1,$$

so the numbers $\mu(E_j)/\mu(E)$ are convex coefficients and

$$\frac{1}{\mu(E)} \int_E s_n\, d\mu = \sum_{j=1}^{k(n)} \frac{\mu(E_j)}{\mu(E)}\, \lambda_{n,j} \in \operatorname{conv} f(E).$$

Then, using Monotone Convergence,

$$\frac{1}{\mu(E)} \int_E f\, d\mu = \lim_n \frac{1}{\mu(E)} \int_E s_n\, d\mu \in \overline{\operatorname{conv}} f(E).$$

Now if $f : X \to \mathbb{R}$, by as above we construct sequences of simple functions $\{s_n^+\}$ and $\{s_n^-\}$ such that $s_n^+ \nearrow f^+$ and $s_n^- \nearrow f^-$ and choosing the coefficients so that $s_n^+(x) - s_n^-(x) \in f(E)$; this is achieved by using the same x in each interval $\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]$. Then

$$\frac{1}{\mu(E)} \int_E f\, d\mu = \lim_n \frac{1}{\mu(E)} \int_E (s_n^+ - s_n^-)\, d\mu \in \overline{\operatorname{conv}} f(E).$$

When $f : X \to \mathbb{C}$, again we may arrange the simple functions so that when we take the averages we get elements in $\operatorname{conv} f(E)$.

(2.5.22) Let $(X, \mathcal{A}, \mu)$ be a measure space such that $\mu(X) < \infty$. Let $f : X \to \mathbb{C}$ be integrable, and $S \subset \mathbb{C}$ closed, and such that

$$\frac{1}{\mu(E)} \int_E f\, d\mu \in S, \qquad \text{whenever } E \in \mathcal{A},\ \mu(E) > 0.$$

Show that $f(x) \in S$ a.e.

Answer. As S^c is open, $S^c = \bigcup_n D_n$, a countable union of disks. Fix n and let $D \subset D_n$ be a closed disk. If $\mu(f^{-1}(D)) > 0$, we would have

$$\frac{1}{\mu(f^{-1}(D))} \int_{f^{-1}(D)} f\, d\mu \in S,$$

a contradiction since an average as above is a limit of convex combinations of values of f and so it is in $D \subset S^c$. As this can be done for any disk $D \subset D_n$, we conclude that $\mu(D_n) = 0$. Then $\mu(S^c) = 0$ (countable union of nullsets), and so $f(x) \in S$ a.e.

(2.5.23) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $\{f_n\}$ a sequence of measurable functions such that $f_n \to 0$ uniformly. Does this imply that $\displaystyle\int_X f_n \, d\mu \to 0$?

Answer. In a finite measure space, yes. If $\mu(X) < \infty$, given $\varepsilon < 0$ there exists n_0 such that $|f_n| < \varepsilon/\mu(X)$ for all $n > n_0$. Then, for such n,

$$\Big| \int_X f_n \, d\mu \Big| \leq \int_X |f_n| \, d\mu < \varepsilon.$$

When $\mu(X) = \infty$, the assertion can fail. On the real line, let $f_n = \frac{1}{n}$. Then $f_n \to 0$ uniformly, but $\displaystyle\int_{\mathbb{R}} |f_n| \, dm = \infty$ for all n.

(2.5.24) Let $f : X \to [0, \infty]$ measurable, with $c = \displaystyle\int_X f \, d\mu < \infty$, and $\alpha > 0$. Show that

$$\lim_{n\to\infty} \int_X n \log\left[1 + \left(\frac{f}{n}\right)^\alpha\right] d\mu = \begin{cases} \infty, & 0 < \alpha < 1 \\ c, & \alpha = 1 \\ 0, & 1 < \alpha < \infty \end{cases}$$

Answer. As a preliminary task, let us establish some basic calculus inequalities. If $g(x) = e^x - 1 - x$, then $g(0) = 0$ and $g'(x) = e^x - 1$. So the only critical point is $x = 0$. As $g''(x) = e^x > 0$, said only critical point is a local minimum, anad it has to be a global minimum since the function is smooth everywhere. It follows that

$$1 + x \leq e^x, \qquad x \in \mathbb{R}.$$

When $x \geq -1$, taking logarithm

$$\log(1 + x) \leq x, \qquad x \in (-1, \infty).$$

For $x \geq -1$, $n \in \mathbb{N}$ and applying the above inequality to $\frac{x}{n}$,

$$n \log(1 + \frac{x}{n}) \leq x. \tag{AB.2.3}$$

Exponentiating, we obtain

$$\left(1 + \frac{x}{n}\right)^n \leq e^x \qquad x \in (-1, \infty). \tag{AB.2.4}$$

Applying the Mean Value Theorem to $h(x) = \log(1+x)$,

$$\frac{\log(1+x)}{x} = \frac{h(x)-h(0)}{x} = \frac{h'(\xi)x}{x} = g'(\xi) = \frac{1}{1+\xi}$$

for ξ between 0 and x. As $x \to 0$ we have $\xi \to 0$, and so the limit is 1. A by-product of this is

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

Back to our original problem, we note that the integrand is zero when f is zero. So by restricting to an appropriate subset, we may assume that $f > 0$ everywhere. We also assume that $\mu(X) > 0$, for otherwise all integrals are zero.

Assume first that $0 < \alpha < 1$. We have, for $t > 0$,

$$n\,\log\left[1+\left(\frac{t}{n}\right)^\alpha\right] = t^\alpha\, n^{1-\alpha}\left(\frac{t}{n}\right)^{-\alpha}\log\left[1+\left(\frac{t}{n}\right)^\alpha\right] \xrightarrow[n\to\infty]{} \infty \qquad \text{(AB.2.5)}$$

(using that $\lim_{x\to 0} x^{-1}\log(1+x) = 1$). By Fatou's Lemma,

$$\liminf_{n\to\infty}\int_X n\,\log\left[1+\left(\frac{f}{n}\right)^\alpha\right]d\mu \geq \int_X \lim_{n\to\infty} n\,\log\left[1+\left(\frac{f}{n}\right)^\alpha\right]d\mu = \infty.$$

When $\alpha = 1$, by our Calculus homework (AB.2.3) we have

$$n\,\log\left[1+\frac{f}{n}\right] \leq f.$$

By Dominated Convergence,

$$\lim_{n\to\infty}\int_X n\,\log\left[1+\frac{f}{n}\right]d\mu = \int_X \lim_n n\,\log\left[1+\frac{f}{n}\right]d\mu = \int_X f = c.$$

And when $\alpha > 1$, from $\log(1+x) \leq x$ we obtain $\alpha\,\log(1+x) \leq \alpha x$; applying the exponential,

$$(1+x)^\alpha \leq e^{\alpha x}.$$

Then

$$1 + x^\alpha \leq (1+x)^\alpha \leq e^{\alpha x},$$

which gives

$$\log(1+x^\alpha) \leq \alpha x, \qquad x > 0.$$

Thus

$$n\,\log\left[1+\left(\frac{f}{n}\right)^\alpha\right] \leq n\,\alpha\left(\frac{f}{n}\right) = \alpha f.$$

Then, Dominated Convergence applies. We can use (AB.2.5), where now $\alpha > 1$ gives us that the limit is 0. Thus

$$\lim_{n\to\infty}\int_X n\,\log\left[1+\left(\frac{t}{n}\right)^\alpha\right]d\mu = \int_X \lim_n n\,\log\left[1+\left(\frac{t}{n}\right)^\alpha\right]d\mu = 0.$$

(2.5.25) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $f : X \to [0, 1]$ integrable. Find

$$\lim_{n\to\infty} \int_X f^n \, d\mu.$$

Answer. Let $A = \{f = 1\}$. Since f is measurable, $A \in \mathcal{A}$. Because $f(X) \subset [0, 1]$, we have that $|f^n| \leq |f|$, and so Dominated convergence applies. We can then write

$$\int_X f^n \, d\mu = \int_A f^n \, d\mu + \int_{X\setminus A} f^n \xrightarrow[n\to\infty]{} \int_A 1 \, d\mu = \mu(A).$$

(2.5.26) Let $(X, \mathcal{A})$ be a measurable space and $f : X \to [0, \infty]$ measurable. Show that

$$\int_0^\infty 1_{f^{-1}[t,\infty)}(s) \, dt = f(s), \qquad s \in X. \tag{2.26}$$

This is sometimes called the *layer-cake representation* of f.

Answer. Let $g(t) = 1_{f^{-1}[t,\infty)}(s)$. We have

$$g(t) = 1_{f^{-1}[t,\infty)}(s) = 1_{[t,\infty)}(f(s)) = 1_{[0,f(s)]}(t).$$

Now we can see that g is measurable, since

$$g^{-1}[r, \infty) = \{r : \ g(r) \geq t\} = \begin{cases} \varnothing, & r > 1 \\ f^{-1}[t, \infty), & 0 < r \leq 1 \\ [0, \infty), & r = 0 \end{cases}$$

Knowing that g is measurable and nonnegative, the integral exists and

$$\int_0^\infty 1_{f^{-1}[t,\infty)}(s) \, dt = \int_0^\infty 1_{[0,f(s)]}(t) \, dt = \int_0^{f(s)} 1 \, dt = f(s).$$

(2.5.27) Find a sequence $\{f_n\}$ of continuous functions $f_n : [0, 1] \to \mathbb{R}$ such that

- $0 \leq f_n \leq 1$;
- for each $x \in [0, 1]$, $\lim_n f_n(x)$ does not exist;

- $\lim_n \int_0^1 f_n = 0.$

Answer. We are going to construct a "travelling bump" that goes along $[0,1]$ over and over again (to make the limit fail to exist) while getting thinner (to make the integrals go to zero). Initially we index our sequence with two positive integers m, n, with $1 \leq m \leq 2^{n+1} - 3$.

So we define, for $m \geq 2$,

$$g_{m,n} = \begin{cases} 0, & \frac{m-1}{2^{n+1}} \leq x \leq \frac{m}{2^{n+1}} \\ 2^{n+1}x - m, & \frac{m}{2^{n+1}} \leq x \leq \frac{m+1}{2^{n+1}} \\ 1, & \frac{m+1}{2^{n+1}} \leq x \leq \frac{m+2}{2^{n+1}} \\ -2^{n+1}x + m + 3, & \frac{m+2}{2^{n+1}} \leq x \leq \frac{m+3}{2^{n+1}} \\ 0, & \frac{m+3}{2^{n+1}} \leq x \leq 1 \end{cases}$$

and for $m = 1$,

$$g_{1,n} = \begin{cases} 1 - 2^{n+1}x, & 0 \leq x \leq \frac{1}{2^{n+1}} \\ 0, & \frac{1}{2^{n+1}} \leq x \leq 1 - \frac{1}{2^{n+1}} \\ 2^{n+1}x + 1 - 2^{n+1}, & 1 - \frac{1}{2^{n+1}} \leq x \end{cases}$$

By construction each $g_{m,n}$ is continuous and $0 \leq g_{m,n} \leq 1$. We have (thinking of the rectangle that contains the trapezoid)

$$\int_0^1 g_{m,m} \leq \frac{3}{2^{n+1}}$$

so the sequence of integrals go to zero as the indices increase. Finally, given any $x \in (0,1)$, we can choose a subsequence $\{m_k\}$ such that $\frac{m_k+1}{2^{k+1}} \leq x \leq \frac{m_k+2}{2^{k+1}}$; then $g_{m_k,k}(x) = 1$ for all x, so the limit of the subsequence at x is 1. But we can also choose m_k such that, for large enough n, $x \leq \frac{m_k}{2^{k+1}}$, and we get another subsequence with $g_{m_k,k}(x) = 0$ for all sufficiently large k. For the cases $x = 0$ and $x = 1$ we can consider the subsequence $g_{1,n}$, with $g_{1,n}(x) = 1$ for all n; and the subsequence $g_{2^n,n}$ with $g_{2^n,n}(x) = 0$. So the limit doesn't exist either at 0 nor 1.

(2.5.28) Find a sequence of continuous functions $f_n : [0,1] \to [0,\infty)$ such that

- $\lim_n f_n(x) = 0$ for all x;

- $\lim_n \int_0^1 f_n = 0$;
- $\lim_n \int_0^1 \sup f_n = \infty$.

(note that these functions fail the hypotheses of DCT but they do satisfy the conclusion)

Answer. In this case we want functions that grow higher while losing area. We may choose

$$f_n(x) = \begin{cases} \sqrt{n} - n^{3/2}|x - \frac{1}{n}|, & 0 \leq x \leq \frac{2}{n} \\ 0, & \frac{2}{n} \leq x \end{cases}$$

Then $\max\{f_n(x) : x \in [0,1]\} = \sqrt{n}$, so $\int_0^1 \sup f_n = \sqrt{n}$. We have $f_n(0) = 0$ for all n, and if $x > 0$, we get $f_n(x) = 0$ for all $n > 2/x$; so $\lim_n f_n(x) = 0$. Finally,

$$\begin{aligned} \int_0^1 f_n &= \int_0^{1/n} f_n + \int_{1/n}^{2/n} f_n \\ &= \int_0^{1/n} \left(\sqrt{n} - n^{3/2}(\frac{1}{n} - x)\right) dx + \int_{1/n}^{2/n} \left(\sqrt{n} - n^{3/2}(x - \frac{1}{n})\right) dx \\ &= \frac{2}{\sqrt{n}}. \end{aligned}$$

2.6. Some More Topology

(2.6.1) Show that If X is a Hausdorff topological space and $K_1, K_2 \subset X$ are disjoint compact sets, there exist disjoint open sets V_1, V_2 with $K_1 \subset V_1$ and $K_2 \subset V_2$. Conclude that a compact Hausdorff space is normal.

Answer. By Lemma 2.6.3, for each $x \in K_2$ there exists V_x open with $K_1 \subset V_x$ and W_x open with $x \in W_x$ and $V_x \cap W_x = \varnothing$. We have $K_1 \subset \bigcup_{x \in K_2} V_x$, so by compactness there exist $x_1, \dots, x_n$ with $K_1 \subset \bigcup_{j=1}^n V_{x_j}$. Let $V = \bigcup_{j=1}^n V_{x_j}$ and $W = \bigcap_{j=1}^n W_{x_j}$. These two sets are open, disjoint, with $K_1 \subset V$ and $K_2 \subset W$.

When X is compact Hausdorff and $C_1, C_2 \subset X$ are closed and disjoint, they are compact by Lemma 1.8.16, so the above applies.

(2.6.2) Let T be a Hausdorff topological space and $f : T \to \mathbb{C}$ continuous. Show that the following statements are equivalent:

(i) f vanishes at infinity;

(ii) for each $\varepsilon > 0$, the set $\{|f| \geq \varepsilon\}$ is compact.

Answer. Suppose first that f vanishes at infinity and fix $\varepsilon > 0$. By definition there exists $K \subset T$, compact, with $|f| < \varepsilon$ outside of K. This means that the closed set $\{|f| \geq \varepsilon\}$ lies inside K. Being a closed subset of a compact set, it is compact.

Conversely, suppose that $\{|f| \geq \varepsilon\}$ is compact for all $\varepsilon > 0$. If we fix one such $\varepsilon > 0$, we can take $K = \{|f| \geq \varepsilon\}$, and then $|f| < \varepsilon$ outside of K.

(2.6.3) Show that if $X = [0, 1]$ with the usual topology, $\mathcal{A} = \mathcal{P}(X)$, and μ is the counting measure, there exists measurable $f : X \to \mathbb{C}$ such that the locally compact version of Lusin's Theorem fails. Discuss which hypothesis of the theorem is not satisfied in this example.

Answer. Take $f = 1_{\mathbb{Q}}$. For any $\varepsilon < 1$, if $\mu(B) < \varepsilon$ then $B = \varnothing$. So we need $f = g$ everywhere, but f is not continuous anywhere. The hypothesis that is not satisfied is that μ is not outer regular: a singleton $\{x_0\}$ has $\mu(\{x_0\}) = 1$, but any open set that contains $\{x_0\}$ has infinite measure.

(2.6.4) Show that if $X = \mathbb{R}$ with the usual topology, $\mathcal{A} = \mathcal{B}(X)$ is the Borel σ-algebra, and μ is the Lebesgue measure, there exists measurable $f : X \to \mathbb{C}$ (taking nonzero values in a set of infinite measure) such that Lusin's Theorem fails.

Answer. Take $f = 1$. If g has compact support, then the set where $f \neq g$ has infinite measure.

(2.6.5) In the proof of Theorem 2.6.9, write the details of the argument that the inequalities in (2.30) show that the series converges to f on K, and that it converges uniformly on T.

Answer. We can rewrite the first inequality as

$$\Big|f - \sum_{k=1}^{n} f_k\Big| \leq \Big(\frac{2}{3}\Big)^n.$$

This shows that $f = \sum_n f_n$, uniformly, on K. For the everywhere uniform convergence, the tails of the series satisfy

$$\Big|\sum_{k>n} f_n\Big| \leq \sum_{k>n} |f_n| \leq \frac{1}{3}\sum_{k>n}\Big(\frac{2}{3}\Big)^{k-1} = \frac{1}{3}\,\frac{\big(\frac{2}{3}\big)^n}{1-\frac{2}{3}} = \Big(\frac{2}{3}\Big)^n.$$

So the sequence of partial sums is uniformly Cauchy, thus showing that the series converges uniformly everywhere on T.

(2.6.6) Let $K \subset \mathbb{R}$ be closed. Is K the support of a continuous real or complex-valued function? If it is, show how to construct such a function; if it isn't, describe which closed sets are supports of continuous functions. Does your answer apply to other topological spaces?

Answer. In general, no. The following reasoning applies to any metric space. If $K = \operatorname{supp} f$, then by definition $K = \overline{\{|f|>0\}}$, so K is the closure of its interior (the set $\{|f|>0\}$ is open by continuity of f). Conversely, if V is an open set with closure K, define

$$f(x) = d(x, V^c) = \inf\{d(x,y):\ y \in V^c\}.$$

We have for any $y \in V^c$ and any z

$$d(x,y) + d(x,z) \geq d(y,z) \geq f(z), \quad d(x,z) + d(z,y) \geq d(x,y) \geq f(x)$$

As we can do this for any $y \in V^c$, we obtain

$$f(x) + d(x,z) \geq f(z), \quad d(x,z) + f(z) \geq f(x),$$

which combine into

$$|f(x) - f(z)| \leq d(x,z).$$

So f is continuous. For any $x \in V$, since V is open, $f(x) > 0$. So $V \subset \operatorname{supp} f$. As $f|_{V^c} = 0$, we get that $\operatorname{supp} f = \overline{V}$.

(2.6.7) Let X be compact Hausdorff, and $f \in C(X)$. Suppose that $f(x) \neq 0$ for all $x \in X$. Show that $1/f \in C(X)$.

Answer. Let $g = |f|$. Then $g(X) \subset (0, \infty)$ is compact (Exercise 1.8.37); in particular, there exists $c > 0$ with $g(x) \geq c$ for all $x \in X$. That is, $|f(x)| \geq c > 0$ for all $x \in X$. Then

$$\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| = \frac{|f(y) - f(x)|}{|f(x)f(y)|} \leq \frac{1}{c^2}\,|f(y) - f(x)|.$$

As f is continuous, given $\varepsilon > 0$ there exists a neighbourhood V of x such that $|f(y) - f(x)| < c^2\varepsilon$ for all $y \in V$. Then

$$\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| \leq \frac{1}{c^2}\,|f(y) - f(x)| < \varepsilon$$

for all $y \in V$, and so $1/f$ is continuous.

(2.6.8) Let X be a locally compact Hausdorff space. Show that the following statements are equivalent:

(i) every $f \in C_0(X)$ is constant;

(ii) X is a singleton.

Answer. If $X = \{x_0\}$, then $f = f(x_0)$ for all $x \in X$. Conversely, suppose that $x_1, x_2 \in X$ with $x_1 \neq x_2$. Applying Urysohn's Lemma to $K = \{x_1\}$ and V open with $x_1 \in V$ and $x_2 \notin V$, there exists $f \in C_0(X)$ with $f(x_1) = 1$ and $f(x_0) = 0$; hence f is not constant.

(2.6.9) The goal of this exercise is to consider an alternative proof of Lusin's Corollary 2.6.13 in the case where X is a metric space. Concretely, we want to show that if $f : X \to \mathbb{C}$ is measurable then for every $\varepsilon > 0$ there exists $K \subset X$ compact, with $\mu(X \setminus K) < \varepsilon$, and $g \in C(X)$ such that $g|_K = f|_K$ and $\|f\|_\infty = \|g\|_\infty$.

(i) Assume that f is real-valued. Show that because μ is inner regular, the proof Theorem 2.6.11 can be repeated with "compact" in place of "closed".

(ii) Fix $\varepsilon > 0$ and let K be the compact subset obtained in the modified version of Theorem 2.6.11; that is, $K \subset X$ is compact with $\mu(X \setminus K) < \varepsilon$ and such that $f|_K$ continuous. Conclude that f is uniformly continuous.

(iii) Define the *modulus of continuity* ω of f, to be the function $\omega_1 : [0,\infty) \to [0,\infty)$ with

$$\omega_1(t) = \sup\{|f(y) - f(x)| : \ x, y \in K, \ d(x,y) \le t\}.$$

Show that $|f(x) - f(y)| \le \omega_1(d(x,y))$ for all $x, y \in K$, that $\omega_1(0) = 0$, and that ω is non-decreasing and continuous at 0.

(iv) Let

$$\omega(t) = \frac{1}{t}\int_t^{2t} \omega_1(s)\, ds.$$

Show ω is continuous, that we can define $\omega(0) = 0$ while maintaining continuity, and $|f(x) - f(y)| \le \omega(d(x,y))$ for all $x, y \in K$.

(v) Define

$$g(x) = \inf\{f(y) + \omega(d(x,y)) : \ y \in K\}. \tag{2.31}$$

Show that g is continuous and that $g|_K = f|_K$.

(vi) Show that if f is complex-valued, a continuous function g as above exists for f.

(vii) If $\|f\|_\infty < \infty$, show that g can be replaced with $h \circ g$, in the sense that there exists a continuous function $h : \mathbb{C} \to \mathbb{C}$ such that $\|h \circ g\|_\infty = \|f\|_\infty$, and $h \circ g$ is still a continuous function that agrees with f on K.

Answer.

(i) Because μ is inner regular, the inner regularity by closed sets in Theorem 2.6.11 can be replaced with inner regularity by compacts. Then the closed subsets produced in the proof will be compact.

(ii) A real-valued function f on a metric compact set K is uniformly continuous by Exercise 1.8.17.

(iii) By definition, if we take $t = d(x,y)$ then $\omega_1(t) \ge |f(x) - f(y)|$. When $t = 0$ the inequality $d(x,y) \le 0$ forces $x = y$, and then $\omega_1(0) = 0$. When we increase t more pairs x, y become possibly available, so the supremum is greater; therefore ω_1 is non-decreasing. As for the continuity, fix $\varepsilon > 0$. As f is uniformly continuous on K, there exists $\delta > 0$ such that $d(x,y) < \delta$ implies $|f(x) - f(y)| < \varepsilon$. So if $0 \le t < \delta$, then $\omega_1(t) < \varepsilon$; that is, ω_1 is continuous at 0.

(iv) Since ω_1 is continuous at 0, given $\varepsilon > 0$ there exists $\delta > 0$ with $\omega_1(t) < \varepsilon$ when $2t < \delta$. Then, for such t,

$$\omega(t) = \frac{1}{t}\int_t^{2t} \omega_1(s)\, ds \leq \varepsilon \frac{1}{t}\int_t^{2t} 1\, ds = \varepsilon.$$

This shows that $\lim_{t\to 0} \omega(t) = 0$. For $t > 0$, as f is bounded on K we have that ω_1 is also bounded, and then ω depends continuously on t. Namely,

$$\omega(t) = \frac{1}{t}\left(\int_0^{2t} \omega_1 - \int_0^{t} \omega_1\right),$$

and the integrals are continuous because

$$\left|\int_0^{t+h} \omega_1 - \int_0^{t} \omega_1\right| = \left|\int_t^{t+h} \omega_1\right| \leq h\, \|\omega_1\|_\infty.$$

Finally,

$$\omega(d(x,y)) = \frac{1}{d(x,y)}\int_{d(x,y)}^{2d(x,y)} \omega_1(s)\, ds \geq \omega_1(d(x,y)) \geq |f(x) - f(y)|.$$

(v) By definition, $g(x) \leq f(x)$ when $x \in K$. And since

$$f(y) + \omega(d(x,y)) \geq f(y) + f(x) - f(y) = f(x),$$

so $f \leq g$ and hence $g = f$ on K. As for the continuity of g, we proceed as follows. Since K is compact, the function $K \ni y \longmapsto d(x,y)$ is uniformly continuous and hence bounded, say $d(x,y) \leq c$ for all $y \in K$. As the interval $[0,c]$ is compact, the function ω is uniformly continuous on $[0,c]$. Fix $x \in X$ and $\varepsilon > 0$ and choose $\delta > 0$ such that

$$|s - t| < \delta \implies |\omega(s) - \omega(t)| < \varepsilon, \qquad s,t \in [0,c]$$

and

$$d(y,y') < \delta \implies |f(y) - f(y')| < \varepsilon, \qquad y,y' \in K.$$

Suppose that $d(x,z) < \delta$. Then

$$|d(z,y) - d(x,y)| \leq d(x,z) < \delta, \qquad y \in X.$$

For any $y \in K$,

$$|f(y) + \omega(d(x,y)) - [f(y) + \omega(d(z,y))]| = |\omega(d(z,y)) - \omega(d(x,y))| < \varepsilon.$$

By definition of $g(x)$ there exists $y \in K$ such that $f(y) + \omega(d(x,y)) < g(x) + \varepsilon$. From (2.31) we have

$$g(z) \leq f(y) + \omega(d(z,y)) < f(y) + \omega(d(x,y)) + \varepsilon < g(x) + 2\varepsilon.$$

So $g(x) - g(z) < 2\varepsilon$. Exchanging roles and combining both inequalities we get that $|g(x) - g(z)| < 2\varepsilon$, and hence g is continuous. .

(vi) Suppose that $f = f_1 + if_2$, with f_1, f_2 real. Fix $\varepsilon > 0$. By the real part of the theorem, there exist K_1, K_2 compact with $\mu(X \setminus K_j) < \frac{\varepsilon}{2}$, $j = 1, 2$, and functions $\tilde{f}_1, \tilde{f}_2 \in C(X)$ with $\tilde{f}_1|_{K_1} = f_1|_{K_1}$ and $\tilde{f}_2|_{K_2} = f_2|_{K_2}$. Let $K = K_1 \cap K_2$. Then K is compact, and

$$X \setminus K = X \setminus (K_1 \cap K_2) = (X \setminus K_1) \cup (X \setminus K_2).$$

So

$$\mu(X \setminus K) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Put $g = \tilde{f}_1 + i\tilde{f}_2$. Then $g \in C(X)$ and $g|_K = f|_K$.

(vii) If $\|f\|_\infty < \infty$, consider the function $h : \mathbb{C} \to \mathbb{C}$ given by

$$h(re^{i\theta}) = \begin{cases} re^{i\theta}, & r \le \|f\|_\infty \\ \|f\|_\infty e^{i\theta}, & r > \|f\|_\infty \end{cases}.$$

The function h is continuous, so we may replace g with $h \circ g$. This keeps all requirements for g, and it satisfies $\|h \circ g\|_\infty = \|f\|_\infty$.

(2.6.10) This exercise sketches the constructive proof of Lusin's Theorem (Corollary 2.6.13) given in [Rud87, Theorem 2.24]. Properly, the statement to be proven is that in the situation of Corollary 2.6.13, when the measure space is complete and the measure μ is outer and inner regular, if there exists $A \in \mathcal{A}$ with $\mu(A) < \infty$ and $f|_{X \setminus A} = 0$, then for every $\varepsilon > 0$ there exists $g \in C_c(X)$ and $B \in \mathcal{A}$, such that $\mu(X \setminus B) < \varepsilon$, $f = g$ on B. If $\|f\|_\infty < \infty$ we can choose g with $\|f\|_\infty = \|g\|_\infty$.

(i) Assume A compact. Show that there exists V open with $\overline{V}$ compact and $A \subset V$.

(ii) Show that if $0 \le f \le 1$ there exists a sequence $\{s_n\}$ of simple functions with $s_n \nearrow f$ uniformly, and such that

$$s_n(X) \subset \left\{ \frac{m}{2^n},\ m \in \mathbb{N} \right\}$$

and

$$s_n(x) - s_{n-1}(x) \in \{0, 2^{-n}\},\ x \in X.$$

(iii) Define simple functions $\{t_n\}$ by

$$t_1 = s_1, \qquad t_n = s_n - s_{n-1}, \qquad n \in 1 + \mathbb{N}.$$

Show that, with $s_0 = 0$,

$$\sum_{n=1}^{\infty} t_n = f.$$

(iv) For each $n \in \mathbb{N}$, let $T_n = t_n^{-1}(\{2^{-n}\})$. Show that T_n is measurable, and $T_n \subset A \subset V$ for all n.

(v) Show that for each n there exist K_n compact and V_n open, with $K_n \subset T_n \subset V_n \subset V$ and with $\mu(V_n \setminus K_n) < \frac{\varepsilon}{2^n}$, and $g_n \in C_c(X)$ with $g_n|_{K_n} = 1$, $0 \leq g \leq 1$, and $\operatorname{supp} g_n \subset V_n$.

(vi) Define

$$g = \sum_{n=1}^{\infty} 2^{-n} g_n.$$

Show that $0 \leq g_n \leq 1$, $g \in C_c(X)$, and there exists $B \in \mathcal{A}$ with $\mu(X \setminus B) < \varepsilon$ and $f = g$ on B.

(vii) Extend the result to arbitrary A measurable and arbitrary f.

Answer.

(i) Lemma 2.6.4 gives us V open, with $\overline{V}$ compact and $A \subset V$.

(ii) Because f is bounded and non-negative, Theorem 2.4.13 provides a sequence $\{s_n\}$ of simple functions with $0 \leq s_n \leq s_{n+1} \leq f$ for all n, and with $s_n \nearrow f$ uniformly. The construction in the proof guarantees that s_n only takes values of the form $m/2^n$, and also that $s_n - s_{n-1}$ can only take the values 0 and 2^{-n}.

(iii) We have

$$\sum_{n=1}^{\infty} t_n = \lim_{N\to\infty} \sum_{n=1}^{N} (s_n - s_{n-1}) = \lim_{N\to\infty} s_N = f.$$

(iv) T_n is measurable because t_n is and singletons are closed in $\mathbb{R}$. Also, $T_n \subset A \subset V$ since $t_n(x) > 0$ implies $f(x) > 0$.

(v) For each n, by the regularity of the measure, where exist K_n compact and V_n open, with $K_n \subset T_n \subset V_n$ and $\mu(V_n \setminus K_n) < \frac{\varepsilon}{2^n}$. We may also assume that $V_n \subset V$, by replacing V_n with $V_n \cap V$. By Urysohn's Lemma there exists $g_n \in C_c(X)$ with $g_n|_{K_n} = 1$, $0 \leq g \leq 1$, and $\operatorname{supp} g_n \subset V_n$.

(vi) We have

$$g = \sum_{n=1}^{\infty} 2^{-n} g_n.$$

As $0 \leq g_n \leq 1$, the series converges uniformly and so g is continuous. Since $g_n = 0$ on V^c, we have that $\operatorname{supp} g \subset V$, and so $g \in C_c(X)$. Let

$$B' = \bigcup_n V_n \setminus K_n.$$

Then $\mu(B') < \varepsilon$. If $x \in X \setminus B' = \bigcap_n V_n^c \cup K_n$, then for each n we have that $x \in V_n^c$ or $x \in K_n$. If $x \in V_n^c$, then $g_n = 0$ and $t_n = 0$; and if $x \in K_n$, then $g_n = 1$ and $t_n = 2^{-n}$; in both cases, the corresponding summands in the series for f and g agree. So $f = g$ on $X \setminus B'$. We put $B = X \setminus B'$.

(vii) Now we assume that A is just measurable, and still $0 \leq f \leq 1$. Since $\mu(A) < \infty$, by the outer regularity there exists V open with $V \subset A$ and $\mu(V \setminus A) < \varepsilon/2$. By the inner regularity there exists K compact with $K \subset V$ and $\mu(V \setminus K) < \varepsilon/2$. By the first part of the proof, applied to $f\,1_K$, there exists $g \in C_c(X)$ and B_0 measurable with $\mu(B_0) < \varepsilon/2$ and $g = f$ on $K \cap B_0^c$. As $f = g = 0$ on V^c,

$$\{f \neq g\} \subset B_0 \cup (V \setminus K).$$

Taking $B' = B_0 \cup (V \setminus K)$, we showed that $f = g$ on $X \setminus B'$ and

$$\mu(B') \leq \mu(B_0) + \mu(V \setminus K) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Then we put $B = X \setminus B'$.

When f is non-negative and bounded, $0 \leq f \leq c$, we apply the above to f/c.

When $f \geq 0$ and unbounded, let $H_n = \{f \geq n\}$. The measurable sets $\{H_n\}$ form a decreasing sequence of sets of finite measure (since $H_n \subset A$); and on $\bigcap_n H_n$, f would have to take the value ∞. As this is not possible, the measure of the intersection is 0, and by continuity of the measure $\mu(H_n) \to 0$. This allows us to choose n so that $\mu(H_n) < \varepsilon/2$ and f is bounded on the complement. We apply the previous part of the proof to f on H_n: so there exists $g \in C_c(X)$ and $B_0 \subset H_n^c$ with $\mu(B_0) < \varepsilon/2$ and $f = g$ on $H_n^c \setminus B_0$. Then $B' = H_n \cup B_0$ satisfies $\mu(B') < \varepsilon$ and $f = g$ on $X \setminus B'$. We put $B = X \setminus B'$.

When f is real valued, we can write $f = f^+ - f^-$, and use the previous part of the proof to find $g_1, g_2 \in C_c(X)$ and B_1, B_2 measurable, with $\mu(B_j) < \varepsilon/2$, and $f^+ = g_1$ on $X \setminus B_1$ and $f^- = g_2$ on $X \setminus B_2$. Then if $B' = B_1 \cup B_2$, we have $\mu(B') < \varepsilon$ and $f = g_1 - g_2$ on $X \setminus B'$. And again we put $B = X \setminus B'$.

When $f : X \to \mathbb{C}$, we write $f = f_1 + if_2$, with f_1, f_2 real-valued, and apply the previous paragraph.

If $\|f\|_\infty < \infty$ then $|f| \le c$ a.e. for some c. Define $\gamma : \mathbb{C} \to \mathbb{C}$ by

$$\gamma(z) = \begin{cases} z, & |z| \le c \\ cz/|z|, & |z| > c \end{cases}$$

and replace the g previously obtained with $\gamma \circ g$. By construction γ is continuous, so $\gamma \circ g \in C_c(X)$. Whenever $g = f$, we have $\gamma \circ g = g = f$; and clearly $|h \circ g| \le c$. As we can do this with $c = \|f\|_\infty$, we get that $\|h \circ g\|_\infty = \|f\|_\infty$.

(2.6.11) Determine where in the proof of Egorov's Theorem 2.6.16 is the condition $\mu(E) < \infty$ used.

Answer. When using continuity of the measure in the proof, what one gets is that $\mu(E) = \lim_n \mu(E_n^m)$; the complete computation is then

$$\mu(E \setminus E_n^m) = \mu(E) - \mu(E_n^m) \xrightarrow[n\to\infty]{} 0.$$

(2.6.12) Let $(X, \mathcal{M}, \mu)$ be a positive measure space. We use the following notation, that will be introduced formally soon: $L^1(\mu)$ is the set of integrable functions; and $\|f\|_1 = \displaystyle\int_X |f|\, d\mu$. A set $\Phi \subset L^1(\mu)$ is said to be *uniformly integrable* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \int_E f\, d\mu \right| < \varepsilon \tag{2.32}$$

whenever $f \in \Phi$ and $\mu(E) < \delta$.

(i) Prove that every finite subset of $L^1(\mu)$ is uniformly integrable.

(ii) Prove the following convergence theorem of Vitali: if
   (a) $\mu(X) < \infty$;
   (b) $\{f_n\}$ is uniformly integrable;
   (c) $f_n(x) \to f(x)$ a.e.;
   (d) $|f(x)| < \infty$ a.e.;
   then $f \in L^1(\mu)$ and $\|f_n - f\|_1 \to 0$.
   Suggestion: Egorov.

(iii) Show that (ii) fails for the Lebesgue measure on $\mathbb{R}$, even if $\|f_n\|_1 \leq c$ for all n. So the finite-measure hypothesis cannot be omitted.

(iv) Show that for a measure space X with finite measure, Vitali's Theorem (ii) implies the Dominated Convergence Theorem.

Answer. Note first that we can replace the condition (2.32) with

$$\int_E |f|\,d\mu < \varepsilon. \tag{AB.2.6}$$

Indeed, if f is uniformly integrable, we may apply the definition to the sets $E \cap \{f \geq 0\}$ and $E \cap \{f < 0\}$ to obtain that both f^+ and f^- are uniformly integrable. And then $|f| = f^+ + f^-$ is also uniformly integrable by using $\varepsilon/2$ and the least δ between the one from f^+ and the one from f^-.

(i) Consider first a single $f \in L^1(\mu)$. If $X_n = \{|f| \leq n\}$, then $f - f\,1_{X_n} \to 0$ pointwise (because $\bigcup X_n = X$ up to a nullset by (d)). By Dominated Convergence,
$$\|f - f\,1_{X_n}\|_1 \to 0.$$
Fix $\varepsilon > 0$. By the above there exist $n \in \mathbb{N}$, $g \in L^1(\mu)$, with $|g| \leq n$, and with $\|f - g\|_1 \leq \frac{\varepsilon}{2}$. Let $\delta = \varepsilon/(2n)$. If $\mu(E) < \delta$, then
$$\int_E |f|\,d\mu \leq \int_E |f-g|\,d\mu + \int_E |g|\,d\mu \leq \frac{\varepsilon}{2} + n\mu(E) < \varepsilon.$$
So f is uniformly integrable. If $\Phi = \{f_1, \ldots, f_m\}$, for each j there is a δ_j as above. Now choose $\delta = \min\{\delta_1, \ldots, \delta_m\}$.

(ii) Fix $\varepsilon > 0$. Since $\{f_n\}$ is uniformly integrable, there exists $\delta > 0$ such that $\int_E |f_n| < \varepsilon/3$ when $\mu(E) < \delta$. By Egorov, there exists measurable $B \subset X$ with $\mu(X \setminus B) < \delta$ and $f_n \to f$ uniformly on B. So choose n_0 such that $|f_n - f| < \varepsilon/(3\mu(X))$ on B when $n \geq n_0$. Then, for $n \geq n_0$ (and using

Fatou at the end),

$$\begin{aligned}\|f_n - f\|_1 = \int_X |f_n - f|\,d\mu &= \int_B |f_n - f|\,d\mu + \int_{X\setminus B} |f_n - f|\,d\mu \\ &\le \frac{\varepsilon}{3\mu(X)} \int_B 1\,d\mu + \int_{X\setminus B} |f_n|\,d\mu + \int_{X\setminus B} |f|\,d\mu \\ &\le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \liminf_n \int_{X\setminus B} |f_n|\,d\mu \\ &\le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

As ε was arbitrary, this shows that $\|f_n - f\|_1 \to 0$ and that $f \in L^1$ (since L^1 is complete).

(iii) Let

$$f_n(x) = \frac{1}{n} 1_{[0,n]}.$$

Then $\lim_n f_n(x) = 0$ for all x. Given $\varepsilon > 0$, if $\mu(E) < \varepsilon$ then

$$\int_E f_n(x)\,dx = \frac{\mu(E \cap [0,n])}{n} \le \mu(E) < \varepsilon.$$

So the $\{f_n\}$ are uniformly integrable. Also, $\|f_n\|_1 = 1$ for all n. Thus, with $f = 0$, we have $\|f_n - f\|_1 = 1$ for all n, contradicting Vitali's Theorem.

(iv) The situation for Dominated Convergence is that $|f_n| \le g$ for some $g \in L^1(\mu)$. By (a), the function g is uniformly integrable. Then since

$$\int_E |f_n|\,d\mu \le \int_E g\,d\mu$$

for all E and all n, the sequence $\{f_n\}$ is uniformly integrable. From $|f_n| \le g$ for all n we obtain $|f| \le g$. As g is integrable, we get that $|f| < \infty$ a.e. Then the four conditions in Vitali's Theorem apply and we obtain that $f \in L^1(\mu)$ and $\|f_n - f\|_1 \to 0$.

2.7. Product Measures

(2.7.1) Let $E, E_k \subset X \times Y$ for all $k \in \mathbb{N}$, and $x \in X$. Show that

$$\Big(\bigcup_k E_k\Big)_x = \bigcup_k (E_k)_x, \qquad \Big(\bigcap_k E_k\Big)_x = \bigcap_k (E_k)_x,$$

and

$$(E^c)_x = (E_x)^c$$

Answer. If $y \in (E_k)_x$ for some k, this means that $(x,y) \in E_k$. Thus $(x,y) \in \bigcup_k E_k$, showing that $y \in \Big(\bigcup_k E_k\Big)_x$. All the implications we just did are reversible, so this proves that

$$\Big(\bigcup_k E_k\Big)_x = \bigcup_k (E_k)_x.$$

Now if $y \in (E^c)_x$, then $(x,y) \in E^c$, so $(x,y) \notin E$, and hence $y \notin E_x$. Thus $(E^c)_x \subset (E_x)^c$. Again the implications are reversible, and this shows that $(E^c)_x = (E_x)^c$.

As for the intersections, combining the other two properties

$$\Big(\bigcap_k E_k\Big)_x = \Big[\Big(\bigcap_k E_k\Big)^{cc}\Big]_x = \Big[\Big(\bigcup_k E_k^c\Big)^c\Big]_x = \Big(\bigcup_k (E_k^c)_x\Big)^c$$
$$= \Big(\bigcup_k (E_k)_x^c\Big)^c = \Big(\bigcap_k (E_k)_x\Big)^{cc} = \bigcap_k (E_k)_x.$$

(2.7.2) Write another proof for Proposition 2.7.2 by showing the set

$$\mathcal{S} = \{E \subset X \times Y : \ E_x \in \mathcal{B}, \ E^y \in \mathcal{A} \ \text{ for all } x \in X, \ y \in Y\}.$$

is a σ-algebra that contains the measurable rectangles.

Answer. Since $(A \times B)_x$ is either B or $\varnothing$ and $(A \times B)^y$ is either A or $\varnothing$, $A \times B \in \mathcal{S}$ whenever $A \in \mathcal{A}$ and $B \in \mathcal{B}$. The fact that $\mathcal{S}$ is a σ-algebra now follows directly from Exercise 2.7.1. Thus $\mathcal{S}$ is a σ-algebra that contains all measurable rectangles and so $\mathcal{A} \boxdot \mathcal{B} \subset \mathcal{S}$. In particular, $E_x \in \mathcal{B}$ and $E^y \in \mathcal{A}$ for all measurable E.

(2.7.3) Show that every σ-finite measure μ is semifinite (see Exercise 2.3.26). Give an example of a semifinite measure that is not σ-finite.

Answer. Suppose that μ is σ-finite. This means that we can write $X = \bigcup_n X_n$, with $\mu(X_n) < \infty$ for all n, and $X_n \subset X_{n+1}$ for all n. Let E be measurable with $\mu(E) = \infty$. Since $E = \bigcup_n (E \cap X_n)$, by continuity of the measure we have $\mu(E) = \lim_n \mu(E \cap X_n)$. So there exists n such that $\mu(E \cap X_n) > 0$, and also $\mu(E \cap X_n) \leq \mu(X_n) < \infty$.

As an example that the reverse implication fails, let $X = \mathbb{R}$ and μ the counting measure. Then every nonempty set contains a point, which has positive finite measure. But the space is not σ-finite, since $\mathbb{R}$ is uncountable.

(2.7.4) Let X and Y be separable metric spaces. Show that $\mathcal{B}(X \times Y) = \mathcal{B}(X) \boxdot \mathcal{B}(Y)$.

Answer. We may generate the product topology with the metric

$$d((a,b),(c,d)) = d_x(a,c) + d_Y(b,d).$$

This way we can see that $X \times Y$ is a separable metric space, and that for any open set $Z \subset X \times Y$ and any $(x,y) \in Z$ there exists an open rectangle $V \times W$ with $(x,y) \in V \times W \subset Z$.

Let $\{q_n\} \subset X$, $\{p_n\} \subset Y$ be dense. Then $\{q_n \times p_m\}_{n,m} \subset X \times Y$ is dense. With the same argument as in Exercise 1.8.46 we deduce that Z is a countable union of open rectangles. This shows that $\mathcal{B}(X) \boxdot \mathcal{B}(Y)$ is a σ-algebra that contains all open sets, and thus $\mathcal{B}(X \times Y) \subset \mathcal{B}(X) \boxdot \mathcal{B}(Y)$.

The reverse inclusion holds without the separability (nor metric) requirements. If $V \subset X$ and $W \subset Y$ are open, then $V \times W$ is open in $X \times Y$ and hence $V \times W \in \mathcal{B}(X \times Y)$. Let

$$\mathcal{S} = \{E \subset X : \ E \times Y \in \mathcal{B}(X \times Y)\}.$$

Since $\mathcal{B}(X \times Y)$ is a σ-algebra and all set operations will happen in the first coordinate, it follows easily that $\mathcal{S}$ is a σ-algebra. As $V \times Y \in \mathcal{B}(X \times Y)$ for all open $V \subset X$, we have that $\mathcal{S}$ is a σ-algebra that contains all open sets in X. Thus $\mathcal{B}(X) \subset \mathcal{S}$. In particular, $E \times Y \in \mathcal{B}(X \times Y)$ for all $E \in \mathcal{B}(X)$. We can similarly show that $X \times F \in \mathcal{B}(X \times Y)$ for all $F \in \mathcal{B}(Y)$. Then $E \times F = (E \times Y) \cap (X \times F) \in \mathcal{B}(X \times Y)$. Thus the σ-algebra $\mathcal{B}(X \times Y)$ contains all rectangles $E \times F$ with $E \in \mathcal{B}(X)$, $F \in \mathcal{B}(Y)$; this shows that $\mathcal{B}(X) \boxdot \mathcal{B}(Y) \subset \mathcal{B}(X \times Y)$.

(2.7.5) Show that for any $\varepsilon > 0$ there exists $V \subset \mathbb{R}^n$, open, dense, and with $m(V) < \varepsilon$. Then answer the following questions:

(i) Does the measure of an open subset of $\mathbb{R}^n$ agree with the measure of its closure?

(ii) Is the measure of the boundary of every open subset of $\mathbb{R}^n$ zero?

Answer. The set $\mathbb{Q}^n$ is dense and countable. Write $\mathbb{Q}^n = \{q_k\}_{k\in\mathbb{N}}$. Given $\varepsilon > 0$, let

$$V = \bigcup_{k=1}^{\infty} B_{2^{(-k-1)/n}\,\varepsilon^{1/n}/c^{1/n}}(q_k),$$

where $c = m(B_1(0))$. Then V is open, dense (since it contains $\mathbb{Q}^n$) and $m(V) < \varepsilon$. As the closure of V is all of $\mathbb{R}^n$, we have $m(V) < \varepsilon$ and $m(\overline{V}) = \infty$. And from $\partial V = \overline{V} \setminus V$ we get $m(\partial V) = 0$.

(2.7.6) Show that, in $\mathbb{R} \times \mathbb{R}$, the set

$$\bigcap_n \bigcup_{k=1}^{n} \left[\frac{k-1}{n}, \frac{k}{n}\right]^2 = \{(x,x) : \ 0 \leq x \leq 1\} \tag{2.41}$$

is not a countable union of rectangles (measurable or not). Conclude that the set of countable unions of rectangles does not form a σ-algebra.

Answer. If $A \times B$ is a rectangle with both A and B having at least two elements $a_1, a_2 \in A$, $b_1, b_2 \in B$, then $\{(a_1,b_1),(a_1,b_2),(a_2,b_1),(a_2,b_2)\} \subset A \times B$ has different points that share a coordinate. That does not happen on $\{(x,x) : \ x\}$, so it cannot contain a non-trivial rectangle. Hence we are left with only rectangles of the form $\{x\} \times \{x\}$, but then we need uncountably many of these to cover the whole diagonal.

The expression (2.41) will show up in any σ-algebra that contains the rectangles $\left[\frac{k-1}{n}, \frac{k}{n}\right]$, so the set of countable unions of rectangles cannot form a σ-algebra in $\mathbb{R} \times \mathbb{R}$.

(2.7.7) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces and $\{A_k \times B_k\}$ a countable family of rectangles. Use Lemma 2.7.4 to show that there exists a pairwise disjoint countable family of rectangles

$\{A'_k \times B'_k\}$ such that

$$\bigcup_k A'_k \times B'_k = \bigcup_k A_k \times B_k$$

and

$$\sum_k \mu(A'_k)\nu(B'_k) \le \sum_k \mu(A_k)\nu(B_k).$$

Answer. Let $S_1 = A_1 \times B_1$ and

$$S_k = (A_k \times B_k) \setminus \bigcup_{j=1}^{k} S_j.$$

By Lemma 2.7.4 each S_k is a finite disjoint union of rectangles $S_k = \bigcup_j C_{k,j} \times D_{k,j}$ with $\{C_{k,j}\}_j$ pairwise disjoint for each k. Then

$$\bigcup_k A_k \times B_k = \bigcup_k S_k = \bigcup_{k,j} C_{k,j} \times D_{k,j}$$

is a disjoint union. Because S_k is a disjoint union of rectangles and $S_k \subset A_k \times B_k$, it follows that $C_{k,j} \subset A_k$, $D_{k,j} \subset B_k$ for all j. Then, because they are pairwise disjoint,

$$\sum_j \mu(C_{k,j}) = \mu\big(\bigcup_j C_{k,j}\big) \le \mu(A_k).$$

All terms are non-negative, so

$$\sum_j \mu(C_{k,j})\nu(D_{k,j}) \le \sum_j \mu(C_{k,j})\nu(B_k) \le \mu(A_k)\nu(B_k)$$

and hence

$$\sum_{k,j} \mu(C_{k,j})\nu(D_{k,j}) \le \sum_k \mu(A_k)\nu(B_k).$$

Finally, we relabel $\{C_{k,j} \times D_{k,j}\}$ as $\{A'_k \times B'_k\}$.

(2.7.8) Show that $m \times m$ is outer regular.

Answer. Let $E \in \mathcal{M}(\mathbb{R} \times \mathbb{R})$. Fix $\varepsilon > 0$. By Lemma 2.7.7 there exists a countable union of rectangles R such that $E \subset R$ and $(m \times m)(R \setminus E) < \varepsilon/2$. We may assume without loss of generality that $R = \bigcup_k R_k$ with all rectangles disjoint (we did this in the proof of Proposition 2.7.5). Now $R_k = A_k \times B_k$, with $A_k, B_k \in \mathcal{M}(\mathbb{R})$. By the outer regularity of m (Corollary 2.3.26) there exist open sets V_k, W_k with $A_k \subset V_k$, $B_k \subset W_k$, and $m(V_k \setminus A_k) < \sqrt{\varepsilon/2^{k+1}}$,

$m(W_k \setminus B_k) < \sqrt{\varepsilon/2^{k+1}}$. Then $V = \bigcup_k V_k \times W_k$ is open and

$$\begin{aligned}(m\times m)(V\setminus R) &= (m\times m)\Big(\bigcup_k (V_k\times W_k)\setminus\bigcup_k R_k\Big)\\ &\le (m\times m)\Big(\bigcup_k (V_k\times W_k)\setminus R_k\Big)\\ &= (m\times m)\Big(\bigcup_k (V_k\setminus A_k)\times(W_k\setminus B_k)\Big)\\ &\le \sum_k \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2}.\end{aligned}$$

Finally, we have that $E \subset V$ and

$$(m\times m)(V\setminus E) = (m\times m)((V\setminus R)\cup(R\setminus E)) \le \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.$$

(2.7.9) Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be measure spaces. Let $f : X\times Y \to \mathbb{C}$ be $\mathcal{A}\boxdot\mathcal{B}$-measurable. Show that the sections f_x and f^y are measurable.

Answer. Let $V \subset \mathbb{C}$ be open. Then

$$\begin{aligned}f_x^{-1}(V) &= \{y\in Y:\ f(x,y)\in V\} = \{y\in Y:\ (x,y)\in f^{-1}(V)\}\\ &= \big(f^{-1}(V)\big)_x.\end{aligned}$$

By Proposition 2.7.2 we conclude that $f_x^{-1}(V)$ is measurable. The argument for y is entirely the same, up to switching roles.

(2.7.10) Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be complete measure spaces, and $f : X\times Y\to\mathbb{C}$ measurable. Show that the sections f_x and f^y are measurable a.e.

Answer. Given $V \subset \mathbb{C}$ open, we have that $f^{-1}(V) \in \mathcal{M}(X\times Y)$. By Lemma 2.7.7 we can write $f^{-1}(V) = F\cup G$ with $F\in\mathcal{A}\boxdot\mathcal{B}$ and G a nullset. Then

$$(f_x)^{-1}(V) = \{y\in Y:\ f(x,y)\in V\} = (f^{-1}(V))_x = (F\cup G)_x = F_x\cup G_x.$$

Now F_x is measurable by Proposition 2.7.2 and, by the completeness, so is G_x by Lemma 2.7.9. Then f_x is measurable μ-a.e., precisely where G_x is measurable.

(2.7.11) Show an example of an $\mathcal{M}(X \times Y)$-measurable function f such that f_x is not measurable for x in a set of positive measure.

Answer. Since every measure can be completed, there is no way to make this exciting, as we can only play with sets that are equal a.e. with measurable sets.

Take for instance $X = Y = \mathbb{R}$ with Borel measure; that is, we consider the Lebesgue measure but we take $\mathcal{B}(\mathbb{R})$ to be our σ-algebra in each of X and Y. Choose $H \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ with $m(H) > 0$, and put $E = [0,1] \times H$. Then $E \in \mathcal{M}(\mathbb{R} \times \mathbb{R})$; to see this, note that by Proposition 2.3.28 we can write $H = B \cup H_0$, with $B \in \mathcal{B}(\mathbb{R})$ and $m(H_0) = 0$. Also, $m(E) > 0$. Then

$$E = ([0,1] \times B) \cup ([0,1] \times H_0) \in \mathcal{M}(\mathbb{R} \times \mathbb{R}),$$

since both sets in the union are measurable (the first one is Borel while the second one is a nullset). Now take $f = 1_E$. Then f is measurable, as E is. But for $x \in [0,1]$,

$$(f_x)^{-1}(0,2) = \{y : (x,y) \in E\} = H,$$

so f_x is not measurable for all $x \in [0,1]$.

(2.7.12) Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that the following statements are equivalent:

(i) there exist $\{X_n\} \subset \mathcal{A}$, pairwise disjoint, with $\mu(X_n) < \infty$ for all n and $X = \bigcup_n X_n$;

(ii) there exist $\{X_n\} \subset \mathcal{A}$, with $X_n \subset X_{n+1}$ and $\mu(X_n) < \infty$ for all n, and with $X = \bigcup_n X_n$.

Answer. Suppose first that $\{X_n\}_n$ are pairwise disjoint, with finite measure, and $X = \bigcup_n X_n$. Let $Y_n = \bigcup_{k=1}^{n} X_k \in \mathcal{A}$. Then $\mu(Y_n) \leq \sum_{k=1}^n \mu(X_k) < \infty$, and

$$X \supset \bigcup_n Y_n \supset \bigcup_n X_n = X,$$

so $X = \bigcup_n Y_n$.

Conversely, suppose that $X = \bigcup_n X_n$ with $X_n \subset X_{n+1}$ and $\mu(X_n) < \infty$ for all n. Let $Y_1 = X_1$ and inductively let

$$Y_n = X_n \setminus \bigcup_{k=1}^{n-1} X_k.$$

Then the Y_n are pairwise disjoint by construction. Also, $\mu(Y_n) \leq \mu(X_n) < \infty$ for all n. And, given any $x \in X$ there exists n such that $x \in X_n$, which implies that $x \in Y_1 \cup \cdots \cup Y_n$. Thus $X = \bigcup_n Y_n$.

(2.7.13) Let $f : \mathbb{R}^d \to \mathbb{C}$ be integrable. Show that

$$\int_{\mathbb{R}^d} f(t)\, dm(t) = \int_{\mathbb{R}^d} f(x-t)\, dm(t), \qquad x \in \mathbb{R}^d. \tag{2.42}$$

Answer. This can be done using Exercise 2.5.7, but we will write an explicit argument.

Assume first that $f = 1_E$ for some measurable E with $m(E) < \infty$. Then, using that Lebesgue measure is translation and reflection invariant,

$$\int_{\mathbb{R}^d} 1_E(x-t)\, dt = \int_{\mathbb{R}^d} 1_{x-E}(t)\, dt = m(x-E) = m(E) = \int_{\mathbb{R}^d} 1_E(t)\, dt.$$

By linearity, it follows that (2.42) holds for f simple. We can then use Monotone Convergence to obtain the equality for $f \geq 0$, and then it will hold for $f = f^+ - f^-$ by linearity of the integral.

(2.7.14) Use Fubini's Theorem and the equality

$$\int_0^\infty e^{-xt}\, dt = \frac{1}{x}, \quad x > 0,$$

to show that

$$\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

Note that $x \longmapsto \frac{\sin x}{x}$ is not integrable.

Answer. There is a bit of a subtlety in that the integral exists as an improper Riemann integral, and not as a Lebesgue integral. By definition,

$$\int_0^\infty \frac{\sin x}{x}\, dx = \lim_{k\to\infty} \int_0^{2k\pi} \frac{\sin x}{x}\, dx.$$

Looking at the iterated integral of the absolute value,

$$\int_0^{2k\pi}\int_0^\infty |\sin x\, e^{-xt}|\,dt\,dx = \int_0^{2k\pi} |\sin x| \int_0^\infty e^{-xt}\,dt\,dx$$
$$= \int_0^{2k\pi} \frac{|\sin x|}{x}\,dx < \infty,$$

since $\frac{\sin x}{x}$ is continuous on $[0, 2k\pi]$. As this iterated integral of the absolute value converges, by Fubini the double integral exists and is equal to the iterated integrals. Thus

$$\int_0^{2k\pi} \frac{\sin x}{x}\,dx = \int_0^{2k\pi}\int_0^\infty \sin x\, e^{-xt}\,dt\,dx = \int_0^\infty\int_0^{2k\pi} \sin x\, e^{-xt}\,dx\,dt$$
$$= \int_0^\infty \frac{1-e^{-2k\pi t}}{1+t^2}\,dt.$$

Since the integrand is nonnegative and bounded by the integrable function $t \longmapsto \frac{1}{1+t^2}$, by Dominated Convergence we get

$$\int_0^\infty \frac{\sin x}{x}\,dx = \lim_{k\to\infty}\int_0^\infty \frac{1-e^{-2k\pi t}}{1+t^2}\,dt = \int_0^\infty \frac{1}{1+t^2}\,dt = \frac{\pi}{2}.$$

(2.7.15) Let $f : X \to [0,\infty)$ be measurable. Use the layer-cake representation (2.26) to conclude that

$$\int_X f\,d\mu = \int_0^\infty \mu(\{f \geq t\})\,dt.$$

Answer. We use (2.26) and Tonelli to get

$$\int_X f\,d\mu = \int_X\int_0^\infty 1_{f^{-1}[t,\infty)}(s)\,dt\,d\mu(s) = \int_0^\infty\int_X 1_{f^{-1}[t,\infty)}(s)\,d\mu(s)\,dt$$
$$= \int_0^\infty \mu(f^{-1}[t,\infty))\,dt = \int_0^\infty \mu(\{f\geq t\})\,dt.$$

(2.7.16) Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. Let $f : X \to [0,\infty)$ be measurable.

(i) Show that the graph of f is a $(\mu \times m)$-nullset.

(ii) Show that if

$$B = \{(x,t) \in X \times \mathbb{R} :\ 0 < t < f(x)\}$$

then $(\mu \times m)(B) = \int_X f\, d\mu$.

(iii) Does the above work if we replace the codomain with an arbitrary measure space?

Answer.

(i) The graph of f is the set $G = \{(x, f(x)) : \ x \in X\}$. It is measurable because the function $g(x,t) = f(x) - t$ is measurable (since it is a linear combination of measurable) and then

$$G = g^{-1}(\{0\})$$

is measurable.

Note that $1_G(x,t) = 1$ if $t = f(x)$, and zero otherwise; so the set $\{t : \ 1_G(x,t) = 1\}$ consists of the single point $f(x)$. Using Tonelli, we have

$$\begin{aligned}(\mu \times m)(G) &= \int_{X\times\mathbb{R}} 1_G\, d(\mu \times m) = \int_X \int_{\mathbb{R}} 1_G(x,t)\, dm(t)\, d\mu(x) \\ &= \int_X 0\, d\mu = 0.\end{aligned}$$

Below is a second argument without Tonelli. If $X = \bigcup_n X_n$ with $\mu(X_n) < \infty$, then $G = \bigcup_n G_n$, where G_n is the graph of $f|_{X_n}$. Hence we can assume without loss of generality that $\mu(X) < \infty$. Similarly, we can partition $[0,\infty) = \bigcup_n [n, n+1)$, which allows us—again without loss of generality—to assume that $f(X) \subset [0,1]$. Fix $k \in \mathbb{N}$ and let $\{I_j\}$ be a dyadic partition of $[0,1]$; that is, $m(I_j) = 2^{-k}$ for all j. Then

$$\begin{aligned}(\mu \times m)(G) &= (\mu \times m)\Big(\bigcup_j (G \cap (X \times I_j))\Big) = \sum_j (\mu \times m)(G \cap (X \times I_j)) \\ &\le \sum_j (\mu \times m)(f^{-1}(I_j) \times I_j) = \sum_j \mu(f^{-1}(I_j))\, m(I_j) \\ &= 2^{-k} \sum_j \mu(f^{-1}(I_j)) = 2^{-k}\mu(X).\end{aligned}$$

As this can be done for any k, $(\mu \times m)(G) = 0$.

(ii) With the same g as above, $B = g^{-1}[0,\infty)$ is measurable. Then we can use Tonneli to see

$$\begin{aligned}(\mu \times m)(B) &= \int_{X\times\mathbb{R}} 1_B\, d(\mu \times m) = \int_X \int_{\mathbb{R}} 1_B(x,t)\, dm(t)\, d\mu(x) \\ &= \int_X m(\{t : \ 0 \le t \le f(x)\}\, d\mu(x) = \int_X f(x)\, d\mu(x).\end{aligned}$$

And here is an argument without Tonelli. Suppose for a moment that $\mu(X) < \infty$ and $f = 1_E$. Then

$$B = \{(x,t) : \ 0 < t < 1_E(x)\} = E \times (0,1),$$

and

$$(\mu \times m)(B) = \mu(E) = \int_X 1_E \, d\mu.$$

When $f = \sum_j \alpha_j 1_{E_j}$ is simple,

$$B = \bigcup_j E_j \times (0, \alpha_j),$$

and then

$$(\mu \times m)(B) = \sum_j \alpha_j \mu(E_j) = \int_X f \, d\mu.$$

Now assume that f is bounded. Given $\varepsilon > 0$ there exists $s = \sum_j \alpha 1_{E_j}$ with $0 \le s \le f$ and $|s(x) - f(x)| < \varepsilon$ for all x. Then $B_s \subset B_f$ and

$$\begin{aligned}(\mu \times m)(B_f) - (\mu \times m)(B_s) &= (\mu \times m)(B_f \setminus B_s) \\ &= (\mu \times m)(\{(x,t) : \ s(x) < t < f(x)\} \\ &\le (\mu \times m)\Big(\bigcup_j E_j \times (\alpha_j, \alpha_j + \varepsilon)\Big) \\ &= \varepsilon \sum_j \mu(E_j) = \varepsilon \, \mu(X).\end{aligned}$$

It follows that if $\{s_n\}$ is an increasing sequence of simple functions that converge uniformly to f, then by Monotone Convergence

$$(\mu \times m)(B_f) = \lim_n (\mu \times m)(B_{s_n}) = \lim_n \int_X s_n \, d\mu = \int_X f \, d\mu.$$

When f is unbounded, let $f_n = \min\{f, n\}$. Then f_n is measurable and $f_n \nearrow f$. We have $B_{f_n} \subset B_{f_{n+1}}$ for all n, and $B_f = \bigcup_n B_{f_n}$. By continuity of the measure and Monotone Convergence,

$$(\mu \times m)(B_f) = \lim_n (\mu \times m)(B_{f_n}) = \lim_n \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Finally, for arbitrary σ-finite X we have $X = \bigcup_n X_n$ with $\mu(X_n) < \infty$ for all n and $X_n \cap X_m = \varnothing$ if $n \neq m$. Then

$$(\mu \times m)(B) = \sum_n (\mu \times m)(B_n) = \sum_n \int_{X_n} f \, d\mu = \int_X f \, d\mu.$$

(iii) Consider $X = \mathbb{R}$ with $\mu = m$, and the counting measure on the codomain. Let $f(x) = x$. Then

$$(m \times \mu)(B) \ge (m \times \mu)([1/2, 1] \times [1/2, 1]) = \frac{1}{2} \mu([1/2, 1]) = \infty,$$

but $\displaystyle\int_X f\,dm = \frac{1}{2}$.

(2.7.17) *(Polar Coordinates in $\mathbb{R}^n$).* Let S_{n-1} be the unit sphere on $\mathbb{R}^n$ (i.e., those u with $\|u\| = 1$). Show that any nonzero $x \in \mathbb{R}^n$ can be written $x = ru$, with $r > 0$ and $u \in S_{n-1}$. Thus $\mathbb{R}^n \setminus \{0\}$ can be seen as the Cartesian product $(0, \infty) \times S_{n-1}$.

Let m be the Lebesgue measure on $\mathbb{R}^n$, and define a measure σ on the Borel sets of S_{n-1} by

$$\sigma(A) = n\, m(\tilde{A}),$$

where $\tilde{A} = \{ru : \ 0 < r < 1,\ u \in A\}$. Show that for every Borel $f \geq 0$,

$$\int_{\mathbb{R}^n} f\,dm = \int_0^\infty \int_{S_{n-1}} r^{n-1} f(ru)\,d\sigma(u)\,dr. \tag{2.43}$$

Hint: if $A \subset S_{n-1}$ is open and $0 < r_1 < r_2$, let $E = \{ru : r_1 < r < r_2,\ u \in A\}$. Show the equality for 1_E, and then pass to characteristics of Borel sets.

Answer. For any nonzero $x \in \mathbb{R}^n$, we have $x = \|x\|\,(x/\|x\|)$, with $x/\|x\| \in S_{n-1}$. Note also that for any measurable $X \subset \mathbb{R}^n$ and any $r > 0$ we have $m(rX) = r^n m(X)$. This can be seen by calculating the outer measure (the covers for rX are precisely r^n times the covers for X) or by using the corresponding property for the 1-dimensional Lebesgue measure and considering the n-dimensional Lebesgue measure as the product measure.

Let A and E as in the hint. We have, since $r_1\tilde{A} \subset r_2\tilde{A}$ when $r_1 < r_2$,

$$\int_{\mathbb{R}^n} 1_E\,dm = m(E) = m(r_2\tilde{A} \setminus r_1\tilde{A}) = m(r_2\tilde{A}) - m(r_1\tilde{A}) = (r_2^n - r_1^n)m(\tilde{A}).$$

On the other hand, since $ru \in E$ if and only if $r_1 < r < r_2$ and $u \in A$,

$$\begin{aligned}
\int_0^\infty \int_{S_{n-1}} r^{n-1}\, 1_E(ru)\,d\sigma(u)\,dr &= \int_{r_1}^{r_2} r^{n-1} \int_A 1\,d\sigma\,dr \\
&= \sigma(A) \int_{r_1}^{r_2} r^{n-1}\,dr = \frac{\sigma(A)}{n}\,(r_2^n - r_1^n) \\
&= (r_2^n - r_1^n)m(\tilde{A}).
\end{aligned}$$

So (2.43) holds for such E when A is open.

For any $x \in E$ we have $x = ru$, $u \in A$. As A is open, there is a ball B_0 (of dimension $n-1$) in A with $u \in B_0$; so $x \in (r_1, r_2) \times B_0$, an open set.

Then there exists a ball B in $\mathbb{R}^n$ with $x \in B \subset (r_1, r_2) \times B_0$; showing that E is open.

Let $V \subset \mathbb{R}^n$ be open and bounded, say $V \subset B_R(0)$; assume also that $m(\overline{V}) = m(V)$. This last condition happens for instance when V is a finite union of balls and can fail for countable unions (proof of both facts at the end).

Given $0 < r_1 < r_2$ let

$$U(r_1, r_2) = \bigcup_{t \in [r_1, r_2)} \big(\frac{1}{t} V\big) \cap S_{n-1}.$$

Each $U(r_1, r_2)$ is open in the relative topology of S_{n-1} since it is a union of open sets. For each $n \in \mathbb{N}$ define

$$E_n = \bigcup_{k=1}^{2^n} \Big\{ ru : \ \frac{(k-1)R}{2^n} \le r < \frac{kR}{2^n}, \ u \in U\Big(\frac{(k-1)R}{2^n}, \frac{kR}{2^n}\Big)\Big\}.$$

Each E_n is a finite disjoint union of sets where (2.43) holds, so by additivity of the integral we get (2.43) for each E_n. Since $U_I \subset U_J$ when $I \subset J$, the sequence E_n is decreasing. Now we claim that, with X_0 either $\{0\}$ or the empty set

$$V \subset X_0 \cup \bigcap_n E_n \subset \overline{V} \tag{AB.2.7}$$

(proof at the end). Our hypothesis about the boundary of V implies that $1_V = 1_{\overline{V}}$ a.e. By (AB.2.7) and continuity of the measure we also have

$$\int 1_{E_n} = m(E_n) \xrightarrow[n\to\infty]{} m(\overline{V}) = \int 1_{\overline{V}}. \tag{AB.2.8}$$

As $1_{E_n} - 1_{\overline{V}}$ is monotone non-increasing a.e., its pointwise limit has to be 0 a.e. (because on the set where the limit is not zero (AB.2.8) fails). Thus $1_{\overline{V}} = \lim_n 1_{E_n}$ a.e. As 1_{E_n} satisfies (2.43), Dominated Convergence (justified by V being bounded, so we are integrating inside $B_R(0)$ and thus any bounded measurable function is integrable) gives us (2.43) for $1_{\overline{V}}$ and thus for V.

For arbitrary V open we can write $V = \bigcup_n B_n$, a countable union of balls. Putting $F_n = \bigcup_{k=1}^{n} B_n$ allows us to write $V = \bigcup_n F_n$, an increasing union of sets which are finite unions of balls. Each F_n satisfies (2.43) by the previous paragraph; and then so does V via Monotone Convergence.

Finally, let $B \subset \mathbb{R}^n$ be Borel. By the outer regularity of the Lebesgue measure, for each $h \in \mathbb{N}$ there exists V_h open with $B \subset V_h$ and $m(V_h \setminus B) < 1/h$. Replacing V_{h+1} with $V_{h+1} \cap V_h$, we can get the sequence V_h to be decreasing. If $m(B) < \infty$, we can apply Dominated Convergence twice to get

$$\int_{\mathbb{R}^n} 1_B \, dm = \lim_h \int_{\mathbb{R}^n} 1_{V_h} \, dm = \lim_h \int_0^\infty \int_{S_{n-1}} r^{n-1} \, 1_{V_h}(ru) \, d\sigma(u) \, dr$$
$$= \int_0^\infty \int_{S_{n-1}} r^{n-1} \, 1_B(ru) \, d\sigma(u) \, dr.$$

When $m(B) = \infty$ we can write it as an increasing union of Borel sets of finite measure, and we get the equality (with infinity on both sides) via Monotone Convergence.

So (2.43) holds for all characteristics of Borel sets. Then by linearity it holds for all simple Borel functions; and by Monotone Convergence the equality holds for all measurable $f \geq 0$.

Proof that for a finite union of balls, the Lebesgue measure of the closure of their union is equal to the measure of the union. For a single ball centered at the origin take $A = S_{n-1}$ and $r_1 = 1 - \frac{1}{k}$, $r_2 = 1 + \frac{1}{k}$, which gives us

$$m\big(\overline{B_1(0)} \setminus B_1(0)\big) \leq m(B_{1+\frac{1}{k}}(0) \setminus B_{1-\frac{1}{k}}(0))$$
$$= \Big[\big(1 + \frac{1}{k}\big)^n - \big(1 - \frac{1}{k}\big)^n\Big] \, m(B_1(0)) \xrightarrow[k \to \infty]{} 0.$$

For a finite union of balls we obtain $m(\overline{V}) = m(V)$ from the unit ball case by translation, scaling, and finite subadditivity of the Lebesgue measure.

To see that this cannot hold in general for countable unions, we use the idea from Exercise 2.7.5. Let $\{q_n\}$ be an enumeration of $\mathbb{Q} \cap [0,1]$, fix $\varepsilon > 0$, and let $V_n = (q_n - \frac{\varepsilon}{2^n}, q_n + \frac{\varepsilon}{2^n})$. Then if $V = \bigcup_n V_n$ we have

$$m(V) \leq \sum_n m(V_n) = \sum_n \frac{\varepsilon}{2^{n-1}} = 2\varepsilon,$$

while $m(\overline{V}) = m([0,1]) = 1$.

Proof of (AB.2.7). If $tu \in V$ with $u \in S_{n-1}$ and $t > 0$, then for all n we have $u \in U\big(\frac{(k_n-1)R}{2^n}, \frac{k_n R}{2^n}\big)$, where $\{k_n\}$ is a sequence of positive integers such that $t \in \big[\frac{(k_n-1)R}{2^n}, \frac{k_n R}{2^n}\big)$. So $tu \in E_n$ for all n. Thus $V \subset X_0 \cup \bigcap\limits_n E_n$. Conversely, if $tu \in \bigcap\limits_n E_n$ then for each n there exists $k_n \in \mathbb{N}$ with $\frac{(k_n-1)R}{2^n} \leq t < \frac{k_n R}{2^n}$ and $u_n \in U\big(\frac{(k_n-1)R}{2^n}, \frac{k_n R}{2^n}\big)$ with $tu = t_n u_n$. For each u_n there exists $s_n \in \big[\frac{(k_n-1)R}{2^n}, \frac{k_n R}{2^n}\big)$ with $s_n u_n \in V$. Note that $s_n \to t$ since the intervals squeeze to t. By compactness of S_{n-1} there exists a convergent subsequence $\{u_{n_j}\}$ and so $tu = \lim_j s_{n_j} u_{n_j} \in \overline{V}$. If $0 \in V$ we will also get $X_0 = \{0\} \subset V$.

(2.7.18) Show that, when $n=2$, the equality (2.43) becomes the usual Calculus polar coordinate change of variable; that is, for $f \in C_c(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} f\,dm = \int_0^\infty \int_0^{2\pi} r\,f(r\cos t, r\sin t)\,dt\,dr.$$

Answer. We have

$$\int_{\mathbb{R}^2} f\,dm = \int_0^\infty \int_{S_1} r\,f(ru)\,du\,dr.$$

We can parametrize S_1, the unit circle, as $(\cos t, \sin t)$ for $t \in [0, 2\pi]$. For an arc $A = (\cos t, \sin t)$ with $a \leq t \leq b$, we have

$$\sigma(A) = 2\,m(\tilde{A}) = 2\,\pi\,\frac{b-a}{2\pi} = b-a.$$

So σ can be seen as the Lebesgue measure on the arc length. As $f(ru) = f(r\cos t, r\sin t)$,

$$\int_{\mathbb{R}^2} f\,dm = \int_0^\infty \int_0^{2\pi} r\,f(r\cos t, r\sin t)\,dt\,dr.$$

(2.7.19) *(Volume of a ball in $\mathbb{R}^n$)* Let $B_r(0)$ be the ball of radius r, in $\mathbb{R}^n$, centered at the origin.

(i) Show that $m(B_1(0)) = \sigma(S_{n-1})/n$.

(ii) Calculate $\displaystyle\int_{\mathbb{R}^n} e^{-|x|^2}\,dm$ using Fubini.

(iii) Calculate $\displaystyle\int_{\mathbb{R}^n} e^{-|x|^2}\,dm$, using (2.43), in terms if $\sigma(S_{n-1})$ and the Gamma Function from Exercise 3.1.6.

(iv) Show that $m(B_r(0)) = r^n\,m(B_1(0))$.

(v) Use what you found to write a formula for $m(B_r(0))$.

(vi) Let $C_n \subset \mathbb{R}^n$ denote the hypercube of side 2 centered at the origin, and $B_n \subset \mathbb{R}^n$ the ball of radius 1 centered at the origin. Show that $\lim_{n\to\infty} \frac{m(B_n)}{m(C_n)} = 0$. This result is not so anti-intuitive when you notice that $\operatorname{diam} B_n = 2$, while $\operatorname{diam} C_n = 2\sqrt{n}$. Another way of seeing it as natural is that the unit ball cannot touch the areas near the corners of the cube. The square has four corners, the cube has eight corners, and in dimension n the hypercube has 2^n corners. So as dimension grows, there are more an more parts of the cube that cannot be touched by the ball.

Answer.

(i) We have, using the previous question,

$$m(B_1(0)) = \int_{\mathbb{R}^n} 1_{B_1(0)}\, dm = \int_0^\infty r^{n-1} \int_{S_{n-1}} 1_{B_1(0)}(ru)\, d\sigma\, dr$$
$$= \int_0^1 r^{n-1} \int_{S_{n-1}} 1\, d\sigma\, dr = \frac{\sigma(S_{n-1})}{n}.$$

Alternatively, we can note that $B_1(0) = \tilde{S}_{n-1}$.

(ii) By Fubini, and using that $e^{-|x|^2} = e^{-x_1^2} \cdots e^{-x_n^2}$,

$$\int_{\mathbb{R}^n} e^{-|x|^2}\, dm = \left(\int_{\mathbb{R}} e^{-t^2}\, dt\right)^n = \pi^{n/2}.$$

(iii) Now, using the substitution $s = r^2$,

$$\int_{\mathbb{R}^n} e^{-|x|^2}\, dm = \int_0^\infty r^{n-1} \int_{S_{n-1}} e^{-r^2|u|^2}\, d\sigma(u)\, dr$$
$$= \int_0^\infty r^{n-1} \int_{S_{n-1}} e^{-r^2}\, d\sigma(u)\, dr$$
$$= \sigma(S_{n-1}) \int_0^\infty r^{n-1}\, e^{-r^2}\, dr = \frac{\sigma(S_{n-1})}{2} \int_0^\infty s^{n/2-1}\, e^{-s}\, ds$$
$$= \frac{\sigma(S_{n-1})}{2}\, \Gamma\left(\frac{n}{2}\right).$$

(iv) This is the same as in the previous question:

$$m(B_r(0)) = m(r\, B_1(0)) = r^n\, m(B_1(0)),$$

using the outer measure (or approximating from within by boxes).

(v) Now

$$m(B_r(0)) = r^n\, m(B_1(0)) = \frac{r^n}{n}\, \sigma(S_{n-1}) = \frac{2\pi^{n/2}\, r^n}{n\, \Gamma(n/2)} = \frac{\pi^{n/2}\, r^n}{\Gamma(\frac{n}{2}+1)}.$$

Using the computation from Exercise 3.10.20 we get

$$\Gamma\left(\frac{1}{2}\right)^2 = \frac{\pi}{\sin \pi/2} = \pi,$$

so $\Gamma(1/2)=\sqrt{\pi}$. Therefore, when $n=2k+1$ is odd,

$$\begin{aligned}\Gamma\Big(\frac{n}{2}+1\Big)=\Gamma\Big(k+1+\frac{1}{2}\Big)&=\Big(k+\frac{1}{2}\Big)\Big(k-1+\frac{1}{2}\Big)\cdots\Big(1+\frac{1}{2}\Big)\Gamma\Big(\frac{1}{2}\Big)\\&=\Big(k+\frac{1}{2}\Big)\Big(k-1+\frac{1}{2}\Big)\cdots\Big(1+\frac{1}{2}\Big)\pi^{1/2}\\&=\Big(\frac{2k+1}{2}\Big)\Big(\frac{2(k-1)+1}{2}\Big)\cdots\frac{2(k-(k-1))+1}{2}\pi^{1/2}\\&=\frac{(2k+1)!}{2\,4^k k!}\pi^{1/2}.\end{aligned}$$

This gives

$$m(B_r^{2k+1})=\frac{2(4\pi)^k\,k!}{(2k+1)!}\,r^{2k+1}.$$

When $n=2k$ is even,

$$\Gamma\Big(\frac{n}{2}+1\Big)=\Gamma(k+1)=k!.$$

So

$$m(B_r^{2k}(0))=\frac{\pi^k}{(k!}\,r^{2k}.$$

(vi) Since $m(C_n)=2^n$, we are doing

$$\lim_{n\to\infty}\frac{\pi^{n/2}}{2^n\Gamma(\frac{n}{2}+1)}\le\lim_{n\to\infty}\frac{\pi^{n/2}}{2^n}=0$$

as $\sqrt{\pi}<2$.

(2.7.20) Let $\alpha>0$, $n\in\mathbb{N}$. In $\mathbb{R}^n$, evaluate

$$\int_{|x|\le 1}\frac{1}{|x|^\alpha}\,dx,\qquad\text{and}\qquad\int_{|x|\ge 1}\frac{1}{|x|^\alpha}\,dx.$$

Answer. We use (2.43). Let $f(x)=\frac{1}{|x|^\alpha}1_{B_1(0)}$. Then

$$\begin{aligned}\int_{|x|\le 1}\frac{1}{|x|^\alpha}\,dx&=\int_0^\infty\int_{S_{n-1}}r^{n-1}\frac{1}{|ru|^\alpha}1_{B_1(0)}(ru)\,d\sigma(u)\,dr\\&=\int_0^1\int_{S_{n-1}}r^{n-1-\alpha}\,d\sigma(u)\,dr=n\,m(B_1(0))\int_0^1 r^{n-1-\alpha}\,dr.\end{aligned}$$

Using the formula for the volume of the ball from Exercise 2.7.19 we get

$$\int_{|x|\le 1} \frac{1}{|x|^\alpha}\,dx = \begin{cases} \dfrac{2n\,k!\,4^k\pi^k}{(n-\alpha)\,(2k+1)!}, & \alpha<n,\ n=2k+1\\ \dfrac{n\,\pi^k}{(n-\alpha)\,k!}, & \alpha<n,\ n=2k\\ \infty, & \alpha\ge n\end{cases}$$

For the integral from 1 to ∞, all that changes is the convergence of the integral on r at the end. So

$$\int_{|x|\ge 1} \frac{1}{|x|^\alpha}\,dx = \begin{cases} \infty, & \alpha\le n\\ \dfrac{2n\,k!\,4^k\pi^k}{(n-\alpha)\,(2k+1)!}, & \alpha>n,\ n=2k+1\\ \dfrac{n\,\pi^k}{(n-\alpha)\,k!}, & \alpha>n,\ n=2k\end{cases}$$

(2.7.21) Let $(X,\mathcal{A},\mu)$ be a σ-finite measure space and $f:X\to\mathbb{C}$ measurable. Let $\omega_f:[0,\infty)\to[0,\infty)$ be the distribution function

$$\omega_f(t)=\mu(\{|f|>t\}).$$

Prove that

$$\int_X |f|^p = \int_0^\infty p\,t^{p-1}\,\omega_f(t)\,dt, \qquad 1\le p<\infty.$$

More generally, show that if $\gamma:[0,\infty)\to[0,\infty)$ is increasing, differentiable, and $\gamma(0)=0$, then

$$\int_X \gamma\circ f\,d\mu = \int_0^\infty \omega_f(t)\,\gamma'(t)\,dt.$$

Answer. We have, using Tonelli since we have iterated integrals of nonnegative functions,

$$\begin{aligned}\int_X \gamma\circ|f|\,d\mu &= \int_X\int_0^{|f(x)|}\gamma'(t)\,dt\,d\mu(x) = \int_X\int_0^\infty \gamma'(t)\,1_{\{|f|>t\}}(t)\,dt\,d\mu(x)\\ &= \int_0^\infty\int_X \gamma'(t)\,1_{\{|f|>t\}}(t)\,d\mu(x)\,dt\\ &= \int_0^\infty \gamma'(t)\,\mu(\{|f|>t\})\,d\mu(x)\,dt.\end{aligned}$$

Note that Exercise 2.7.15 is the particular case $\gamma(t)=t$.

2.8. L^p-Spaces

(2.8.1) Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that the relation defined in (2.46) is an equivalence relation. Show that addition, multiplication, and p-norms of classes are well-defined by means of their representatives.

Answer. Since $f = f$ everywhere, we have $f \sim f$. If $f = g$ on A and $\mu(A^c) = 0$, then $g = f$ on A and $\mu(A^c) = 0$, so $g \sim f$. If $f \sim g$ and $g \sim h$, there exist A, B with $\mu(A^c) = \mu(B^c) = 0$ with $f = g$ on A and $g = h$ on B. Let $C = A \cap B$. On C we have $f = g = h$; and $\mu(C^c) = \mu(A^c \cup B^c) \leq \mu(A^c) + \mu(B^c) = 0$, so $f \sim h$.

If $f \sim f'$ and $g \sim g'$ then there exist A, B with $f = f'$ on A, $g = g'$ on B, and $\mu(A^c) = \mu(B^c) = 0$. Let $C = A \cap B$. On C, $f + g = f' + g'$ and $fg = f'g'$. And $\mu(C^c) = \mu(A^c \cup B^c) \leq \mu(A^c) + \mu(B^c) = 0$.

Finally, if $f = g$ on A and $\mu(A^c) = 0$,

$$\int_X |f|^p \, d\mu = \int_A |f|^p \, d\mu = \int_A |g|^p \, d\mu = \int_X |g|^p \, d\mu.$$

(2.8.2) Let $p \in [1, \infty]$ and $f \in L^p(X)$. Show that $\{|f| = \infty\}$ is a nullset.

Answer. Suppose first that $p = \infty$. Then $\mu(\{|f| = \infty\}) \leq \mu(\{|f| > \|f\|_\infty\}) = 0$.

When $p < \infty$, if $\mu(\{|f| = \infty\}) > 0$ then

$$\|f\|_p^p = \int_X |f|^p \, d\mu \geq \int_{\{|f|=\infty\}} |f|^p \, d\mu = \infty,$$

a contradiction.

(2.8.3) Show that $|f| \leq \|f\|_\infty$ a.e.

Answer. By definition, for each $n \in \mathbb{N}$ there exists A_n such that $|f| \leq \|f\|_\infty + \frac{1}{n}$ on A_n and $\mu(A_n^c) = 0$. Let $A = \bigcap_n A_n$. Then

$$\mu(A^c) = \mu\Big(\bigcap_n A_n^c\Big) \leq \sum_n \mu(A_n^c) = 0.$$

And on A we have $|f| \leq \|f\|_\infty + \frac{1}{n}$ for all n, so $|f| \leq \|f\|_\infty$ a.e.

(2.8.4) Show that if $(X, \mathcal{A}, \mu)$ is a measure space and $\{f_n\} \subset L^p(X)$, then

$$\Big\| \sum_{k=1}^\infty |f_k| \Big\|_p \leq \sum_{k=1}^\infty \|f_k\|_p,$$

and the inequality still holds even if one or both sides are infinite.

Answer. We have, using Monotone Convergence and both continuity and monotonicity of the exponential functions,

$$\begin{aligned}\Big\| \sum_{k=1}^\infty |f_k| \Big\|_p &= \left(\int_X \Big(\sum_{k=1}^\infty |f_k|\Big)^p d\mu\right)^{1/p} = \left(\int_X \lim_{K\to\infty} \Big(\sum_{k=1}^K |f_k|\Big)^p d\mu\right)^{1/p} \\ &= \lim_{K\to\infty} \left(\int_X \Big(\sum_{k=1}^K |f_k|\Big)^p d\mu\right)^{1/p} = \lim_{K\to\infty} \Big\| \sum_{k=1}^K |f_k| \Big\|_p \\ &\leq \lim_{K\to\infty} \sum_{k=1}^K \|f_k\|_p = \sum_{k=1}^\infty \|f_k\|_p.\end{aligned}$$

(2.8.5) Let $p \in [1, \infty)$. Show that if $a \in \ell^p(\mathbb{N})$, then a is bounded. Can we say the same for $f \in L^p(\mathbb{R})$?

Answer. We have

$$|a(n)| \leq \Big(\sum_{k=1}^\infty |a(k)|^p\Big)^{1/p} = \|a\|_p.$$

So $\|a\|_\infty \leq \|a\|_p$.

On the other hand, in $L^p(\mathbb{R})$ there are unbounded functions. For instance we can consider

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x^{-1/(2p)}, & x \in (0,1) \\ 0, & x \geq 1 \end{cases}$$

or

$$f(x) = \sum_{n=1}^{\infty} n\, 1_{\left[n, n+\frac{1}{n^{p+2}}\right]}.$$

(2.8.6) Show that the hypothesis that $\|f\|_r < \infty$ for some r in Proposition 2.8.11 cannot be omitted.

Answer. Let $X = \mathbb{R}$ with Lebesgue measure, and take $f = 1$. Then $\|f\|_\infty = 1$, while $\|f\|_p = \infty$ for all $p < \infty$.

(2.8.7) Show that a Cauchy sequence $\{f_n\} \subset L^p(X)$ is uniformly bounded in $L^p(X)$: that is, there exists $c > 0$ such that $\|f_n\|_p < c$ for all n.

Answer. This holds in any metric space; we did it in Exercise 1.8.27. Let us still write the argument in the particular case at hand. Let $\varepsilon > 0$. Since $\{f_n\}$ is Cauchy, there exists n_0 such that for all $n, m \geq n_0$ we have

$$\|f_n - f_m\|_p < \varepsilon.$$

Then

$$\big|\|f_n\|_p - \|f_m\|_p\big| < \varepsilon,$$

showing that the sequence of real numbers $\{\|f_n\|_p\}_n$ is Cauchy. As Cauchy sequences in $\mathbb{R}$ are bounded, there exists $c > 0$ with $\|f_n\|_p < c$ for all n.

(2.8.8) Show that $\ell^p(\mathbb{N})$, $1 \leq p < \infty$ is separable, while $\ell^\infty(\mathbb{N})$ is not separable.

Answer. With $\{e_j\}$ as usual the canonical basis, let

$$X = \Big\{\sum_{j=1}^{n} c_j e_j : \ n \in \mathbb{N},\ c_j \in \mathbb{Q} + i\mathbb{Q}\Big\}.$$

Then X is countable; let us show it is dense. Given $f \in \ell^p(\mathbb{N})$, we have

$$f = \sum_{j=1}^{\infty} a_j e_j, \quad \text{where} \quad \sum_{j=1}^{\infty} |a_j|^p < \infty.$$

Fix $\varepsilon > 0$. Choose n such that $\sum_{j=n+1}^{\infty} |a_j|^p < (\varepsilon/2)^p$. Choose $c_j \in \mathbb{Q} + i\mathbb{Q}$ with $|c_j - a_j| < (\varepsilon^p/2^{j+1})^{1/p}$. Then $y = \sum_{j=1}^{n} c_j e_j \in X$, and

$$\begin{aligned}\|f - y\|_p &= \left\| \sum_{j=1}^{n} (a_j - c_j) e_j + \sum_{j=n=1}^{\infty} a_j e_j \right\|_p \\ &\le \Big(\sum_{j=1}^{n} |a_j - c_j|^p \Big)^{1/p} + \Big(\sum_{j=n+1}^{\infty} |a_j|^p \Big)^{1/p} \\ &\le \Big(\sum_{j=1}^{n} \frac{\varepsilon^p}{2^{j+1}} \Big)^{1/p} + \frac{\varepsilon}{2} \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

On the other hand, $\ell^\infty(\mathbb{N})$ is not separable. Consider the uncountable set $\mathcal{P}(\mathbb{N})$, and let $\alpha : \mathcal{P}(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ be given by $\alpha(R) = 1_R$. That is, we map R to the sequence x that has $x_k = 1$ if $k \in R$, and $x_k = 0$ otherwise. For any two sets $R, S \in \mathcal{P}(\mathbb{N})$ with $R \ne S$, there exists $k \in (R \setminus S) \cup (S \setminus R)$. Then $|(1_R - 1_S)(k)| = 1$ and so $\|1_R - 1_S\|_\infty = 1$. We have uncountable many points all at distance 1 from each other, and so $\ell^\infty(\mathbb{N})$ cannot be separable.

Here is another argument to show that $\ell^\infty(\mathbb{N})$ is not separable. Assume that $A \subset \ell^\infty(\mathbb{N})$ is countable. Write $A = \{a_n : n \in \mathbb{N}\}$. Now construct $x \in \ell^\infty$ by

$$x(m) = \begin{cases} 0, & |a_m(m)| \ge 1 \\ 2, & |a_m(m)| < 1 \end{cases}$$

Then $|x(m) - a_m(m)| \ge 1$, giving us $\|x - a_m\| \ge 1$ for all m. Thus A is not dense.

(2.8.9) Show that the subspace

$$c_{00} = \{x : \mathbb{N} \to \mathbb{C}, \text{ with finite support}\}$$

is dense in $\ell^p(\mathbb{N})$, $1 \le p < \infty$. What about the case $p = \infty$?

Answer. Taking the set X from Exercise 2.8.8, $X \subset c_{00}$, so

$$\ell^p(\mathbb{N}) = \overline{X} \subset \overline{c_{00}} \subset \ell^p(\mathbb{N}).$$

.

When $p = \infty$ it cannot be dense, again by Exercise 2.8.8, since c_{00} is separable, so we would have that $\ell^\infty(\mathbb{N})$ is separable.

(2.8.10) Show that if $f \in L^1(\mu)$, then for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\displaystyle\int_E |f| d\mu < \varepsilon$ whenever $\mu(E) < \delta$. Can δ be chosen independently of f?

Answer. The argument we need was already used to answer Exercise 2.6.12.

Fix $\varepsilon > 0$. For $n \in \mathbb{N}$, let $A_n = \{|f| > n\}$, $B_n = \{|f| \leq n\}$. We have

$$\mu(A_n) = \int_{A_n} 1\, d\mu = \frac{1}{n}\int_{A_n} n\, d\mu \leq \frac{1}{n}\int_{A_n} |f|\, d\mu \leq \frac{\|f\|_1}{n}.$$

So

$$\int_E |f|\, d\mu = \int_{E\cap A_n} |f|\, d\mu + \int_{E\cap B_n} |f|\, d\mu \leq \frac{\|f\|_1}{n} + n\,\mu(E).$$

Choose n so that $n > 2\|f\|_1/\varepsilon$, and let $\delta = \frac{\varepsilon}{2n}$. Then, if $\mu(E) < \delta$,

$$\int_E |f|\, d\mu \leq \frac{\|f\|_1}{n} + n\,\mu(E) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

The choice of δ is intrinsically dependent on f. Consider the functions $f_n = n\,1_{[0,\frac{1}{n}]}$. Then for $E_n = [0, \frac{1}{n}]$ we have $\int_{E_n} f_n = 1$, while $\mu(E_n) = \frac{1}{n}$. If we fix $\varepsilon > 0$ with $\varepsilon < 1$ and we fix $\delta > 0$, then for $n > \frac{1}{\delta}$ we have $\mu(E_n) < \delta$ while $\int_{E_n} f_n > \varepsilon$.

(2.8.11) Let $(X, \mathcal{A}, \mu)$ be a measure space, $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, and $f \in L^p(X)$, $g \in L^q(X)$. Show that the following statements are equivalent:

(i) $\|fg\|_1 = \|f\|_p\|g\|_q$;

(ii) either $f = 0$, $g = 0$, or there exists $\alpha \in (0, \infty)$ with $|f|^p = \alpha|g|^q$.

(Hint: for the nontrivial part, try to undo the proof of Young's Inequality)

Answer. If $f = 0$ or $g = 0$, the equality holds trivially. If $|f|^p = \alpha|g|^q$, then

$$\|fg\|_1 = \int_X |fg|\, d\mu = \int_X \alpha^{1/p}|g|^{q/p+1}\, d\mu = \alpha^{1/p}\int_X |g|^q\, d\mu,$$

and

$$\|f\|_p\|g\|_q = \left(\int \alpha|g|^q\, d\mu\right)^{1/p}\left(\int_X |g|^q\, d\mu\right)^{1/q} = \alpha^{1/p}\int_X |g|^q\, d\mu$$

Conversely, assume that $\|fg\|_1 = \|f\|_p\|g\|_q$. If $\|f\|_p\|g\|_q = 0$, then either $f = 0$ or $g = 0$. So we assume that $\|f\|_p > 0$ and $\|g\|_q > 0$. By replacing f with $f/\|f\|_p$ and g with $g/\|g\|_q$, we may assume that $\|fg\|_1 = 1$ and $\|f\|_p = \|g\|_q = 1$. So we have

$$\int_X |f|\,|g| = 1 = \frac{1}{p}\int_X |f|^p + \frac{1}{q}\int_X |g|^q.$$

We may rewrite this as

$$0 = \int_X \left(\frac{1}{p}\,|f|^p + \frac{1}{q}\,|g|^q - |fg|\right).$$

By Young's Inequality, the integrand above is nonnegative, so from the equality we conclude that

$$|fg| = \frac{1}{p}\,|f|^p + \frac{1}{q}\,|g|^q \qquad \text{a.e.}$$

Now fix an x where the equality holds; writing $a = |f(x)|^p$, $b = |g(x)|^q$, we have

$$a^{1/p}b^{1/q} = \frac{1}{p}\,a + \frac{1}{q}\,b.$$

Applying logarithm, we obtain

$$\frac{1}{p}\,\log a + \frac{1}{q}\,\log b = \log\left(\frac{1}{p}\,a + \frac{1}{q}\,b\right).$$

Because of the concavity of log, this equality can only happen if $a = b$. So $|f(x)|^p = |g(x)|^q$ for all such x; that is, $|f|^p = |g|^q$ a.e. Going back to the original f and g, we get $|f|^p/\|f\|_p^p = |g|^q/\|g\|_q^q$ a.e., and we can take $\alpha = \|f\|_p^p/\|g\|_q^q$.

(2.8.12) Let $(X, \mathcal{A}, \mu)$ be a measure space and $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $f \in L^p(X)$, $g \in L^q(X)$. Show that the following statements are equivalent:

(i) $\|f + g\|_p = \|f\|_p + \|g\|_p$;

(ii) there exist $\alpha, \beta \geq 0$, not both zero, with $\alpha f = \beta g$ a.e.

Answer.

If $f + g = 0$ then the equality $0 = \|f\|_p + \|g\|_p$ implies $f = g = 0$ a.e. If $f = 0$ a.e. we can take $\beta = 0$, $\alpha = 1$. Similarly, if $g = 0$ a.e. we can take $\alpha = 0$, $\beta = 1$.

Otherwise, both f, g are nonzero a.e. and we can rewrite $\|f + g\|_p = \|f\|_p + \|g\|_p$ as

$$\|f + g\|_p^p = \big(\|f\|_p + \|g\|_p\big)\,\|f + g\|_p^{p-1}.$$

If we now look into the inequalities used to prove Minkowski's inequality, we get equality in the two Hölder inequalities in between. By Exercise 2.8.11 there exist numbers $a', b' \in \mathbb{C}$, with at least one of them nonzero, such that $|f|^p = b' \, |f+g|^p$, $|g|^p = a' \, |f+g|^p$ a.e. In fact, since $|f| \neq 0$ and $|g| \neq 0$, we have $a'b' > 0$. Thus $\alpha, \beta > 0$ with

$$|f| = \beta \, |f+g|, \qquad\qquad |g| = \alpha \, |f+g|.$$

This implies that $\alpha \, |f| = \beta \, |g|$. The equalities in the proof of Minkowsky's inequality also give

$$\int_X \big[|f+g| - |f| - |g|\big] \, |f+g|^{p-1} \, d\mu = 0.$$

This means that $|f|+|g| = |f+g|$ whenever $|f+g| \neq 0$; but when $|f+g| = 0$ we have $|f| = |g| = 0$, and thus $|f+g| = |f|+|g|$ a.e. This last equality occurs if and only if $f\bar{g} \geq 0$ a.e. Indeed, by squaring and cancelling we get $|fg| = \operatorname{Re} f\bar{g}$; and the equality $|z| = \operatorname{Re} z$ for $z \in \mathbb{C}$ implies $z \geq 0$.

So whenever both $f \neq 0$ and $g \neq 0$, if $h = f\bar{g}$,

$$f = \frac{h}{|g|} \, g.$$

Taking absolute value (recall that $h \geq 0$), $\frac{h}{|g|} = \frac{\beta}{\alpha}$, constant. Then $\alpha f = \beta g$.

The converse is trivial: if $\alpha f = \beta g$ with $\alpha, \beta \geq 0$ and $\alpha \neq 0$ we have

$$\|f+g\|_p = \left\| \frac{\beta}{\alpha} g + g \right\|_p = \frac{\beta}{\alpha} \, \|g\|_p + \|g\|_p = \|f\|_p + \|g\|_p.$$

If $\alpha = 0$ then $\beta \neq 0$ and we can exchange roles.

(2.8.13) Show by example that the p-norm is not a norm when $p < 1$.

Answer. We want to show that the triangle inequality fails. In $L^p(\{1,2\})$, let $f = (1,0)$, $g = (0,1)$. Then

$$\|f\|_p = \|g\|_p = 1, \qquad \|f+g\|_p = 2^{1/p}.$$

As $0 < p < 1$, we get

$$\|f+g\|_p = 2^{1/p} > 2 = \|f\|_p + \|g\|_p.$$

(2.8.14) Prove the following generalization of Hölder's Inequality for functions $f_1, \ldots, f_n$. Namely, if $p_1, \ldots, p_n \geq 1$ with $\sum_{j=1}^n \frac{1}{p_j} =$

1, and $f_j \in L^{p_j}(X)$, $j = 1, \dots, n$. Then

$$\|f_1 \cdots f_n\|_1 \le \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}. \tag{2.53}$$

Answer. We will do induction on the usual Hölder inequality. As a base case, we can take the case $n = 2$. So now assume that (2.53) holds for n. Fix $p_1, \dots, p_{n+1} \ge 1$ with $\sum_j 1/p_j = 1$ and let

$$r = \frac{1}{\frac{1}{p_1} + \cdots + \frac{1}{p_n}}.$$

From $\frac{1}{r} + \frac{1}{p_{n+1}} = 1$ we get $r \ge 1$. Then, using Hölder first with conjugate exponents r and p_{n+1} and later with exponents $p_1/r, \dots, p_n/r$ (note that $r \le p_j$, so $p_j/r \ge 1$), we get

$$\begin{aligned}
\|f_1 \cdots f_{n+1}\|_1 &= \int_X |f_1 \cdots f_n|\,|f_{n+1}|\,d\mu \le \|f_1 \cdots f_n\|_r\,\|f_{n+1}\|_{p_{n+1}} \\
&= \left(\int_X |f_1|^r \cdots |f_n|^r\,d\mu\right)^{1/r} \|f_{n+1}\|_{p_{n+1}} \\
&\le \left(\|f_1\|_{p_1}^r \cdots \|f_n\|_{p_n}^r\right)^{1/r} \|f_{n+1}\|_{p_{n+1}} \\
&= \|f_1\|_{p_1} \cdots \|f_n\|_{p_{n+1}}.
\end{aligned}$$

The argument can be simplified slightly by using the inequality from Exercise 2.8.15.

The proof can also be made by mimicking the proof of the original Hölder inequality, but using a version of Young's inequality with n terms, that also follows from the convexity of the log function.

(2.8.15) Use Hölder's Inequality to prove the following more general inequality: if $f \in L^p(X)$, $g \in L^q(X)$, where $p, q \ge 1$ and r is such that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, then

$$\|fg\|_r \le \|f\|_p \|g\|_q. \tag{2.54}$$

Answer. Note that r cannot be infinite. We apply the usual Hölder inequality, with exponents p/r and q/r. Then

$$\|fg\|_r = \|\,|f|^r\,|g|^r\,\|_1^{1/r} \le \|\,|f|^r\,\|_{p/r}^{1/r}\,\|\,|g|^r\,\|_{q/r}^{1/r} = \|f\|_p\,\|g\|_q.$$

This works even when one of p, q is infinite, as long as the original relation $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ is satisfied.

(2.8.16) Show that if f, g have compact support, then $f*g$ has compact support.

Answer. Let $F = \operatorname{supp} f$ and $G = \operatorname{supp} g$. Then

$$(f*g)(x) = \int_X f(t)\, g(x-t)\, dt = \int_F f(t)\, g(x-t)\, dt = \int_{F\cap(x-G)} f(t)\, g(x-t)\, dt.$$

For $y \in G$, $x-y \in F$ if and only if $x \in y+F \subset G+F$. Thus $\operatorname{supp} f*g \subset G+F$.

(2.8.17) *(Young's Convolution Inequality)* Prove Young's Convolution Inequality (2.50). *(Hint: for non-negative f, g and p, q, r finite,*

$$(f*g)(x) = \int_{\mathbb{R}^d} f(t)^{p/r} g(x-t)^{q/r}\, f(t)^{p/p_1}\, g(x-t)^{q/p_2}\, dt;$$

then use (2.53) for the exponents r, p_1, p_2, where $\frac{1}{p_1} = \frac{1}{p} - \frac{1}{r}$ and $\frac{1}{p_2} = \frac{1}{q} - \frac{1}{r}$)

Answer. We have $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$, and

$$\frac{1}{r} + \frac{1}{p_1} + \frac{1}{q_1} = \frac{1}{r} + \frac{1}{p} - \frac{1}{r} + \frac{1}{q} - \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - \frac{1}{r} = 1.$$

As with the original Hölder inequality, we may assume without loss of generality that $f, g \geq 0$, for they always appear inside an absolute value. Following the hint,

$$\begin{aligned}(f*g)(x) &= \int_{\mathbb{R}^d} f(t)^{p/r} g(x-t)^{q/r}\, f(t)^{p/p_1}\, g(x-t)^{q/p_2}\, dt \\ &\leq \Big(\int_{\mathbb{R}^d} f(t)^p g(x-t)^q dt\Big)^{1/r} \Big(\int_{\mathbb{R}^d} f(t)^p dt\Big)^{1/p_1} \Big(\int_{\mathbb{R}^d} g(x-t)^q dt\Big)^{1/p_2} \\ &= \Big(\int_{\mathbb{R}^d} f(t)^p g(x-t)^q\, dt\Big)^{1/r} \|f\|_p^{p/p_1}\, \|g\|_q^{q/p_2}.\end{aligned}$$

Then, as $rp/p_1 = r - p$ and $rq/p_2 = r - q$,

$$\begin{aligned}\|f * g\|_r^r &= \int_{\mathbb{R}^d} |(f*g)(x)|^r \, dx \\ &\le \|f\|_p^{rp/p_1} \|g\|_q^{rq/p_2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(t)^p g(x-t)^q \, dt \, dx \\ &= \|f\|_p^{r-p} \|g\|_q^{r-q} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(t)^p g(x-t)^q \, dx \, dt \\ &= \|f\|_p^{r-p} \|g\|_q^{r-q} \|f\|_p^p \|g\|_q^q = \|f\|_p^r \|g\|_q^r.\end{aligned}$$

Taking the r^{th} root we get (2.50).

It remains to address the case where at least one of p, q, r is ∞. When $r = \infty$ we have $\frac{1}{p} + \frac{1}{q} = 1$, and

$$|(f*g)(x)| \le \int_{\mathbb{R}^d} |f(t)| \, |g(x-t)| \, dx \le \|f\|_p \|g\|_q$$

by Hölder's Inequality; so $\|f * g\|_\infty \le \|f\|_p \|g\|_q$. If $p = \infty$ this forces $q = 1$ and $r = \infty$ and again we can apply the Hölder. Same when $q = \infty$.

(2.8.18) *(Interpolation).* Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $r, s \in [1, \infty]$ with $r < s$. Show that if $f \in L^r(X) \cap L^s(X)$, then $f \in L^p(X)$ for all $p \in (r, s)$.

Answer. There exists $t \in (0,1)$ with $\frac{1}{p} = \frac{t}{r} + \frac{1-t}{s}$ (this works even if $s = \infty$). Using Exercise 2.8.15,

$$\|f\|_p = \||f|^t \, |f|^{1-t}\|_p \le \| \, |f|^t \, \|_{r/t} \, \| \, |f|^{1-t} \, \|_{s/(1-t)} = \|f\|_r^t \, \|f\|_s^{1-t} < \infty.$$

(2.8.19) Let $f_n : X \to [0, \infty)$ be measurable for all n, with $f_1 \ge f_2 \ge \cdots \ge 0$, with $f_n \to f$ pointwise, and such that there exists n_0 with $f_{n_0} \in L^1(\mu)$. Prove that

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Show that the assertion fails if the L^1 condition is omitted.

Answer. The limit is the same if the sequence begins at n_0; so without loss of generality we may assume that $f_1 \in L^1(\mu)$. Then Dominated Convergence applies, with $g = f_1$.

If no integrability is required, consider $X = \mathbb{R}$ with Lebesgue measure, and $f_n = \frac{1}{n}$. Then $f_n \searrow 0$, and

$$\int_X f_n = \infty \quad \text{for all } n, \qquad \int_X f = 0.$$

(2.8.20) Let X be a set. Show that if $1 \le p < q < \infty$, then $\ell^q(X) \subset \ell^p(X) \subset \ell^\infty(X)$.

Answer. Since $1 \le p < q$ we have $p/q < 1$. Fix $a \in \ell^p(X)$. Then

$$\Big(\sum_{n=1}^m |a_n|^q\Big)^{p/q} \le \sum_{n=1}^m |a_n|^{qp/q} = \sum_{n=1}^m |a_n|^p.$$

This idea works for any finite subset of X, and taking limit works for any countable subset. Even if X is uncountable, $a \in \ell^p(X)$ implies that $a_n \neq 0$ only on a countable subset. So $\|a\|_q \le \|a\|_p$, showing that $\ell^q(X) \subset \ell^p(X)$. Finally, any element of $\ell^p(X)$ is bounded, so $\ell^p(X) \subset \ell^\infty(X)$.

(2.8.21) For some measures, $1 \le r < s$ implies $L^r(\mu) \subset L^s(\mu)$; for others, the reverse inclusion holds; for others, $L^r(\mu) = L^s(\mu)$; and still for others, no inclusion holds if $r \neq s$. Show examples of all these situations, and find conditions on μ under which each case occurs.

Answer. If $r < s$, then $\ell^r(\mathbb{N}) \subset \ell^s(\mathbb{N})$ by Exercise 2.8.20. This will happen whenever μ is atomic.

If $\mu(X) < \infty$, then $r < s$ implies $L^s(\mu) \subset L^r(\mu)$. Indeed, if $\int_X |f|^r = \infty$, then $\int_{|f|\ge 1} |f|^r = \infty$ (otherwise, as $\mu(X) < \infty$, we would have that the whole integral is finite). Then $\int_{|f|\ge 1} |f|^s \ge \int_{|f|\ge 1} |f|^r = \infty$.

If $\mathcal{A}$ is finite (in particular, if X is finite), then $L^r(X) = L^s(X)$ for all r, s, as $f \in L^p(X)$ if and only if $|f| < \infty$ a.e. and $f|_A = 0$ a.e. if $\mu(A) = \infty$; independently of p.

On the real line with the Lebesgue measure, there is no inclusion $L^r(\mathbb{R}) \subset L^s(\mathbb{R})$ if $r \neq s$, as shown in Exercise 2.8.22 below.

(2.8.22) Given $p \geq 1$, find $f \in L^p(\mathbb{R})$ such that $f \notin L^q(\mathbb{R})$ for any $q \neq p$.

Answer. Let $f_n : (0,1] \to \mathbb{R}$ be $f_n(x) = x^{(-1+1/n)/p}$. Then $f_n \in L^p[0,1]$ and $f_n \notin L^q[0,1]$ for all $q \geq p/(1-\frac{1}{n})$. Similarly, let $g_n(x) = x^{(-1-\frac{1}{n})/p}$. Then $g_n \in L^p[1,\infty)$ while $g_n \notin L^q[1,\infty)$ for all $q \leq p/(1+\frac{1}{n})$.

An easy computation shows that

$$\int_0^1 |f_n|^p = \int_1^\infty |g_n|^p = n.$$

The idea is to use the infinitely many intervals available to us to patch things and "use all n". Let $\gamma : \mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ be a bijection. Define intervals

$$I_{m,n} = [\gamma(m,n), \gamma(m,n)+1).$$

As γ is bijective, the intervals $I_{m,n}$ are pairwise disjoint and cover all of $\mathbb{R}$. Let

$$f(x) = \begin{cases} 2^{-n/p} n^{-1/p}\, f_n(x - \gamma(1,n)), & x \in I_{1,n} \\ 2^{-n/p} n^{-1/p}\, g_n(x - \gamma(m,n) + m - 1), & x \in I_{m,n},\ m > 1 \end{cases}$$

Now

$$\begin{aligned} \int_{\mathbb{R}} |f|^p &= \sum_{m,n} \int_{I_{m,n}} |f|^p \\ &= \sum_n 2^{-n} n^{-1} \int_{I_{1,n}} |f_n(x-\gamma(1,n))|^p \\ &\qquad + \sum_n 2^{-n} n^{-1} \sum_{m=2}^\infty \int_{I_{m,n}} |g_n(x - \gamma(m,n) + m - 1)|^p \\ &= \sum_n 2^{-n} n^{-1} \int_0^1 |f_n(x)|^p + \sum_n 2^{-n} n^{-1} \sum_{m=2}^\infty \int_0^1 |g_n(x+m-1)|^p \\ &= \sum_n 2^{-n} + \sum_n 2^{-n} n^{-1} \sum_{m=2}^\infty \int_{m-1}^m |g_n(x)|^p \\ &= 1 + \sum_n 2^{-n} n^{-1} \int_1^\infty |g_n(x)|^p \\ &= 2. \end{aligned}$$

On the other hand, if $q > p$ there exists n with $q > p/(1-\frac{1}{n})$. Then

$$\int_{\mathbb{R}} |f|^q \geq \int_{I_{1,n}} |f|^q = 2^{-nq/p} n^{-q/p} \int_0^1 \frac{1}{x^{\frac{q}{p}(1-\frac{1}{n})}}\, dx = \infty.$$

And if $q < p$ there exists n with $q < p/(1+\frac{1}{n})$. Then
$$\int_{\mathbb{R}} |f|^q \geq \int_{\bigcup_m I_{m,n}} |f|^q = \int_1^\infty \frac{1}{\frac{q}{p}(1+\frac{1}{n})}\, dx = \infty.$$

For a fancier and less intuitive example, let
$$f(x) = \begin{cases} \dfrac{1}{x^{1/p}(\log^2 x + 1)}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Then
$$\int_{\mathbb{R}} |f|^p\, dx \leq \int_0^{1/2} \frac{1}{x\log^2 x}\, dx + \int_{1/2}^2 \frac{1}{x(\log^2 x + 1)^p}\, dx + \int_2^\infty \frac{1}{x\log^2 x}\, dx < \infty.$$
If $q > p$, using the substitution $t = -\log x$,
$$\int_{\mathbb{R}} |f|^q\, dx \geq \int_0^1 \frac{1}{x^{q/p}(\log^2 x + 1)^q}\, dx = \int_0^\infty \frac{e^{t(\frac{q}{p}-1)}}{(t^2+1)^q}\, dt = \infty.$$
And if $q < p$,
$$\int_{\mathbb{R}} |f|^q\, dx \geq \int_1^\infty \frac{1}{x^{q/p}(\log^2 x + 1)^q}\, dx = \int_1^\infty \frac{e^{t(1-\frac{q}{p})}}{(t^2+1)^q}\, dt = \infty.$$

(2.8.23) Do Propositions 2.8.14, 2.8.16 and 2.8.18 hold in $\ell^\infty(\mathbb{N})$? Give proofs or counterexamples.

Answer. All three fail in $\ell^\infty(\mathbb{N})$. When $\mathbb{N}$ is considered with the discrete topology, the counting measure is a Radon measure, and so the compactly supported continuous functions are precisely the measurable functions with finite support, which in turn are just the functions with finite support. If a has finite support, then $a(n) = 0$ for some n; this gives $\|1 - a\|_\infty \geq 1$; that is, the constant function 1 is at distance 1 from the sets of finitely supported functions and of compactly supported continuous functions.

(2.8.24) Prove Lemma 2.8.19.

Answer. Fix $\varepsilon > 0$. By Proposition 2.8.18 there exists $g \in C_c(X)$ with $\|f - g\|_p < \varepsilon/3$. As g has compact support E, we will show at the end of the proof that there exists an open neighbourhood V of 0 such that $|g(x - t) - g(x)| < \varepsilon/(3\mu(E)^{1/p})$ for all x, whenever $t \in V$. From this estimate we obtain $\|g_t - g\|_p < \varepsilon/3$, and then

$$\begin{aligned}\|f_t - f\|_p &\le \|f_t - g_t\|_p + \|g_t - g\|_p + \|g - f\|_p \\ &= 2\|g - f\|_p + \|g_t - g\|_p \\ &< \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

We now prove the inequality for $|g(x-t) - g(x)|$. Let $\varepsilon' = \varepsilon/(3\mu(E)^{1/p})$; since g is continuous for each $x \in E$ there exists an open set V_x, with $0 \in V_x$ and such that $y - x \in V_x$ implies $|g(y) - g(x)| < \varepsilon'/2$. By replacing V_x with $V_x \cap (-V_x)$ we may assume that $-V_x = V_x$. The continuity of addition guarantees that there exists an open neighbourhood W_x of 0 with $W_x + W_x \subset V_x$, and again we may assume that $-W_x = W_x$. As E is compact and $E \subset \bigcup_{x \in E} (x + W_x)$, there exist $x_1, \ldots, x_n$ with

$$E \subset (x_1 + W_{x_1}) \cup \cdots \cup (x_n + W_{x_n}). \tag{AB.2.9}$$

Let $V = \bigcap_{j=1}^{n} W_{x_j}$. Then V is open, $-V = V$, and $0 \in V$. If $t \in V$ and $x \in E$ from (AB.2.9) there exists j with $x \in x_j + W_{x_j}$; so $|g(x) - g(x_j)| < \varepsilon'/2$. As $x - t - x_j = (x - x_j) - t \in W_{x_j} - W_{x_j} \subset V_{x_j}$, we also have $|g(x-t) - g(x_j)| < \varepsilon'/2$. Then

$$\begin{aligned}|g(x-t) - g(x)| &\le |g(x-t) - g(x_j)| + |g(x_j) - g(x)| \\ &< \frac{\varepsilon'}{2} + \frac{\varepsilon'}{2} = \varepsilon' = \frac{\varepsilon}{3\mu(E)^{1/p}}.\end{aligned}$$

When $x \notin E$, we have $g(x) = 0$. Since $X \setminus E$ is open there exists a neighbourhood W of 0 such that $x + W \subset X \setminus E$ and $-W = W$. For any $t \in W$, $x - t \in X \setminus E$, and so $g(x-t) = 0$. So for $t \in V \cap W$ the estimate $|g(x-t) - g(x)| < \varepsilon'$ holds for all $x \in X$.

(2.8.25) Show that (2.48) can fail when $p = \infty$, even if $\mu(X) < \infty$.

Answer. Let $X = [0,1]$ with Lebesgue measure, and $f = 1_{[0,\frac{1}{2}]}$. Then $f_t - f = 1_{(t,\frac{1}{2}+t)}$, so $\|f_t - f\|_\infty = 1$ for all $t > 0$.

(2.8.26) Let (X, μ) and (Y, ν) be σ-finite complete measure spaces such that $L^2(X)$ and $L^2(Y)$ are separable. Fix orthonormal bases

$\{f_n\}$ and $\{g_n\}$ for $L^2(X)$ and $L^2(Y)$ respectively. Show that $\{f_n(x)g_m(y)\}_{n,m}$ is an orthonormal basis for $L^2(X \times Y)$.

Answer. First,

$$\int_X \int_Y |f_n(x)f_s(x)|\,|g_m(y)g_t(y)|\,d\nu(y)\,d\mu(x) = \Big(\int_Y |g_m(y)g_t(y)|\,d\nu(y)\Big) \Big(\int_X |f_n(x)f_s(x)|\,d\mu(x)\Big) \le \|g_m\|_2\|g_t\|_2\|f_n\|_2\|f_s\|_2 < \infty.$$

Then Fubini (Theorem 2.7.16) guarantees that the double integral exists and agrees with the iterated integrals. Thus

$$\langle f_n g_m, f_s g_t\rangle = \int_{X\times Y} f_n(x)\overline{f_s(x)}\, g_m(y)\overline{g_t(y)}\, d(\mu\times\nu)(x,y) = \langle f_n, f_s\rangle\langle g_m, g_t\rangle = \delta_{(n,s),(m,t)}$$

and thus the set $\{f_n(x)g_m(y)\}_{n,m}$ is orthonormal. It remains to see that it is total. If $\langle h, f_n \times g_m\rangle = 0$ for all n, m, we have (using Fubini again)

$$0 = \int_X \left(\int_Y h(x,y)\, g_m(y)\, d\nu(y)\right) f_n(x)\, d\mu(x);$$

so, as n is arbitrary, the function $x \mapsto \int_Y h(x,y)\, g_m(y)\, d\nu(y)$ is zero almost everywhere for each m. Let

$$E_m = \Big\{x \in X : \int_Y h(x,y)\, g_m(y)\, d\nu(y) \neq 0\Big\}.$$

Each E_m is a null-set (so measurable, by the completeness), and then so is its (countable) union E. Outside of E,

$$\int_Y h(x,y)\, g_m(y)\, d\nu(y) = 0 \quad \text{for all } m.$$

Thus for each $x \in X \setminus E$, $h(x,y) = 0$ almost everywhere. As $|h|^2$ is integrable, its integral agrees with the iterated integrals, so

$$\int_{X\times Y} |h(x,y)|^2\, d(\mu\times\nu) = \int_X \int_Y |h(x,y)|^2\, d\nu(y)\, d\mu(x) = \int_{X\setminus E} \int_Y |h(x,y)|^2\, d\nu(y)\, d\mu(x) = 0.$$

Hence $h = 0$ in $L^2(X \times Y)$.

(2.8.27) Let (X, Σ, μ) be a measure space. A sequence $\{f_n\}$ of complex measurable functions on X is said to *converge in measure* to the measurable function f if for every $\varepsilon > 0$ there exists N such that

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) < \varepsilon, \quad n > N.$$

Prove:

(i) If $\mu(X) < \infty$ and $f_n \to f$ a.e., then $f_n \to f$ in measure (*Hint: use Egorov's Theorem*).

(ii) For $1 \leq p \leq \infty$, if $f_n \in L^p(\mu)$ for all n and $\|f_n - f\|_p \to 0$, then $f_n \to f$ in measure.

(iii) If $f_n \to f$ in measure, then there exists a subsequence f_{n_k} that converges to f a.e.

Answer.

(i) Let $\varepsilon > 0$. Since $\mu(X) < \infty$, Egorov's Theorem applies. So there exists $E \subset X$, with $\mu(X \setminus E) < \varepsilon$ and such that $f_n \to f$ uniformly on E. Hence there exists $N \in \mathbb{N}$ such that $|f_n - f| < \varepsilon$, on E, whenever $n > N$. Then

$$\mu\{|f_n - f| > \varepsilon\} \leq \mu(E^c) < \varepsilon.$$

(ii) If f_n does not converge in measure to f, then there exists $\varepsilon > 0$ such that for every $N \in \mathbb{N}$ there exists $n > N$ with $\mu(\{|f_n - f| > \varepsilon\} \geq \varepsilon$. This means that we can choose an unbounded sequence $\{n_k\}$ such that $\mu(\{|f_{n_k} - f| > \varepsilon\} \geq \varepsilon$. Then, if $p < \infty$,

$$\begin{aligned}\|f_{n_k} - f\|_p^p &= \int_X |f_n - f|^p \, d\mu \\ &\geq \int_{\{|f_n - f| > \varepsilon\}} |f_n - f|^p \, d\mu \\ &\geq \varepsilon^p \, \mu(\{|f_n - f| > \varepsilon\}) \geq \varepsilon^{p+1}\end{aligned}$$

for all k, so $\|f_n - f\|_p$ does not go to zero. When $p = \infty$ we have $\|f_{n_k} - f\|_\infty \geq \varepsilon$ for all k, so we obtain the same conclusion.

(iii) Let $E_{n,k} = \{|f_n - f| > \frac{1}{k}\}$. By hypothesis there exists $n_k > n_{k-1}$ such that $\mu(E_{n,k}) < \frac{1}{k}$ for all $n \geq n_k$. Then we have $E_{n_{k+1},k+1} \subset E_{n_k,k}$. Let $E = \bigcap_k E_{n_k,k}$. By continuity of the measure, $\mu(E) = 0$. If $x \in X \setminus E = \bigcup_k (X \setminus E_{n_k,k})$, then there exists k_0 such that $x \in X \setminus E_{n_{k_0},k_0}$. As the union is increasing, $x \in X \setminus E_{n_k,k}$ for all $k > k_0$; so $|f_{n_k}(x) - f(x)| < \frac{1}{k}$, for all $k \geq k_0$ showing that $f_{n_k} \to f$ outside of E.

(2.8.28) Let $\{a_n\}_{n\in\mathbb{N}} \subset (0,\infty)$, (X,Σ,μ) a finite measure space, and $f_n : X \to [0,\infty)$, $n \in \mathbb{N}$, measurable functions such that for all $n \in \mathbb{N}$

$$\int_X f_n \, d\mu = n, \qquad \int_X f_n^2 \, d\mu = a_n n^2. \tag{2.55}$$

(i) Show that if the sequence $\{a_n\}$ is bounded, then f_n does not converge to 0 a.e. (*Hint: one possible approach uses Egorov's Theorem*)

(ii) Does the above hold when $\{a_n\}$ is unbounded?

Answer. Suppose that (2.55) holds, that $f_n \to 0$ a.e. and that $a_n \leq c$ for all n. By Egorov's Theorem (2.6.16) there exist sets $E_k \in \Sigma$ with $\mu(E_k) < \frac{1}{k}$ and $f_n \to 0$ uniformly on $X \setminus E_k$. In particular for each k there exists $n(k) \geq 2$ such that $f_n|_{X\setminus E_k} < \frac{1}{2\mu(X\setminus E_k)}$ for all $n \geq n(k)$. Then (writing $n = n(k)$ from now on, for simplicity)

$$\int_{E_k} f_n \, d\mu = n - \int_{X\setminus E_k} f_n \, d\mu > n - \frac{1}{2} > n - 1.$$

This gives us, using Cauchy Schwarz,

$$n - 1 < \int_{E_k} f_n \, d\mu \leq \mu(E_k)^{1/2} \Big(\int_X f_n^2 \, d\mu \Big)^{1/2} = \mu(E_k)^{1/2} \, a_n^{1/2} \, n < \frac{c^{1/2}\, n}{k^{1/2}}.$$

Thus

$$\frac{1}{2} \leq 1 - \frac{1}{n} < \frac{c^{1/2}}{k^{1/2}}.$$

As the inequality $\frac{1}{2} < \frac{c^{1/2}}{k^{1/2}}$ is impossible for k big enough, we conclude that $f_n \to 0$ a.e. is not possible.

For the case where a_n is unbounded, let $X = [0,1]$ with Lebesgue measure, $f_n = n^2 \, 1_{[0,\frac{1}{n}]}$ and $a_n = n$. Then

$$\int_X f_n \, dm = n, \qquad \int_X f_n^2 \, dm = n^3 = a_n \, n^2,$$

and $f_n \to 0$ a.e.

(2.8.29) Suppose that $\mu(X) < \infty$, $f \in L^\infty(\mu)$, $\|f\|_\infty > 0$, and

$$\alpha_n = \int_X |f|^n \, d\mu, \qquad n \in \mathbb{N}.$$

Prove that

$$\lim_{n\to\infty} \frac{\alpha_{n+1}}{\alpha_n} = \|f\|_\infty.$$

Answer. Since $\mu(X) < \infty$,

$$\frac{\alpha_{n+1}}{\alpha_n} = \frac{\int_X |f|^{n+1} \, d\mu}{\int_X |f|^n \, d\mu} \leq \|f\|_\infty \frac{\int_X |f|^n \, d\mu}{\int_X |f|^n \, d\mu} = \|f\|_\infty.$$

So

$$\limsup_n \frac{\alpha_{n+1}}{\alpha_n} \leq \|f\|_\infty.$$

Now we use Hölder (with $p = (n+1)/n$, $q = n+1$) to obtain

$$\|f\|_n^n = \int_X |f|^n \leq \left(\int_X |f|^{n+1}\right)^{n/(n+1)} \mu(X)^{1/(n+1)} = \|f\|_{n+1}^n \, \mu(X)^{1/(n+1)}.$$

Then

$$\frac{\alpha_{n+1}}{\alpha_n} = \frac{\|f\|_{n+1}^{n+1}}{\|f\|_n^n} \geq \frac{\|f\|_{n+1}^{n+1}}{\|f\|_{n+1}^n \, \mu(X)^{1/(n+1)}} = \|f\|_{n+1} \, \mu(X)^{-1/(n+1)}.$$

Then, as the right-hand-side converges to $\|f\|_\infty$ (Proposition 2.8.11),

$$\liminf_{n\to\infty} \frac{\alpha_{n+1}}{\alpha_n} \geq \|f\|_\infty.$$

(2.8.30) Let h_1 as in (2.52). Suppose that $g \in C^\infty(-\delta, \delta)$ for some $\delta > 0$ and $0 \leq g(x) \leq h_1(x)$ for all $x \in (0, \delta)$. Show that $g^{(k)}(0) = 0$ for all $k \in \mathbb{N}$.

Answer. Since g is C^∞ we know by hypothesis that its derivatives exist at all points in $(-\delta, \delta)$. So it is enough to show that the right derivatives at 0 are 0. The inequality $0 \leq g(x) \leq h_1(x)$, after taking limit as $x \to 0^+$, gives us directly that $g(0) = 0$. Suppose for induction that $g^{(k)}(0) = 0$ for $k = 0, 1, \ldots, n$. Then we can write the Taylor polynomial of g as

$$g(x) = \frac{g^{(n+1)}(\xi(x))}{(n+1)!} \, x^{n+1}, \qquad x \in (-\delta, \delta)$$

with $|\xi(x)| \leq |x|$. So

$$0 \leq |g^{(n+1)}(\xi(x))| = \frac{(n+1)! g(x)}{|x|^{n+1}} \leq (n+1)! \, \frac{h_1(x)}{x^{n+1}}.$$

Taking limit as $x \to 0$, we get $\xi(x) \to 0$ and $0 \le g^{(n+1)}(0) \le 0$, so $g^{(n+1)}(0) = 0$. We then get by induction that $g^{(n+1)}(0) = 0$ for all $n \in \mathbb{N}$.

The limit for h_1 comes (with the substitution $t = 1/x^2$) from

$$\lim_{x\to 0^+} \frac{e^{-1/x^2}}{x^n} = \lim_{t\to\infty} t^{n/2}e^{-t} = 0.$$

(2.8.31) *(this is not an easy one; the topic of which L^p spaces are separable is subtle)* Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space such that $\mathcal{A}$ is countably generated. Show that if μ is σ-finite, then $L^p(\mu)$ is separable for $p \in [1, \infty)$.

Answer. By hypothesis $X = \bigcup_n X_n$, pairwise disjoint, with $\mu(X_n) < \infty$ for all n. This produces a decomposition of $L^p(\mu)$ into summands $L^p(\mu_{X_n})$. So we may assume without loss of generality that μ is finite .

Let $\{A_n\}$ be a countable family that generates $\mathcal{A}$. For each $k \in \mathbb{N}$ let $\mathcal{A}_k = \Sigma(\{A_1, \ldots, A_k\})$, which is finite. Let $\tilde{\mathcal{A}} = \bigcup_k \mathcal{A}_k$. Then $\tilde{\mathcal{A}}$ is countable. Define an outer measure on $\mathcal{P}(X)$ by

$$\mu^*(E) = \inf\Big\{\sum_r \mu(A_{k_r}) : \ A_{k_1}, \ldots, A_{k_s} \in \tilde{\mathcal{A}}, \ E \subset \bigcup_r A_{k_r}\Big\}.$$

This is indeed, an outer measure; because we are using a measure and σ-algebra to define it, we don't need to go through the cumbersome kind of argument in Proposition 2.3.12. Indeed, if $E \in \tilde{\mathcal{A}}$ then $\mu^*(E) \le \mu(E)$ by definition. Given $\varepsilon > 0$ there exist $A_1, \ldots, A_n \in \tilde{\mathcal{A}}$ with $E \subset \bigcup_j A_j$ and $\mu\big(\bigcup_j A_j\big) \le \mu^*(E) + \varepsilon$. Then

$$\mu(E) \le \mu\Big(\bigcup_j A_j\Big) \le \mu^*(E) + \varepsilon.$$

This can be done for any $\varepsilon > 0$, so $\mu^*(E) = \mu(E)$.

By Carathéodory's Theorem there exists a σ-algebra $\mathcal{E} \subset \mathcal{P}(X)$ such that μ^* is a measure on $\mathcal{E}$. As $\mathcal{A}_k \subset \mathcal{E}$, we have that $\mathcal{A} \subset \mathcal{E}$. And, as mentioned, $\mu^*(E) = \mu(E)$ for all $E \in \tilde{\mathcal{A}}$, so μ^* extends μ to $\mathcal{E}$. What this gives us is that for any $E \in \mathcal{A}$ there exists $E' \in \tilde{\mathcal{A}}$ with $E \subset E'$ and $\mu(E' \setminus E)$ arbitrarily small. So, by Proposition 2.8.17

$$\operatorname{span}_{\mathbb{Q}}\{1_E, i\,1_F : \ E, F \in \tilde{\mathcal{A}}\}$$

is a countable dense set in $L^p(\mu)$.

2.9. The Riesz–Markov Theorem

(2.9.1) Show that under the hypotheses of Riesz–Markov, if a measure μ satisfies (2.56) then $\mu(K) < \infty$ for all K compact.

Answer. Since T is locally compact and K is compact, by Urysohn's Lemma there exists $f \in C_c(T)$, $0 \leq f \leq 1$ and $f|_K = 1$. Then

$$\mu(K) = \int_T 1_K \, d\mu \leq \int_T f \, d\mu = \varphi(f) < \infty.$$

(2.9.2) Use Riesz–Markov to construct Lebesgue measure in $\mathbb{R}^n$, $n \in \mathbb{N}$. Prove that the measure you constructed is **the** Lebesgue measure, by showing that it agrees with the Lebesgue outer measure on boxes. Show that the σ-algebra $\mathcal{M}$ from the theorem is $\mathcal{M}(\mathbb{R}^n)$.

Answer. As mentioned in Remark 2.9.5, we apply Riesz–Markov to $C_c(\mathbb{R}^n)$ and the linear functional

$$f \longmapsto \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) \, dx_1 \cdots dx_n,$$

where $\operatorname{supp} f \subset [a_1, b_1] \times \cdots \times [a_n, b_n]$. If B_1 and B_2 are two boxes such that each contains $\operatorname{supp} f$, we know from (2.33)—together with the fact that intersection of intervals is an interval—that the intersection $B_1 \cap B_2$ is a box, and of course it contains $\operatorname{supp} f$. By (2.34) we know that the complement of $B_1 \cap B_2$ is a union of boxes. So we can write $B_1 = (B_1 \cap B_2) \cup \bigcup_j C_j$, where each C_j is a box and $f = 0$ on each C_j. Then, as B_2 admits a similar decomposition,

$$\int_{B_1} f \, dx = \int_{B_1 \cap B_2} f \, dx = \int_{B_2} f \, dx.$$

So the linear functional is well-defined. Positivity is clear, as any Riemann sum of a nonnegative function will be nonnegative, and so will their limits.

Next we need to show that the measure μ such that

$$\int_B f \, dx = \int_B f \, d\mu$$

satisfies $\mu(B) = \prod_j (b_j - a_j)$. Let

$$h_{k,a,b}(t) = \begin{cases} 0, & t < a - \frac{1}{k} \text{ or } t > b + \frac{1}{k} \\ k(t-a)+1, & a - \frac{1}{k} \le t < a \\ 1, & a \le t \le b \\ k(b-t)+1, & b < t < b + \frac{1}{k} \end{cases}$$

and

$$f_k(x_1, \ldots, x_n) = \prod_{j=1}^n h_{k,a_j,b_j}(x_j).$$

Then each f_k is continuous, $f_k \searrow 1_B$, and by Monotone Convergence

$$\begin{aligned} \mu(B) &= \lim_k \int_{\mathbb{R}^n} f_k \, dx \\ &= \lim_k \prod_{j=1}^n \left[\int_{a_j - \frac{1}{k}}^{a_j} \big(k(t-a_j)+1\big)\, dt + \int_{a_j}^{b_j} 1 \, dt \right. \\ &\qquad \left. + \int_{b_j}^{b_j + \frac{1}{k}} \big(k(b_j - t)+1\big)\, dt \right] \\ &= \lim_k \prod_{j=1}^n \left[\frac{1}{k} + b_j - a_j \right] = \lim_k o\Big(\frac{1}{k}\Big) + \prod_{j=1}^n (b_j - a_j) \\ &= \prod_{j=1}^n (b_j - a_j). \end{aligned}$$

So μ agrees with m on open boxes. As the open boxes generate $\mathcal{B}(\mathbb{R}^n)$, the two measures agree on Borel sets.

The uniqueness of the measure is guaranteed by Riesz–Markov, so the only remaining question is the comparison between $\mathcal{M}$ and $\mathcal{M}(\mathbb{R}^n)$. And these are equal by (iv) in Proposition 2.9.10 and the outer regularity of μ and m, which guarantees that both measures have the same nullsets.

(2.9.3) Let X be a topological space and μ a Borel measure. The **support** of μ is the set

$$\operatorname{supp} \mu = \{x \in X : \ \mu(V) > 0 \ \text{ for all } V \text{ open with } x \in V\}.$$

Prove that $\operatorname{supp} \mu$ is closed, and that its complement is the largest open nullset.

Answer. If $x \notin \operatorname{supp}\mu$, then there exists V open with $x \in V$ and $\mu(V) = 0$. Since V is open, for any $y \in V$ there exists $W \subset V$, open, with $y \in W$. Then $\mu(W) \leq \mu(V) = 0$, so $\mu(W) = 0$ and $y \notin \operatorname{supp}\mu$. Thus $(\operatorname{supp}\mu)^c$ is open, and so $\operatorname{supp}\mu$ is closed. The argument shows that if $W \subset X$ is any open set with $\mu(W) = 0$ then $W \subset (\operatorname{supp}\mu)^c$, showing that $(\operatorname{supp}\mu)^c$ is the largest open nullset.

(2.9.4) Let X be a compact Hausdorff space, and μ a Borel measure with $\mu(X) = 1$.

(i) Show that $\mu(\operatorname{supp}\mu) = 1$.

(ii) If $H \subsetneq \operatorname{supp}\mu$ is compact, show that $\mu(H) < 1$.

Answer.

We know from Exercise 2.9.3 that $(\operatorname{supp}\mu)^c$ is the largest nullset. Then

$$1 = \mu(X) = \mu(\operatorname{supp}\mu \cup (\operatorname{supp}\mu)^c) = \mu(\operatorname{supp}\mu) + \mu((\operatorname{supp}\mu)^c) = \mu(\operatorname{supp}\mu).$$

As X is compact Hausdorff, it is normal (Exercise 2.6.1). If $H \subsetneq \operatorname{supp}\mu$, let $x \in \operatorname{supp}\mu \setminus H$. As H is compact, there exist V, W open with $V \cap W = \varnothing$ and $x \in V$, $K \subset W$ (Lemma 2.6.3). As $x \in \operatorname{supp}\mu$ we have that $\mu(V) > 0$, and so $\mu(H) \leq \mu(W) \leq 1 - \mu(V) < 1$.

(2.9.5) On $X = \mathbb{R}^2$, define

$$d((x_1, y_1), (x_2, y_2)) = \begin{cases} |y_1 - y_2|, & x_1 = x_2 \\ 1 + |y_1 - y_2|, & x_1 \neq x_2 \end{cases}$$

(i) Show that d is a metric.

(ii) Show that (X, d) is locally compact.

(iii) For $f \in C_c(X)$, show that there are only finitely many $x_1, \ldots, x_n$ such that $f(x_j, y) \neq 0$ for at least one y.

(iv) Let

$$\Lambda f = \sum_{j=1}^{n} \int_{-\infty}^{\infty} f(x_j, y)\, dy,$$

and show that $\Lambda : C_c(X) \to \mathbb{C}$ is linear and positive.

(v) Let μ be the measure corresponding to Λ via Riesz–Markov. Let $E = \{(x, 0) : \ x \in \mathbb{R}\}$. Show that $\mu(E) = \infty$ and that

$\mu(K) = 0$ for all $K \subset E$ compact. Does this contradict the Riesz–Markov theorem?

Answer.

(i) Let $(x_j, y_j) \in X$, $j = 1, 2, 3$. If $x_1 = x_2$, then

$$\begin{aligned} d((x_1, y_1), (x_2, y_2)) &= |y_1 - y_2| \le |y_1 - y_3| + |y_3 - y_2| \\ &\le d((x_1, y_1), (x_3, y_3)) + d((x_3, y_3), (x_2, y_2)). \end{aligned}$$

And if $x_1 \ne x_2$, then either $x_1 \ne x_3$ or $x_2 \ne x_3$. Then

$$\begin{aligned} d((x_1, y_1), (x_2, y_2)) &= 1 + |y_1 - y_2| \le 1 + |y_1 - y_3| + |y_3 - y_1| \\ &\le d((x_1, y_1), (x_3, y_3)) + d((x_3, y_3), (x_2, y_2)). \end{aligned}$$

So the triangle inequality holds. If $d((x_1, y_1), (x_2, y_2)) = 0$ then since $1 + |y_1 - y_2| > 0$ we get that $x_1 = x_2$ and $y_1 = y_2$ directly from the definition, so that $(x_1, y_1) = (x_2, y_2)$.

(ii) Let us first identify the balls. If $x_1 \ne x_2$ then $d((x_1, y_1), (x_2, y_2)) \ge 1$. So for a fixed (x_1, y_1) and $r > 0$, if $r < 1$ we have

$$B_r((x_1, y_1) = \{(x_1, y) : \ |y - y_1| < r\},$$

a vertical segment containing the point (x_1, y_1). If $r \ge 1$, then

$$B_r((x_1, y_1) = \{(x_1, y) : \ |y - y_1| < r\} \cup \{(x, y) : \ |y - y_1| \le r - 1\}.$$

It is important to notice that the second coordinates of the open balls are open intervals on the y-axis.

Now given $(x_1, y_1) \in X$, consider the neighbourhood $B_{1/2}((x_1, y_1))$. Its closure $\{(x_1, y) : \ |y - y_1| \le 1/2$ is compact: indeed, if $\{V_\alpha\}$ is an open cover, the sets $W_\alpha = \pi_2(V_\alpha)$ give an open cover of the segment $[y_1 - 1/2, y_1 + 1/2]$. By compactness, there is a finite subcover with indices $\alpha_1, \ldots, \alpha_n$. Then

$$\overline{B_{1/2}((x_1, y_1))} = \{x_1\} \times [y_1 - 1/2, y_1 + 1/2] \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}.$$

So $\overline{B_{1/2}((x_1, y_1))}$ is compact, and thus (X, d) is locally compact.

(iii) If $K \subset X$ is compact, then $\pi_1(K)$ is finite (because if $\pi_1(K)$ is infinite, then $\{\{x\} \times \mathbb{R} : \ x \in \pi_1(K)\}$ is an infinite open cover that does not admit a finite subcover).

On the second components, we need $\pi_2(K)$ to be compact (a continuous function maps a compact to a compact, and it is easy to check that π_2 is continuous). So the compact sets in (X, d) are precisely those of the form

$$K = \bigcup_{j=1}^{n} \{x_j\} \times L_j, \tag{AB.2.10}$$

where $L_j \subset \mathbb{R}$ is compact.

Therefore, if f has compact support, it can only be nonzero on finitely many $x_1, \ldots, x_n$.

(iv) The functional Λ is well-defined by the previous item, since the integral will occur in a compact subset of $\mathbb{R}$. More importantly, the definition is ok if we enlarge the set of x_j to include other values of x where $f(x, y) = 0$ for all y. This makes linearity trivial, since we may work with the same set $\bigcup_{j=1}^{n} \{x_j\} \times L_j$ for both f and g. Positivity is obvious.

(v) Since μ comes from Riesz–Markov, it is outer regular. We have

$$\mu(E) = \inf\{\mu(V) : \ E \subset V, \ V \ \text{ open}\}.$$

And for an open set V,

$$\mu(V) = \sup\{\Lambda f : \ f \prec V\}.$$

Let V be open with $E \subset V$. For each $x \in \mathbb{R}$, there exists $\delta_x > 0$ such that $\{x\} \times (-\delta_x, \delta_x) \subset V$. Since there are uncountably many x, there exists $\delta > 0$ such that there are infinitely many x with $\delta_x > \delta$; denote the set of such x as D.

Now, for any $n \in \mathbb{N}$ let $x_1, \ldots, x_n \in D$. Let

$$g_j(y) = \begin{cases} 1, & y \in (-\frac{\delta_{x_j}}{2}, \frac{\delta_{x_j}}{2}), \\ 0, & |y| \geq \delta_{x_j} \\ \text{linear segment in between} \end{cases}$$

and

$$f(x, y) = \sum_{j=1}^{n} \delta_{x_j}\, g_j(y).$$

Then $f \in C_c(X)_+$ with $f \prec V$. And

$$\Lambda f = \sum_{j=1}^{n} \int_{\mathbb{R}} f(x_j, y)\, dy \geq \sum_{j=1}^{n} \frac{\delta_{x_j}}{2} \geq \frac{n\delta}{2}.$$

It follows that $\mu(V) \geq n\delta$ for all n, so $\mu(V) = \infty$. As V was an arbitrary open with $E \subset V$, we get that $\mu(E) = \infty$.

For any compact $K \subset E$, we have $K = \{(x_1, 0), \ldots, (x_n, 0)\}$. Fix $\varepsilon > 0$, and let $V = \bigcup_{j=1}^{n} \{x_j\} \times (-\varepsilon/(2n), \varepsilon/(2n))$. If $f \prec V$, then

$$\Lambda f = \sum_{j=1}^{n} \int_{-\varepsilon/2}^{\varepsilon/2} f(x_j, y)\, dy \leq \sum_{j=1}^{n} \frac{\varepsilon}{n} = \varepsilon.$$

As this occurs with any $f \prec V$, we get that $\mu(V) \leq \varepsilon$. Then $\mu(K) \leq \mu(V) \leq \varepsilon$. And so $\mu(K) = 0$.

There is no contradiction because Riesz–Markov only promises inner regularity for finite-measure sets.

(2.9.6) Consider the same topological space X from Exercise 2.9.5.

(i) Show that if $E \in \mathcal{B}(X)$, then each vertical slice E_x is Borel.

(ii) Define

$$\mu(E) = \sum_x m(E_x), \qquad E \in \mathcal{B}(X).$$

Show that μ is a measure, inner regular.

(iii) Show that the μ-nullsets are those Borel sets that intersect each vertical line in a nullset. Show that the diagonal $D = \{(x,x) : \ x \in \mathbb{R}\}$ is a nullset, and that every open V with $D \subset V$ intersects each vertical line in a set of positive measure. Conclude that μ is not outer regular.

(iv) Define

$$\nu(E) = \begin{cases} \mu(E), & E \text{ intersects countably many vertical lines} \\ \infty, & \text{otherwise} \end{cases}$$

Show that ν is inner regular (on open sets), and outer regular on Borel sets.

(v) Show that

$$\int_X f\,d\mu = \int_X f\,d\nu, \qquad f \in C_c(X).$$

Answer.

(i) Let

$$\mathcal{S} = \{E \in \mathcal{B}(X) : \ E_x \in \mathcal{B}(\mathbb{R}) \ \text{ for all } x\}.$$

As shown in the proof of Proposition 2.7.2 sections preserve all σ-algebra operations. So $\mathcal{S}$ is a σ-algebra. If $0 < r \leq 1$, then

$$B_r((x_1, y_1)) = \{(x_1, y) : \ |y - y_1| < r\}.$$

So $B_r((x_1, y_1))_x = \{y : \ |y - y_1| < r\} = B_r(y_1) \in \mathcal{B}(\mathbb{R})$. And if $r > 1$,

$$B_r((x_1, y_1)) = \{(x_1, y) : \ |y - y_1| < r\} \cup \{(x, y) : \ |y - y_1| < r - 1\},$$

giving $B_r((x_1, y_1))_x = \{y : |y - y_1| < r\} = B_r(y_1) \in \mathcal{B}(\mathbb{R})$. So $\mathcal{S}$ contains the open balls and thus all the Borel σ-algebra $\mathcal{B}(\mathbb{R})$.

(ii) Since every summand is nonnegative, μ makes sense even if $m(E_x) > 0$ for uncountably many x. If $\{E_k\}$ is a countable pairwise disjoint family in $\mathcal{B}(X)$ then for fixed x the sets $\{(E_k)_x\}$ are pairwise disjoint; so

$$\mu\Big(\bigcup_k E_k\Big) = \sum_x m\bigg(\Big(\bigcup_k E_k\Big)_x\bigg) = \sum_x m\Big(\bigcup_k (E_k)_x\Big) = \sum_x \sum_k m((E_k)_x)$$
$$= \sum_k \sum_x m((E_k)_x) = \sum_k \mu(E_k).$$

First equality is the definition of μ. The third equality is the σ-additivity of m. The exchange of sums is a direct application of Tonelli's Theorem. So μ is a measure.

The compact sets are of the form (AB.2.10). Let $E \in \mathcal{B}(X)$ and $R = \{x : m(E_x) > 0\}$. We consider cases:

- R **is finite.** Say, $R = \{x_1, \ldots, x_n\}$. Fix $\varepsilon > 0$. Since $E_{x_j} \in \mathcal{B}(\mathbb{R})$, by Proposition 2.3.25 there exists $L_j \subset E_{x_j}$, compact, with $m(E_{x_j} \setminus L_j) < \varepsilon/2^j$. Then $K = \bigcup_{j=1}^{n} \{x_j\} \times L_j$ is compact, $K \subset E$, and $\mu(E \setminus K) < \varepsilon$.
- R **is countably infinite.** If $\mu(E) = \infty$, we proceed as in the uncountable case below. Otherwise, $R = \{x_1, x_2, \ldots\}$. Fix $\varepsilon > 0$. There exists n such that $\sum_{j>n} m(E_{x_j}) < \varepsilon/2$. On $x_1, \ldots, x_m$ we proceed as in the finite case to obtain K compact, $K \subset E$, and $\mu(E \setminus K) < \varepsilon$.
- R **is uncountable.** Now $\mu(E) = \infty$. Given $s > 0$ there exist $x_1, \ldots, x_n$ such that $\sum_{j=1}^{n} m(E_{x_j}) > 2s$. Choosing K as in the cases above we get $K \subset E$ and $\mu(K) > s$.

So μ is inner regular.

(iii) If $\mu(E) = 0$, then $m(E_x) = 0$ for all x. So E intersects the vertical line at x in the nullset E_x. For the diagonal,

$$\mu(D) = \sum_x m(\{x\}) = 0.$$

If V is open and $D \subset V$, then there exist numbers $r_x > 0$ such that $B_{r_x}(\{x, x\}) \subset V$ for all x. By reducing them if needed, we may assume that $r_x < 1$ for all x. Then $W = \bigcup_x B_{r_x}(x)$ is open, $D \subset W \subset V$, and

$$\mu(V) \geq \mu(W) = \sum_x m((x - r_x, x + r_x)) = \sum_x 2r_x = \infty$$

since any series with uncountably many nonzero terms is divergent. So $\mu(D) = 0$ and $\mu(V) = \infty$ for any open V with $D \subset V$, showing that μ is not outer regular.

(iv) Let V be open and nonempty. If $\nu(V) < \infty$, as $\nu(V) = \mu(V)$ we proceed as in the proof of the inner regularity of μ. When $\nu(V) = \infty$, there is an uncountable set $R \subset \mathbb{R}$ and intervals J_x, $x \in R$, with $\{x\} \times J_x \subset V$. Since there are uncountable many J_x, there exists $\delta > 0$ and $R' \subset R$, infinite, such that $m(J_x) > \delta$ for all $x \in R'$. Fix $K_x \subset J_x$, compact, with $m(K_x) > \delta$. Given $x_1, \ldots, x_n \in R'$, the set $K = \bigcup_{j=1}^n \{x_j\} \times K_{x_j} \subset V$ is compact, and $\nu(K) = \mu(K) = \sum_{j=1}^n m(K_{x_j}) > n\delta$. So we can produce $K \subset V$, compact, with $\nu(K)$ arbitrarily big.

Outer regularity is automatic when $\nu(E) = \infty$. If $\nu(E) < \infty$, then E cuts countably many vertical lines and $\mu(E) < \infty$. So there exist $\{x_1, x_2, \ldots\}$ such that $\nu(E) = \sum_j m(E_{x_j})$. Fix $\varepsilon > 0$. By Proposition 2.3.25 there exist open sets $V_j \subset \mathbb{R}$ such that $E_{x_j} \subset V_j$ and $m(V_j \setminus E_{x_j}) < \varepsilon/2^j$. In turn we can decompose each V_j as a countable union of $V_{j,k}$ with $V_{j,k}$ an interval of length less than 1. Then $V = \bigcup_{j,k} \{x_j\} \times V_{j,k}$ is a union of balls, so open, $E \subset V$, and

$$\nu(V \setminus E) = \sum_{j=1}^{\infty} m(V_j \setminus E_{x_j}) < \varepsilon.$$

We cannot have full inner regularity for ν. For instance let $E = \{(x, x+q) : \ x \in \mathbb{R},\ q \in \mathbb{Q}\}$. Then E crosses uncountably many vertical lines and so $\nu(E) = \infty$. If $K \subset E$ is compact, $K = \bigcup_{j=1}^n \{x_j\} \times L_j$ with L_j compact. As $K \subset E$ this means that each L_j is countable. Then $\mu(K) = \sum_j \mu(K_{x_j}) = \sum_j m(L_j) = 0$. Hence $\nu(E) = \infty$, while $\mu(E) = 0$.

(v) If $f \in C_c(X)$, let $K = \operatorname{supp} T$. As K is compact, $\mu(K) < \infty$, and so $\nu(K) = \mu(K)$. For any Borel set $E \subset K$, we also have $\nu(E) = \mu(E)$. So the integrals will agree on any simple function that approximates f, and thus

$$\int_X f\, d\mu = \int_X f\, d\nu.$$

2.10. Complex Measures and Differentiation

(2.10.1) Let $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ such that $\sum_n c_n$ converges. Show that $\sum_n c_n$ converges absolutely if and only if the limit does not change under permutations; concretely, show that the following statements are equivalent:

(i) $\sum_n c_n$ converges absolutely;

(ii) $\sum_n c_{\sigma(n)} = \sum_n c_n$ for all permutations $\sigma : \mathbb{N} \to \mathbb{N}$;

(iii) $\sum_n c_{\sigma(n)}$ converges for all permutations $\sigma : \mathbb{N} \to \mathbb{N}$.

Answer. (i) $\implies$ (ii) Let $\sigma : \mathbb{N} \to \mathbb{N}$ be a bijection. Write $L = \sum_n c_n$. Suppose that $\sum_n |c_n| < \infty$ and let $\varepsilon > 0$. Then there exists n_0 such that $\sum_{n>n_0} |c_n| < \varepsilon$. Let $n_1 = \max\{\sigma^{-1}(k) : k = 1, \dots, n_0\}$. If $n > n_1$, then $n \neq \sigma^{-1}(k)$ for all $k = 1, \dots, n_0$; so $\sigma(n) > n_0$. Then

$$\left| L - \sum_{n=1}^{n_1} c_{\sigma(n)} \right| \leq \sum_{n>n_1} |c_{\sigma(n)}| \leq \sum_{n>n_0} |c_n| < \varepsilon,$$

showing that $\sum_n c_{\sigma(n)} = \sum_n c_n$.

(ii) $\implies$ (iii) Trivial.

(iii) $\implies$ (i) Suppose first that $c_n \in \mathbb{R}$ for all n. Then we can write $c_n = c_n^+ - c_n^-$ as a difference of non-negative numbers. Because $\sum_n c_n$ converges we have that $c_n \to 0$, so $c_n^+ \to 0$ since $c_n^+ \leq |c_n|$. If $\sum_n c_n^+ = \infty$, then we can choose $1 = n_0 < n_1 < n_2 < \cdots$ such that $\sum_{n_k \leq n < n_{k+1}} c_n^+ > 1 + c_k^-$ for all k. So if σ is the permutation that gives the order

$$c_1^+, \dots, c_{n_1-1}^+, c_1^-, c_{n_1}^+, \dots, c_{n_2-1}^+, c_2^-, \dots$$

For all k we have

$$\sum_{j=1}^{n_{k+1}} c_{\sigma(j)} > \sum_{r=1}^{k} 1 = k,$$

and this would make $\sum_n c_{\sigma(n)} = \infty$, a contradiction. So $\sum_n c_n^+ < \infty$. Repeating the argument with the series $\sum_n (-c_n)$ gives us $\sum_n c_n^- < \infty$. Then $\sum_n |c_n| \leq \sum_n c_n^+ + \sum_n c_n^- < \infty$ and the series converges absolutely. In the general case, the convergence of $\sum_n c_n$ is equivalent to that of $\sum_n \operatorname{Re} c_n$ and $\sum_n \operatorname{Im} c_n$, so by the above these latter two series converge absolutely and then from $|c_n| \leq |\operatorname{Re} c_n| + |\operatorname{Im} c_n|$ we get that $\sum_n |c_n| < \infty$.

(2.10.2) Prove Proposition 2.10.6.

Answer.

(i) Suppose that λ is concentrated on A. Since $\lambda(E) = \lambda(E \cap A)$ for all E, the same happens for each partition in the definition of $|\lambda|$; so $|\lambda|$ is concentrated in A.

(ii) If $\lambda_1 \perp \lambda_2$ then $|\lambda_1| \perp |\lambda_2|$, as they will be concentrated respectively in the same disjoint sets.

(iii) If $\lambda_1, \lambda_2 \perp \mu$, say λ_1 is concentrated on A_1, λ_2 on A_2, and μ on B. Then $A_1 \cap B = A_2 \cap B = \varnothing$. Thus $(A_1 \cup A_2) \cap B = \varnothing$. As $\lambda_1 + c\lambda_2$ is concentrated on $A_1 \cup A_2$, we get that $\lambda_1 + c\lambda_2 \perp \mu$.

(iv) If $\lambda_1, \lambda_2 \ll \mu$ and $\mu(E) = 0$, then $(\lambda_1 + c\lambda_2)(E) = \lambda_1(E) + c\lambda_2(E) = 0 + c0 = 0$. So $\lambda_1 + c\lambda_2 \ll \mu$.

(v) If $\lambda \ll \mu$ and $\mu(E) = 0$, then $\mu(E_j) = 0$ for all j and any partition of E, giving us $|\lambda|(E) = 0$. So $|\lambda| \ll \mu$.

(vi) If $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$, let A_2 be a measurable set where λ_2 is concentrated, and B a measurable set where μ is concentrated. Then $A_2 \cap B = \varnothing$. Since $\mu(B^c) = 0$, we have $\lambda_1(B^c) = 0$, so $\lambda_1(E) = \lambda_1(E \cap B)$ for all E. So λ_1 is concentrated on B, and thus $\lambda_1 \perp \lambda_2$.

(vii) If $\lambda \ll \mu$ and $\lambda \perp \mu$, by the previous paragraph we have that $\lambda \perp \lambda$. This can only happen if λ is concentrated on the empty set: that is, $\lambda = 0$.

(viii) Since E is itself a partition, $|\lambda(E)| \leq |\lambda|(E)$. If $|\lambda|(E) = 0$, then $\lambda(E) = 0$. Thus $\lambda \ll |\lambda|$.

(2.10.3) Let μ be a complex measure on a σ-algebra $\mathcal{A}$. For $E \in \mathcal{A}$, define

$$\lambda(E) = \sup\left\{ \sum_{j=1}^m |\mu(E_j)| : \ E_j \in \mathcal{A} \ \text{ disjoint, } \bigcup_j E_j = E \right\}.$$

Show that $\lambda = |\mu|$.

Answer. Since the definition of $|\mu|$ allows countable partitions, we have $\lambda \leq |\mu|$. By Theorem 6.4, $|\mu|(E) < \infty$. Now fix $\varepsilon > 0$ and let $\{E_j\}_{j=1}^\infty$ be a partition of E such that $\sum_j |\mu(E_j)| \geq |\mu|(E) - \varepsilon/2$. As $|\mu|(E) \geq$

$\sum_{j=1}^{\infty} |\mu(E_j)|$, there exists m such that

$$\sum_{j=1}^{m} |\mu(E_j)| \geq \sum_{j=1}^{\infty} |\mu(E_j)| - \varepsilon/2 \geq |\mu|(E) - \varepsilon/2 - \varepsilon/2.$$

So $\lambda(E) \geq |\mu|(E) - \varepsilon$. As ε was arbitrary, $\lambda(E) \geq |\mu|(E)$, showing the equality.

(2.10.4) Let λ_1, λ_2 be mutually singular complex measures on a σ-algebra $\mathcal{A}$ over X. Show that

$$|\lambda_1 + \lambda_2| = |\lambda_1| + |\lambda_2|.$$

Answer. Fix $E \in \mathcal{A}$ and $\{E_n\} \subset \mathcal{A}$ a countable partition of E. Then

$$\sum_k |(\lambda_1 + \lambda_2)(E_k)| \leq \sum_k |\lambda_1(E_k)| + \sum_k |\lambda_2(E_k)| \leq |\lambda_1|(E) + |\lambda_2|(E).$$

So $|\lambda_1 + \lambda_2|(E) \leq |\lambda_1|(E) + |\lambda_2|(E)$.

By definition there exist disjoint $A_1, A_2 \in \mathcal{A}$ with $A_1 \cup A_2 = X$ and such that λ_1 is concentrated on A_1 and λ_2 is concentrated on A_2. Fix $E \in \mathcal{A}$ and $\varepsilon > 0$. Then there exist partitions $\{E_n\}$ and $\{F_n\}$ of E such that

$$\sum_k |\lambda_1(E_k)| + \varepsilon > |\lambda_1|(E), \qquad \sum_k |\lambda_2(E_k)| + \varepsilon > |\lambda_2|(E).$$

We have

$$\sum_k |\lambda_1(E_k)| = \sum_k |\lambda_1(E_k \cap A_1)| = \sum_k |(\lambda_1 + \lambda_2)(E_k \cap A_1)|,$$

and

$$\sum_k |\lambda_1(E_k)| = \sum_k |\lambda_2(F_k \cap A_2)| = \sum_k |(\lambda_1 + \lambda_2)(F_k \cap A_2)|.$$

Since $\{E_k \cap A_1\}$ and $\{F_k \cap A_2\}$ are partitions of $E \cap A_1$ and $E \cap A_2$ respectively, their union is a partition of E, and hence

$$\begin{aligned}(|\lambda_1| + |\lambda_2|)(E) &\leq 2\varepsilon + \sum_k |\lambda_1(E_k)| + \sum_k |\lambda_1(E_k)| \\ &= 2\varepsilon + \sum_k |(\lambda_1 + \lambda_2)(E_k \cap A_1)| + \sum_k |(\lambda_1 + \lambda_2)(F_k \cap A_2)| \\ &\leq 2\varepsilon + |\lambda_1 + \lambda_2|(E).\end{aligned}$$

As the inequality holds for all $\varepsilon > 0$ and the reverse inequality was already shown, we have $|\lambda_1 + \lambda_2|(E) = |\lambda_1|(E) + |\lambda_2(E)|$.

(2.10.5) Use Proposition 2.10.7 for an alternative proof of Exercise 2.5.8.

Answer. Define a measure $\tilde{\gamma}$ on $[0,1]$ by $\tilde{\gamma}(E) = \int_E |g|\,dm$. Then $\tilde{\gamma} \ll m$. Fix $s \in [0,1]$ and $\varepsilon > 0$. Let $\delta > 0$ be as in Proposition 2.10.7. If $t > s$ and $|t-s| < \delta$, then $m([s,t]) < \delta$ and so

$$|\gamma(t) - \gamma(s)| = \left| \int_s^t |g| \right| = |\tilde{\gamma}([s,t])| < \varepsilon.$$

For $t < s$ the estimate is entirely similar. Hence γ is continuous.

(2.10.6) *(Radon–Nikodym and Lebesgue's Decomposition can fail when X is not σ-finite)* Let μ be the Lebesgue measure on $(0,1)$ and λ the counting measure on $\mathcal{M}((0,1))$. Show that λ has no Lebesgue decomposition relative to μ, and that $\mu \ll \lambda$ and μ is bounded, but there is no measurable function h such that $d\mu = h\,d\lambda$.

Answer. Suppose that $\lambda = \lambda_a + \lambda_s$ with $\lambda_a \ll \mu$. As $\lambda_s \perp \mu$, there exists a μ-nullset A with λ_s supported on A. On $B = (0,1) \setminus A$, $\lambda_s(B) = 0$, $\mu(B) = 1$, $\lambda_a(B) = \lambda(B) = \infty$. But for any $b \in B$ we get

$$1 = \lambda(\{b\}) = \lambda_a(\{b\}), \qquad \text{while} \qquad \mu(\{b\}) = 0,$$

a contradiction.

For the second part, we trivially have $\mu \ll \lambda$, since λ is zero only on the empty set. Now suppose that $\mu(E) = \int_E h\,d\lambda$ for all $E \in \mathcal{M}((0,1))$. Then, for any $t \in [0,1]$,

$$0 = \mu(\{t\}) = \int_{\{t\}} h\,d\lambda = h(t).$$

So we would have $h = 0$, a contradiction.

(2.10.7) Suppose that $\{g_n\}$ is a sequence of positive continuous functions on $[0,1]$. Let μ be a positive Borel measure on $[0,1]$, and suppose also that

(i) $\lim_{n\to\infty} g_n(x) = 0$ a.e.[m]

(ii) $\displaystyle\int_{[0,1]} g_n\,dm = 1$ for all n

(iii) $\lim_{n\to\infty} \int_{[0,1]} f g_n \, dm = \int_{[0,1]} f \, d\mu$ for all $f \in C[0,1]$.

Does it follow that $\mu \perp m$?

(Hint: Egorov's Theorem. There are probably other ways to attack this)

Answer. Yes. Let $E = \{x : \lim_n g_n(x) = 0\}$. By hypothesis, $m(E) = 1$. So m is concentrated on E: if $E \cap A = \varnothing$, then $m(A) = 0$. It remains to show that μ is concentrated on $[0,1] \setminus E$.

Because the g_n are continuous, E is Borel:

$$E = \bigcap_m \bigcup_n \bigcap_{k \geq n} \left\{ g_k < \frac{1}{m} \right\}.$$

Fix $\varepsilon > 0$. By Egorov's Theorem, there exists a Borel set $B \subset E$ such that $\mu(E \setminus B) < \varepsilon$ and $g_n \to 0$ uniformly on B; so there exists n_0 such that, on B, $g_k < \varepsilon$ whenever $k \geq n_0$. Then, for any $f \in C[0,1]$ and $n \geq n_0$

$$\int_B f \, g_n \, dm \leq \varepsilon \int_B |f| \, dm.$$

We can do this for any ε, so we conclude that $\lim_n \int_B f \, g_n \, dm = 0$. Then

$$\int_{[0,1]} f \, d\mu = \lim_{n\to\infty} \int_B f g_n \, dm + \int_{B^c} f g_n \, dm = \lim_{n\to\infty} \int_{B^c} f g_n \, dm.$$

If we look carefully at the proof of Egorov's Theorem, it is clear that we can assume B to be open: that's because the sets in the proof can be defined in terms of strict inequality which makes them open, and then B is obtained as a (countable, although not important for this) union of open sets.

Since $B \subset [0,1]$ is open, it can be written as a countable union of open intervals, and from there we see that we can construct $\{f_j\} \subset C[0,1]$ such that $f_j \nearrow 1_B$. Then

$$\mu(B) = \int_{[0,1]} 1_B \, d\mu = \lim_j \int_{[0,1]} f_j \, d\mu = \lim_j \lim_n \int_{B^c} f_j \, g_n \, dm = 0,$$

since each f_j is supported in B.

We have shown that for each $\varepsilon > 0$ there exists $B \subset E$ with $\mu(E \setminus B) < \varepsilon$ and $\mu(B) = 0$. Then

$$\mu(E) = \mu(E \setminus B) + \mu(B) = \mu(E \setminus B) < \varepsilon$$

for all $\varepsilon > 0$, and thus $\mu(E) = 0$. So μ is supported in E^c, which is a nullset for m.

(2.10.8) Let $(X, \mathcal{A})$ be a measurable space, and μ, λ positive measures such that μ is σ-finite and $\lambda \ll \mu$. Let $g = d\lambda/d\mu$. Show that for any measurable $f : X \to \mathbb{C}$ such that $\int_X f\, d\lambda$ exists,

$$\int_X f\, d\lambda = \int_X fg\, d\mu. \tag{2.63}$$

Answer. The Radon–Nikodym derivative definition gives us that

$$\int_X 1_E\, d\lambda = \lambda(E) = \int_E g\, d\mu = \int_X 1_E\, g\, d\mu.$$

By linearity we obtain that

$$\int_X s\, d\lambda = \int_X s\, g\, d\mu$$

for all simple functions s. For $f \geq 0$ measurable, there exists a sequence $\{s_n\}$ with $0 \leq s_n \nearrow f$. Then Monotone Convergence gives us

$$\int_X f\, d\lambda = \int_X s\, f\, d\mu.$$

Now for general f, we can write $f = f_1 - f_2 + i(f_3 - f_4)$ with $f_j \geq 0$ and at least three of their integrals are finite; for each f_j the above applies, and at least three of the four integrals against $g\, d\mu$ are finite, so (2.63) occurs by linearity.

(2.10.9) Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $f \in L^1(X)$. Define $\nu : \mathcal{A} \to \mathbb{C}$ by

$$\nu(E) = \int_E f\, d\mu.$$

(i) Show that ν is a complex measure on $(X, \mathcal{A})$ with $\nu \ll \mu$.

(ii) Show that for any measurable function $g : X \to \mathbb{C}$,

$$\int_X g\, d\nu = \int_X gf\, d\mu \tag{2.64}$$

if the left integral exists.

(iii) Conclude that $f = d\nu/d\mu$ a.e. (μ).

Answer.

(i) By writing $f = f_1 - f_2 + i(f_3 - f_4)$ with $f_j \geq 0$ for all j we get that $\nu = \nu_1 - \nu_2 + i(\nu_3 - \nu_4)$, where $\nu_j(E) = \int_E f_j\, d\mu$. These are positive measures, all finite since

$$|\nu_j(E)| \le \int_E |f_j|\, d\mu \le \int_E |f|\, d\mu \le \|f\|_1 < \infty.$$

By Exercise 2.5.4, if $\{E_n\} \subset \mathcal{A}$ is a pairwise disjoint sequence,

$$\nu_j\Big(\bigcup_n E_n\Big) = \int_{\bigcup_n E_n} f_j\, d\mu = \sum_n \int_{E_n} f_j\, d\mu = \sum_n \nu_j(E_n).$$

By linearity, we get that ν is a complex measure.

If $\mu(E) = 0$, then $0 = \int_E f\, d\mu = \nu(E)$. So $\nu \ll \mu$.

(ii) This is Exercise 2.10.8 .

(iii) Let $h = d\nu/d\mu$. Then by Exercise 2.10.8 we have

$$\int_X g(h - f)\, d\mu = 0$$

for all g measurable such that $\int_X g\, d\nu$ exists. Using simple functions and Dominated Convergence we get

$$\int_X |h - f|^2\, d\mu = 0,$$

and so $h = f$ a.e. (μ).

(2.10.10) Let $(X, \mathcal{A})$ be a measurable space, and let μ, λ be σ-finite positive measures such that $\lambda \ll \mu$ and $\mu \ll \lambda$. Show that

$$\frac{d\lambda}{d\mu} = \left(\frac{d\mu}{d\lambda}\right)^{-1} \qquad \text{a.e. } (\mu, \lambda).$$

Answer. Write $f = d\lambda/d\mu$ and $g = d\mu/d\lambda$. Then, for any $E \in \mathcal{A}$,

$$\int_E 1\, d\lambda = \lambda(E) = \int_E f\, d\mu = \int_E fg\, d\lambda.$$

As E is arbitrary, by Exercise 2.5.3 $1 - fg = 0$ a.e. (λ). Exchanging roles we get that $1 - gf = 0$ a.e. (μ).

(2.10.11) If $d\nu = f\, d\mu$ and $d\lambda = g\, d\nu$ for positive measures μ and ν, show that $\lambda \ll \mu$ and find the Radon–Nikodym derivative $d\lambda/d\mu$.

Answer. Using Exercise 2.10.8 we have

$$\lambda(E) = \int_E g\,d\nu = \int_E fg\,d\mu.$$

This shows that $d\lambda = fg\,d\mu$, so $\lambda \ll \mu$ and the Radon–Nikodym derivative is fg.

(2.10.12) Let (X, Σ) be a measurable space and $M(X)$ the set of complex measures on Σ. Show that, with the norm $\|\mu\| = |\mu|(X)$ and the obvious addition and multiplication by scalars, $M(X)$ is complete with the metric $d(\mu, \eta) = \|\mu - \eta\|$.

Answer. Since $|\mu|$ is defined in terms of a supremum and sums of absolute values, the triangle inequality for the absolute value gives $\|\mu+\nu\| \leq \|\mu\|+\|\nu\|$. Similarly, $\|c\,\mu\| = |c|\,\|\mu\|$. If $\|\mu\| = 0$, then $|\mu|(X) = 0$, so $|\mu| = 0$ and thus $\mu = 0$. So the norm is indeed a norm.

It remains to check completeness. Let $\{\mu_k\}$ be a Cauchy sequence. For any $E \in \mathcal{B}(X)$, we have

$$\begin{aligned}|\mu_k(E) - \mu_j(E)| = |(\mu_k - \mu_j)(E)| &\leq |\mu_k - \mu_j|(E)\\ &\leq |\mu_k - \mu_j|(X) = \|\mu_k - \mu_j\|.\end{aligned}$$

So the number sequence $\{\mu_k(E)\}$ is (uniformly) Cauchy in $\mathbb{C}$, and thus uniformly convergent to a number $\mu(E)$. Linearity of the limit gives us immediately that μ is additive, and so we need to check that it is σ-additive. Let $\{E_n\} \subset \Sigma$, pairwise disjoint. Put $E = \bigcup_n E_n$ and

$$F_m = E \setminus \bigcup_{n=1}^{m} E_n.$$

Then $\bigcap_n F_n = \varnothing$. Given $\varepsilon > 0$ there exists k such that $\|\mu_j - \mu_k\| < \varepsilon$ for all $j \geq k$. Because μ_k is a complex measure there exists n_0 such that $\mu_k(F_n) < \varepsilon/2$ for all $n \geq n_0$. Using the additivity of μ, for $m \geq n_0$

$$\begin{aligned}\left|\mu(E) - \sum_{n=1}^{m} \mu(E_n)\right| = |\mu(F_m)| &\leq |\mu(F_m) - \mu_k(F_m)| + |\mu_k(F_m)|\\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

This shows that $\mu(E) = \sum_n \mu(E_n)$ and so μ is a complex measure.

Here is a different argument to show the completeness. Given a Cauchy sequence $\{\mu_k\}$ in $M(X)$, define

$$\lambda = \sum_{k=1}^{\infty} \frac{1}{n^2}\,|\mu_k|.$$

This is a finite positive measure (the σ-additivity follows easily from the σ-additivity of each μ_k, the uniform boundedness of the $\{\mu_k\}$, and the uniform convergence of the series. Clearly $\mu_k \ll \lambda$ for each k, so by Theorem 2.10.10 there exist functions $f_k \in L^1(\lambda)$ with $d\mu_k = f_k\,d\lambda$. We have, via Proposition 2.10.12,

$$\|f_k - f_j\|_1 = \int_X |f_k - f_j|\,d\lambda = \|\mu_k - \mu_j\|.$$

Therefore the sequence $\{f_k\} \subset L^1(\lambda)$ is Cauchy. Let $f = \lim_k f_k$ and put $\mu = f\,d\lambda$. Then $\mu \in M(X)$ and

$$\|\mu_k - \mu\| = \|f_k - f\|_1 \to 0.$$

(2.10.13) Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, and λ_1, λ_2 positive σ-finite measures with $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$. Give a necessary and sufficient condition, in terms of the Radon–Nikodym derivatives, for $\lambda_1 \ll \lambda_2$. In such case, express $d\lambda_1/d\lambda_2$ in terms of $d\lambda_1/d\mu$ and $d\lambda_2/d\mu$.

Answer. Let f, g be the respective Radon–Nikodym derivatives, so that $\lambda_1 = f\,d\mu$, $\lambda_2 = g\,d\mu$. The necessary and sufficient condition is that

$$\{g = 0\} \cap E \subset \{f = 0\} \cap E \quad \text{for all } E \quad \text{such that} \quad \mu(E) > 0.$$

If $\lambda_1 \ll \lambda_2$ and $\mu(E) > 0$, we have that $g|_E = 0$ a.e. (μ) implies $f|_E = 0$ a.e. (μ). So $\{g = 0\} \cap E \subset \{f = 0\} \cap E$ for all E such that $\mu(E) > 0$.

Conversely, assume $\{g = 0\} \cap E \subset \{f = 0\} \cap E$ for all E such that $\mu(E) > 0$ and suppose $\lambda_2(E) = 0$. If $\mu(E) = 0$, then $\lambda_1(E) = 0$. If $\mu(E) > 0$, then $g|_E = 0$ a.e. (μ) and so by the assumption we have $f|_E = 0$ a.e. (μ). So $\lambda_1(E) = 0$ and therefore $\lambda_1 \ll \lambda_2$.

For the Radon–Nikodym derivatives, note that from $d\lambda_2 = g\,d\mu$ we get, if $E \subset \{g \neq 0\}$,

$$\mu(E) = \int_E 1\,d\mu = \int_E \frac{1}{g}\,g\,d\mu = \int_E \frac{1}{g}\,d\lambda_2.$$

So $d\mu = \frac{1}{g}\,d\lambda_2$ on those sets where $g \neq 0$.

When $\lambda_1 \ll \lambda_2$ we have, with $h = d\lambda_1/d\lambda_2$ and $E \subset \{g \neq 0\}$,

$$\int_E h\, d\lambda_2 = \lambda_1(E) = \int_E f\, d\mu = \int_E \frac{f}{g}\, d\lambda_2.$$

So

$$\frac{d\lambda_1}{d\lambda_2} = \frac{f}{g}\, 1_{\{g\neq 0\}} = \frac{d\lambda_1/d\mu}{d\lambda_2/d\mu}\, 1_{\{d\lambda_2/d\mu\neq 0\}}.$$

(2.10.14) Let (X, Σ, μ) be a σ-finite measure space. Show that there exists a finite measure ν on Σ such that $\nu \ll \mu$ and $\mu \ll \nu$. Does such a ν always exist when μ is not necessarily σ-finite?

Answer. By hypothesis we have $X = \bigcup_n X_n$ with $\{X_n\} \subset \Sigma$, pairwise disjoint and $\mu(X_n) < \infty$ for all n. Let

$$\nu(E) = \sum_n \frac{1}{2^n \mu(X_n)}\, \mu(E \cap X_n), \qquad E \in \Sigma.$$

Then $\nu(X) = 1$, $\nu(\varnothing) = 0$, and if $\{E_k\} \subset \Sigma$ are pairwise disjoint we have, using Tonelli,

$$\begin{aligned}\nu\Big(\bigcup_k E_k\Big) &= \sum_n \frac{1}{2^n \mu(X_n)} \sum_k \mu(E_k \cap X_n) \\ &= \sum_k \sum_n \frac{1}{2^n \mu(X_n)} \mu(E_k \cap X_n) = \sum_k \nu(E_k).\end{aligned}$$

So ν is a finite measure on Σ. If $\mu(E) = 0$ then $\nu(E) = 0$, so $\nu \ll \mu$. And if $\nu(E) = 0$, then $\mu(E \cap X_n) = 0$ for all n, and therefore

$$\mu(E) = \sum_n \mu(E \cap X_n) = 0.$$

Thus $\mu \ll \nu$.

The σ-finiteness is crucial. Let $X = [0, 1]$ and μ the counting measure. This is not σ-finite. Suppose that ν is a finite measure on $\Sigma = \mathcal{P}([0, 1])$. We will always have $\nu \ll \mu$, for $\mu(E) = 0$ if and only if $E = \varnothing$. If we had $\mu \ll \nu$, this means that ν cannot be zero on any nonempty set. Let $R_n = \{t \in [0, 1] : \ \nu(\{t\}) > 1/n\}$. If R_n were finite for all $n \in \mathbb{N}$, then the set $\{t : \ \nu(\{t\} \neq 0\}$ would be countable, contradicting that ν is nonzero on all nonempty sets. Hence there exists n such that R_n is infinite. Then $\nu(R_n) \geq \sum_{t\in R_n} 1/n = \infty$, showing that ν cannot be finite.

(2.10.15) Show that a linear combination of absolutely continuous functions is again absolutely continuous.

Answer. It is enough to show that if $f_1, f_2 : [a,b] \to \mathbb{C}$ are absolutely continuous and $c \in \mathbb{C}$ then $cf_1 + f_2$ is absolutely continuous. Let $\varepsilon > 0$. Then there exist δ_k, $k = 1, 2$, such that

$$\sum_{j=1}^{n} |f_k(b_j) - f_k(a_j)| < \frac{\varepsilon}{2(|c|+1)}$$

for any partition $a \leq a_1 < b_1 < a_2 < b_2 < \cdots < b_n \leq b$ with $\sum_j |b_j - a_j| < \delta_k$. Put $\delta = \min\{\delta_1, \delta_2\}$. Then, for a partition

$$a \leq a_1 < b_1 < a_2 < b_2 < \cdots < b_n \leq b$$

with $\sum_j |b_j - a_j| < \delta$, we have

$$\begin{aligned}\sum_{j=1}^{n} |cf_1(b_j) + f_2(b_j) - cf_1(a_j) - f_2(a_j)| &\leq \sum_{j=1}^{n} |c|\,|f_1(b_j) - f_1(a_j)| \\ &\quad + \sum_{j=1}^{n} |f_2(b_j) - f_2(a_j)| \\ &< |c| \frac{\varepsilon}{2(|c|+1)} + \frac{\varepsilon}{2(|c|+1)} \leq \varepsilon.\end{aligned}$$

(2.10.16) Assume that both f and Mf are in $L^1(\mathbb{R}^n)$. Prove that $f = 0$ a.e.

Answer. Assume, without loss of generality, that $f \geq 0$. Suppose that f is not zero a.e. Then there exists some ball $B_r(x_0)$ with $c_0 = \int_{B_r(x_0)} f > 0$ (otherwise, $f = 0$ a.e. by Theorem 2.11.13). For any x with $\|x\| \geq \|x_0\| + r$, we have the inclusion $B_r(x_0) \subset B_{2\|x\|}(x)$ (since, for y with $\|y\| \leq r$, $\|x_0 + y - x\| \leq \|x_0\| + r + \|x\| \leq 2\|x\|$). Then, for x with $\|x\| \geq \|x_0\| + r$,

$$\begin{aligned}Mf(x) &\geq \frac{1}{m(B_{2\|x\|}(x))} \int_{B_{2\|x\|}(x)} f \\ &\geq \frac{1}{m(B_{2\|x\|}(x))} \int_{B_r(x_0)} f \\ &\geq \frac{c_0}{m(B_{2\|x\|}(x))} = \frac{c}{\|x\|^n}\end{aligned}$$

for some constant c. Note that one can use the definition of the Lebesgue measure, via boxes, to deduce that $m(\alpha E) = \alpha^n m(E)$, so we don't need to actually know the formula for the volume of a ball above.

Then, with $R = \|x_0\| + r$, using Tonelli, and writing $m(B_s(0)) = s^n d$,

$$\begin{aligned}\int_{\mathbb{R}^n} Mf &\geq \int_{\|x\|\geq R} \frac{c}{\|x\|^n}\,dx = c\int_{\|x\|\geq R}\int_{\|x\|}^{\infty} \frac{n}{s^{n+1}}\,ds\,dx \\ &= cn\int_R^{\infty}\int_{\|x\|\leq s} \frac{1}{s^{n+1}}\,dx\,ds = cn\int_R^{\infty} \frac{m(B_s(0))}{s^{n+1}}\,ds \\ &= cdn\int_R^{\infty} \frac{1}{s}\,ds = \infty.\end{aligned}$$

(2.10.17) Let X be a locally compact Hausdorff space. We say that a complex measure μ on $\mathcal{B}(X)$ is **regular** if $|\mu|$ is regular, and **Radon** if $|\mu|$ is Radon. Let γ be a positive Radon measure on $\mathcal{B}(X)$, $g \in L^\infty(X)$, and define a complex Borel measure μ by

$$\mu(E) = \int_E g\,d\gamma.$$

Show that μ is regular if γ is regular, and Radon if γ is Radon.

Answer. By Proposition 2.10.12,

$$|\mu|(E) = \int_E |g|\,d\gamma.$$

If $E \in \mathcal{B}(X)$ and $\varepsilon > 0$, by the outer regularity of γ there exists V open with $E \subset V$ and $\gamma(V \setminus E) < \varepsilon$. Then

$$|\mu|(V \setminus E) = \int_{V\setminus E} |g|\,d\gamma \leq \|g\|_\infty\,\gamma(V \setminus E) < \|g\|_\infty\,\varepsilon.$$

Thus $|\mu|(E) = \inf\{|\mu|(V):\ V\ \text{open and } E \subset V\}$. Similarly, if $K \subset E$ and $\gamma(E \setminus K) < \varepsilon$, then with the same inequality $|\mu|(E \setminus K) \leq \|g\|_\infty\,\varepsilon$. Finally, if K is compact, $|\mu|(K) \leq \|g\|_\infty\,\gamma(K) < \infty$.

(2.10.18) Let $X \subset \mathbb{C}$ be compact and μ a complex Radon measure on X. Show that if $\int_X f\,d\mu \geq 0$ for all polynomials f such that $f(X) \subset [0, \infty)$, then μ is a positive measure.

Answer. Since X is compact we may approximate any continuous function uniformly by polynomials (Stone–Weierstrass: Theorem 7.4.20). So we obtain that $\int_X f\,d\mu \geq 0$ for all $f \in C(X)$. Given any closed $E \subset X$, choose a decreasing sequence $\{V_n\}$ of open sets with $m(V_n \setminus E) < 1/n$. By Urysohn's Lemma (Theorem 2.6.5) there exist continuous functions f_n, with $0 \leq f_n \leq 1$, $f_n|_E = 1$, and supported in V_n. We have $f_n \to 1_E$ a.e., since

$$m\Big(E \cap \bigcap_n V_n\Big) = m\Big(\bigcap_n V_n \setminus E\Big) = \lim_n m(V_n \setminus E) = 0.$$

Then by Dominated Convergence we get

$$\mu(E) = \int_X 1_E\,d\mu = \lim_n \int_X f_n\,d\mu \geq 0.$$

In particular $\mu(X) \geq 0$. For any $V \subset X$ open, $\mu(V) = \mu(X) - \mu(X \setminus E) \geq 0$. And then, by regularity, $\mu(E) \geq 0$ for any Borel set E.

(2.10.19) Let X be a compact Hausdorff space and μ a regular complex Borel measure on $\mathcal{A}$. Let $\varphi : C(X) \to \mathbb{C}$ be given by $\varphi(f) = \int_X f\,d\mu$. Show that φ is multiplicative if and only if $\mu = \delta_{x_0}$ for some $x_0 \in X$.

Answer. We have $\delta_{x_0}(fg) = (gf)(x_0) = f(x_0)g(x_0) = \delta_{x_0}(f)\delta_{x_0}(g)$.

Conversely, assume that φ is multiplicative. There is a fairly sleek proof of this in Proposition 7.4.6 (that does not even require the existence of the measure μ, let alone its regularity), but we will provide an ad hoc proof here. We have the advantage that we know that the multiplicative functional is given by a measure.

Given $E \subset X$ Borel, by Corollary 2.6.14 there exists a sequence $\{g_n\} \subset C(X)$ such that $g_n \to 1_E$. We may assume that the sequence is non-negative and uniformly bounded by 2, as we can replace g_n with $p \circ q \circ g_n$, where $p(x) = \max\{x, 0\}$ and $q(x) = \min\{x, 2\}$. Then Dominated convergence applies (as $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$ with μ_j positive finite measures) and so

$$\mu(E) = \int_X 1_E\,d\mu = \lim_n \int_X g_n\,d\mu = \lim_n \varphi(g_n).$$

If F is another Borel subset of X and $\{h_n\} \subset C(X)$ are bounded positive functions that converge pointwise to h, then $g_n h_n \to 1_E 1_F = 1_{E\cap F}$. Hence

$$\mu(E \cap F) = \lim_n \varphi(g_n h_n) = \lim_n \varphi(g_n)\varphi(h_n) = \mu(E)\mu(F).$$

This gives $\mu(E) = \mu(E \cap E) = \mu(E)^2$. So $\mu(E) \in \{0, 1\}$ for all Borel sets E; in particular, μ is a positive measure. We can write $1 = \mu(X) = \mu(E \cup E^c) = \mu(E) + \mu(E^c)$ so $\mu(E) = 1$ if and only if $\mu(E^c) = 0$. We deduce that if $E_1, E_2, \ldots$ are Borel and disjoint with $\bigcup_n E_n = X$, then there exists k such that $\mu(E_k) = 1$ and $\mu(E_j) = 0$ for all $j \neq k$.

Let

$$X_0 = \bigcap \{E \subset X : \quad \text{compact}, \quad \mu(E) = 1\}. \tag{AB.2.11}$$

Let $E_1, \dots, E_n \subset X$ be compact with $\mu(E_j) = 1$ for all j. We will show by induction on n that $\mu(E_1 \cap \cdots \cap E_n) = 1$. When $n = 1$, there is nothing to prove. Suppose that $\mu(E_1 \cap \cdots \cap E_k) = 1$. Then

$$E_{k+1} = \big(E_{k+1} \setminus (E_1 \cap \cdots \cap E_k)\big) \cup (E_1 \cap \cdots \cap E_k \cap E_{k+1}).$$

As $\mu(E_1 \cap \cdots \cap E_k) = 1$, we get that $\mu(E_{k+1} \setminus (E_1 \cap \cdots \cap E_k)) = 0$, as this last set lies in $X \setminus (E_1 \cap \cdots \cap E_k)$. Then $\mu(E_1 \cap \cdots \cap E_k \cap E_{k+1}) = 1$.

The family of compact sets in (AB.2.11) has the finite intersection property and so $X_0 \neq \varnothing$ by Proposition 1.8.19. If $x_0, x_1 \in X_0$ are distinct, by X being Hausdorff there exist disjoint open sets $V_0', V_1' \subset X$ with $x_0 \in V_0'$, $x_1 \in V_1'$. Using Lemma 2.6.4 there exist $V_0, V_1 \subset X$, open, disjoint, with disjoint compact closure, and $x_0 \in V_0$, $x_1 \in V_1$. For any $E \subset X$ compact with $\mu(E) = 1$, we have $x_0, x_1 \in X_0 \subset E$. If $\mu(V_0) = 0$, then $E \setminus V_0$ is compact and $\mu(E \setminus V_0) = 1$, giving us $X_0 = X_0 \setminus V_0$, contradicting that $x_0 \in X_0$. Thus $\mu(V_0) = 1$, and similarly $\mu(V_1) = 1$. But as V_0 and V_1 are disjoint, only one of them can have measure 1 (indeed, if $\mu(V_0) = \mu(V_1) = 1$, we get $1 \geq \mu(V_0 \cup V_1) = 1 + 1 = 2$). So $X_0 = \{x_0\}$. Now is the time to use that μ is regular. If $\mu(\{x_0\}) = 0$, then by the regularity there exists V open with $x_0 \in V$ and $\mu(V) < 1$, so $\mu(V) = 0$. Then for every $E \subset X$ compact with $\mu(E) = 1$, $E \setminus V$ is compact and $\mu(E \setminus V) = 1$; this implies that $X_0 \subset E \setminus V$, and so $V \cap X_0 = \varnothing$, a contradiction. So $\mu(\{x_0\}) = 1$ and therefore $\mu = \delta_{x_0}$.

2.11. Differentiation

(2.11.1) Consider $\mathbb{R}$ with Lebesgue measure, and $E \subset \mathbb{R}$ measurable. If it exists, the number

$$d_E(x) = \lim_{\varepsilon \to 0} \frac{m(E \cap (x - \varepsilon, x + \varepsilon))}{2\varepsilon}$$

is the *density* of E at x. Show that $d_E(x) = 1_E(x)$ a.e. Can you formulate and prove an analog result in $\mathbb{R}^n$?

Answer. In $\mathbb{R}^n$, we can define

$$d_E(x) = \lim_{\varepsilon\to 0} \frac{m(E\cap B_\varepsilon(x))}{m(B_\varepsilon(x))}.$$

We have

$$\frac{m(E\cap B_\varepsilon(x))}{m(B_\varepsilon(x))} = \frac{1}{m(B_\varepsilon(x))}\int_{B_\varepsilon(x)} 1_E\,dm \xrightarrow[\varepsilon\to 0]{} 1_E \qquad \text{a.e.}$$

by Theorem 2.11.9. Note that while 1_E might not be integrable, we only care about its behaviour on balls, where it is integrable. So we may replace 1_E in the integral with $1_{E\cap B_n(0)}$ and we get the equality a.e. over an increasing countable union of sets.

(2.11.2) Let $f\in L^1[0,\infty)$ be such that

$$\int_0^x f\,dm = 0, \qquad x>0.$$

Show that $f=0$. This was already done in Exercise 2.5.6 but now a much much shorter proof is available.

Answer. By Lebesgue differentiation (Theorem 2.11.13) we have

$$f(x) = \frac{1}{2x}\int_0^{2x} f = 0$$

almost everywhere.

(2.11.3) For a closed square and a closed disk in $\mathbb{R}^2$, calculate the density at every point.

Answer. The density is

$$d_E(x) = \lim_{r\to 0} \frac{m(E\cap B_r(x))}{m(B_r(x))}.$$

For any interior point, a small enough ball will be entirely within the set, and so $d_E(x)=1$. In the boundary of the square, for any point not a vertex a small enough ball will have precisely half in the square and half outside, so $d_E(x)=\frac12$. In each of the four vertices, for a small enough ball precisely a quarter of the ball will be inside the square, so $d_E(x)=\frac14$. For points in the boundary of the disk, for small enough r the boundary of the disk will be basically a straight line, so it divides the small ball almost in half: thus $d_E(x)=\frac12$.

(2.11.4) For $E \subset \mathbb{R}^2$, the boundary ∂E is the closure of E minus the interior of E.

(i) Show that E is Lebesgue measurable if $m(\partial E) = 0$.

(ii) Suppose that E is an arbitrary union of a collection of closed disks with radii at least c for some $c > 0$. Show that E is measurable.

(iii) Show that the radii above don't need to be restricted.

(iv) Show that some unions of closed disks of radius 1 are not Borel sets.

(v) Can disks be replaced by triangles, rectangles, arbitrary polygons, etc.? What is the relevant geometric property?

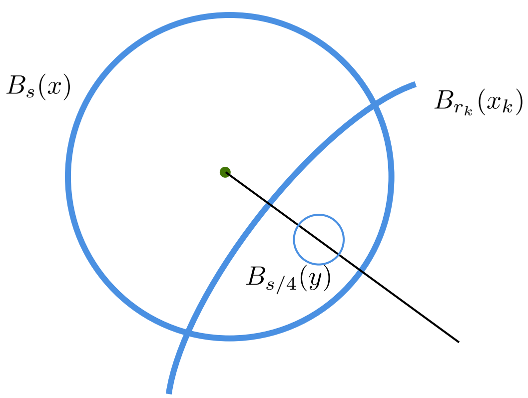
Answer. Write E^o for the interior of E.

(i) If $m(\partial E) = 0$, then ∂E is measurable. More importantly, $E \cap \partial E \subset \partial E$ is a null-set. Then

$$E = E^o \cup (E \cap \partial E)$$

is measurable as E^o is open.

(ii) Write $E = \bigcup_j \overline{B_{r_j}(x_j)}$. Let $x \in \overline{E} \setminus E^o$. As $\bigcup_j B_{r_j}(x_j) \subset E^o$, we get that $x \in \overline{E} \setminus \bigcup_j B_{r_j}(x_j)$. Let $s > 0$ with $s < c/4$. As $x \in \overline{E}$, there exists k such that $\operatorname{dist}(x, B_{r_k}(x_k)) < s/2$. Looking at the segment that joins x and x_k, because $r_k \geq c > 4s$ we can fit a ball $B_{s/4}(y)$ of radius $s/4$ in the intersection $B_s(x) \cap B_{r_k}(x_k)$. As this little ball lies outside of $\overline{E} \setminus \bigcup_j B_{r_j}(x_j)$, it lies outside of $\overline{E} \setminus E^o$.



Then

$$\frac{m((\overline{E}\setminus E^o)\cap B_s(x))}{m(B_s(x))}\leq\frac{m(B_s(x)\setminus B_{s/4}(y))}{m(B_s(x))}=1-\frac{\pi s^2/16}{\pi s^2}=1-\frac{1}{16}.$$

As s was arbitrary, this shows that $d_{\overline{E}\setminus E^o}(x)<1$ for all $x\in\overline{E}\setminus E^o$. Since by Exercise 2.11.1 the density is 1 a.e., this implies that $m(\overline{E}\setminus E^o)=0$. That is, $m(\partial E)=0$ and so E is measurable.

(iii) Let $J_n=\{j:\ r_j\geq 1/n\}$. Then

$$\bigcup_j B_{r_j}(x_j)=\bigcup_n\bigcup_{j\in J_n}B_{r_j}(x_n)$$

and the case with minimum radii applies to each union of J_n, so we get a countable union of measurable sets.

(iv) Let $V\subset\mathbb{R}$ be measurable but not Borel. Let $E=\bigcup_{v\in V}B_1(v,0)$. By the above, E is measurable. But if E is Borel, so is

$$E\cap(\mathbb{R}\times\{1\})=V\times\{1\}.$$

As $V=f^{-1}(V\times\{1\})$ with $f:x\longmapsto(x,1)$ continuous, this would imply that V is Borel (Proposition 2.4.3), a contradiction.

(v) The key feature seems to be convexity. That is what guarantees that we can put the smallest ball in the intersection, as in the picture.

(2.11.5) Let $\{f_n\}$ be a sequence of non-decreasing functions $f_n:\mathbb{R}\to[0,\infty)$, such that $f(x)=\sum_n f_n(x)<\infty$ for all x. Show that $f'(x)=\sum_n f_n'(x)$ a.e.

Answer. Derivatives are local, so we may restrict the domain to an interval $[a,b]$. Since we can do this for any interval and a countable union of nullsets is a nullset, we do not lose generality.

We know that f_n' exists a.e. (because f_n is monotone). By removing a countable union of nullsets, we may consider only those x such that $f_n'(x)$ exists for all n. Fix one such x. Let

$$g(x)=\sum_n f_n'(x).$$

This exists (possibly as infinity, for now) because $f_n'(x)\geq 0$ for all n. The monotonicity of f_n makes all Newton quotients non-negative. Then

$$\sum_{n=1}^N\frac{f_n(x+h)-f_n(x)}{h}\leq\sum_{n=1}^\infty\frac{f_n(x+h)-f_n(x)}{h}=\frac{f(x+h)-f(x)}{h}.$$

Taking the limit as $h \to 0$, and then as $N \to \infty$,

$$g(x) = \sum_{n=1}^{\infty} f_n'(x) \le f'(x) \tag{AB.2.12}$$

Let $h_N = \sum_{n=1}^{N} f_n$. Then $h_N \nearrow f$. Choose numbers $\{N_k\}_k \subset \mathbb{N}$, with $N_{k+1} > N_k$ and $f(b) - h_{N_k}(b) < 2^{-k}$. Define

$$s(x) = \sum_k \big(f(x) - h_{N_k}(x)\big) = \sum_k \sum_{n > N_k+1} f_n(x).$$

This function s is monotone on $[a, b]$, since the f_n are; so $0 \le s(x) \le s(b) \le 1$, where the last inequality is guaranteed by the choice of the N_k. It follows that s is differentiable a.e., and then applying (AB.2.12) to s

$$0 \le \sum_k \big(f'(x) - h_{N_k}'(x)\big) \le s'(x) \qquad \text{a.e.}$$

This in particular implies that $h_{N_k}'(x) \to f'(x)$ a.e., which is $\sum_n f_n'(x) = f'(x)$ a.e.

(2.11.6) Suppose that $E \subset [a, b]$, $m(E) = 0$. Construct an absolutely continuous monotonic function f on $[a, b]$ so that $f'(x) = \infty$ for all $x \in E$.

Answer. Since $m(E) = 0$, for each n choose a covering $\{I_{n,k}\}$ of open intervals such that

$$E \subset \bigcup_k I_{n,k}, \qquad \sum_k m(I_{n,k}) < 4^{-n}.$$

Via Proposition 1.8.1 we may assume that for each n the intervals $\{I_{n,k}\}$ are pairwise disjoint. Hence for each $x \in E$ and $n \in \mathbb{N}$ there exists an interval $I_n(x)$ with $x \in I_n$ and $m(I_n(x)) < 4^{-n}$. For an interval $I = (a, b)$ consider the function

$$\psi_I(t) = \begin{cases} 0, & t \le a \\ \dfrac{t-a}{b-a}, & a < t < b \\ 1, & t \ge b \end{cases}$$

Then ψ_I is monotone, with $0 \le \psi_I \le 1$ and $\psi_I'(t) = 1/|I|$ for $t \in I$. Now define

$$f(t) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} 2^{-n}\, \psi_{I_{n,k}}.$$

The fact that for fixed n the intervals $\{I_{n,k}\}$ are pairwise disjoint guarantees that $\sum_k \psi_{I_{n,k}} \le 1$ and so $0 \le f \le 1$ and the series converges uniformly, so

f is continuous. As f is monotone, continuous, and bounded, it is absolutely continuous.

Now fix $x \in E$. For each n there exists a unique $k(n)$ with $x \in I_{n,k(n)} = (a_n, b_n)$. Fix $M > 0$ and n such that $2^n > M$. If $0 < h < b_n - x$,

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &\geq 2^{-n}\frac{\psi_{I_{n,k(n)}}(x+h)-\psi_{I_{n,k(n)}}(x)}{h} \\ &= \frac{2^{-n}}{m(I_{n,k(n)})} = 2^n > M.\end{aligned}$$

This shows that $f'(x) = \infty$ for all $x \in E$.

(2.11.7) Let $f : [a, b] \to \mathbb{C}$ be of bounded variation. Show that f admits side-limits at all points.

Answer. Let $x_0 \in [a, b]$. Suppose that $\lim_{x\to x_0^+} f(x)$ does not exist. This means that there exists $\varepsilon > 0$ and a monotone sequence $\{x_n\} \subset [a, b]$ such that $x_n \searrow x$ and $|f(x_{n+1}) - f(x_n)| \geq \varepsilon$. Indeed, the non-existence of the limit means that there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exist $y, z \in (x_0, x_0 + \delta)$ that satisfy $|f(y) - f(z)| \geq \varepsilon$. Start with $\delta_1 > 0$ with $x_0 + \delta < b$ and choose y_1, z_1 with $x_0 < z_1 < y_1 < x_0 + \delta$ and $|f(y_1) - f(z_1)| \geq \varepsilon$. Inductively, given $y_1, \ldots, y_m$ and $z_1, \ldots, z_m$ with

$$x_0 < z_m < y_m < z_{m-1} < y_{m-1} < \cdots < z_1 < y_1 < x_0 + \delta_1,$$

let $\delta_{m+1} = z_m - x_0$. Since the limit does not exist, there exist z_{m+1}, y_{m+1} with $x_0 < z_{m+1} < y_{m+1} < z_m$ and $|f(y_{m+1}) - f(z_{m+1})| \geq \varepsilon$ and the induction is complete.

Define

$$x_n = \begin{cases} y_{(n+1)/2}, & n \text{ odd} \\ z_{n/2}, & n \text{ even} \end{cases}$$

Then $x_n \searrow x_0$, and

$$\sum_{k=1}^{m} |f(x_{k+1}) - f(x_k)| \geq m\varepsilon.$$

As this can be done for any m, the total variation of f is infinite. The contradiction implies that the right-limit exists. An analog argument shows that the left-limits also exist.

(2.11.8) Let $f : [a, b] \to \mathbb{R}$. Show that f is of bounded variation if and only if there exist $g, h : [a, b] \to \mathbb{R}$, both monotone, and such that $f = g - h$.

Answer. With F the total variation of f, it is proven in Proposition 2.11.16 that both $g = F + f$ and $h = F$ are monotone, so $f = g - h$.

Conversely, if $f = g - h$ with g, h monotone, let $a = t_0 < t_1 < \cdots < t_n = b$. Then

$$\begin{aligned}\sum_{j=1}^{n} |f(t_j) - f(t_{j-1})| &\leq \sum_{j=1}^{n} g(t_j) - g(t_{j-1}) + h(t_j) - h(t_{j-1}) \\ &= g(b) - g(a) + h(b) - h(a)\end{aligned}$$

for any partition of $[a, b]$. Thus $F(b) \leq g(b) - g(a) + h(b) - h(a) < \infty$ and f is of bounded variation.

(2.11.9) Let $f : [a, b] \to \mathbb{C}$ be a function of bounded variation. Show that f is Riemann integrable.

Answer. By Exercise 2.11.8 we can write $f = g - h$ with g, h monotone non-decreasing. So the assertion reduces to arguing that a monotone function g is Riemann-integrable. We may assume without loss of generality that $g \geq 0$ (replacing g with $g - g(a)$).

If $\varepsilon > 0$, let $P = \{a_0, \ldots, a_n\}$ be a partition of $[a, b]$ with $\Delta_j = a_j - a_{j-1} < \frac{\varepsilon}{g(b) - g(a)}$. Then, using that g is monotone,

$$\begin{aligned}U(g, P) - L(g, P) &= \sum_{k=1}^{n} \big[g(a_j) - g(a_{j-1})\big] \Delta_j \\ &\leq \frac{\varepsilon}{g(b) - g(a)} \sum_{k=1}^{n} g(a_j) - g(a_{j-1}) = \varepsilon.\end{aligned}$$

Thus g is Riemann-integrable.

CHAPTER

3

A Bit of Complex Analysis

3.1. Analytic and Holomorphic Functions

(3.1.1) Construct an example of a convergent series of positive terms $\sum_k b_k$ such that $\alpha > 1$ and $\beta < 1$.

Answer. Choose $s, t \in (0,1)$ with $s < t$. Define

$$b_k = \begin{cases} s^k, & k \text{ even} \\ t^k, & k \text{ odd} \end{cases}$$

Then $\sum_k b_k = \sum_k s^k + \sum_k t^k < \infty$. Also, when k is even,

$$\frac{|b_{k+1}|}{|b_k|} = \frac{t^{k+1}}{s^k} \to \infty.$$

And when k is odd,

$$\frac{|b_{k+1}|}{|b_k|} = \frac{s^{k+1}}{t^k} \to 0.$$

So $\alpha = \infty$, $\beta = 0$.

Let us now get an example of a convergent series with finite $\alpha > 1$. Choose any convergent series $\sum_k a_k$ with $\frac{|a_{k+1}|}{|a_k|} \to 1$, for instance $a_k = 1/k^2$. Fix $c > 0$ and let

$$b_k = \begin{cases} a_{(k-1)/2}, & k \text{ odd} \\ c\, a_{k/2}, & k \text{ even} \end{cases}$$

Then, when $k = 2h+1$ is odd,

$$\frac{|b_{k+1}|}{|b_k|} = \frac{c\, a_{h+1}}{a_h} \to c.$$

And when $k = 2h$ is even,

$$\frac{|b_{k+1}|}{|b_k|} = \frac{a_h}{c\, a_h} = \frac{1}{c}.$$

Then $\alpha = c$ and $\beta = 1/c$.

(3.1.2) Construct an example of a divergent series of positive terms $\sum_k b_k$ such that $\alpha > 1$ and $\beta < 1$.

Answer. As in the previous exercise, choose any $c > 1$ and any divergent series $\sum_k a_k$ with $\frac{|a_{k+1}|}{|a_k|} \to 1$; for instance $a_k = 1/k$. Form

$$b_k = \begin{cases} a_{(k-1)/2}, & k \text{ odd} \\ c\, a_{k/2}, & k \text{ even} \end{cases}$$

Then, when $k = 2h+1$ is odd,

$$\frac{|b_{k+1}|}{|b_k|} = \frac{c\, a_{h+1}}{a_h} \to c.$$

And when $k = 2h$ is even,

$$\frac{|b_{k+1}|}{|b_k|} = \frac{a_h}{c\, a_h} = \frac{1}{c}.$$

So $\alpha = c$ and $\beta = 1/c$.

(3.1.3) Show that the function $f(z) = \bar{z}$ is not holomorphic anywhere in the complex plane.

Answer. When $h = t$ is real,

$$\frac{f(z+h) - f(z)}{h} = \frac{t}{t} = 1.$$

But when $h = it$ is imaginary,

$$\frac{f(z+h)-f(z)}{h} = \frac{-it}{it} = -1.$$

Then f is not differentiable at z and thus not holomorphic at z.

(3.1.4) Prove the **Dirichlet Criterion**: If $\{a_n\}, \{b_n\} \subset \mathbb{C}$ and

(i) $\lim_{n\to\infty} a_n = 0$;

(ii) $\sum_{n=1}^{\infty} |a_{n+1} - a_n| < \infty$;

(iii) there exists $M > 0$ such that $\left|\sum_{n=1}^{m} b_n\right| < M$ for all m.

Show that $\sum_n a_n b_n$ converges.

Answer. The trick is to use **summation by parts**. If $s_k = \sum_{n=1}^k a_n b_n$ and $t_k = \sum_{n=1}^k b_n$, then

$$s_k = a_{k+1}t_k + \sum_{n=1}^{k} t_n(a_n - a_{n+1}).$$

As $|t_n| < M$ by hypothesis, the series above converges absolutely, and so

$$\lim_{k\to\infty} s_k = \sum_{n=1}^{\infty} t_n(a_n - a_{n+1}).$$

(3.1.5) Show that in Exercise 3.1.4 the hypothesis "$\sum_{n=1}^{\infty} |a_{n+1} - a_n| < \infty$" can be replaced, when it makes sense, by "$\{a_n\}$ is monotone".

Answer. Suppose that $\{a_n\}$ is non-decreasing. Then

$$\sum_{n=1}^{k} |a_{n+1} - a_n| = \sum_{n=1}^{k} a_{n+1} - a_n = a_{k+1} - a_1 \xrightarrow[k\to\infty]{} -a_1,$$

so the series converges. Similarly, if $\{a_n\}$ is non-increasing,

$$\sum_{n=1}^{k} |a_{n+1} - a_n| = \sum_{n=1}^{k} a_n - a_{n+1} = a_1 - a_{k+1} \xrightarrow[k\to\infty]{} a_1,$$

In both cases, the hypothesis in Exercise 3.1.4 is satisfied.

(3.1.6) Show that the **Gamma Function**

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

defines a holomorphic function on the semiplane $\operatorname{Re} z > 0$. Show also that $\Gamma(z+1) = z\Gamma(z)$ for every z in its domain, and that $\Gamma(n) = (n-1)!$ for all $n \in \mathbb{N}$.

Answer. First we check that the integral exists. If $z = a + ib$ with $a > 0$ then

$$t^{z-1} = t^{a-1+ib} = t^{a-1}\, e^{ib\log t} = e^{(a-1)\log t}\, e^{ib\log t}.$$

Then

$$\left| \int_M^\infty t^{z-1} e^{-t}\, dt \right| \leq \int_M^\infty e^{(a-1)\log t - t}\, dt.$$

Taking M big enough so that $(a-1)\log t < t/2$ for all $t \geq M$,

$$\int_M^\infty e^{(a-1)\log t - t}\, dt \leq \int_M^\infty e^{-t/2}\, dt = 2e^{-M/2} \xrightarrow{M\to\infty} 0,$$

so the integral converges.

Now we look at the Newton quotients. Given a sequence $\{h_n\}$ with $|h_n| \leq 1$ for all n and $h_n \to 0$, we have by Dominated Convergence

$$\begin{aligned}
\lim_{n\to\infty} \frac{\Gamma(z+h_n) - \Gamma(z)}{h_n} &= \int_0 \lim_{n\to\infty} \frac{t^{z-1+h_n} - t^{z-1}}{h_n}\, e^{-t}\, dt \\
&= \int_0^\infty t^{z-1} \log t\, e^{-t}\, dt.
\end{aligned}$$

As this can be done for any sequence that converges to 0, the limit of the Newton quotients exists (the integral converges as the exponential wins over the power of t and the logarithm). So Γ is holomorphic.

Integrating by parts,

$$\Gamma(z+1) = \int_0^\infty t^z\, e^{-t}\, dt = \int_0^\infty z t^{z-1}\, e^{-t}\, dt = z\Gamma(z).$$

The expression for n follows by induction, since $\Gamma(1) = 1$ and $\Gamma(n) = (n-1)\Gamma(n-1)$.

3.2. Inverses of Holomorphic Functions and the Logarithm

3.3. Line Integrals

(3.3.1) Show that if $\gamma : [a, b] \to \mathbb{C}$ is piecewise continuously differentiable, then it is of bounded variation.

Answer. Assume first that γ' is continuous. Let $a = a_0 < a_1 < \cdots < a_n = b$ be a partition of $[a, b]$. We have

$$\begin{aligned}\sum_{k=1}^{n} |\gamma(a_k) - \gamma(a_{k-1})| &= \sum_{k=1}^{n} \Big| \int_{a_{k-1}}^{a_k} \gamma'(t)\, dt \Big| \\ &\leq \sum_{k=1}^{n} \int_{a_{k-1}}^{a_k} |\gamma'(t)|\, dt \\ &= \int_{a}^{b} |\gamma'(t)|\, dt \\ &\leq (b-a)\, \|\gamma'\|_\infty .\end{aligned}$$

We know that γ' is bounded because it is continuous on each of the finitely many intervals $[a_{k-1}, a_k]$.

In the general case, by definition of piecewise continuity γ' has jump discontinuities (the left/right continuity of γ' at its discontinuity points prevents worse discontinuities from arising) at points $c_1, \ldots, c_m$. The only terms where the first paragraph does not apply are those of the form

$$|\gamma(c_r) - \gamma(a_{k-1})| + |\gamma(a_k) - \gamma(c_r)|.$$

We have

$$\begin{aligned}|\gamma(c_r) - \gamma(a_{k-1})| &\leq |\gamma(c_r) - \gamma(c_r^-)| + |\gamma(c_r^-) - \gamma(a_{k-1})| \\ &\leq |\gamma(c_r) - \gamma(c_r^-)| + (c_r - a_{k-1})\|\gamma'\|_\infty ,\end{aligned}$$

The sum of all terms of the form $|\gamma(c_r) - \gamma(c_r^-)|$ is bounded by the sum of the height d of all the height of the jump discontinuities. The other term is dealt similarly. Hence

$$\sum_{k=1}^{n} |\gamma(a_k) - \gamma(a_{k-1})| \leq 2d + (b-a)\|\gamma'\|_\infty .$$

(3.3.2) Let $\gamma(t)$, $t \in [a,b]$ be a curve. Show that it is natural to define the length of γ as

$$L(\gamma) = \int_a^b |\gamma'(t)| \, dt$$

Answer. If we partition the interval $[a,b]$ as $a = t_0 < t_1 < \cdots < t_n = b$, an approximation to the length of the curve would be $\sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1}|$. Then we can do, using the Mean Value Theorem,

$$\sum_{j=1}^{n} |\gamma(t_j) - \gamma(t_{j-1}| = \sum_{j=1}^{n} \frac{|\gamma(t_j) - \gamma(t_{j-1}|}{t_j - t_{j-1}} \, (t_j - t_{j-1}) = \sum_{j=1}^{n} |\gamma'(t_j')| \, (t_j - t_{j-1})$$

for some $t_j' \in [t_{j-1}, t_j]$, and now the sum is a Riemann sum for the integral $\int_a^b |\gamma'|$.

3.4. The Index

3.5. Cauchy's Theorem

3.6. Zeros of Holomorphic Functions

(3.6.1) Show that the order of a zero is well-defined, in the sense that if $(z-w)^n g(z) = (z-w)^m h(z)$ with g, h holomorphic at w and $g(w)h(w) \neq 0$, then $n = m$.

Answer. Assume without loss of generality that $m \geq n$. Since both g and h are nonzero at w and they are continuous, there exists a disk $w + r\,\mathbb{D}$ around w where both g and h are nonzero. For any $z \in w + r\,\mathbb{D}$, we have $(z-w)^n g(z) = (z-w)^m h(z)$, and since $g(z) \neq 0$ and $h(z) \neq 0$ this gives us $(z-w)^n = (z-w)^m$. That is, $(z-w)^{m-n} - 1 = 0$ for all $z \in w + r\,\mathbb{D}$. Then all derivatives will be zero on the disk. If $m - n > 0$, this gives us, after differentiating $m-n-1$ times, $z - w = 0$ for all $z \in w + r\,\mathbb{D}$, a contradiction. Thus $m = n$.

(3.6.2) Let V be open and f holomorphic on V. Show that f has at most countably many zeros.

Answer. Since V is an open subset of the plane, we can cover it with countably many closed disks (for instance, $\overline{B_n(0)}$, $n \in \mathbb{N}$. By Corollary 3.6.3 each of the closed disks can only have finitely many zeros, so the union of all of them can have at most countably many.

(3.6.3) Let V be a simply connected region and f holomorphic on V such that it has no zeros on V. Show if $a \in V$ and

$$g(z) = \log f(a) + \int_{\gamma_{a,z}} \frac{f'(w)}{f(w)}\,dw,$$

then $f(z) = e^{g(z)}$. Note that we know from the proof of Proposition 3.3.8 how to differentiate the integral.

Answer. We have that

$$g'(z) = \frac{f'(z)}{f(z)}.$$

Then

$$(f(z)e^{-g(z)})' = \big(f'(z) - f(z)g'(z)\big)e^{-g(z)} = 0,$$

so there exists $c \in \mathbb{C}$ such that $f(z) = ce^{g(z)}$. Evaluating at $z = a$, $f(a) = cf(a)$, so $c = 1$ (note that $f(a) \neq 0$ by hypothesis).

3.7. Maximum Modulus Principle and Liouville's Theorem

3.8. Consequences of Cauchy's Theorem

3.9. The General Cauchy Theorem

(3.9.1) Let f be entire and such that there exist $c, r > 0$ and $n \in \mathbb{N}$ such that $|f(z)| \le c|z|^n$ for all z with $|z| > r$. Prove that f is a polynomial with $\deg f \le n$.

Answer. Consider z with $r < |z| < 2r$, $k \in \mathbb{N}$ with $k > 2r$, and let $\gamma(t) = z + ke^{it}$, $t \in [0, 2\pi]$. By Corollary 3.9.4,

$$|f^{(n+1)}(z)| = \frac{(n+1)!}{2\pi}\Big|\int_0^{2\pi} \frac{f(z+ke^{it})}{k^{n+2}e^{(n+2)it}}\, ki\, e^{it}\, dt\Big|$$
$$\le \frac{c\,(n+1)!\,(2r+k)^n}{k^{n+1}}.$$

As we are free to choose k as big as we want, we conclude that $f^{n+1}(z) = 0$. So the holomorphic function f^{n+1} agrees with zero in a set with a cluster point; thus $f^{n+1} = 0$ by Corollary 3.6.2. By writing f as its Taylor series centered at 0, it follows that f is a polynomial of degree at most n.

(3.9.2) Let f be entire and such that, for z big enough, $\operatorname{Re} f(z) \le c|z|^s$ for some $c > 0$ and $s > 0$. Show that f is a polynomial of degree at most $m = \lfloor s \rfloor$.

Answer. Let $n \geq m+1$. By hypothesis there exists $r_0 > 0$ with $\operatorname{Re} f(z) \leq c|z|^s$ for all z with $|z| \geq r_0$. Using Corollary 3.9.4,

$$\begin{aligned} f^{(n)}(0) &= \frac{n!}{2\pi i} \oint_{r\mathbb{T}} \frac{f(z)}{z^{n+1}}\, dz = \frac{n!}{2\pi i} \int_0^{2\pi} \frac{f(re^{i\theta})}{r^{n+1} e^{i(n+1)\theta}}\, rie^{i\theta}\, d\theta \\ &= \frac{n!}{2\pi r^n} \int_0^{2\pi} f(re^{i\theta})\, e^{-in\theta}\, d\theta. \end{aligned}$$

Applying Cauchy's Theorem to $z^{n-1} f(z)$ we get

$$\begin{aligned} 0 &= \oint_{r\mathbb{T}} f(z) z^{n-1}\, dz = \int_0^{2\pi} f(re^{i\theta}) r^{n-1} e^{i(n-1)\theta}\, rie^{i\theta}\, d\theta \\ &= ir^n \int_0^{2\pi} f(re^{i\theta}) e^{in\theta}\, d\theta. \end{aligned}$$

Complex conjugation then gives

$$0 = \int_0^{2\pi} \overline{f(re^{i\theta})} e^{-in\theta}\, d\theta.$$

This allows us to write, since $2\operatorname{Re} f = f + \overline{f}$,

$$\begin{aligned} |f^{(n)}(0)| &= \frac{n!}{\pi r^n} \left| \int_0^{2\pi} \operatorname{Re} f(re^{i\theta})\, e^{-in\theta}\, d\theta \right| \\ &= \frac{n!}{\pi r^n} \left| \int_0^{2\pi} \left(cr^s - \operatorname{Re} f(re^{i\theta}) \right) e^{-in\theta}\, d\theta \right| \\ &\leq \frac{n!}{\pi r^n} \int_0^{2\pi} \left(cr^s - \operatorname{Re} f(re^{i\theta}) \right) d\theta \\ &= 2cn! r^{s-n} - 2n! \operatorname{Re} f(0) r^{-n}. \end{aligned}$$

As we are free to choose $r > r_0$, it follows that $f^{(n)}(0) = 0$ for all $n > m$ and hence f is a polynomial of degree at most m.

3.10. Meromorphic Functions and Residues

(3.10.1) Let $V \subset \mathbb{C}$ be open, $z_0 \in V$, and $f : V \to \mathbb{C}$ a function. Show that the following statements are equivalent:

(i) there exists $n \in \mathbb{N}$ such that $(z - z_0)^n f(z)$ has a removable singularity at z_0;

(ii) there exist $a_1, \ldots, a_n \in \mathbb{C}$, with $a_n \neq 0$, such that on some disk around z_0

$$f(z) - \sum_{k=1}^{n} \frac{a_k}{(z - z_0)^k}$$

has a removable singularity at z_0;

(iii) there exist coefficients $\{c_k\}_{k=-n}^{\infty} \subset \mathbb{C}$ such that, on some disk around z_0,

$$f(z) = \sum_{k=-n}^{\infty} c_k (z - z_0)^k.$$

Answer. (i) $\implies$ (ii) Suppose that $(z - z_0)^n f(z)$ is holomorphic at z_0. By Corollary 3.5.2 there is a disk around z_0 such that

$$(z - z_0)^n f(z) = \sum_{k=0}^{\infty} b_k (z - z_0)^k.$$

We can rewrite this, splitting the sum as convenient and defining $a_k = b_{n-k}$,

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty} b_k (z - z_0)^{k-n} = \sum_{k=0}^{n-1} b_k (z - z_0)^{k-n} + \sum_{k=n}^{\infty} b_k (z - z_0)^{k-n} \\ &= \sum_{k=1}^{n} a_k (z - z_0)^{-k} + \sum_{k=0}^{\infty} b_{n-k} (z - z_0)^k, \end{aligned}$$

which gives the desired expression for f.

(ii) $\implies$ (iii) If we now assume that $f(z) - \sum_{k=1}^{n} \frac{a_k}{(z - z_0)^k}$ has a removable singularity at z_0, on some disk we can write

$$f(z) - \sum_{k=1}^{n} \frac{a_k}{(z - z_0)^k} = \sum_{k=0}^{\infty} b_k (z - z_0)^k.$$

This is

$$f(z) = \sum_{k=-n}^{\infty} c_k(z - z_0)^k, \tag{AB.3.1}$$

if we put $c_k = a_{-k}$ for $k < 0$, and $c_k = b_k$ for a$k \geq 0$.

(iii) $\implies$ (i) If f is as in (AB.3.1), then

$$(z - z_0)^n f(z) = \sum_{k=-n}^{\infty} c_k(z - z_0)^{k+n} = \sum_{k=0}^{\infty} c_{k-n}(z - z_0)^k$$

and so $(z - z_0)^n f(z)$ is analytic at z_0.

(3.10.2) For each of the following functions, classify its singularities.

(i) $f(z) = \frac{\sin z}{z}$, $z \in \mathbb{C} \setminus \{0\}$;

(ii) $f(z) = \frac{e^z}{z^2}$, $z \in \mathbb{C} \setminus \{0\}$;

(iii) $f(z) = e^{1/z}$, $z \in \mathbb{C} \setminus \{0\}$.

Answer.

(i) The function is bounded on any disk that does not contain 0. Since we have

$$\frac{\sin z}{z} = 1 - \frac{z^2}{3} + o(z^5),$$

the singularity is removable.

(ii) We have

$$f(z) = \frac{1}{z^2} + \frac{1}{z} + \frac{1}{2} + \frac{z}{6} + o(z^2).$$

By Exercise 3.10.1, f has a pole of order 2 at 0.

(iii) The singularity at 0 is essential. For

$$z^n f(z) = \sum_{k=0}^{\infty} \frac{z^{n-k}}{k!}$$

is always unbounded on disks around 0. This can be seen for instance by taking $z = \frac{1}{m(n+1)!}$ and then $z^n f(z) \geq m$ (by considering the $(n+1)^{\text{th}}$ term of the series).

(3.10.3) Use the Residue Theorem to evaluate $\displaystyle\oint_{|z|=2} \frac{1}{z^2(z-1)}\, dz$.

Answer. Since we have a circle, the index is 1. The integrand has poles at $z = 0$ (order 2) and $z = 1$ (order 1). Since

$$f(z) = -\frac{1}{z^2} - \frac{1}{z} + \frac{1}{z-1},$$

we see that $\operatorname{Res}(f, 0) = -1$ and $\operatorname{Res}(f, 1) = 1$. Then

$$\oint_{|z|=2} \frac{1}{z^2(z-1)}\, dz = 2\pi i(-1+1) = 0.$$

(3.10.4) Find $\displaystyle\oint_{|z|=3} \frac{e^z}{z^2+\pi^2}\, dz$.

Answer. Since e^z is nonzero and entire, the only poles are $\pm i\pi$, of order 1. These lie outside the simple curve $|z| = 3$, so our integrand is analytic on the interior of the curve, and hence the integral is zero by Cauchy's Theorem.

(3.10.5) Find $\displaystyle\oint_{|z-i|=3} \frac{e^z}{z^2+\pi^2}\, dz$.

Answer. As mentioned in the answer to Exercise 3.10.4, the poles are $\pm i\pi$. Now the curve has $i\pi$ inside of it, and $-i\pi$ outside (so with index 0). We have

$$\operatorname{Res}(f, i\pi) = \lim_{z\to i\pi} \frac{e^z}{z+i\pi} = \frac{e^{i\pi}}{2i\pi} = -\frac{1}{2i\pi}.$$

Therefore

$$\oint_{|z-i|=3} \frac{e^z}{z^2+\pi^2}\, dz = -2\pi i\, \frac{1}{2\pi i} = -1.$$

(3.10.6) Find $\displaystyle\int_0^{2\pi} \frac{1}{5+4\cos t}\, dt$.

Answer. If we put $z = e^{it}$, then $\cos t = \frac{1}{2}(z + z^{-1})$. Hence

$$\int_0^{2\pi} \frac{1}{5 + 4\cos t}\, dt = \int_0^{2\pi} \frac{ie^{it}}{ie^{it}(5 + 2(e^{it} + e^{-it}))}\, dt$$
$$= \oint_{|z|=1} \frac{1}{iz(5 + 2(z + z^{-1}))}\, dz$$
$$= -i \oint_{|z|=1} \frac{1}{2z^2 + 5z + 2}\, dz.$$

The function $f(z) = (2z^2 + 5z + 2)^{-1}$ has simple poles at $z = -2$ (outside the curve) and $z = -1/2$ (inside the curve). The residue is

$$\operatorname{Res}\left(f, -\frac{1}{2}\right) = \lim_{z \to -\frac{1}{2}} \left(z + \frac{1}{2}\right) \frac{1}{(2z + 1)(z + 2)} = \frac{1}{2\left(-\frac{1}{2} + 2\right)} = \frac{1}{3}.$$

Thus

$$\int_0^{2\pi} \frac{1}{5 + 4\cos t}\, dt = -i\, 2\pi i\, \frac{1}{3} = \frac{2\pi}{3}.$$

(3.10.7) Compute $\displaystyle\int_{-\infty}^{\infty} \frac{1}{x^2 + 4x + 5}\, dx.$

Answer. Since $x^2 + 4x + 5 = (x + 2)^2 + 1$, its two roots are $-2 \pm i$. Let $\gamma_R = \gamma_1 + \gamma_2$, where

$$\gamma_1(t) = -R + t, \qquad t \in [0, R]$$

and

$$\gamma_2(t) = Re^{it}, \qquad t \in [0, \pi].$$

So γ_R is a closed curve, going from $-R$ to R on the real line, and then coming back as an arc towards $-R$. Write $f(z) = (z^2 + 4z + 5)^{-1}$. Since $-2 + i$ is the only (simple) pole of f inside γ_R and

$$\operatorname{Res}(f, -2 + i) = \lim_{z \to -2+i} \frac{(z + 2 - i)}{(z + 2 - i)(z + 2 + i)} = \frac{1}{2i},$$

we have

$$\int_{\gamma_R} f(z)\, dz = 2i\pi\, \frac{1}{2i} = \pi.$$

Also, when $|z| = R$,

$$|z^2 + 4z + 5| \geq |z^2| - |4z| - 4 = R^2 - 4R - 5$$

so

$$\left| \int_{\gamma_2} f(z)\, dz \right| = \left| \int_0^{\pi} \frac{iRe^{it}}{(Re^{it})^2 + 4Re^{it} + 5}\, dt \right| \leq \int_0^{\pi} \frac{R}{R^2 - 4R - 5}\, dt \leq \frac{2\pi}{R}.$$

Thus

$$\int_{-\infty}^{\infty} \frac{1}{x^2+4x+5}\,dx = \lim_{R\to\infty} \int_{-R}^{R} \frac{1}{x^2+4x+5}\,dx$$
$$= \lim_{R\to\infty} \int_{\gamma_R} \frac{1}{z^2+4z+5}\,dz = \pi.$$

(3.10.8) Compute $\displaystyle\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}\,dx.$

Answer. Consider $\gamma_1, \gamma_2, \gamma_R$ as in the answer to Exercise 3.10.7. Since $\cos z$ is not bounded on an arc, we will instead use that

$$\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}\,dx = \operatorname{Re} \int_{-\infty}^{\infty} \frac{e^{ix}}{1+x^2}\,dx.$$

So let $f(z) = e^{iz}/(1+z^2)$. This function has simple poles $\pm i$, and

$$\operatorname{Res}(f,i) = \lim_{z\to i} \frac{(z-i)e^{iz}}{(z-i)(z+i)} = \frac{e^{i^2}}{2i} = \frac{e^{-1}}{2i}.$$

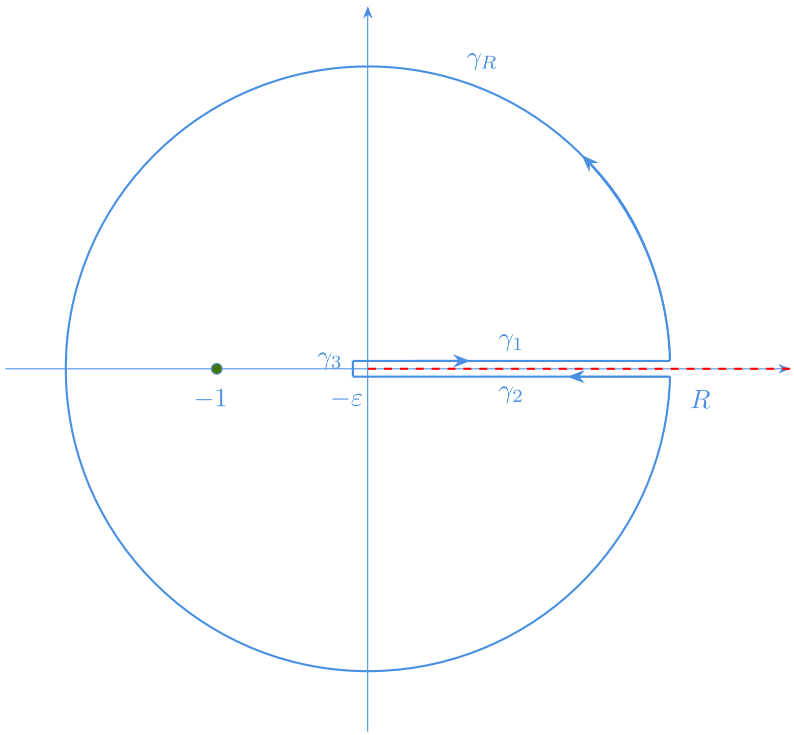
When $|z| = R$ we have $|1+z^2| \geq R^2 - 1$, which shows that $|Re^{it}/(1+(Re^{it})^2)| \leq R/(R^2-1)$ and thus $\int_{\gamma_2} f(z)\,dz \to 0$ with R. Then

$$\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}\,dx = \lim_{R\to\infty} \operatorname{Re} \int_{\gamma_R} \frac{e^{iz}}{1+z^2}\,dz = 2\pi i \operatorname{Res}(f,i) = \frac{\pi}{e}.$$

(3.10.9) Show that if $s \in (0,1)$ then $\displaystyle\int_0^{\infty} \frac{x^{s-1}}{x+1}\,dx = \frac{\pi}{\sin \pi s}.$

Answer.

Consider $f(z) = z^{s-1}/(z+1)$. We use a "keyhole" contour that leaves out the positive x-axis:



This allows us to choose the branch of the logarithm where $0 < \theta < 2\pi$. Which is crucial because we need the logarithm to calculate z^{s-1}. It is common to use a small circle around the origin (which justifies better the name "keyhole") but using straight lines simplifies the computations a bit; most sources gloss over the estimates below to avoid the effort.

Over $\gamma_1(t) = t + i\varepsilon$, $t \in [-\varepsilon, \sqrt{R^2 - \varepsilon^2}]$,

$$\begin{aligned}\int_{\gamma_1} f(z)\,dz &= \int_{-\varepsilon}^{\sqrt{R^2-\varepsilon^2}} \frac{(t+i\varepsilon)^{s-1}}{t+1+i\varepsilon}\,dt \\ &= \int_{-\varepsilon}^{\sqrt{R^2-\varepsilon^2}} \frac{(t^2+\varepsilon^2)^{(s-1)/2}\, e^{i(s-1)\arctan\frac{\varepsilon}{t}}}{t+1+i\varepsilon}\,dt \\ &\xrightarrow[\substack{R\to\infty \\ \varepsilon\to 0}]{} \int_0^\infty \frac{x^{s-1}}{x+1}\,dx.\end{aligned}$$

The limit is taken via Dominated convergence: for big enough t and small ε

$$\left| \frac{(t^2+\varepsilon^2)^{(s-1)/2}\, e^{i(s-1)\arctan\frac{\varepsilon}{t}}}{t+1+i\varepsilon} \right| \le 4^{s-1} t^{s-2}$$

which is integrable since $0 < s < 1$.

Over $\gamma_2(t) = \sqrt{R^2-\varepsilon^2} - \varepsilon - t - i\varepsilon$, $t \in [-\varepsilon, \sqrt{R^2-\varepsilon^2}]$, we need to look carefully at the log branch we are using. Concretely,

$$(a - i\varepsilon)^{s-1} = (a^2+\varepsilon^2)^{(s-1)/2}\, e^{i(s-1)\left(2\pi - \arctan\frac{\varepsilon}{a}\right)}.$$

Then

$$\begin{aligned}\int_{\gamma_2} f(z)\,dz &= \int_{-\varepsilon}^{\sqrt{R^2-\varepsilon^2}} \frac{(\sqrt{R^2-\varepsilon^2}-\varepsilon-t-i\varepsilon)^{s-1}}{\sqrt{R^2-\varepsilon^2}-\varepsilon-t+1-i\varepsilon}\,(-1)\,dt\\ &= -\int_{-\varepsilon}^{\sqrt{R^2-\varepsilon^2}} \frac{(t-i\varepsilon)^{s-1}}{t+1-i\varepsilon}\,dt\\ &= -e^{2\pi(s-1)i}\int_{-\varepsilon}^{\sqrt{R^2-\varepsilon^2}} \frac{(t^2+\varepsilon^2)^{(s-1)/2}\,e^{-i(s-1)\arctan\frac{\varepsilon}{t}}}{t+1-i\varepsilon}\,dt\\ &\xrightarrow[\substack{R\to\infty\\ \varepsilon\to 0}]{} -e^{2\pi(s-1)i}\int_0^\infty \frac{x^{s-1}}{x+1}\,dx.\end{aligned}$$

Over $\gamma_R(t) = Re^{it}$, $t \in \left[\arctan\frac{\varepsilon}{R}, 2\pi - \arctan\frac{\varepsilon}{R}\right]$ we have, as long as $R > 2$,

$$|f(Re^{it})| \le \frac{R^{s-1}}{|Re^{it}+1|} \le \frac{R^{(s-1)}}{R-1} \le \frac{2R^{s-1}}{R} = 2R^{s-2}.$$

From Lemma 3.3.4 we get

$$\left|\int_{\gamma_R} f(z)\,dz\right| \le 2R^{s-2}\,2\pi R = 4\pi\,R^{s-1} \xrightarrow[R\to\infty]{} 0.$$

And over $\gamma_3(t) = -\varepsilon + it$, $t \in [-\varepsilon, \varepsilon]$, as long as $\varepsilon < \frac{1}{2}$

$$\left|\int_{\gamma_\varepsilon} f(z)\,dz\right| = \left|\int_{-\varepsilon}^{\varepsilon} \frac{(-\varepsilon+it)^{s-1}}{-\varepsilon+1+it}\,i\,dt\right| \le \int_{-\varepsilon}^{\varepsilon} \frac{|t-\varepsilon|^{s-1}}{1-\varepsilon}\,dt \le 4\varepsilon(2\varepsilon)^{s-1} \xrightarrow[\varepsilon\to 0]{} 0.$$

Therefore we have shown that

$$(1-e^{2\pi(s-1)i})\int_0^\infty \frac{x^{s-1}}{x+1}\,dx = \lim_{\substack{R\to\infty\\ \varepsilon\to 0}} \int_\gamma f(z)\,dz.$$

Now we calculate the integral using its residue at the only pole $z = -1$. The pole is simple, and we have

$$\operatorname{Res}(f,-1) = \lim_{z\to -1} \frac{(z+1)z^{s-1}}{z+1} = (-1)^{s-1} = e^{i\pi(s-1)}.$$

Therefore

$$(1-e^{2\pi(s-1)i})\int_0^\infty \frac{x^{s-1}}{x+1}\,dx = 2\pi i e^{i\pi(s-1)}.$$

Using that $e^{i\pi(s-1)} = e^{i\pi s}e^{-i\pi} = -e^{i\pi s}$,

$$\int_0^\infty \frac{x^{s-1}}{x+1}\,dx = \frac{2\pi i e^{i\pi(s-1)}}{1-e^{2\pi(s-1)i}} = \frac{\pi}{\dfrac{1-e^{2\pi(s-1)i}}{2i\,e^{i\pi(s-1)}}} = \frac{\pi}{\dfrac{e^{-i\pi(s-1)}-e^{i\pi(s-1)}}{2i}}$$
$$= \frac{\pi}{\dfrac{-e^{-i\pi s}+e^{i\pi s}}{2i}} = \frac{\pi}{\sin \pi s}.$$

(3.10.10) Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{x\sin x}{x^2+4x+20}\,dx$.

Answer. We use the same curve γ_R from 3.10.7 and again we work with the real part:

$$\int_{-\infty}^{\infty} \frac{x\sin x}{x^2+4x+20}\,dx = \operatorname{Re}\int_{-\infty}^{\infty} \frac{xe^{ix}}{x^2+4x+20}\,dx.$$

The function $f(z) = ze^{iz}/(z^2+4z+20)$ has poles at $z = -2\pm 4i$. Only $-2+4i$ lies inside γ_R, and its residue is

$$\operatorname{Res}(f,-2+4i) = \frac{(-2+4i)e^{i(-2+4i)}}{-2+4i+2+4i} = \frac{(-2+4i)e^{-2i}}{8e^4 i}.$$

Then

$$\int_{-\infty}^{\infty} \frac{x\sin x}{x^2+4x+20}\,dx = \operatorname{Re}\left(2\pi i\,\frac{(-2+4i)e^{-4-2i}}{8i}\right)$$
$$= \frac{\pi}{4e^4}\operatorname{Re}\left((-2+4i)e^{-2i}\right)$$
$$= \frac{\pi(2\cos 2+\sin 2)}{2e^4}.$$

(3.10.11) Show that, for $0<s<1$, $\displaystyle\int_{-\infty}^{\infty} \frac{e^{sx}}{1+e^x}\,dx = \frac{\pi}{\sin s\pi}$.

Answer. The poles of $f(z) = \frac{e^{sz}}{1+e^z}$ occur when $e^z = -1$; that is, when $z = (2k+1)\pi i$, $k\in\mathbb{Z}$. The residue at πi is

$$\operatorname{Res}(f,\pi i) = \lim_{z\to\pi i}\frac{(z-\pi i)e^{sz}}{1+e^z} = \lim_{\omega\to 0}\frac{\omega e^{s(\omega+\pi i)}}{1+e^{w+\pi i}}$$
$$= \lim_{\omega\to 0}\frac{\omega e^{s(\omega+\pi i)}}{1-e^w} = -e^{s\pi i}$$

We let γ_r be the curve describing the rectangle with vertices

$$-R, R, R+2\pi i, -R+2\pi i$$

(so only πi lies inside the curve). That is, $\gamma_r = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$, where

$$\begin{aligned} \gamma_1(t) &= t, \qquad t \in [-R, R]; \\ \gamma_2(t) &= R + 2\pi i t, \qquad t \in [0,1]; \\ \gamma_3(t) &= -t + 2\pi i, \qquad t \in [-R, R]; \\ \gamma_4(t) &= -R + 2\pi i(1-t), \qquad t \in [0,1]. \end{aligned}$$

Then

$$\int_{\gamma_1} f(z)\,dz = \int_{-R}^{R} \frac{e^{sx}}{1+e^x}\,dx$$

and

$$\int_{\gamma_3} f(z)\,dz = -\int_{-R}^{R} \frac{e^{-sx+2\pi i s}}{1+e^{-x+2\pi i}}\,dx = -e^{2\pi i s}\int_{-R}^{R} \frac{e^{sx}}{1+e^x}\,dx.$$

We have the estimates (using that $0 < s < 1$ for the limit)

$$\Big|\int_{\gamma_2} f(z)\,dz\Big| \le \int_0^1 \frac{2\pi e^{sR}}{|1+e^{R+2\pi i t}|}\,dt \le \frac{2\pi e^{sR}}{e^R - 1} \xrightarrow[R\to\infty]{} 0,$$

and

$$\Big|\int_{\gamma_4} f(z)\,dz\Big| \le \int_0^1 \frac{2\pi e^{-sR}}{|1+e^{-R+2\pi i(1-t)}|}\,dt \le \frac{2\pi e^{-sR}}{1-e^{-R}} \xrightarrow[R\to\infty]{} 0.$$

Then

$$\lim_{R\to\infty} \oint_{\gamma_R} f(z)\,dz = (1-e^{2\pi i s})\int_{-\infty}^{\infty} \frac{e^{sx}}{1+e^x}\,dx.$$

Therefore

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{e^{sx}}{1+e^x}\,dx &= \frac{2\pi i}{1-e^{2\pi i s}}\ \mathrm{Res}(f, \pi i) \\ &= \frac{-2\pi i e^{s\pi i}}{1-e^{s\pi i}} = \frac{-\pi}{\frac{e^{-s\pi i}-e^{s\pi i}}{2i}} \\ &= \frac{\pi}{\sin s\pi}. \end{aligned}$$

(3.10.12) Show that if $-1 < \alpha < 3$, then $\displaystyle\int_0^\infty \frac{x^\alpha}{(1+x^2)^2}\,dx = \frac{\pi(1-\alpha)}{4\cos\frac{\alpha\pi}{2}}$.

Answer. We have to deal with the expression z^α. This is $e^{\alpha \log z}$, but depends on the branch of the logarithm we choose. Since we will only work in the semiplane $\mathrm{Im}\, z \ge 0$, we may use the branch of the logarithm where $0 \le \theta \le \pi$.

Then, when $z = Re^{i\theta}$ with $\operatorname{Im} z \geq 0$ we have $0 \leq \theta$ and then

$$(R^{i\theta})^\alpha = e^{\alpha \log(Re^{i\theta})} = e^{\alpha(\log R + i\theta)} = R^\alpha\, e^{i\alpha\theta}.$$

With this interpretation we have, when $x < 0$, $x^\alpha = (-x)^\alpha\, e^{i\alpha\pi}$. Then

$$\begin{aligned}\int_{-R}^{R} \frac{x^\alpha}{(1+x^2)^2}\,dx &= \int_0^R \frac{x^\alpha}{(1+x^2)^2}\,dx + \int_{-R}^{0} \frac{x^\alpha}{(1+x^2)^2}\,dx \\ &= \int_0^R \frac{x^\alpha}{(1+x^2)^2}\,dx + e^{i\alpha\pi} \int_{-R}^{0} \frac{(-x)^\alpha}{(1+x^2)^2}\,dx \\ &= (1+e^{i\alpha\pi}) \int_0^R \frac{x^\alpha}{(1+x^2)^2}\,dx.\end{aligned}$$

We consider the upper semicircle given by $\gamma_1(t) = -R + t$, $t \in [0, 2R]$ and $\gamma_2(t) = Re^{it}$, $t \in [0, \pi]$. We have, using that when $R \geq \sqrt{2}$

$$|1 + R^2 e^{2it}|^2 = R^4 + 1 - 2R^2 \cos 2t \geq R^4 + 1 - 2R^2 = (R^2-1)^2 \geq \frac{R^4}{4},$$

the estimate

$$\begin{aligned}\left| \int_{\gamma_2} \frac{z^\alpha}{(1+z^2)^2}\,dz \right| &= \left| \int_0^\pi \frac{(Re^{it})^\alpha Ri\, e^{it}}{(1+R^2e^{2it})^2}\,dt \right| \leq \int_0^\pi \frac{R^{\alpha+1}}{|1+R^2e^{2it}|^2}\,dt \\ &\leq \int_0^\pi \frac{4R^{\alpha+1}}{R^4}\,dt = 4\pi R^{\alpha-3} \xrightarrow[R\to\infty]{} 0.\end{aligned}$$

It follows that

$$\begin{aligned}\int_0^\infty \frac{x^\alpha}{(1+x^2)^2}\,dx &= \frac{1}{(1+e^{i\alpha\pi})} \lim_{R\to\infty} \int_{\gamma_1+\gamma_2} \frac{z^\alpha}{(1+z^2)^2}\,dz \\ &= \frac{2\pi i}{(1+e^{i\alpha\pi})} \operatorname{Res}(f(z), i).\end{aligned}$$

It remains to calculate the residue. We have

$$\begin{aligned}\operatorname{Res}\left(\frac{z^\alpha}{(1+z^2)^2}, i\right) &= \lim_{z\to i} \frac{d}{dz} \frac{(z-i)^2 z^\alpha}{(1+z^2)^2} = \lim_{z\to i} \frac{d}{dz} \frac{z^\alpha}{(z+i)^2} \\ &= \lim_{z\to i} \frac{z^{\alpha-1}}{(z+i)^3} \left(\alpha(z+i) - 2z\right) = \frac{1-\alpha}{4}\, e^{(\alpha-1)\frac{\pi}{2}\, i}.\end{aligned}$$

Hence

$$\begin{aligned}\int_0^\infty \frac{x^\alpha}{(1+x^2)^2}\,dx &= \frac{2\pi i}{(1+e^{i\alpha\pi})} \frac{1-\alpha}{4}\, e^{(\alpha-1)\frac{\pi}{2}\, i} \\ &= \frac{2\pi}{e^{i\alpha\pi/2}(e^{-i\alpha\pi/2} + e^{i\alpha\pi/2})} \frac{1-\alpha}{4}\, e^{\alpha\frac{\pi}{2}\, i} \\ &= \frac{2\pi}{2e^{i\alpha\pi/2}\cos\alpha\frac{\pi}{2}} \frac{1-\alpha}{4}\, e^{\alpha\frac{\pi}{2}\, i} = \frac{\pi(1-\alpha)}{4\cos\frac{\alpha\pi}{2}}.\end{aligned}$$

(3.10.13) Using the ideas in Example 3.10.6 show that, for any meromorphic function f with no integer poles and such that there exist $s, c > 0$ with $N^{1+s}\,|f(\gamma_N)| \leq c$ for all N, where γ_N is the curve from Example 3.10.6,

$$\sum_{n\in\mathbb{Z}} f(n) = -\pi \sum_{w\in P} \operatorname{Res}\Big(f(z)\frac{\cos\pi z}{\sin\pi z}, w\Big), \tag{3.30}$$

where P is the set of poles of f.

Answer. Let $g(z) = \frac{f(z)\cos\pi z}{\sin\pi z}$. Since f has no integer poles, the poles of g are $P\cup\mathbb{Z}$. The integer poles are of order one, reasoning as in Example 3.10.6. We have

$$\begin{aligned}\operatorname{Res}\left(\frac{f(z)\cos\pi z}{\sin\pi z}, n\right) &= \lim_{z\to n}\frac{(z-n)f(z)\cos\pi z}{\sin\pi z} = \lim_{\omega\to 0}\frac{\omega f(\omega+n)\cos\pi(\omega+n)}{\sin\pi(\omega+n)}\\ &= \lim_{\omega\to 0}\frac{\omega f(\omega+n)\cos\pi\omega}{\sin\pi\omega} = \frac{f(n)}{\pi}.\end{aligned}$$

So

$$\frac{1}{2\pi i}\oint_{\gamma_N}\frac{f(z)\cos\pi z}{\sin\pi z}\,dz = \sum_{|w|<N}\operatorname{Res}\left(\frac{f(z)\cos\pi z}{\sin\pi z}, w\right) + \frac{1}{\pi}\sum_{n<|N|} f(n). \tag{AB.3.2}$$

We have that γ_N is the square with vertices $\big(N+\frac14\big)(\pm 1 \pm i)$ as in Example 3.10.6. We ran the estimates as in the example, only that now instead of (3.26) (where we had $f(z) = z^{-2}$) we obtain

$$\oint_{\gamma_N}\left|\frac{\cos\pi z}{z^2\,\sin\pi z}\right| dz \leq \frac{(3-e^{-\frac{\pi}{2}})(8N+2)}{(1-e^{-\frac{\pi}{2}})}\,c\,N^{-1-s}\xrightarrow[N\to\infty]{} 0.$$

Then, taking the limit on (AB.3.2),

$$\sum_{n\in\mathbb{Z}} f(n) = -\pi\sum_{w\in P}\operatorname{Res}\left(\frac{f(z)\cos\pi z}{\sin\pi z}, w\right)$$

(3.10.14) Show that $\displaystyle\sum_{n\in\mathbb{Z}}\frac{1}{\left(n-\frac12\right)^2} = \pi^2.$

Answer. Here we are taking $f(z) = \frac{1}{(z-\frac12)^2}$, with a single pole of order 2 at $z = \frac12$. So

$$\operatorname{Res}(f,\tfrac12) = \lim_{z\to\frac12}\frac{d}{dz}\left[\frac{(z-\frac12)^2\cos\pi z}{(z-\frac12)^2\sin\pi z}\right] = \lim_{z\to\frac12}\frac{d}{dz}\left[\frac{\cos\pi z}{\sin\pi z}\right] = \lim_{z\to\frac12}\left[\frac{-\pi}{\sin^2\pi z}\right] = -\pi.$$

Then (3.30) gives us

$$\sum_{n\in\mathbb{Z}} \frac{1}{\left(n-\frac{1}{2}\right)^2} = \pi^2.$$

(3.10.15) Show that, for any $r>0$, $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2+r^2} = \frac{\pi}{2r} + \frac{\pi}{r(e^{2\pi r}-1)} - \frac{1}{2r^2}$.

Answer. We apply (3.30) to the function $f(z) = \frac{1}{z^2+r^2}$. This function has poles at $\pm ri$. The residues are

$$\begin{aligned}\operatorname{Res}\left(\frac{\cos \pi z}{(z^2+r^2)\sin \pi z}, ri\right) &= \lim_{z\to ri} \frac{(z-ri)\cos\pi z}{(z-ri)(z+ri)\sin\pi z} = \frac{\cos \pi ri}{2ri \sin \pi ri} \\ &= -\frac{e^{r\pi}+e^{-r\pi}}{2r(e^{r\pi}-e^{-r\pi})} = -\frac{1+e^{-2r\pi}}{2r(1-e^{-2r\pi})}\end{aligned}$$

and

$$\begin{aligned}\operatorname{Res}\left(\frac{\cos \pi z}{(z^2+r^2)\sin \pi z}, -ri\right) &= \lim_{z\to -ri} \frac{(z+ri)\cos\pi z}{(z-ri)(z+ri)\sin\pi z} = \frac{\cos \pi ri}{2ri \sin \pi ri} \\ &= -\frac{e^{r\pi}+e^{-r\pi}}{2r(e^{r\pi}-e^{-r\pi})} = -\frac{1+e^{-2r\pi}}{2r(1-e^{-2r\pi})}.\end{aligned}$$

Therefore

$$\sum_{n\in\mathbb{Z}} f(n) = \pi\,\frac{1+e^{-2r\pi}}{r(1-e^{-2r\pi})}.$$

Since $f(0) = \frac{1}{r^2}$ and $f(-z) = f(z)$,

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n^2+r^2} &= \frac{1}{2}\Big(-f(0) + \sum_{n\in\mathbb{Z}} f(n)\Big) = \frac{1}{2}\left(-\frac{1}{r^2} + \pi\,\frac{1+e^{-2r\pi}}{r(1-e^{-2r\pi})}\right) \\ &= \frac{\pi}{2r} + \frac{\pi}{r(e^{2r\pi}-1)} - \frac{1}{2r^2}.\end{aligned}$$

(3.10.16) Let $r\in\mathbb{R}\setminus\mathbb{Z}$. Show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2-r^2} = \frac{1}{2r^2} - \frac{\pi\cos\pi r}{2r\sin\pi r}$.

Answer. We apply Exercise 3.10.13 to $f(z) = (z^2-r^2)^{-1}$. This has two simple poles $\pm r$. The residues are

$$\operatorname{Res}\left(\frac{\cos \pi z}{(z^2-r^2)\sin \pi z}, r\right) = \lim_{z\to r} \frac{\cos\pi z}{(z+r)\sin\pi z} = \frac{\cos \pi r}{2r\sin \pi r}$$

and

$$\operatorname{Res}\left(\frac{\cos \pi z}{(z^2 - r^2)\sin \pi z}, -r\right) = \lim_{z\to r} \frac{\cos \pi z}{(z-r)\sin \pi z} = \frac{\cos \pi r}{2r \sin \pi r}.$$

As $f(0) = -\frac{1}{r^2}$,

$$\sum_{n=1}^{\infty} \frac{1}{n^2 - r^2} = -\frac{f(0)}{2} + \frac{1}{2}\sum_{n\in\mathbb{Z}} f(n) = \frac{1}{2r^2} - \frac{\pi \cos \pi r}{2r \sin \pi r}.$$

(3.10.17) Show that the function $\frac{1}{\sin \pi z}$ is bounded on the square with vertices $(N + \frac{1}{4})(\pm 1 \pm i)$, independently of N. Conclude that if f is meromorphic, it has no integer poles, and there exist $s, c > 0$ such that $N^{1+s}\,|f(\gamma_N)| \le c$ for all N, where γ_N is the curve from Example 3.10.6, then

$$\sum_{n\in\mathbb{Z}} (-1)^n f(n) = -\pi \sum_{w\in P} \operatorname{Res}\Big(\frac{f(z)}{\sin \pi z}, w\Big), \tag{3.31}$$

where P is the set of poles of f.

Answer. We have

$$\frac{1}{|\sin \pi z|} = \frac{2}{|e^{iz} - e^{-iz}|}.$$

We consider the same square γ_N as in Example 3.10.6. This time we need estimates for $\frac{1}{\sin \pi z}$. On the vertical line $(N + \frac{1}{4})(1 + it)$ we have

$$\begin{aligned}\Big|2\sin\Big(\big(N+\tfrac{1}{4}\big)\pi + i\pi t\big(N+\tfrac{1}{4}\big)\Big)\Big|^2 &= \Big|2\sin\Big(\tfrac{\pi}{4} + i\pi t\big(N+\tfrac{1}{4}\big)\Big)\Big|^2 \\ &= \Big|e^{i\pi/4}e^{-\pi(N+\frac{1}{4})t} - e^{-i\pi/4}e^{\pi\left(N+\frac{1}{4}\right)t}\Big|^2 \\ &= e^{-2\pi\left(N+\frac{1}{2}\right)t} + e^{2\pi\left(N+\frac{1}{2}\right)t} \ge 1.\end{aligned}$$

For the other vertical line the estimates are the same as we can exchange the roles of z and $-z$ without changing the absolute values. On the horizontal line $\left(N + \frac{1}{4}\right)(t + i)$,

$$\begin{aligned}\Big|2\sin \pi\Big(N+\tfrac{1}{4}\Big)(t+i)\Big|^2 &= \Big|e^{i\left(N+\frac{1}{4}\right)t\pi}e^{-\left(N+\frac{1}{4}\right)\pi} - e^{-i\left(N+\frac{1}{4}\right)t\pi}e^{\left(N+\frac{1}{4}\right)\pi}\Big|^2 \\ &= e^{-2\left(N+\frac{1}{4}\right)\pi} + e^{2\left(N+\frac{1}{4}\right)\pi} - 2\cos 2\Big(N+\tfrac{1}{4}\Big)t\pi \\ &\ge \frac{1}{2}\end{aligned}$$

(by looking at the function $t \longmapsto t+t^{-1}-2\cos t$ on the positive x-axis). The estimate for the other horizontal line is entirely similar. In all, we have shown that there exists $c'>0$, independent of N, such that

$$\frac{1}{|\sin \pi z|} \le c'$$

on γ_N. By hypothesis, there exists $c''>0$ with $|f(n)| \le \frac{c''}{|n|^{1+s}}$ for sufficiently big n. Then, with $c = \max\{c', c''\}$ and N big enough, and using Lemma 3.3.4,

$$\left| \oint_{\gamma_N} \frac{f(z)}{\sin \pi z}\, dz \right| \le \frac{c(8N+4)}{N^{1+s}} \le \frac{12c}{N^s} \xrightarrow[N\to\infty]{} 0.$$

The residue of $\frac{1}{\sin \pi z}$ at $z=n$ is (using that $\sin(z+\pi n) = (-1)^n \sin z$)

$$\lim_{z\to n} \frac{(z-n)}{\sin \pi z} = \lim_{z\to\ n} \frac{(-1)^n}{\pi} \frac{1}{\frac{\sin \pi(z-n)}{\pi(z-n)}} = \frac{(-1)^n}{\pi}.$$

Thus

$$\operatorname{Res}\left(\frac{f(z)}{\sin \pi z}, n\right) = \frac{(-1)^n f(n)}{\pi}.$$

As the poles of $\frac{f(z)}{\sin \pi z}$ occur at $n \in \mathbb{Z}$ and at the poles of f, Theorem 3.10.4 gives (3.31).

(3.10.18) Show that for any $r \in \mathbb{R} \setminus \mathbb{Z}$ we have $\displaystyle\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(n+r)^2} = \frac{\pi^2 \cos \pi r}{\sin^2 \pi r}$.

Answer. We apply Exercise 3.10.17 to the function $f(z) = \frac{1}{(z+r)^2}$. The only pole of f occurs at $z=-r$, with order 2. So

$$\operatorname{Res}(f,-r) = \lim_{z\to -r} \frac{d}{dz} \frac{(z+r)^2}{(z+r)^2 \sin \pi z} = \lim_{z\to -r} \frac{\pi \cos \pi z}{\sin^2 \pi z} = -\frac{\pi \cos \pi r}{\sin^2 \pi r},$$

and then (3.31) gives us

$$\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(n+r)^2} = (-\pi) \frac{-\pi \cos \pi r}{\sin^2 \pi r} = \frac{\pi^2 \cos \pi r}{\sin^2 \pi r}$$

(3.10.19) Find $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2}$.

Answer. We apply Exercise 3.10.17 to the function $f(z) = \frac{1}{1+z^2}$. The poles are $\pm i$, both simple. The residues are

$$\operatorname{Res}\left(\frac{f(z)}{\sin \pi z}, i\right) = \lim_{z\to i} \frac{1}{(z+i)\sin \pi z} = \frac{1}{2i \sin \pi i} = \frac{2i}{2i(e^{-\pi} - e^{\pi})} = -\frac{1}{e^{\pi} - e^{-\pi}}$$

and

$$\operatorname{Res}\left(\frac{f(z)}{\sin \pi z}, -i\right) = \lim_{z\to -i} \frac{1}{(z-i)\sin \pi z} = \frac{2i}{-2i(e^{\pi} - e^{-\pi})} = -\frac{1}{e^{\pi} - e^{-\pi}}.$$

By Exercise 3.10.17

$$\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{1+n^2} = \frac{2\pi}{e^{\pi} - e^{-\pi}}.$$

Therefore

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2} &= \frac{1}{2}\Big(-1 + \sum_{n\in\mathbb{Z}} \frac{(-1)^n}{1+n^2}\Big)\\ &= \frac{1}{2}\left(-1 + \frac{2\pi}{e^{\pi} - e^{-\pi}}\right) = \frac{\pi}{e^{\pi} - e^{-\pi}} - \frac{1}{2}.\end{aligned}$$

(3.10.20) Consider the Gamma Function from Exercise 3.1.6. It is only defined for $\operatorname{Re} z > 0$. But it can be continued analytically (doing **analytic continuation**) in the following way. We know that $\Gamma(z+1) = z\Gamma(z)$. This we can write as

$$\Gamma(z) = \frac{\Gamma(z+1)}{z},$$

which suggests a way to extend the function "to the left", first to the strip $-1 < \operatorname{Re} z \leq 0$ (avoiding $z = 0$), and subsequently to each strip $-(n+1) < \operatorname{Re} z \leq -n$ (avoiding $z = -n$). We have to avoid 0 because otherwise it appears on the denominator, and this makes the extension undefined on all non-positive integers. The formula, using that Γ is holomorphic, shows that this extension is holomorphic. So we get a meromorphic function with poles at $-n+1$ for $n \in \mathbb{N}$.

(i) Show that these poles are simple.

(ii) Show that, for all $z \in \mathbb{C} \setminus \mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}.$$

Exercise 3.10.9, via a substitution of the form $y = t/x$, should be useful.

(iii) Show that there exist $a, b > 0$ such that for all $z \in \mathbb{C}$

$$|\Gamma(z)| \leq ae^{b|z|\log|z|}.$$

(iv) Show that there exist $a, b > 0$ such that for all $z \in \mathbb{C}$

$$\frac{1}{|\Gamma(z)|} \leq ae^{b|z|\log|z|}.$$

Answer.

(i) For each pole $-n$ with $n \in \{0\} \cup \mathbb{N}$,

$$\begin{aligned}\lim_{z\to -n}(z+n)\Gamma(z) &= \lim_{z\to -n}(z+n)\frac{\Gamma(z+1)}{z} \\ &= \lim_{z\to -n}(z+n)\frac{\Gamma(z+2)}{z(z+1)} \\ &= \cdots = \lim_{z\to -n}(z+n)\frac{\Gamma(z+n+1)}{z(z+1)\cdots(z+n)} \\ &= \frac{(-1)^n\Gamma(1)}{n!} = \frac{(-1)^n}{n!}.\end{aligned}$$

So the residue of first order exists and hence each pole is simple.

(ii) For $s \in (0,1)$, using Tonelli and the substitution $y = t/x$,

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \int_0^\infty t^{s-1}e^{-t}\,dt \int_0^\infty x^{-s}e^{-x}\,dx \\ &= \int_0^\infty \int_0^\infty t^{s-1}x^{-s}e^{-x-t}\,dt\,dx \\ &= \int_0^\infty \int_0^\infty \left(\frac{t}{x}\right)^s e^{-x-t}\,\frac{1}{t}\,dt\,dx \\ &= \int_0^\infty \int_0^\infty y^s e^{-x(y+1)}\,\frac{1}{y}\,dy\,dx \\ &= \int_0^\infty \int_0^\infty y^{s-1}e^{-x(y+1)}\,dy\,dx \\ &= \int_0^\infty \int_0^\infty y^{s-1}e^{-x(y+1)}\,dx\,dy\end{aligned}$$

$$\int_0^\infty \frac{y^{s-1}}{y+1}\,dy = \frac{\pi}{\sin \pi s},$$

the last equality coming from Exercise 3.10.9. By Corollary 3.6.2 the equality $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$ extends to all z where the function is holomorphic, which is $z \in \mathbb{C} \setminus \mathbb{Z}$.

(iii) Suppose first that $\operatorname{Re} z > \frac{1}{2}$. Let $k = \lfloor |z| \rfloor$. Then

$$\begin{aligned}|\Gamma(z)| &= \left| \int_0^\infty t^{z-1}\, e^{-t}\, dt \right| = \left| \int_0^\infty t^{\operatorname{Re} z - 1}\, e^{i(\operatorname{Im} z)\log t}\, e^{-t}\, dt \right| \\ &\le \int_0^\infty t^{\operatorname{Re} z-1}\, e^{-t}\, dt = \int_0^1 t^{\operatorname{Re} z-1}\, e^{-t}\, dt + \int_1^\infty t^{\operatorname{Re} z-1}\, e^{-t}\, dt \\ &\le \int_0^1 t^{-1/2} e^{-t}\, dt + \int_1^\infty t^k e^{-t}\, dt = c_1 + \Gamma(k+1) \\ &= c_1 + k! \le c_1 + k^k = c_1 + e^{k \log k}\end{aligned}$$

Next consider the case where $|\operatorname{Re} z| \le \frac{1}{2}$. When $|\operatorname{Im} z| \le 1$ we have a continuous function on a compact set, so there exists $c_2 > 0$ with $|\Gamma(z)| \le c_2$ there. When $|\operatorname{Im} z| > 1$, noting that $|z| \ge |\operatorname{Im} z| > 1$ and $\operatorname{Re} z + 1 \ge \frac{1}{2}$,

$$|\Gamma(z)| = \frac{|\Gamma(z+1)|}{|z|} \le |\Gamma(z+1)| \le c_1 + e^{k \log k}$$

for $k = \lfloor |z| \rfloor$. So in this case $|\Gamma(z)| \le c_1 + c_2 + e^{k \log k}$.

Finally, when $-n - \frac{1}{2} \le \operatorname{Re} z \le -n + \frac{1}{2}$ for $n \in \mathbb{N}$, we reduce to the previous case by

$$|\Gamma(z)| = \frac{|\Gamma(z+n)|}{|z(z+1)\cdots(z+n-1)|} \le 4\,|\Gamma(z+n)| \le 4\left(c_1 + c_2 + e^{k \log k}\right),$$

where we used that $|z+m| \ge -\operatorname{Re} z - m \ge n - \frac{1}{2} - m$ for $m = 0, \ldots, n-1$.

Now combining all the estimates and relabelling constants we have that $|\Gamma(z)| \le c_1 + e^{k \log k}$, where $k = \lfloor |z| \rfloor$. So

$$|\Gamma(z)| \le c_1 + e^{|z| \log |z|} \le a e^{b|z|\, \log |z|},$$

where $a = c_1 + 1$ and $b = 1$.

(iv) By (ii) we know that

$$\frac{1}{|\Gamma(z)|} = \frac{|\Gamma(1-z)\sin \pi z|}{\pi}.$$

We also know that

$$|\sin \pi z| = \frac{1}{2}\left| e^{i\pi z} - e^{-i\pi z} \right| \le \frac{1}{2}\sqrt{e^{|z|} + 1} \le e^{2|z|}.$$

Therefore, using that $|1-z| \le 1 + |z|$ and the inequality

$$(1+t)\log(1+t) \le 1 + 2\log t, \qquad t > 0,$$

we get

$$\frac{1}{|\Gamma(z)|} \leq \frac{1}{\pi} ae^{b|1-z|\log|1-z|}e^{2|z|} \leq \frac{1}{\pi} ae^{b(1+|z|)\log(1+|z|)}e^{2|z|}$$
$$\leq \frac{1}{\pi} ae^{b(1+2|z|\log|z|)}e^{2|z|} = \frac{1}{\pi} ae^{b}e^{2b|z|\log|z|}e^{2|z|}$$
$$\leq \frac{1}{\pi} ae^{b}e^{2(b+1)|z|\log|z|}$$

Renaming the constants we get

$$\frac{1}{|\Gamma(z)|} \leq ae^{b|z|\log|z|}.$$

3.11. Weierstrass' Factorization Theorem

(3.11.1) Let $h, k : V \to \mathbb{C}$ be functions defined on a region V. Show that if both functions are nonzero and differentiable at $z = z_0$, then

$$\frac{(hk)'(z_0)}{h(z_0)k(z_0)} = \frac{h'(z_0)}{h(z_0)} + \frac{k'(z_0)}{k(z_0)}$$

and

$$\frac{(h/k)'(z_0)}{h(z_0)/k(z_0)} = \frac{h'(z_0)}{h(z_0)} - \frac{k'(z_0)}{k(z_0)}.$$

Answer. Using the product rule,

$$\frac{(hk)'(z_0)}{h(z_0)k(z_0)} = \frac{h(z_0)k'(z_0) + h'(z_0)k(z_0)}{h(z_0)k(z_0)} = \frac{h'(z_0)}{h(z_0)} + \frac{k'(z_0)}{k(z_0)}.$$

Similarly,

$$\frac{(h/k)'(z_0)}{h(z_0)/k(z_0)} = \frac{h'(z_0)k(z_0) - h(z_0)k'(z_0)}{k(z_0)^2 h(z_0)/k(z_0)} = \frac{h'(z_0)}{h(z_0)} - \frac{k'(z_0)}{k(z_0)}.$$

(3.11.2) Find the Weierstrass factorization of the entire function $f(z) = \sin \pi z$.

Answer. The zeros of f are $z_n = n$, $n \in \mathbb{Z}$. We can take $p_n = 1$ for all n to satisfy (3.34). We also note that $z = 0$ is a zero of order 1. Then

$$\sin \pi z = z\, e^{g(z)} \prod_{n \neq 0} \Big(1 - \frac{z}{n}\Big) e^{z/n}.$$

If we index by only the positive integers, the terms for $-n$ have corresponding factors of the form $(1 + \frac{z}{n})e^{-z/n}$. Therefore

$$\sin \pi z = z\, e^{g(z)} \prod_{n>0} \Big(1 - \frac{z^2}{n^2}\Big). \tag{AB.3.3}$$

Next we use logarithmic differentiation repeatedly. We can also do it on the limit because both the product and the series and their derivatives converge uniformly (the product, by Proposition 3.11.5). So

$$\frac{\pi \cos \pi z}{\sin \pi z} = \frac{1}{z} + g'(z) + \sum_{n=1}^{\infty} \frac{-2z/n^2}{\Big(1 - \frac{z^2}{n^2}\Big)} = \frac{1}{z} + g'(z) + \sum_{n=1}^{\infty} \frac{-2z}{n^2 - z^2}.$$

We know from (3.28) that

$$\frac{\pi \cos \pi z}{\sin \pi z} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{-2z}{n^2 - z^2}.$$

Comparing the two expressions we conclude that $g'(z) = 0$, so g is constant. From (AB.3.3) we have, writing $c = e^{g(z)}$,

$$\frac{\sin \pi z}{z} = c \prod_{n>0} \Big(1 - \frac{z^2}{n^2}\Big).$$

Taking limit as $z \to 0$, we get $\pi = c$. Thus

$$\sin \pi z = \pi z \prod_{n>0} \Big(1 - \frac{z^2}{n^2}\Big).$$

(3.11.3) Show that Theorem 3.11.9 holds for entire functions with the convention that the empty product is equal to 1. That is, given f entire with no zeros, there exists g entire with $f = e^g$.

Answer. We can repeat the corresponding part of the argument in the proof of the Weierstrass factorization. Let f be entire with no zeros. Then f'/f is entire. By Proposition 3.3.8 there exists an entire function g with $g' = f'/f$. Then

$$(fe^{-g})' = f'e^{-g} - f\,g'\,e^{-g} = 0.$$

So there exists $c \in \mathbb{C}$ with $f = ce^g$. Choose $c_0 \in \mathbb{C}$ with $e^{c_0} = c$, and then $f = e^{c_0 g}$.

(3.11.4) Show that the Weierstrass product

$$f(z) = z \prod_{n=1}^{\infty} \left(1 - \frac{z}{n}\right) e^{z/n}$$

converges uniformly on compact sets and defines an entire function.

Answer. Since $\sum_n \frac{1}{n^2} < \infty$, the proof Proposition 3.11.8, with $f_n(z) = 1 - z/n$, gives us that

$$\prod_{n=1}^{\infty} \left(1 - \frac{z}{n}\right) e^{z/n}$$

converges uniformly on compact sets. As compact sets are bounded, multiplication by z does not affect the uniform convergence.

(3.11.5) Find an entire function with simple zeros at $z = n^2$ for all $n \in \mathbb{N}$. Ensure that you choose the minimal k in each factor $E_k(z)$.

Answer. Proposition 3.11.8 and the condition $\displaystyle\sum_n \frac{1}{n^4} < \infty$ guarantee that

$$f(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{n^2}\right) e^{z/n^2}$$

works. The choice $p_n = 0$ for all n works, since $\sum_n \frac{1}{n^2} < \infty$, so we can also consider

$$g(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{n^2}\right),$$

and that would be the minimal choice.

(3.11.6) Considering the Gamma function as a meromorphic function like in Exercise 3.10.20, show that $f(z) = 1/\Gamma(z)$ is entire and

$$\frac{1}{\Gamma(z)} = z e^{\gamma z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n}$$

where γ is the **Euler constant**

$$\gamma = \sum_{n=1}^{\infty} \frac{1}{n} - \log\left(1 + \frac{1}{n}\right).$$

Answer. From Exercise 3.10.20 we know that $\Gamma(z)$ is defined everywhere on $\mathbb{C}$ with the exception of $-n$ for $n \in \{0\} \cup \mathbb{N}$. We also know from the same exercise that

$$\lim_{z\to -n} (z+n)\Gamma(z) = \frac{(-1)^n}{n!}.$$

In particular for a fixed $n \in \mathbb{N}$ there exists $r > 0$ such that for all $z \in B_r(-n)$ we have

$$|z+n|\,|\Gamma(z)| > \frac{1}{2n!}.$$

That is,

$$\left|\frac{1}{\Gamma(z)}\right| = \frac{1}{|\Gamma(z)|} < 2n!\,|z+n|, \qquad |z+n| < r.$$

This shows that $1/\Gamma(z)$ has a removable singularity at $z = -n$, and that it can be extended as 0 at the point. As Γ takes finite values for all other choices of z, it follows that $1/\Gamma(z)$ extends to an entire function with zeros $z_n = -n$, $n \in \mathbb{N} \cup \{0\}$. The estimates from Exercise 3.10.20 allow us to use Theorem 3.11.12 with $k = 1$, so

$$\frac{1}{\Gamma(z)} = z\,e^{az+b} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n}$$

for certain $a, b \in \mathbb{C}$. As $z\Gamma(z) = \Gamma(z+1)$, we get that $\lim_{z\to 0} z\Gamma(z) = 1$. Then

$$1 = \lim_{z\to 0} \frac{1}{z\Gamma(z)} = \lim_{z\to 0} e^{az+b} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n} = e^b$$

since the product is holomorphic (hence continuous) by Proposition 3.11.8. So $b = 0$. Now evaluating at $z = 1$,

$$1 = e^a \prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right) e^{-1/n} = e^a \exp\left(\log\left(\prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right) e^{-1/n}\right)\right)$$

$$= e^a \exp\left(\sum_{n=1}^{\infty} \log\left(1 + \frac{1}{n}\right) - \frac{1}{n}\right) = e^a e^{-\gamma}.$$

So $a = \gamma$ and we are done.

(3.11.7) Show the inequalities

$$|E_n(z)| \geq e^{-2|z|^{n+1}}, \qquad |z| < \frac{1}{2}$$

and

$$|E_n(z)| \geq |1-z|e^{c_n|z|^n}, \qquad |z| \geq \frac{1}{2}$$

for constants $c_n > 0$, $n \in \mathbb{N}$.

Answer. Assume first that $|z| < \frac{1}{2}$. For such z we have the expansion

$$\log(1-z) = -\sum_{k=1}^{\infty} \frac{z^k}{k}.$$

Then

$$\begin{aligned}
|E_n(z)| &= |1-z| \left| e^{\sum_{k=1}^{n} \frac{z^k}{k}} \right| = |1-z| \left| e^{-\log(1-z) - \sum_{k=n+1}^{\infty} \frac{z^k}{k}} \right| \\
&= \left| e^{-\sum_{k=n+1}^{\infty} \frac{z^k}{k}} \right| \geq e^{-\sum_{k=n+1}^{\infty} \frac{|z|^k}{k}} \\
&= e^{-|z|^{n+1} \sum_{k=n+1}^{\infty} \frac{|z|^{k-n-1}}{k}} \geq e^{-|z|^{n+1} \sum_{k=n+1}^{\infty} \frac{2^{n+1-k}}{k}} \\
&\geq e^{-2|z|^{n+1}}.
\end{aligned}$$

When $|z| \geq \frac{1}{2}$,

$$\begin{aligned}
|E_n(z)| &= |1-z| \left| e^{\sum_{k=1}^{n} \frac{z^k}{k}} \right| \geq |1-z|\, e^{-\sum_{k=1}^{n} \frac{|z|^k}{k}} \\
&= |1-z|\, e^{-|z|^n \sum_{k=1}^{n} \frac{|z|^{k-n}}{k}} \\
&\geq |1-z|\, e^{-|z|^n \sum_{k=1}^{n} 2^{n-k}} = |1-z|\, e^{-(2^{n+1}(1-2^{-n})|z|^n}.
\end{aligned}$$

(3.11.8) Show that if $\{z_n\}$ is a sequence with $\sum_n |z_n|^{-s} < \infty$ for any $s > k$, such that $|\{|z_n| < r\}| \leq c_0 r^s$, and $|z - z_n| > |z_n|^{-k-1}$ for all n, then there exist $a, b > 0$ such that if $s \in (k, k+1)$

$$\left| \prod_{n=1}^{\infty} E_k\left(\frac{z}{z_n}\right) \right| \geq a e^{-b|z|^s}.$$

Answer. We can write

$$\Big|\prod_{n=1}^{\infty} E_k\Big(\frac{z}{z_n}\Big)\Big| = \prod_{\frac{|z|}{|z_n|}<\frac{1}{2}} \Big|E_k\Big(\frac{z}{z_n}\Big)\Big| \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} \Big|E_k\Big(\frac{z}{z_n}\Big)\Big|.$$

Using Exercise 3.11.7,

$$\prod_{\frac{|z|}{|z_n|}<\frac{1}{2}} \Big|E_k\Big(\frac{z}{z_n}\Big)\Big| \geq \prod_{\frac{|z|}{|z_n|}<\frac{1}{2}} e^{-2|z/z_n|^s} = \exp\Big(-2\sum_{\frac{|z|}{|z_n|}<\frac{1}{2}} |z/z_n|^s\Big)$$

$$\geq \exp\Big(-2|z|^s \sum_{n=1}^{\infty} \frac{1}{|z_n|^s}\Big) = e^{-c_s|z|^s},$$

where $c_s = 2\sum_{n=1}^{\infty} \frac{1}{|z_n|^s}$.

For the other product, using again Exercise 3.11.7,

$$\prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} \Big|E_k\Big(\frac{z}{z_n}\Big)\Big| \geq \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} \Big|1-\frac{z}{z_n}\Big| \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} e^{-c\left|\frac{z}{z_n}\right|^k} \tag{AB.3.4}$$

(use c for the constant to simplify notation since k is fixed). For the first of these products, using the condition $|z_n| \leq 2|z|$ (which in particular means that the product has finitely many factors, since $|\{|z_n| \leq 2|z|\}| \leq c_0(2|z|)^s$),

$$\prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} \Big|1-\frac{z}{z_n}\Big| = \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} \Big|\frac{z_n - z}{z_n}\Big| \geq \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} |z_n|^{-k-2}$$

$$\geq \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} (2|z|)^{-k-2} \geq (2|z|)^{-(k+2)c_0(2|z|)^s}$$

$$= e^{-(k+2)c_0(2|z|)^s \log 2|z|} \geq e^{-a_0|z|^{s'}}$$

for an appropriate constant a_0 and $s' > s$. But then the inequality holds for s, too (namely, we could have worked with an $s_0 < s$ and use s where we used s').

For the second of the products in (AB.3.4),

$$\prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} e^{-c\left|\frac{z}{z_n}\right|^k} \geq \prod_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} e^{-2c\left|\frac{2z}{z_n}\right|^s}$$

$$= \exp\Big(-2c\sum_{\frac{|z|}{|z_n|}\geq\frac{1}{2}} |z/z_n|^s\Big)$$

$$\geq \exp\Big(-2c|z|^s \sum_{n=1}^{\infty} \frac{1}{|z_n|^s}\Big) = e^{-c'_s|z|^s}$$

Collecting the estimates we have

$$\left|\prod_{n=1}^{\infty} E_k\left(\frac{z}{z_n}\right)\right| \geq e^{-c_s|z|^s} e^{-a_0|z|^s} e^{-c_s'|z|^s} = e^{-b|z|^s}.$$

(3.11.9) Let $R > 0$ and f analytic on $R\overline{\mathbb{D}}$, with $f(0) \neq 0$ and $f \neq 0$ on the boundary. Denote the finitely many roots of f by $\{z_1, \ldots, z_m\}$, counting multiplicities. Show Jensen's formula

$$\log|f(0)| = \sum_n \log\frac{|z_n|}{R} + \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\,d\theta.$$

by going through the following steps.

(i) Show that it is enough to prove the case $R = 1$.

(ii) Prove the case where f has not roots.

(iii) Show that the formula behaves nicely with products.

(iv) Prove the case where $f(z) = z - z_1$.

(v) Conclude the general case.

Answer.

(i) The case for general $R > 0$ follows from applying the case $R = 1$ to $g(z) = f(Rz)$.

(ii) If f has no roots, then $f(z) = e^{g(z)}$ by Exercise 3.6.3. So $|f(z)| = e^{\operatorname{Re} g(z)}$. Then from Cauchy's Theorem

$$\log|f(0)| = \operatorname{Re} g(0) = \operatorname{Re}\frac{1}{2\pi}\int_0^{2\pi} g(e^{i\theta})\,d\theta$$
$$= \frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})|\,d\theta.$$

(iii) We have that the zeros of $f_1 f_2$, counting multiplicities, are precisely the joined list of zeros (counting multiplicities) of f_1 and f_2. Also,

$$\log|f_1(0)f_2(0)| = \log|f_1(0)| + \log|f_2(0)|.$$

and

$$\int_0^{2\pi} \log|f_1(Re^{i\theta})f_2(Re^{i\theta})|\,d\theta = \int_0^{2\pi} \log|f_1(Re^{i\theta})|\,d\theta + \int_0^{2\pi} \log|f_2(Re^{i\theta})|\,d\theta.$$

(iv) We have $f(z) = z - z_1$. The hypotheses on the root guarantee that $0 < |z_1| < 1$. We may write $z - z_1 = z_1(z/z_1 - 1)$. The formula holds trivially for a constant and it works nicely with products as we just proved above, so it is enough to show the formula for $f(z) = 1 - z/z_1$. Then $f(0) = 1$, so the formula to be proven becomes

$$0 = \log|z_1| + \int_0^{2\pi} \log|1 - e^{i\theta}/z_1|\, d\theta.$$

We have

$$\begin{aligned}
\log|z_1| + \int_0^{2\pi} \log|1 - e^{i\theta}/z_1|\, d\theta &= \int_0^{2\pi} \log|z_1 - e^{i\theta}|\, d\theta \\
&= \int_0^{2\pi} \log|e^{-i\theta} z_1 - 1|\, d\theta \\
&= -\int_0^{2\pi} \log|e^{i\theta} z_1 - 1|\, d\theta \\
&= 0,
\end{aligned}$$

the last equality by (ii), since $z_1 z - 1$ has no roots in disk (for $1/|z_1| > 1$).

(v) Given any f with the given conditions, we know from Corollary 3.6.3 that $f(z) = (z - z_1)\cdots(z - z_n)g(z)$, with $g(z)$ holomorphic with no zeros. The previous parts of the exercise show that Jensen's formula applies to each factor, and then it applies to the product.

(3.11.10) Let $R > 0$ and f analytic on $R\overline{\mathbb{D}}$, with $f(0) \neq 0$ and $f \neq 0$ on the boundary. Denote the finitely many roots of f by $\{z_1, \ldots, z_m\}$, counting multiplicities. Let $\mathfrak{n}(r)$ denote the number of nonzero roots of f which have absolute value less than r. Show the following variation of Jensen's formula:

$$\log|f(0)| = -\int_0^R \frac{\mathfrak{n}(x)}{x}\, dx + \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\, d\theta. \quad (3.36)$$

Answer. We need to show that

$$-\int_0^R \frac{\mathfrak{n}(x)}{x}\, dx = \sum_n \log\frac{|z_n|}{R}.$$

The formula reduces immediately to the case $R = 1$ as in the previous exercise. Let $h_n(x) = 1$ if $x > |z_n|$ and 0 otherwise. So $\mathfrak{n}(x) = \sum_n h_n(x)$. Then

$$-\int_0^1 \frac{\mathfrak{n}(x)}{x}\, dx = -\sum_n \int_0^1 \frac{h_n(x)}{x}\, dx = -\sum_n \int_{z_n}^1 \frac{1}{x}\, dx = \sum_n \log|z_n|.$$

(3.11.11) Let f be entire with $\{z_n\}$ its nonzero roots and such that $|f(z)| \le ae^{b|z|^s}$ for certain constants a, b and all $s > s_0 > 0$.

(i) Show that there exists $c > 0$ with $\mathfrak{n}(r) \le cr^{s_0}$ for large enough r.

(ii) Show that
$$\sum_{n=1}^{\infty} \frac{1}{|z_n|^s} < \infty$$
for all $s > s_0$.

Answer.

(i) If $f(0) = 0$, we replace f with $g(z) = f(z)/z^m$ for appropriate m. The function g has the same zeros as f with the exception of 0, and outside a small disk around the origin we still have $|g(z)| \le ae^{b|z|^s}$ after possibly redefining a. So we may assume that $f(0) \ne 0$. Fix $r > 0$ and let $R = 2r$. Using (3.36) and that $\mathfrak{n}(r) \le \mathfrak{n}(x)$ if $r \le x$,
$$\begin{aligned}\mathfrak{n}(r)\log 2 = \mathfrak{n}(r)\int_r^R \frac{1}{x}\,dx &\le \int_r^R \frac{\mathfrak{n}(x)}{x}\,dx \le \int_0^R \frac{\mathfrak{n}(x)}{x}\,dx \\ &= -\log|f(0)| + \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\,d\theta \\ &\le c_1 + \frac{1}{2\pi}\int_0^{2\pi} \log(ae^{bR^s})\,d\theta \\ &= c_1 + \log a + 2^s br^s.\end{aligned}$$
Taking limit as $s \to s_0$ we get
$$\mathfrak{n}(r) \le \frac{c_1 + \log a}{\log 2} + \frac{2^{s_0} b}{\log 2} r^{s_0}.$$
So as long as $r \ge (c_1 + \log a)^{1/s_0}$,
$$\mathfrak{n}(r) \le \frac{2^{s_0}b + 1}{\log 2} r^{s_0}.$$

(ii) Now the series. Choose n_0 such that $\mathfrak{n}(r) \le cr^{s_0}$ for all $r \ge 2^{n_0}$. Then for $s > s_0$
$$\begin{aligned}\sum_{|z_n| \ge 2^{n_0}} |z_n|^{-s} = \sum_{k=n_0}^{\infty} \sum_{2^k \le |z_n| < 2^{k+1}} |z_n|^{-s} &\le \sum_{k=n_0}^{\infty} \frac{\mathfrak{n}(2^{k+1})}{2^{ks}} \\ &\le c \sum_{k=n_0}^{\infty} \frac{2^{(k+1)s_0}}{2^{ks}} = 2^{s_0} c \sum_{k=n_0}^{\infty} 2^{k(s_0 - s)} < \infty\end{aligned}$$

since $2^{s_0-s} < 1$. Thus $\sum_{n=1}^{\infty} \frac{1}{|z_n|^s} < \infty$.

(3.11.12) Use Exercises 3.9.2, 3.11.7, 3.11.8 and 3.11.11 to write the proof of Theorem 3.11.12.

Answer. Fix $s > s_0$. Using Theorem 3.11.9 we can write

$$f(z) = z^m\, e^{g(z)} \prod_{n=1}^{\infty} E_n\Big(\frac{z}{z_n}\Big)$$

The condition $|f(z)| \le ae^{b|z|^s}$ gives us, via Exercise 3.11.11, that

$$\sum_{n=1}^{\infty} \frac{1}{|z_n|^s} < \infty.$$

Then Proposition 3.11.8 allows us to take $p_n = \lfloor s_0 \rfloor$ for all n; we denote this number by k.

Using the hypothesis on f and Exercise 3.11.8 there exist c, d positive with

$$e^{\operatorname{Re} g(z)} = |e^{g(z)}| = \left| \frac{f(z)}{\prod_{n=1}^{\infty} E_k\Big(\frac{z}{z_n}\Big)} \right| \le \frac{ae^{b|z|^s}}{ce^{-d|z|^s}} = ac^{-1}e^{(b+d)|z|^s}.$$

It follows, after renaming constants, that for each $s > k$ there exist $a, b > 0$ such that $\operatorname{Re} g(z) \le a + b|z|^s$. Then Exercise 3.9.2 shows that g is a polynomial of degree at most k.

(3.11.13) Consider the entire function $s(z) = \sum_{n=0}^{\infty} \frac{z^n}{(2n+1)!}$.

(i) Show that

$$s(z^2) = \frac{\sinh z}{z}.$$

(ii) Show that

$$s(z) = \prod_{n=1}^{\infty} \Big(1 + \frac{z}{n^2\pi^2}\Big).$$

Answer.

(i) We have

$$s(z^2) = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n+1)!} = \frac{1}{z} \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} = \frac{\sinh z}{z}.$$

(ii) We can estimate, for $|z| \geq \frac{1}{4}$,

$$|s(z)| \leq \sum_{n=0}^{\infty} \frac{|z|^n}{(2n+1)!} = \frac{\sinh |z|^{1/2}}{|z|^{1/2}} \leq \frac{2e^{|z|^{1/2}}}{|z|^{1/2}} \leq 4e^{|z|^{1/2}}.$$

The roots of $s(z^2)$ are the roots of sinh. If $\sinh(a+ib) = 0$, this is

$$0 = e^a e^{ib} - e^{-a} e^{-ib}.$$

This gives $e^{2a} = e^{2ib}$. This can only be satisfied with $a = 0$, $b = in\pi$. It follows that the nonzero roots of s are $\big\{ -n^2\pi^2 \ : \ n \in \mathbb{N} \big\}$. Then Theorem 3.11.12 with $k = \lfloor 1/2 \rfloor = 0$ gives us

$$s(z) = c \prod_{n \in \mathbb{N}} \Big(1 + \frac{z}{n^2\pi^2}\Big).$$

Evaluating at $z = 0$ we get that $c = 1$.

CHAPTER

4

Hilbert spaces

4.1. Basic Definitions

(4.1.1) Prove that each of the spaces in Examples 4.1.2 is actually a Hilbert space.

Answer.

(i) Since $\mathbb{C}^n$ is finite-dimensional, it is complete. Indeed, the inner product makes $\mathbb{C}^n$ a metric space with the metric $d(x,y) = \big(\sum_{k=1}^n |x(k) - y(k)|^2\big)^{1/2}$. If $\{x_n\} \subset \mathbb{C}^n$ is Cauchy, then for each k

$$|x_n(k) - x_m(k)|^2 \leq \sum_{j=1}^m |x_n(j) - x_m(j)|^2 = d(x_n, x_m)^2.$$

This shows that the sequence $\{x_n(k)\} \subset \mathbb{C}$ is Cauchy. As $\mathbb{C}$ is complete, there exists $x(k) \in \mathbb{C}$ with $x_n(k) \to x(k)$. Then

$$d(x_n, x)^2 = \sum_{j=1}^m |x_n(j) - x(j)|^2 \xrightarrow[n\to\infty]{} 0,$$

proving that any Cauchy sequence converges and hence $\mathbb{C}^n$ is complete.

The sesquilinearity is obtained directly from doing arithmetic on the expression y^*x.

(ii) The formula for the inner product is a series version of y^*x, so the sesquilinearity is automatic. The series for $\langle x, y\rangle$ converges by the Cauchy–Schwarz inequality:

$$\sum_n |x_n y_n| \leq \Big(\sum_n |x_n|^2\Big)^{1/2}\Big(\sum_n |y_n|^2\Big)^{1/2} < \infty,$$

showing that the series converges absolutely. The completeness is a consequence of Riesz–Fischer (Theorem 2.8.12).

(iii) Again the inner product is sesquilinear by elementary computations. And again the Cauchy–Schwarz inequality guarantees that the integral for the inner product converges. Completeness is given by Riesz–Fischer (Theorem 2.8.12). When $\langle f, f\rangle = 0$, then $[f] = [0]$ by definition, so the inner product is definite.

(iv) Finite-dimensional, so complete. The expression Y^*X is sesquilinear and the trace is linear, so the inner product is indeed sesquilinear. When $\langle X, X\rangle = 0$, this is $\mathrm{Tr}(X^*X) = 0$, and in turn this is

$$0 = \mathrm{Tr}(X^*X) = \sum_{k,j} |x_{kj}|^2,$$

so $X = 0$.

(4.1.2) Show that c_{00} is dense in $\ell^2(\mathbb{N})$.

Answer. Let $x \in \ell^2(\mathbb{N})$. Fix $\varepsilon > 0$ and let $n_0 \in \mathbb{N}$ such that $\sum_{n>n_0} \|x_n\|^2 < \varepsilon^2$. Let $z \in c_{00}$ be given by $z_k = x_k$ for $k \leq n_0$ and $z_k = 0$ for $k > n_0$. Then

$$\|x - z\|_2^2 = \sum_{n>n_0} \|x_n\|^2 < \varepsilon^2.$$

As ε was arbitrary, this shows that $\overline{c_{00}} = \ell^2(\mathbb{N})$.

(4.1.3) Find examples of pre-Hilbert spaces which are not Hilbert spaces.

Answer. An already mentioned example is c_{00} in $\ell^2(\mathbb{N})$. Another example is $C[0,1]$ in $L^2[0,1]$.

4.2. The role of the inner product in the topological structure

(4.2.1) Prove (4.3).

Answer. This simply requires expanding:

$$\begin{aligned}\|\xi+\eta\|^2 &= \langle \xi+\eta, \xi+\eta\rangle = \langle \xi,\xi\rangle + \langle \eta,\eta\rangle + \langle \xi,\eta\rangle + \langle \eta,\xi\rangle \\ &= \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\,\langle \xi,\eta\rangle.\end{aligned}$$

(4.2.2) Let $n \in \mathbb{N}$ and $\mathcal{H} = \mathbb{C}^n$. Show that $\mathcal{H}$ admits infinitely many inner products that are not multiples of each other, and that they all induce the same topology (for this, show that the induced norms are equivalent).

Answer. Given $t_1, \ldots, t_n \in (0, \infty)$, define $\bar{t} = (t_1, \ldots, t_n)$ and

$$\langle x, y\rangle_{\bar{t}} = \sum_{k=0}^{n} t_k\, x_k\, \overline{y_k}.$$

The inner product properties are proven with the exact same proofs as in the usual case (which is $t_1 = \cdots = t_n = 1$), since $t_k > 0$ for all k. Then

$$\|x\|_{\bar{t}} = \Big(\sum_{k=0}^{n} |t_k|^2\, |x_k|^2\Big)^{1/2},$$

and with respect to the Euclidean norm we have

$$\min\{\bar{t}\}\, \|x\|_2 \le \|x\|_{\bar{t}} \le \max\{\bar{t}\}\, \|x\|_2$$

More generally, for $A \in M_n(\mathbb{C})$ positive definite we can define $\langle x, y\rangle_A = \langle Ax, y\rangle$. This generalizes the above construction. Indeed, using Proposition 1.7.14 one can show that A, being positive-definite, can be written $A = U^*DU$, with U a unitary and D diagonal with positive diagonal entries. Then

$$\langle x, y\rangle_A = \langle Ux, Uy\rangle_{\mathrm{diag}(D)}.$$

(4.2.3) Prove the Polarization Identity (4.6).

Answer. Since $\sum_{k=0}^3 i^k = \sum_{k=0}^3 i^{2k} = 0$,

$$\begin{aligned}\sum_{k=0}^3 i^k\psi(\xi+i^k\eta,\xi+i^k\eta) &= \sum_{k=0}^3 i^k\big(\psi(\xi,\xi)+\psi(\eta,\eta)+i^k\psi(\eta,\xi)+i^{-k}\psi(\xi,\eta)\big)\\ &= (\psi(\xi,\xi)+\psi(\eta,\eta))\sum_{k=0}^3 i^k\\ &\qquad + \psi(\eta,\xi)\sum_{k=0}^3 i^{2k} + \psi(\xi,\eta)\sum_{k=0}^3 1\\ &= 4\psi(\xi,\eta).\end{aligned}$$

(4.2.4) Prove the Parallelogram Identity (4.2).

Answer. We have

$$\begin{aligned}\|\xi+\eta\|^2+\|\xi-\eta\|^2 &= \langle\xi+\eta,\xi+\eta\rangle+\langle\xi-\eta,\xi-\eta\rangle\\ &= 2\langle\xi,\xi\rangle+2\langle\eta,\eta\rangle+2\mathrm{Re}\,\langle\xi,\eta\rangle-2\mathrm{Re}\,\langle\xi,\eta\rangle\\ &= 2\|\xi\|^2+2\|\eta\|^2.\end{aligned}$$

(4.2.5) Use the Parallelogram Identity (4.2) to show that none of the examples on Section 5.1 is a Hilbert space, with the exception of Example 5.1.4 and the case $p=2$ in Examples 5.1.7 and 5.1.8.

Answer. Consider first $\ell^p(\mathbb{N})$. Let $\xi=(1,0,0,0,\ldots)$, $\eta=(0,1,0,0,\ldots)$. Then

$$\|\xi+\eta\|_p^2+\|\xi-\eta\|_p^2 = 2(1^p+1^p)^{2/p} = 2^{1+\frac{2}{p}},$$

while

$$2\|\xi\|_p^2+2\|\eta\|_p^2 = 2+2 = 4..$$

The equality is only possible when $p=2$ (since $2^{1+\frac{2}{p}}=2^2$ implies $p=2$), so $\|\cdot\|_p$ is never a Hilbert space norm when $p\neq 2$.

For $L^p(X)$ we can use the same idea as above.

For c_0, use the same ξ and η from the ℓ^p example: now we have

$$\|\xi+\eta\|_\infty^2+\|\xi-\eta\|_\infty^2=1+1=2,$$
$$2\|\xi\|_\infty^2+2\|\eta\|_\infty^2=2+2=4.$$

(4.2.6) Fix $n \geq 3$, and let $\omega \in \mathbb{C}$ be an n^{th} primitive root of unity. Let $\psi : \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ be a sesquilinear form. Show that

$$\psi(\xi,\eta)=\frac{1}{n}\sum_{k=0}^{n-1} w^k\psi(\xi+\omega^k\eta,\xi+\omega^k\eta). \tag{4.7}$$

Answer. We have, since $\sum_{k=0}^{n-1}\omega^k=\sum_{k=0}^{n-1}\omega^{2k}=0$,

$$\begin{aligned}\sum_{k=0}^{n-1}\omega^k\psi(\xi+\omega^k\eta,\xi+\omega^k\eta)&=\sum_{k=0}^{n-1}\omega^k\big(\psi(\xi,\xi)+\psi(\eta,\eta)\\&\qquad+\omega^k\psi(\eta,\xi)+\overline{\omega}^k\psi(\xi,\eta)\big)\\&=(\psi(\xi,\xi)+\psi(\eta,\eta))\sum_{k=0}^{n-1}\omega^k\\&\qquad+\psi(\eta,\xi)\sum_{k=0}^{n-1}\omega^{2k}+\psi(\xi,\eta)\sum_{k=0}^{n-1}1\\&=n\psi(\xi,\eta).\end{aligned}$$

(4.2.7) Does (4.7) work for $n=2$? Why?

Answer. No, it doesn't. When $n=2$, the roots are 1 and -1, so there are no conjugates; conjugates were essential in the sum corresponding to $\psi(\xi,\eta)$ being nonzero. Concretely, the root -1, which is the only primitive root of unity of order 2, does not satisfy $(-1)^2+1^2=0$.

Explicitly,

$$\begin{aligned}\sum_{k=0}^{1}(-1)^k\psi(\xi+(-1)^k\eta,\xi+(-1)^k\eta)&=\psi(\xi+\eta,\xi+\eta)-\psi(\xi-\eta,\xi-\eta)\\&=2\psi(\xi,\eta)+2\psi(\eta,\xi).\end{aligned}$$

In other words, when $n=2$ the expression in (4.7) gives $\operatorname{Re}\psi(\xi,\eta)$.

(4.2.8) Let $\psi : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form. Show that

$$\psi(\xi, \eta) = \frac{1}{2\pi} \int_0^{2\pi} e^{it} \psi(\xi + e^{it}\eta, \xi + e^{it}\eta)\, dt. \tag{4.8}$$

Answer. This is exactly the same computation as with the roots of unity, only now using the identities $\int_0^{2\pi} e^{it}\, dt = \int_0^{2\pi} e^{2it}\, dt = 0$. So

$$\begin{aligned}
\int_0^{2\pi} e^{it} \psi(\xi + e^{it}\eta, \xi + e^{it}\eta) &= \int_0^{2\pi} e^{it} \big(\psi(\xi, \xi) + \psi(\eta, \eta) \\
&\qquad + e^{it}\psi(\eta, \xi) + e^{-it}\psi(\xi, \eta)\big)\, dt \\
&= (\psi(\xi, \xi) + \psi(\eta, \eta)) \int_0^{2\pi} e^{it}\, dt \\
&\qquad + \psi(\eta, \xi) \int_0^{2\pi} e^{2it}\, dt + \psi(\xi, \eta) \int_0^{2\pi} 1\, dt \\
&= 2\pi\psi(\xi, \eta).
\end{aligned}$$

(4.2.9) Let $\mathcal{H}$ be a Hilbert space and $\xi, \eta, \nu \in \mathcal{H}$ with $\|\xi\| = 1$ and $\|\eta\| \leq 1$, $\|\nu\| \leq 1$. Fix $\varepsilon > 0$. Show that

$$\left\| \xi - \frac{\eta + \nu}{2} \right\| < \varepsilon \quad \text{implies} \quad \|\xi - \eta\| < 2\sqrt{\varepsilon} \text{ and } \|\xi - \nu\| < 2\sqrt{\varepsilon}.$$

Answer. From $\|\xi\| = 1$ we get

$$1 = \|\xi\| \leq \left\| \xi - \frac{\eta + \nu}{2} \right\| + \left\| \frac{\eta + \nu}{2} \right\| < \varepsilon + \left\| \frac{\eta + \nu}{2} \right\|,$$

giving us

$$\left\| \frac{\eta + \nu}{2} \right\|^2 > (1 - \varepsilon)^2.$$

Also,

$$\operatorname{Re} \langle \xi, \eta \rangle \leq |\langle \xi, \eta \rangle| \leq \|\xi\| \, \|\eta\| \leq 1.$$

So $1 - \mathrm{Re}\,\langle \xi, \eta\rangle \geq 0$ and similarly $1 - \mathrm{Re}\,\langle \xi, \nu\rangle \geq 0$. Now

$$\begin{aligned}
\|\xi - \eta\|^2 &= \|\xi\|^2 + \|\eta\|^2 - 2\mathrm{Re}\,\langle \xi, \eta\rangle \\
&\leq 2(1 - \mathrm{Re}\,\langle \xi, \eta\rangle) \\
&\leq 2(1 - \mathrm{Re}\,\langle \xi, \eta\rangle + 1 - \mathrm{Re}\,\langle \xi, \nu\rangle)) = 2(2 - \mathrm{Re}\,\langle \xi, \eta + \nu\rangle) \\
&= 2(1 + (1-\varepsilon)^2 - \mathrm{Re}\,\langle \xi, \eta + \nu\rangle + 1 - (1-\varepsilon)^2) \\
&\leq 2\Big(\|\xi\|^2 + \Big\|\frac{\eta+\nu}{2}\Big\|^2 - 2\mathrm{Re}\,\Big\langle \xi, \frac{\eta+\nu}{2}\Big\rangle + 2\varepsilon - \varepsilon^2 \\
&= 2\Big(\Big\|\xi - \frac{\eta+\nu}{2}\Big\|^2 + 2\varepsilon - \varepsilon^2\Big) \\
&< 2(\varepsilon^2 + 2\varepsilon - \varepsilon^2) = 4\varepsilon.
\end{aligned}$$

Taking square root we get $\|\xi - \eta\| < 2\sqrt{\varepsilon}$. Exchanging the roles of η and ν gives us the other inequality.

4.3. Orthogonality

(4.3.1) *(Definition of convexity)* Let $K \subset \mathcal{H}$. Show that the following statements are equivalent:

(i) for any $\xi, \eta \in K$ and $t \in [0,1]$, $t\xi + (1-t)\eta \in K$;

(ii) for any $\xi_1, \ldots, \xi_n \in K$ and $t_1, \ldots, t_n \in [0,1]$ with $\sum_j t_j = 1$, we have $\sum_j t_j \xi_j \in K$.

Answer. (i) $\Longrightarrow$ (ii) We proceed by induction. Assume that for any elements $\xi_1, \ldots, \xi_k \in K$ and $t_1, \ldots, t_k \in [0,1]$ with $\sum_j t_j = 1$, we have $\sum_j t_j \xi_j \in K$. Then, if $\xi_1, \ldots, \xi_{k+1} \in K$ and $t_1, \ldots, t_{k+1} \in [0,1]$ with $\sum_j t_j = 1$ and $t_{k+1} \neq 1$, let

$$s = \sum_{j=1}^k t_j, \qquad s_j = \frac{t_j}{s}, \qquad\qquad j = 1, \ldots, k.$$

Then $0 < s \leq 1$, $s_j \geq 0$ for all j, and $\sum_{j=1}^{k} s_j = 1$. By the inductive hypothesis, $\sum_{j=1}^{k} s_j \xi_j \in K$. Then, since $1 - s = t_{k+1}$,

$$\sum_{j=1}^{k+1} t_j \xi_j = s\left(\sum_{j=1}^{k} s_j \xi_j\right) + (1-s)\xi_{k+1} \in K.$$

When $t_{k+1} = 1$ we have $t_1 = \cdots = t_k = 0$, so $\displaystyle\sum_{j=1}^{k+1} t_j \xi_j = \xi_{k+1} \in K$.

(ii) $\Longrightarrow$ (i) Trivial, taking $n = 2$.

(4.3.2) Prove Lemma 4.3.15.

Answer. Suppose $\xi_1, \ldots, \xi_m$ are orthonormal and $c_1, \ldots, c_n \in \mathbb{C}$. We proceed by induction. For the base case, we have $\|c\,\xi_1\| = |c|\,\|\xi_1\| = |c|$. Assume as inductive hypothesis that

$$\Big\|\sum_{j=1}^{k} c_j \xi_j\Big\|^2 = \sum_{j=1}^{k} |c_j|^2.$$

Then

$$\begin{aligned}
\Big\|\sum_{j=1}^{k+1} c_j \xi_j\Big\|^2 &= \Big\|\sum_{j=1}^{k} c_j \xi_j + c_{k+1}\xi_{k+1}\Big\|^2 \\
&= \Big\|\sum_{j=1}^{k} c_j \xi_j\Big\|^2 + \|c_{k+1}\xi_{k+1}\|^2 + 2\mathrm{Re}\left\langle \sum_{j=1}^{k} c_j \xi_j, c_{k+1}\xi_{k+1}\right\rangle \\
&= \sum_{j=1}^{k} |c_j|^2 + |c_{k+1}|^2 + 2\sum_{j=1}^{k} c_j c_{k+1} \langle \xi_j, \xi_{k+1}\rangle \\
&= \sum_{j=1}^{k+1} |c_j|^2.
\end{aligned}$$

(4.3.3) Let $\eta_1, \eta_2, \ldots, \eta_n$ be linearly independent. Show that if $\xi_1 = \eta_1/\|\eta_1\|$ and

$$\xi_{k+1} = \frac{\eta_{k+1} - \sum_{j=1}^{k} \langle \eta_{k+1}, \xi_j\rangle\, \xi_j}{\|\eta_{k+1} - \sum_{j=1}^{k} \langle \eta_{k+1}, \xi_j\rangle\, \xi_j\|} \tag{4.15}$$

then $\xi_1, \ldots, \xi_n$ is orthonormal and

$$\operatorname{span}\{\xi_1, \ldots, \xi_n\} = \operatorname{span}\{\eta_1, \ldots, \eta_n\}.$$

Show that the same algorithm works if the original set of vectors is countably infinite. This process of "orthonormalizing a basis" is called the **Gram–Schmidt Process**.

Answer. We write

$$\xi_{k+1} = c_{k+1}\left(\eta_{k+1} - \sum_{j=1}^{k}\langle \eta_{k+1}, \xi_j\rangle\, \xi_j\right),$$

where c_{k+1} is the constant given by dividing by the norm; note that $c_{k+1} \neq 0$ by the linear independence. By induction, assuming that $\xi_1, \ldots, \xi_k$ are orthonormal, we have for $j = 1, \ldots, k$

$$\frac{1}{c_{k+1}}\,\langle \xi_{k+1}, \xi_j\rangle = \langle \eta_{k+1}, \xi_j\rangle - \langle \eta_{k+1}, \xi_j\rangle = 0.$$

This shows, without loss of generality, that $\langle \xi_k, \xi_j\rangle = 0$ when $j \neq k$. And $\langle \xi_k, \xi_k\rangle = 1$, since the ξ_k are normalized by construction. Also by (4.15), $\eta_{k+1} \in \operatorname{span}\{\xi_1, \ldots, \xi_{k+1}\}$ for all k. As the n linearly independent vectors $\xi_1, \ldots, \xi_n$ span the n-dimensional subspace $\operatorname{span}\{\eta_1, \ldots, \eta_n\}$, we get the equality

$$\operatorname{span}\{\xi_1, \ldots, \xi_n\} = \operatorname{span}\{\eta_1, \ldots, \eta_n\}.$$

This latter argument does not work when we are dealing with infinitely many vectors. In such case we can prove the equality

$$\operatorname{span}\{\xi_1, \ldots, \xi_n\} = \operatorname{span}\{\eta_1, \ldots, \eta_n\}$$

for all n by induction, which shows that they span the same space.

(4.3.4) Let $\mathcal{K} \subset \mathcal{H}$ be a subspace, and let $\xi \in \mathcal{H}$. Prove that if

$$0 \leq \|\nu\|^2 + 2\operatorname{Re}\langle \xi, \nu\rangle, \quad \nu \in \mathcal{K},$$

then $\xi \in \mathcal{K}^\perp$.

Answer. Fix $\nu \in \mathcal{K}$. For any $z \in \mathbb{C}$, we have

$$0 \leq |z|^2\,\|\nu\|^2 + 2\operatorname{Re}\bar{z}\,\langle \xi, \nu\rangle.$$

Choose $\theta \in \mathbb{R}$ so that $e^{i\theta}\langle \xi, \eta\rangle = |\langle \xi, \nu\rangle|$. For $t \in \mathbb{R}$ we may use $z = -e^{-i\theta}\,t$ to get

$$0 \leq t^2\,\|\nu\|^2 - 2t\,|\langle \xi, \nu\rangle|.$$

We can rewrite this inequality, for $t > 0$, as

$$2|\langle \xi, \nu\rangle| \leq t\,\|\nu\|^2.$$

As t can be chosen arbitrarily, $\langle \xi, \nu \rangle = 0$. And since this can be done for any $\nu \in \mathcal{K}$, we have shown that $\xi \in \mathcal{K}^\perp$.

(4.3.5) Let J be a set and $c : J \to \mathbb{C}$ a function. If $\sum_j |c(j)|^2 < \infty$, show that $\{j \in J : \; c(j) \neq 0\}$ is countable.

Answer. We have

$$\{j \in J : \; c(j) \neq 0\} = \bigcup_{n \in \mathbb{N}} \Big\{ j \in J : \; |c(j)| > \frac{1}{n} \Big\}.$$

If J is uncountable, then at least one of the sets in the union is infinite, say $J_m = \{j : \; |c(j)| > \frac{1}{m}\}$. Then

$$\sum_j |c(j)|^2 \geq \sum_{j \in J_m} \frac{1}{m} = \infty.$$

(4.3.6) Show that any orthonormal basis of $\mathbb{C}^2$ is of the form shown in (4.11).

Answer. Let $\{x, y\} \subset \mathbb{C}^2$ be an orthonormal basis. Then

$$|x_1|^2 + |x_2|^2 = 1, \qquad |y_1|^2 + |y_2|^2 = 1, \qquad x_1\overline{y_1} + x_2\overline{y_2} = 0.$$

From the first two equalities we get that

$$x_1 = e^{ia}\cos t, \; x_2 = e^{ib}\sin t, \qquad y_1 = e^{ic}\sin s, \; y_2 = e^{id}\cos s,$$

with $a, b, c, d, r, s \in \mathbb{R}$. By absorbing the sign of the sine and/or cosine into the exponentials, we can modify these coefficients so that we can assume that $s, t \in [0, \pi/2]$; indeed, $\cos t$ and $\sin t$ are two real numbers with $\cos^2 t + \sin^2 t = 1$, so with some combination of signs they can both be made positive and there exists $t' \in [0, \pi/2]$ with $|\cos t| = |\cos t'|$ and $|\sin t| = |\sin t'|$, and the sign can be absorbed by the exponential like if $\cos t' = -\cos t$ then $e^{ia}\cos t = e^{i(a+\pi)}\cos t'$. Now the orthogonality equality becomes

$$e^{i(a-c)}\cos t \sin s + e^{i(b-d)}\sin t \cos s = 0.$$

If $\sin s = 0$ the $\cos s \neq 0$, which forces $\sin t = 0$, and $s = t$ since both are in $[0, \pi/2]$. We similarly get $s = t$ if $\cos s = 0$. In both these cases we are free to choose two of a, b, c, d, so the relationship certainly works. When $\sin s \cos s \neq 0$, we get $|\tan t| = |\tan s|$, but $t, s \in [0, \pi/2)$ and then $t = s$. Now the equation becomes

$$(e^{i(a-c)} + e^{i(b-d)}) \sin t \cos t = 0, .$$

so $e^{i(a-c)} = -e^{i(b-d)} = e^{i(b-d+\pi)}$. Then $a-c = b-d+\pi$ (possibly plus $2k\pi$, but again we can re-adjust a, say, to absorb this) and thus $d = \pi - a + b + c$.

(4.3.7) For each of the Hilbert spaces considered in Examples 4.3.17, find examples of orthonormal bases other than the canonical ones.

Answer. The possibilities are endless.

In $\mathbb{C}^n$, given the canonical basis $\{e_1, \ldots, e_n\}$, we can form an orthonormal basis by doing

$$f_1 = \frac{1}{\sqrt{2}}(e_1 + e_2),\ f_2 = \frac{1}{\sqrt{2}}(e_1 - e_2), \ldots,$$

$$f_{2k-1} = \frac{1}{\sqrt{2}}(e_{2k-1} + e_{2k}),\ f_{2k} = \frac{1}{\sqrt{2}}(e_{2k-1} - e_{2k}), \cdots$$

(how to finish depends on whether n is even or odd). And the same example works in $\ell^2(\mathbb{N})$ and $L^2(\mathbb{T})$. Orthonormality is easily checked, and since each e_k is in the span of these new elements (like, for example, $e_1 = \frac{1}{\sqrt{2}}(f_1 + f_2)$, and $e_2 = \frac{1}{\sqrt{2}}(f_1 - f_2)$), the new orthonormal system is total.

In general, we can take U to be a unitary and apply it to any orthonormal basis to obtain another orthonormal basis.

An easy way to perturb orthonormal bases while still getting orthonormal bases is to play with phases. For instance, if $\{\xi_j\} \subset \mathcal{H}$ is an orthonormal basis and we choose numbers $\{\theta_j\} \subset \mathbb{R}$, then $\{e^{i\theta_j}\xi_j\}$ is an orthonormal basis.

In $L^2(\mathbb{T})$ we can do any of the above tricks. For instance we can let $g_0 = 1$, $g_k = \sqrt{2}\,\mathrm{Re}\, z^k$, $k \in \mathbb{N}$, and $g_k = \sqrt{2}\,\mathrm{Im}\, z^k$, $k \in \mathbb{N}$.

(4.3.8) Prove that the inner product in a Hilbert space is jointly continuous in its two variables.

Answer. If $\xi_n \to \xi$ and $\eta_n \to \eta$, then

$$|\,\|\eta_n\| - \|\eta\|\,| \leq \|\eta_n - \eta\| \to 0.$$

So $\|\eta_n\| \to \|\eta\|$, implying in particular that there exists $r > 0$ with $\|\eta_n\| \leq r$ for all n. Then

$$\begin{aligned}
|\langle \xi_n, \eta_n\rangle - \langle \xi, \eta\rangle| &= |\langle \xi_n - \xi, \eta_n\rangle + \langle \xi, \eta_n - \eta\rangle| \\
&\leq |\langle \xi_n - \xi, \eta_n\rangle| + |\langle \xi, \eta_n - \eta\rangle| \\
&\leq \|\xi_n - \xi\|\,\|\eta_n\| + \|\xi\|\,\|\eta_n - \eta\| \\
&\leq r\,\|\xi_n - \xi\| + \|\xi\|\,\|\eta_n - \eta\|,
\end{aligned}$$

and the continuity follows.

(4.3.9) Let $\mathcal{K}_1, \mathcal{K}_2 \subset \mathcal{H}$ be subspaces. Show that

$$(\mathcal{K}_1 + \mathcal{K}_2)^\perp = \mathcal{K}_1^\perp \cap \mathcal{K}_2^\perp.$$

Show also that the equality is not true for arbitrary subsets.

Answer. Since $\mathcal{K}_1 \subset \mathcal{K}_1 + \mathcal{K}_2$, we immediately have $(\mathcal{K}_1 + \mathcal{K}_2)^\perp \subset \mathcal{K}_1^\perp$; as we can do the same for $\mathcal{K}_2$, we have shown that $(\mathcal{K}_1 + \mathcal{K}_2)^\perp \subset \mathcal{K}_1^\perp \cap \mathcal{K}_2^\perp$. Conversely, if $\nu \in \mathcal{K}_1^\perp \cap \mathcal{K}_2^\perp$ and $\xi \in \mathcal{K}_1$, $\eta \in \mathcal{K}_2$, then

$$\langle \nu, \xi + \eta \rangle = \langle \nu, \xi \rangle + \langle \nu, \eta \rangle = 0 + 0 = 0,$$

and so $\mathcal{K}_1^\perp \cap \mathcal{K}_2^\perp \subset (\mathcal{K}_1 + \mathcal{K}_2)^\perp$.

When $\mathcal{K}_1, \mathcal{K}_2$ are not subspaces, we can fix nonzero $\xi \in \mathcal{H}$ and take $\mathcal{K}_1 = \{\xi\}$, $\mathcal{K}_2 = \{-\xi\}$. Then

$$(\mathcal{K}_1 + \mathcal{K}_2)^\perp = \{0\}^\perp = \mathcal{H}, \qquad \text{while } \mathcal{K}_1^\perp \cap \mathcal{K}_2^\perp = \{\xi\}^\perp.$$

As long as $\dim \mathcal{H} \geq 1$, $\{\xi\}^\perp \neq \mathcal{H}$.

(4.3.10) Let $A \subset \mathcal{H}$. Use Proposition 4.3.8 to obtain directly that $A^{\perp\perp} = \overline{\operatorname{span}} A$.

Answer. We will apply Proposition 4.3.8 to $\mathcal{K} = \overline{\operatorname{span}} A$. We have $\mathcal{K}^\perp = A^\perp$; indeed, if $\xi \in \mathcal{K}^\perp$, then $\xi \in A^\perp$ (since $A \subset \mathcal{K}$). Conversely, if $\xi \in A^\perp$, then for any $\eta_1, \ldots, \eta_m \in A$ and $a_1, \ldots, a_m \in \mathbb{C}$,

$$\langle \xi, \sum_j a_j \eta_j \rangle = \sum_j \overline{a_j} \langle \xi, \eta_j \rangle = 0.$$

Hence $\xi \in (\overline{\operatorname{span}} A)^\perp = \mathcal{K}^\perp$. So it is enough to show that $\mathcal{K}^{\perp\perp} = \mathcal{K}$.

By Proposition 4.3.8,

$$P_{\mathcal{K}^{\perp\perp}} = I - P_{\mathcal{K}^\perp} = I - (I - P_\mathcal{K}) = P_\mathcal{K}.$$

This shows that if $\xi \in \mathcal{K}^{\perp\perp}$ then $\xi = P_{\mathcal{K}^{\perp\perp}} \xi = P_\mathcal{K} \xi \in \mathcal{K}$. Hence $\mathcal{K} \subset \mathcal{K}^{\perp\perp} \subset \mathcal{K}$ and the equality follows.

(4.3.11) Let $A \subset \mathcal{H}$. Show that $(\overline{\operatorname{span}} A)^\perp = A^\perp$.

Answer. Since $A \subset \overline{\operatorname{span}} A$, we get that $(\overline{\operatorname{span}} A)^\perp \subset A^\perp$. Now, if $\xi \in A^\perp$, then for $\sum_{k=1}^m c_k a_k \in \operatorname{span} A$,

$$Big\langle \xi, \sum_{k=1}^m c_k a_k \Big\rangle = \sum_{k=1}^m c_k \langle \xi, a_k \rangle = 0.$$

Thus $\xi \in (\operatorname{span} A)^\perp$. And if $\eta = \lim_n \eta_n$, with $\eta_n \operatorname{span} A$, then (since the inner product is continuous, Exercise 4.3.8)

$$\langle \xi, \eta \rangle = \langle \xi, \lim_n \eta_n \rangle = \lim_n \langle \xi, \eta_n \rangle = 0.$$

So $\xi \in (\overline{\operatorname{span}} A)^\perp$. Hence $A^\perp \subset \left(\overline{\operatorname{span}} A\right)^\perp$, giving us the equality.

(4.3.12) Let $\{\mathcal{H}_j\}$ be a family of closed subspaces of $\mathcal{H}$. Show that

$$\left(\bigcup_j \mathcal{H}_j\right)^\perp = \bigcap_j \mathcal{H}_j^\perp, \qquad \left(\bigcap_j \mathcal{H}_j\right)^\perp = \overline{\operatorname{span}} \bigcup_j \mathcal{H}_j^\perp.$$

Answer. Let $\xi \in \left(\bigcup_j \mathcal{H}_j\right)^\perp$ and $\eta \in \mathcal{H}_k$. As $\eta \in \bigcup_j \mathcal{H}_j$, we have $\langle \xi, \eta \rangle = 0$, and so $\xi \in \mathcal{H}_k$. As this can be done for all k, $\xi \in \bigcap_j \mathcal{H}_j^\perp$. Conversely, if $\xi \in \bigcap_j \mathcal{H}_j^\perp$ and $\eta \in \bigcup_j \mathcal{H}_j$, there exists k with $\eta \in \mathcal{H}_k$; as $\xi \in \mathcal{H}_k^\perp$, $\langle \xi, \eta \rangle = 0$ and so $\xi \in \left(\bigcup_j \mathcal{H}_j\right)^\perp$. This establishes the first equality.

The second equality follows from the first one by taking orthogonals and using Exercise 4.3.11.

(4.3.13) Show that if $\mathcal{K} \subset \mathcal{H}$ is a closed subspace, then $\ker P_\mathcal{K} = \mathcal{K}^\perp$.

Answer. By Proposition 4.3.8 we have $\mathcal{H} = \mathcal{K} + \mathcal{K}^\perp$ as a direct sum, and if $\xi = \xi_\mathcal{K} + \xi_\perp$, then $P_\mathcal{K} \xi = \xi_\mathcal{K}$. So if $P_\mathcal{K} \xi = 0$, this means that $\xi = \xi_\perp \in \mathcal{K}^\perp$. Conversely, if $\xi \in \mathcal{K}^\perp$, then $\xi = \xi_\perp$ and thus $P_\mathcal{K} \xi = 0$.

(4.3.14) Let $A \subset \mathcal{H}$. Prove that $A^{\perp\perp}$ is equal to the intersection of all closed subspaces of $\mathcal{H}$ that contain A.

Answer. The intersection of all closed subspaces that contain A is $\overline{\operatorname{span}} A$. As $A^{\perp\perp}$ is closed (any orthogonal is closed), we have $\overline{\operatorname{span}} A \subset A^{\perp\perp}$.

As $A^\perp$ is a closed subspace of $\mathcal{H}$, we have $\mathcal{H} = A^\perp + A^{\perp\perp}$, as a direct sum. This in particular shows that $(A^{\perp\perp})^\perp = A^\perp$. From Exercise 4.3.11

we have $(\overline{\operatorname{span}}A)^\perp = A^\perp$. Taking orthogonal on the equality, we get via Proposition 4.3.6

$$A^{\perp\perp} = (\overline{\operatorname{span}}A)^{\perp\perp} = \overline{\operatorname{span}}A.$$

(4.3.15) Let $\mathcal{K} \subset \mathcal{H}$ be a subspace. Show that $\mathcal{K}$ is dense in $\mathcal{H}$ if and only if $\mathcal{K}^\perp = \{0\}$.

Answer. If $\mathcal{K}^\perp = \{0\}$, then $\overline{\mathcal{K}} = \mathcal{K}^{\perp\perp} = \{0\}^\perp = \mathcal{H}$, so $\mathcal{K}$ is dense. Conversely, if $\overline{\mathcal{K}} = \mathcal{H}$, from $\mathcal{H} = \overline{\mathcal{K}} + \mathcal{K}^\perp$ (Proposition 4.3.8 and Exercise 4.3.11) we get $\mathcal{K}^\perp = \{0\}$.

(4.3.16) Prove the equality (4.14).

Answer. Because the inner product is continuous, we can put limits (and hence series, by the sequilinearity) in and out. So

$$\begin{aligned}\langle \xi, \eta \rangle &= \Big\langle \sum_j \langle \xi, \xi_j \rangle\, \xi_j, \sum_k \langle \eta, \xi_k \rangle\, \xi_k \Big\rangle \\ &= \sum_{k,j} \langle \xi, \xi_j \rangle \overline{\langle \eta, \xi_k \rangle}\, \langle \xi_j, \xi_k \rangle = \sum_j \langle \xi, \xi_j \rangle \overline{\langle \eta, \xi_j \rangle} \\ &= \sum_j \langle \xi, \xi_j \rangle \langle \xi_j, \eta \rangle.\end{aligned}$$

(4.3.17) Let $A \subset \mathcal{H}$ be a set and $f : \mathcal{H} \to \mathcal{H}$ a function such that $f(\mathcal{H}) \subset A$ and $\xi - f(\xi) \in A^\perp$ for all $\xi \in \mathcal{H}$. Show that A is a closed subspace and that $f = P_A$.

Answer. Given any $\xi \in H$, we can write

$$x = f(\xi) + (\xi - f(\xi)) \in A + A^\perp.$$

This shows that $\mathcal{H} = A + A^\perp$.

We will show that $A = A^{\perp\perp}$, which makes it a closed subspace. Indeed, if $\xi \in A^{\perp\perp}$, then writing $\xi = \eta + \nu$ with $\eta \in A$ and $\nu \in A^\perp$, we have

$$0 = \langle \xi, \nu \rangle = \|\nu\|^2 + \langle \eta, \nu \rangle = \|\nu\|^2.$$

So $\nu = 0$ and then $\xi = \eta \in A$. Hence $A^{\perp\perp} \subset A$, showing $A^{\perp\perp} = A$.

Given $\xi \in \mathcal{H}$, as $\xi = f(\xi) + [\xi - f(\xi)]$ with $f(\xi) \in A$ and $\xi - f(\xi) \in A^\perp$, we have that $f(\xi) = P_A\xi$.

(4.3.18) Let $\{\xi_n\}_n \subset \mathcal{H}$ be an orthonormal basis, and put

$$M = \Big\{\eta \in \mathcal{H} : \sum_{n=1}^\infty \Big(1 + \frac{1}{n}\Big)^2 \big|\langle \eta, \xi_n\rangle\big|^2 \leq 1\Big\}.$$

Show that M is bounded, closed, convex, and that it has no element with greatest norm. *(This gives us an example of a bounded, closed, convex subset and an $\mathbb{R}$-valued nonnegative continuous function that does not attain its maximum. Of course, M is not compact).*

Answer. For any $\eta \in M$,

$$\|\eta\|^2 = \sum_n |\langle \eta, \xi_n\rangle|^2 \leq \sum_n \Big(1 + \tfrac{1}{n}\Big)^2 |\langle \eta, \xi_n\rangle|^2 \leq 1,$$

so M is bounded. If $\|\eta\| = 1$, then

$$\sum_n \Big(1 + \tfrac{1}{n}\Big)^2 |\langle \eta, \xi_n\rangle|^2 > \sum_n |\langle \eta, \xi_n\rangle|^2 = \|\eta\|^2 = 1,$$

and then $\eta \notin M$. That is, $\|\eta\| < 1$ for all $\eta \in M$. On the other hand,

$$\eta_m = \frac{1}{\left(1+\frac{1}{m}\right)}\,\xi_m \in M,$$

and $\|\eta_m\| \to 1$. This says that the norm does not attain its maximum in M.

Convexity: if $\xi, \eta \in M$, $t \in [0,1]$, and writing $b_n = 1 + 1/n$ to manage the spacing,

$$\begin{aligned}\sum_n b_n^2 \, |\langle t\xi + (1-t)\eta, \xi_n\rangle|^2 &= \sum_n b_n^2 \, \big[t^2 |\langle \xi, \xi_n\rangle|^2 + (1-t)^2 |\langle \eta, \xi_n\rangle|^2 \\ &\quad + 2\mathrm{Re}\, t(1-t)\langle \xi, \xi_n\rangle\langle \eta, \xi_n\rangle\big] \\ &\leq \sum_n b_n^2 \, \big[t^2 |\langle \xi, \xi_n\rangle|^2 + (1-t)^2 |\langle \eta, \xi_n\rangle|^2 \\ &\quad + 2t(1-t)|\langle \xi, \xi_n\rangle|\,|\langle \eta, \xi_n\rangle|\big] \\ &= t^2 \sum_n b_n^2 \, |\langle \xi, \xi_n\rangle|^2 + (1-t)^2 \sum_n b_n^2 \, |\langle \eta, \xi_n\rangle|^2 \\ &\quad + 2t(1-t) \sum_n b_n^2 \, |\langle \xi, \xi_n\rangle|\,|\langle \eta, \xi_n\rangle| \\ &\leq t^2 + (1-t)^2 \\ &\quad + 2t(1-t)\Big[\sum_n b_n^2 \, |\langle \xi, \xi_n\rangle|^2\Big]^{1/2} \Big[\sum_n b_n^2 \, |\langle \eta, \xi_n\rangle|^2\Big]^{1/2} \\ &\leq t^2 + (1-t)^2 + 2t(1-t) = 1.\end{aligned}$$

It remains to check that M is closed. If $\eta_m \in M$ for all m and $\eta_m \to \eta$, then

$$\sum_{n=1}^k \left(1 + \tfrac{1}{n}\right)^2 |\langle \eta, \xi_n\rangle|^2 = \lim_m \sum_{n=1}^k \left(1 + \tfrac{1}{n}\right)^2 |\langle \eta_m, \xi_n\rangle|^2 \leq 1,$$

for all k. Then, with $k \to \infty$, we obtain that $\eta \in M$.

(4.3.19) *(While nothing is wrong with this exercise, the examples known to the author are far from trivial, so not being able to do this exercise does not show any lack of expertise)* Show an example of a non-separable Hilbert space $\mathcal{H}$ with a dense subspace K such that K contains no orthonormal basis.

Answer. Let $\mathcal{H}_1 = L^2[0,1]$. Consider the set $\{g_t : \ t \in (0,1)$, where $g_t = 1_{(0,t)}$. This set is uncountable, and linearly independent. Let $\mathcal{H}_2$ be a Hilbert space with orthonormal basis $\{\eta_t\}_{t\in(0,1)}$; we can achieve this by taking $\mathcal{H}_2 = \ell^2(0,1)$, note the little ℓ. Let $K = \mathrm{span}\{(g_t, \eta_t) : \ t \in (0,1)\} \subset \mathcal{H}_1 \oplus \mathcal{H}_2$, and let $\mathcal{H} = \overline{K}$. Then $\mathcal{H}$ is a Hilbert space with dense subspace K. We claim that K cannot contain an orthonormal basis for $\mathcal{H}$. Let $\{x_a : \ a \in A\}$ be an orthonormal set in K. By definition, each x_a can be written as

$$x_a = \sum_{j=1}^{n_a} c_j\,(g_{t_j}, \eta_{t_j}).$$

Let $\{h_n : \ n \in \mathbb{N}\}$ be an orthonormal basis of $\mathcal{H}_1$. An element of a Hilbert space can have only countably many nonzero coefficients in a given orthonormal basis (Exercise 4.3.5). So for each n there are at most countably many $a \in A$ such that $\langle (h_n, 0), x_a\rangle \neq 0$. Note that $\sum_j c_j g_{t_j} \neq 0$, since the g_{t_j} are linearly independent and $c \neq 0$ since $x_a \neq 0$. Then

$$\Big\langle \Big(\sum_j c_j g_{t_j}, 0\Big), x_a \Big\rangle = \Big\| \sum_j c_j g_{t_j} \Big\|^2 > 0$$

for all $a \in A$, which says that A is countable (because $\{x_a\}$ is orthonormal). So there is no uncountable orthonormal set in K, which precludes K from containing an orthonormal basis for $\mathcal{H}$.

(4.3.20) Let $\mathcal{H}$ be a Hilbert space. Since $\mathbb{R} \subset \mathbb{C}$, we can see $\mathcal{H}$ as a real vector space, that we denote $\mathcal{H}_\mathbb{R}$.

(i) Show that $\mathcal{H}_\mathbb{R}$ is a real Hilbert space with the inner product $\langle \xi, \eta\rangle = \mathrm{Re}\,\langle \xi, \eta\rangle$.

(ii) Show that if $\{\xi_j\}$ is an orthonormal basis for $\mathcal{H}$, then $\{\xi_j\} \cup \{i\xi_j\}$ is an orthonormal basis for $\mathcal{H}_\mathbb{R}$.

(iii) We say that $\xi, \eta \in \mathcal{H}$ are **real orthogonal**, denoted $\xi \perp_\mathbb{R} \eta$, if $\mathrm{Re}\,\langle \xi, \eta\rangle = 0$. Show that $\xi \perp \eta$ if and only if $\xi \perp_\mathbb{R} \eta$ and $i\xi \perp_\mathbb{R} \eta$.

(iv) Show that if $H \subset \mathcal{H}$ is a real subspace, then $(iH)^\perp_\mathbb{R} = iH^\perp_\mathbb{R}$.

Answer.

(i) Since $\langle \xi, \xi\rangle \geq 0$, if $\mathrm{Re}\,\langle \xi, \xi\rangle = 0$ then $\xi = 0$. The real bilinearity follows directly from the sesquilinearity, so we do get a real inner product.

(ii) For any $\xi \in \mathcal{H}$ we have $\mathrm{Re}\,\langle \xi, i\xi\rangle = -\mathrm{Re}\, i\|\xi\|^2 = 0$. So $\mathrm{Re}\,\langle \xi_j, i\xi_k\rangle = 0$ for all k, j. If ξ is real orthogonal to all of $\{\xi_j\}$ and $\{i\xi_j\}$, then for each j we have

$$\begin{aligned} 0 &= \mathrm{Re}\,\langle \xi, \xi_j\rangle + i\mathrm{Re}\,\langle \xi, i\xi_j\rangle = \mathrm{Re}\,\langle \xi, \xi_j\rangle + i\mathrm{Re}\,(-i)\langle \xi, \xi_j\rangle \\ &= \mathrm{Re}\,\langle \xi, \xi_j\rangle + i\mathrm{Im}\,\langle \xi, i\xi_j\rangle = \langle \xi, \xi_j\rangle. \end{aligned}$$

So $\xi \in \{\xi_j\}^\perp = \{0\}$ and hence $\{\xi_j\} \cup \{i\xi_j\}$ is an orthonormal basis for $\mathcal{H}_\mathbb{R}$.

(iii) If $\langle \xi, \eta \rangle = 0$, then $\operatorname{Re}\langle \xi, \eta \rangle = 0$ and

$$\operatorname{Re}\langle i\xi, \eta \rangle = -\operatorname{Im}\langle \xi, \eta \rangle = 0.$$

Conversely, if both $\operatorname{Re}\langle \xi, \eta \rangle = 0$ and $\operatorname{Re}\langle i\xi, \eta \rangle = 0$, then $\langle \xi, \eta \rangle = 0$ as in the previous paragraph.

(iv) We have

$$\begin{aligned}(iH)^{\perp}_{\mathbb{R}} &= \{\xi \in \mathcal{H} : \ \operatorname{Re}\langle \xi, i\eta \rangle = 0, \ \eta \in H\} \\ &= \{\xi \in \mathcal{H} : \ \operatorname{Re}\langle -i\xi, \eta \rangle = 0, \ \eta \in H\} \\ &= \{i\xi \in \mathcal{H} : \ \operatorname{Re}\langle \xi, \eta \rangle = 0, \ \eta \in H\} = iH^{\perp}_{\mathbb{R}}.\end{aligned}$$

4.4. Dimension

(4.4.1) Write a complete proof of Corollary 4.4.5, without using Theorem 4.4.4. *(Hint: show that the linear operator induced by mapping one orthonormal basis to another is an isomorphism)*

Answer. Suppose that $\pi : \mathcal{H}_1 \to \mathcal{H}_2$ is an isomorphism. Let $\{\xi_j\}$ be an orthonormal basis of $\mathcal{H}_1$. We have

$$\langle \pi(\xi_j), \pi(\xi_k) \rangle_2 = \langle \xi_j, \xi_k \rangle_1 = \delta_{k,j}$$

so the set $\{\pi(\xi_j)\}_j \subset \mathcal{H}_2$ is orthonormal. If $\nu \perp \pi(\xi_j)$ for all j, use that π is surjective to get $\eta \in \mathcal{H}_1$ with $\pi(\eta) = \nu$. Then, for all j,

$$0 = \langle \nu, \pi(\xi_j) \rangle_2 = \langle \pi(\eta), \pi(\xi_j) \rangle_2 = \langle \eta, \xi_j \rangle_1.$$

It follows that $\eta = 0$, so $\nu = 0$, and $\{\pi(\xi_j\}_j$ is total, and thus an orthonormal basis. Such basis obviously has the same cardinality as $\{\xi_j\}$, so $\dim \mathcal{H}_2 = \dim \mathcal{H}_1$.

Conversely, if $\dim \mathcal{H}_1 = \dim \mathcal{H}_2$, let $\{\xi_j\}_{j \in J}$ and $\{\nu_j\}_{j \in J}$ be orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Define $\pi : \mathcal{H}_1 \to \mathcal{H}_2$ by

$$\pi\Big(\sum_j c_j \xi_j\Big) = \sum_j c_j \nu_j.$$

This is well-defined, by the uniqueness of the representation of a vector in an orthonormal basis. Clearly

$$\|\pi(\xi)\| = \sum_j |\langle \xi, \xi_j\rangle|^2 = \|\xi\|,$$

so π is an isometry. And π is surjective: given any $\sum_j c_j\nu_j \in \mathcal{H}_2$, we can write it as $\pi(\sum_j c_j\xi_j)$. So $\mathcal{H}_1 \simeq \mathcal{H}_2$.

(4.4.2) Show that a finite-dimensional Hilbert space is separable.

Answer. Let $\{\xi_1, \dots, \xi_m\}$ be an orthonormal basis for $\mathcal{H}$. Let

$$Q = \Big\{ \sum_{j=1}^m (c_j + id_j)\,\xi_j : \ c_j, d_j \in \mathbb{Q} \Big\}.$$

This set is countable, as it can be written as

$$Q = \bigcup_{c_1,d_1,\dots,c_m,d_m \in \mathbb{Q}} \Big\{ \sum_{j=1}^m (c_j + id_j)\,\xi_j \Big\}.$$

And it is dense: given $\xi = \sum_{j=1}^m (a_j + ib_j)\,\xi_j \in \mathcal{H}$, and $\varepsilon > 0$, choose rational numbers $c_1, d_1, \dots, c_m, d_m \in \mathbb{Q}$ with $|a_j - c_j| < \varepsilon/\sqrt{2m}$, $|b_j - d_j| < \varepsilon/\sqrt{2m}$. Then

$$\begin{aligned}\Big\|\xi - \sum_{j=1}^m (c_j + id_j)\,\xi_j\Big\| &= \Big\| \sum_{j=1}^m [(a_j - c_j) + i(b_j - d_j)]\,\xi \Big\| \\ &= \Big(\sum_{j=1}^m (a_j - c_j)^2 + (b_j - d_j)^2 \Big)^{1/2} < \varepsilon.\end{aligned}$$

(4.4.3) Prove Proposition 4.4.6.

Answer. If $\mathcal{H}$ has countable dimension, let $\{\xi_n\}_{n\in\mathbb{N}}$ be an orthonormal basis. Now form the set

$$X = \left\{ \sum_n (a_n + ib_n)\xi_n : \ a_n, b_n \in \mathbb{Q} \right\}.$$

Given $\varepsilon > 0$ and $\xi \in \mathcal{H}$, write $\xi = \sum_n c_n\xi_n$, and choose $a_n, b_n \in \mathbb{Q}$ such that $|c_n - (a_n + ib_n)| < \varepsilon/2^n$. Then

$$\Big\|\xi - \sum_n (a_n + ib_n)\xi_n\Big\|^2 = \sum_n |c_n - (a_n + ib_n)|^2 < \varepsilon^2 \sum_n 2^{-2n} < \varepsilon^2,$$

and X is dense.

Conversely, suppose that $\dim \mathcal{H} = |J|$ with $|J| > |\mathbb{N}|$. Let $\{\xi_j\}_{j\in J}$ be an orthonormal basis for $\mathcal{H}$. Since $\|\xi_j - \xi_k\| = \sqrt{2}$ for all $j \neq k$, the uncountable family of open balls $\{B_{\sqrt{2}/2}(\xi_j)\}_{j\in J}$ is disjoint. So $\mathcal{H}$ is not separable.

(4.4.4) Show that the unit ball of a finite-dimensional Hilbert space is compact.

Answer. Since $\mathcal{H}$ is a metric space, it is enough to show that any sequence in the unit ball admits a convergent subsequence. Let $\{\eta_n\}_n$ be a sequence with $\|\eta_n\| \leq 1$ for all n. Fix an orthonormal basis $\{\xi_1, \dots, \xi_m\}$.

The sequence $\{\langle \eta_n, \xi_1\rangle\}_n$ is inside the closed unit disk $\overline{\mathbb{D}} \subset \mathbb{C}$; so it admits a convergent subsequence $\{\langle \eta_{n_k}, \xi_1\rangle\}_k$. Now consider $\{\langle \eta_{n_k}, \xi_2\rangle\}_k$; again, this sequence admits a convergent subsequence. Repeating the procedure with all ξ_j, we obtain a subsequence $\{\eta_{n_r}\}$ such that each of the sequences $\{\langle \eta_{n_r}, \xi_j\rangle\}_r$ is convergent, $j = 1, \dots, m$. Let $c_j = \lim_r \langle \eta_{n_r}, \xi_j\rangle$, and define $\eta = \sum_j c_j\xi_j$. Then

$$\|\eta - \eta_{n_r}\|^2 = \sum_{j=1}^m |c_j - \langle \eta_{n_r}, \xi_j\rangle|^2 \xrightarrow[r\to\infty]{} 0.$$

4.5. The Riesz Representation Theorem

(4.5.1) Let φ be a bounded functional on $\mathcal{H}$. Fix an orthonormal basis $\{\xi_j\}$. Given any $\xi = \sum_j c_j\xi_j \in \mathcal{H}$, show that $\varphi(\xi) = \sum_j c_j\varphi(\xi_j)$ and use this fact to get an alternative proof of the Riesz Representation Theorem (4.5.4). Explain where the continuity of φ was used.

Answer. Since φ is continuous,

$$\varphi\Big(\sum_j c_j\xi_j\Big) = \varphi\Big(\lim_F \sum_{j\in F} c_j\xi_j\Big) = \lim_F \varphi\Big(\sum_{j\in F} c_j\xi_j\Big)$$
$$= \lim_F \sum_{j\in F} c_j\varphi(\xi_j) = \sum_j c_j\varphi(\xi_j).$$

As for the Riesz Representation Theorem, it is clear that we need

$$\eta = \sum_j \overline{\varphi(\xi_j)}\,\xi_j.$$

The only thing to prove, then, is that such an element exists in $\mathcal{H}$. For any finite subset $F \subset J$,

$$\sum_{j\in F} |\varphi(\xi_j)|^2 = \varphi\Big(\sum_{j\in F} \overline{\varphi(\xi_j)}\,\xi_j\Big) \le \|\varphi\|\,\Big(\sum_{j\in F} |\varphi(\xi_j)|^2\Big)^{1/2}.$$

From this we get that

$$\sum_{j\in F} |\varphi(\xi_j)|^2 \le \|\varphi\|^2.$$

As this can be done for any F, we conclude that $\{\varphi(\xi_j)\}_{j\in J} \in \ell^2(J)$, and thus $\eta = \sum_j \overline{\varphi(\xi_j)}\,\xi_j \in \mathcal{H}$ exists by Lemma 4.3.18. The continuity of φ was used to evaluate φ on the series and evaluate term by term; we first use the continuity to exchange φ with the limit, and then apply linearity.

We now offer a second way to prove the Riesz Representation Theorem. Note that, for any net $\{a_j\}_{j\in J}$, we have

$$\|a\|_2 = \sup\Big\{\Big|\sum_{j\in J} c_j a_j\Big| :\ c \in \ell^2(J),\ \|c\|_2 = 1\Big\}, \tag{AB.4.1}$$

even if we allow for $\|a\|_2 = \infty$. Indeed, we have, by Cauchy–Schwarz (doing it first for finite sums),

$$\Big|\sum_{j\in J} c_j a_j\Big| \le \|c\|_2\,\|a\|_2 = \|a\|_2.$$

If $\|a\|_2 < \infty$, take $c = a/\|a\|_2$ to get equality in (AB.4.1). And if $\|a\|_2 = \infty$, for any finite $F \subset J$ we can take $c = \overline{a|_F}/\|a|_F\|_2$, so

$$\sum_{j\in J} c_j a_j = \sum_{j\in F} |a_j|^2/\|a|_F\|_2 = \|a|_F\|_2.$$

We get that the supremum is infinite by taking larger and larger sets F.

Going back to our first equality, for any $c \in \ell^2(J)$ with $\|c\|_2 = 1$,

$$\Big|\sum_j c_j\varphi(\xi_j)\Big| = \Big|\varphi\Big(\sum_j c_j\xi_j\Big)\Big| \le \|\varphi\|\,\Big\|\sum_j c_j\xi_j\Big\| = \|\varphi\|\,\|c\|_2 = \|\varphi\|.$$

Thus $\{\varphi(\xi_j)\}_j \in \ell^2(J)$. Now let $\eta = \sum_j \varphi(\xi_j)\xi_j \in h$, and we get that $\varphi(\xi) = \langle \xi, \eta\rangle$ for all $\xi \in \mathcal{H}$.

(4.5.2) The vector space of all bounded functionals on $\mathcal{H}$ is called its **dual**, denoted by $\mathcal{H}^*$. Show that

$$\|\varphi\| = \sup\{|\varphi(\xi)| : \ \|\xi\| = 1\}$$

defines a norm on $\mathcal{H}^*$, and that if $\eta \in \mathcal{H}$ is the vector corresponding to φ via the Riesz Representation Theorem, then $\|\varphi\| = \|\eta\|$.

Answer. If $\|\varphi\| = 0$, then $|\varphi(\xi)| = 0$ for all ξ, so $\varphi = 0$. We have

$$\begin{aligned}\|\lambda\varphi\| = \sup\{|\lambda\varphi(\xi)| : \ \|\xi\| = 1\} &= \sup\{|\lambda|\,|\varphi(\xi)| : \ \|\xi\| = 1\} \\ &= |\lambda| \sup\{|\varphi(\xi)| : \ \|\xi\| = 1\} \\ &= |\lambda|\,\|\varphi\|.\end{aligned}$$

And

$$\begin{aligned}\|\varphi + \psi\| &= \sup\{|\varphi(\xi) + \psi(\xi)| : \ \|\xi\| = 1\} \\ &\leq \sup\{|\varphi(\xi)| + |\psi(\xi)| : \ \|\xi\| = 1\} \\ &\leq \sup\{|\varphi(\xi)| : \ \|\xi\| = 1\} + \sup\{|\psi(\xi)| : \ \|\xi\| = 1\} \\ &= \|\varphi\| + \|\psi\|.\end{aligned}$$

Since $\varphi(\xi) = \langle \xi, \eta \rangle$ for all ξ, we have

$$|\varphi(\xi)| = |\langle \xi, \eta \rangle| \leq \|\xi\|\,\|\eta\|,$$

so $\|\varphi\| \leq \|\eta\|$. Also,

$$\varphi(\eta/\|\eta\|) = \frac{\langle \eta, \eta \rangle}{\|\eta\|} = \|\eta\|.$$

So $\|\varphi\| = \|\eta\|$.

(4.5.3) Let $\varphi : \mathcal{H} \to \mathbb{C}$ be nonzero, linear, and bounded. Show that

$$\dim(\ker\varphi)^\perp = 1.$$

Answer. Since $\varphi \neq 0$ we have $\ker\varphi \subsetneq \mathcal{H}$, so $(\ker\varphi)^\perp \neq \{0\}$. Let $\eta_1 \in (\ker\varphi)^\perp$ be nonzero, and let $\eta = \eta_1/\varphi(\eta_1)$; then $\varphi(\eta) = 1$. For any $\xi \in (\ker\varphi)^\perp$ we have that $\xi - \varphi(\xi)\eta \in \ker\varphi$; but since $\xi, \eta \in (\ker\varphi)^\perp$, we also have $\xi - \varphi(\xi)\eta \in (\ker\varphi)^\perp$. Thus $\xi - \varphi(\xi)\eta = 0$; that is, $\xi = \varphi(\xi)\eta \in \mathbb{C}\eta$. So $(\ker\varphi)^\perp = \mathbb{C}\eta$, one-dimensional.

(4.5.4) Let $\varphi : \mathcal{H} \to \mathbb{C}$ be linear. Prove that the following two statements are equivalent:

(i) φ is bounded;

(ii) $\ker\varphi$ is closed.

Use the ideas in your proof to show that if φ is unbounded, then $\ker\varphi$ is dense in $\mathcal{H}$.

Answer. If φ is bounded, then it is continuous. Thus $\ker\varphi = \varphi^{-1}(\{0\})$ is closed, being a continuous pre-image of a closed set.

Conversely, if $\ker\varphi$ is closed, if it is all of $\mathcal{H}$ then $\varphi = 0$; otherwise, we can proceed as in the proof of the Riesz Representation Theorem. Namely, we take a nonzero $\eta_1 \in (\ker\varphi)^\perp$ (assuming that this is nonzero is where we use that $\ker\varphi$ is closed) with $\varphi(\eta_1) = 1$, and put $\eta = \eta_1/\|\eta_1\|^2$. Then for any $\xi \in \mathcal{H}$, $\xi - \varphi(\xi)\eta_1 \in \ker\varphi$, so $\langle \xi - \varphi(\xi)\eta_1, \eta_1\rangle = 0$, which implies $\varphi(\xi) = \langle \xi, \eta\rangle$. Then

$$|\varphi(\xi)| = |\langle \xi, \eta\rangle| \leq \|\xi\|\,\|\eta\|,$$

and φ is bounded.

If φ is unbounded, by the proof above we have that $\ker\varphi$ is not closed. If $\overline{\ker\varphi}$ is a proper subspace, we can still repeat the above argument with $\eta_1 \in (\ker\varphi)^\perp$, and we would have proven that φ is bounded. So $\overline{\ker\varphi} = \mathcal{H}$.

CHAPTER

5 Banach spaces

5.1. Normed spaces

(5.1.1) Prove that in a normed vector space, addition of vectors and multiplication by a scalar are continuous.

Answer. Since we are dealing with metric spaces, sequences are enough for continuity. Suppose that $\mathcal{X}$ is a normed space, $\alpha \in \mathbb{C}$, and that $x_n \to x$, $y_n \to y$ in $\mathcal{X}$. Then

$$\begin{aligned} \|\alpha x + y - (\alpha x_n + y_n)\| &= \|(\alpha x - \alpha x_n) + (y - y_n)\| \\ &\leq |\alpha| \, \|x - x_n\| + \|y - y_n\| \to 0 \end{aligned}$$

Similarly, if $\alpha_n \to \alpha$, there exists $c > 0$ with $|\alpha_n| < c$ for all n, and then

$$\|\alpha_n x_n - \alpha x\| \leq |\alpha_n| \, \|v_x - v\| + |\alpha_n - \alpha| \, \|x\| \leq c \, \|x_n - x\| + |\alpha_n - \alpha| \, \|x\|.$$

(5.1.2) Let $\mathcal{X}$ be a normed space, $x \in \mathcal{X}$, and $\{x_n\}$ a sequence such that $x_n \to x$. Show that $\|x_n\| \to \|x\|$.

Answer. We have, via the reverse triangle inequality,

$$\big|\, \|x_n\| - \|x\| \,\big| \le \|x_n - x\|.$$

Hence $\|x_n\| \to \|x\|$.

(5.1.3) Show that equivalence of norms is an equivalence relation.

Answer. Let p, q, r be norms on $\mathcal{X}$. We have $p(x) \le p(x) \le p(x)$ for all $x \in \mathcal{X}$, so the relation is reflexive. If $p \sim q$, there exist $\alpha, \beta > 0$ with $\alpha p(x) \le q(x) \le \beta p(x)$ for all x. Then $\frac{1}{\beta} q(x) \le p(x) \le \frac{1}{\alpha} q(x)$, so $q \sim p$ and the relation is symmetric.

If $p \sim q$ and $q \sim r$, there exist $\alpha, \beta\gamma, \delta > 0$ with

$$\alpha p(x) \le q(x) \le \beta p(x), \quad \gamma q(x) \le r(x) \le \delta q(x).$$

Then

$$\alpha\gamma p(x) \le r(x) \le \beta\delta p(x),$$

so $p \sim r$ and the relation is transitive.

(5.1.4) Let $\mathcal{X}$ be a normed space and $\{x_n\} \subset \mathcal{X}$ a Cauchy sequence. Show that the sequence is bounded, that is there exists $c > 0$ such that $\|x_n\| \le c$ for all n.

Answer. By Exercise 1.8.27 there exists $x \in X$ and $r > 0$ with $\|x_n - x\| < r$ for all n. Then

$$\|x_n\| \le \|x_n - x\| + \|x\| < r + \|x\|.$$

A second fairly direct argument uses the reverse triangle inequality. Namely, from

$$\big|\, \|x_n\| - \|x_m\| \,\big| \le \|x_n - x_m\|$$

we learn that $\{\|x_n\|\} \subset (0, \infty)$ is a Cauchy sequence. Being a Cauchy sequence in $\mathbb{R}$ it is bounded. This argument is not really distinct from the first one, other than one reduces the problem to showing that a Cauchy sequence is bounded in $\mathbb{R}$; but the proof in $\mathbb{R}$ is likely to be the same as in an arbitrary metric space.

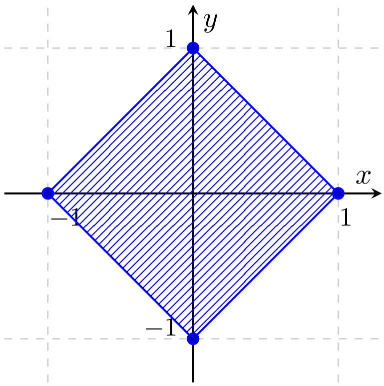
(5.1.5) Consider the real Banach spaces $X_1 = (\mathbb{R}^2, \|\cdot\|_1)$, $X_2 = (\mathbb{R}^2, \|\cdot\|_2)$, $X_3 = (\mathbb{R}^2, \|\cdot\|_\infty)$. Find and describe graphically the unit ball of each space (the use of real spaces is only to allow the possibility of drawing the unit balls).

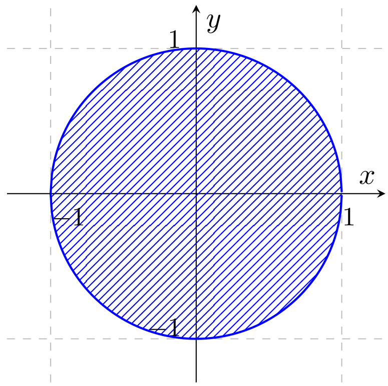
Answer.

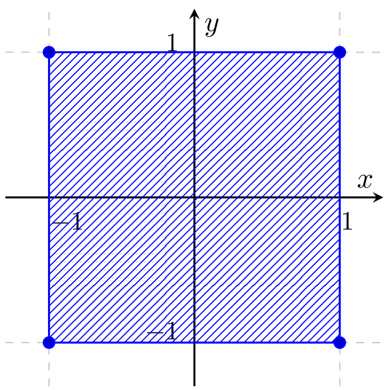
For X_1, the edges of the ball are the lines $|x| + |y| = 1$, so the unit ball is the square with vertices $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-1, -1)$.

For X_2, the unit ball is the usual closed disk of radius one, centered at the origin.

For X_3, the unit ball is the square with vertices $(1, 1)$, $(-1, 1)$, $(-1, -1)$, $(1, -1)$.







(5.1.6) Show that the canonical basis is a Schauder basis for $\ell^p(\mathbb{N})$ when $1 \leq p < \infty$.

Answer. Let $a \in \ell^p(\mathbb{N})$ with $p < \infty$. Given $\varepsilon > 0$, there exists n_0 such that $\sum_{n>n_0} |a_n|^p < \varepsilon^p$. Then for any finite set $F_1 \supset \{1, \ldots, n_0\}$,

$$\Big\| a - \sum_{n \in F_1} a_n e_n \Big\| = \Big\| \sum_{n \in \mathbb{N} \setminus F_1} a_n e_n \Big\|_p = \Big(\sum_{n \in \mathbb{N} \setminus F_1} |a_n|^p \Big)^{1/p}$$
$$\leq \Big(\sum_{n=n_0+1}^{m} |a_n|^p \Big)^{1/p} < \varepsilon.$$

So the sequence of partial sums of the series converges to a. As for the uniqueness, if $\displaystyle\sum_n a_n e_n = 0$, then $\sum_n |a_n|^p = 0$, which implies $a_n = 0$ for all n.

(5.1.7) Show that the canonical basis is not a Schauder basis for $\ell^\infty(\mathbb{N})$.

Answer. For a tail of a series we have

$$\Big\| \sum_{n>m} a_n e_n \Big\|_\infty = \sup\{|a_n| : \ n > m\}.$$

And $\ell^\infty(\mathbb{N})$ contains elements that do not go to zero, like constant functions: if $a_n = 1$ for all n, for instance, the series will not converge.

This exercise follows from the more general Exercise 5.1.8.

(5.1.8) Show that $\ell^\infty(\mathbb{N})$ is not separable and hence it has no Schauder basis.

Answer. Any space with a Schauder basis is separable, so we just need to show that $\ell^\infty(\mathbb{N})$ is not separable. Consider the set $\mathcal{R} = \{1_A : \ A \subset \mathbb{N}\} \subset \ell^\infty(\mathbb{N})$. We know from Proposition 1.6.28 that $\mathcal{P}(\mathbb{N})$ is uncountable, so $\mathcal{R}$ is uncountable. For any distinct $A, B \subset \mathbb{N}$, $\|1_A - 1_B\|_\infty = 1$. If $D \subset \ell^\infty(\mathbb{N})$ is dense, for each ball $B_{1/2}(1_A)$ there exists $d_A \in D$ with $d_A \in B_{1/2}(1_A)$. As the balls are all disjoint we get $d_B \neq d_A$ if $B \neq A$. Hence D is uncountable.

(5.1.9) Show that the canonical basis is a Schauder basis for c_0.

Answer. As in $\ell^\infty(\mathbb{N})$, we have $\Big\| \sum_{n>m} a_n e_n \Big\|_\infty = \sup\{|a_n| : n > m\}$; but now sequences converge to zero, so the tails of the series do too. Concretely, if $a \in c_0$ then

$$\Big\| a - \sum_{n=1}^{m} a_n e_n \Big\| = \Big\| \sum_{n>m} a_n e_n \Big\| = \sup\{|a_n| : n > m\} \xrightarrow[m\to\infty]{} 0.$$

As for the uniqueness of the coefficients, if $\sum_n a_n e_n = 0$, then $0 = \sup\{|a_n| : n\}$, so $a_n = 0$ for all n.

(5.1.10) Let $\mathcal{X}$ be a normed space and $x, z \in \mathcal{X}$, with S_z the reflection of $\mathcal{X}$ on z. That is, $S_z x = -(x - z) + z = 2z - x$. Show that

(i) $S_z^2 = \mathrm{id}_{\mathcal{X}}$;

(ii) S_z is an isometry;

(iii) S_z is affine;

(iv) z is the only fixed point of S_z;

(v) We have the equalities

$$\|z - S_z x\| = \|z - x\|, \qquad \|x - S_z x\| = 2\|x - z\|, \quad x, z \in \mathcal{X}.$$

Answer.

(i) For any $x \in \mathcal{X}$,

$$S_z^2 x = S_z(2z - x) = [2z - (2z - x)] = x.$$

(ii) For $x, y \in \mathcal{X}$,

$$\|S_z x - S_z y\| = \|2z - x - (2z - y)\| = \|x - y\|.$$

(iii) If $x, y \in \mathcal{X}$ and $t \in [0, 1]$,

$$\begin{aligned} S_z(tx + (1-t)y) &= 2z - (tx + (1-t)y) = t(2z - x) + (1-t)(2z - y) \\ &= tS_z x + (1-t)S_z y. \end{aligned}$$

(iv) If $S_z x = x$, this is $2z - x = x$, hence $x = z$.

(v) We have

$$\|z - S_z x\| = \|z - (2z - x)\| = \|x - z\|,$$

and

$$\|x - S_z x\| = \|x - (2z - x)\| = 2\|x - z\|.$$

5.2. Finite-dimensional Banach spaces

(5.2.1) Show that a finite-dimensional normed space $\mathcal{X}$ is separable.

Answer. Let $n = \dim \mathcal{X}$. Using a basis for $\mathcal{X}$ we can construct a linear bijection $\Gamma : \mathcal{X} \to \mathbb{C}^n$. We may consider on $\mathbb{C}^n$ the 2-norm, $\|c\| = \left(\sum_k |c_k|^2\right)^{1/2}$, which makes $\mathbb{C}^n$ a Hilbert space. As all norms are equivalent by Theorem 5.2.2, $\mathcal{X}$ is linearly homeomorphic to $\ell^2(\{1, \ldots, n\})$. The latter is separable by Exercise 4.4.2, and so $\mathcal{X}$ is separable.

(5.2.2) By Theorem 5.2.2, the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ are equivalent on $\mathbb{C}^n$. Find specific constants that realize the inequalities.

Answer. From Cauchy–Schwarz,

$$\|x\|_1 = \sum_k |x_k| \leq \Big(\sum_k |x_k|^2\Big)^{1/2} \Big(\sum_k 1^2\Big)^{1/2} = \sqrt{n}\,\|x\|_2.$$

Conversely,

$$\|x\|_1^2 = \Big(\sum_k |x_k|\Big)^2 \geq \sum_k |x_k|^2,$$

so

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\,\|x\|_2.$$

Both inequalities are sharp, as seen by taking $x = e_1$ for the first one, and $x = \sum_k e_k$ for the second one.

We also have, directly, that

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty.$$

These again are sharp, with the same $x = e_1$ for the first inequality and $x = \sum_k e_k$ for the second one.

Finally,

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\,\|x\|_\infty.$$

These can be seen to be sharp by taking once more the same choices as above.

(5.2.3) Show that $\|\cdot\|_1$, defined in (5.2), is a norm.

Answer. By construction $\|x\|_1 \geq 0$ since the absolute value is non-negative. If $\|x\|_1 = 0$ then $\sum_k |x_k| = 0$ gives us $x_k = 0$ for all k, so $x = 0$. Given $\lambda \in \mathbb{C}$, we have $\|\lambda x\| = \sum_k |\lambda x_k| = |\lambda|\,\|x\|_1$. Finally,

$$\|x+y\|_1 = \sum_k |x_k + y_k| \leq \sum_k |x_k| + |y_k| = \sum_k |x_k| + \sum_k |y_k| = \|x\|_1 + \|y\|_1.$$

(5.2.4) Let $\mathcal{X}$ be a normed vector space of dimension n. Prove that there is a bicontinuous isomorphism between $\mathcal{X}$ and $\mathbb{C}^n$ (considered with the Euclidean norm).

Answer. Fix a basis $\{e_1, \ldots, e_n\}$ for $\mathcal{X}$. For $x \in \mathcal{X}$ we denote by $x_1, \ldots, x_n$ the coefficients of $x = \sum_k x_k e_k$. Then $\gamma(x) = \left(\sum_k |x_k|\right)^{1/2}$ defines a norm on $\mathcal{X}$. By Corollary 5.2.3 there exist $\alpha, \beta > 0$ with

$$\alpha\|x\| \leq \gamma(x) \leq \beta\|x\|. \tag{AB.5.1}$$

Let $\Gamma : \mathcal{X} \to \mathbb{C}^n$ be given by $\Gamma(x) = (x_1, \ldots, x_n)$. Then Γ is linear and bijective from the fact that $\{e_1, \ldots, e_n\}$ is a basis. And we have

$$\|\Gamma(x)\| = \gamma(x) \leq \beta\|x\|,$$

so Γ is continuous. And by (AB.5.1)

$$\|\Gamma^{-1}(x_1, \ldots, x_n)\| \leq \alpha^{-1}\gamma(\Gamma^{-1}(x_1, \ldots, x_n)) = \alpha^{-1}\Big(\sum_k |x_k|^2\Big)^{1/2}.$$

So Γ is a bicontinuous isomorphism.

(5.2.5) Prove Corollary 5.2.4.

Answer. By Corollary 5.2.3 we may choose a norm that suits us. For instance we can fix a basis $f_1, \ldots, f_n$ and choose

$$\Big\|\sum_{j=1}^n x_j f_j\Big\| = \Big(\sum_{j=1}^n |x_j|^2\Big)^2.$$

This norm trivially satisfies the Parallelogram Identity, and hence $\mathcal{X}$ becomes a Hilbert space with this norm. Then the closed bounded sets are compact by Theorem 4.4.8.

The same argument, but without mentioning Hilbert spaces, would look as follows. Fix a basis $f_1, \ldots, f_n$. By Corollary 5.2.3 we can consider the norm

$$\Big\| \sum_j c_j f_j \Big\|_1 = \sum_j |c_j|. \tag{AB.5.2}$$

In this norm, any element of the unit ball will have coefficients in the unit ball of $\mathbb{C}$. By the Heine–Borel Theorem there is a subsequence such that the first coordinate converges. From this subsequence we can then extract a subsequence where the second coordinate converges. After n steps we will obtain a "coordinate wise" convergent subsequence; using the norm in (AB.5.2) we see that the subsequence is norm-convergent. So the unit ball is compact.

A third argument is as follows. Fix a basis $f_1, \ldots, f_n \in \mathcal{X}$ and define $T : \mathcal{X} \to \mathbb{C}^n$ by

$$T\Big(\sum_{j=1}^n x_j f_j\Big) = (x_1, \ldots, x_n).$$

It is easy to check that T is a bijection. By Corollary 5.2.5, T and T^{-1} are bounded. In particular both T and T^{-1} map compact sets to compact sets. We can define a norm on $\mathcal{X}$ by $\|x\| = \|Tx\|_2$. Then $B_1(0)^{\mathcal{X}} = T^{-1}B_1(0)^{\mathbb{C}}$, hence compact. Once we know that the unit ball in $\mathcal{X}$ is compact for some norm, since all norms are equivalent (Theorem 5.2.2) and they generate the same topology (Proposition 5.2.1), it is compact for all norms.

(5.2.6) Let $\mathcal{X}, \mathcal{Y}$ be finite-dimensional vector spaces, and $T : \mathcal{X} \to \mathcal{Y}$ linear. Show that

$$\dim \ker T + \dim \operatorname{ran} T = \dim \mathcal{X}.$$

This equality is usually called the **Rank-Nullity Theorem**.

Answer. Let $n = \dim \mathcal{X}$ and $x_1, \ldots, x_r$ a basis of $\ker T$. Complete it to a basis $x_1, \ldots, x_n$ of $\mathcal{X}$. Let $\mathcal{X}_0 = \operatorname{span}\{x_{r+1}, \ldots, x_n\}$. Then $T|_{\mathcal{X}_0}$ is injective, and $\operatorname{ran} T|_{\mathcal{X}_0} = \operatorname{ran} T$, so $T|_{\mathcal{X}_0}$ is an isomorphism between $\mathcal{X}_0$ and $\operatorname{ran} T$. Then

$$\dim \ker T + \dim \operatorname{ran} T = \dim \ker T + \dim \mathcal{X}_0 = r + (n - t) = n = \dim \mathcal{X}.$$

5.3. Direct Sums and Quotient spaces

(5.3.1) Let $\mathcal{X}$ be a vector space and $M, N \subset \mathcal{X}$ subspaces with $\mathcal{X} = M + N$. Show that the following statements are equivalent:

(i) each element $x \in \mathcal{X}$ can be written as $x = y+z$, with $y \in M$ and $z \in N$, in a unique way;

(ii) $M \cap N = \{0\}$.

Answer. Suppose uniqueness first. If $z \in M \cap N$, then $z = z + 0 = 0 + z$ (first term in M, second term in N) and hence $z = 0$. Conversely, suppose that $M \cap N = \{0\}$ and that $y_1 + z_1 = y_2 + z_2$, with $y_1, y_2 \in M$ and $z_1, z_2 \in N$. Then the element $y_1 - y_2 = z_2 - z_1$ is in $M \cap N = \{0\}$, and so $y_2 = y_1$, $z_2 = z_1$.

(5.3.2) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. Show that the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$—as defined after Definition 5.3.1—are equivalent.

Answer. The estimate for norms in $\mathbb{R}^2$ will serve us here, since

$$\|(x, y)\|_1 = \|(\|x\|, \|y\|)\|_1, \qquad \|(x, y)\|_2 = \|(\|x\|, \|y\|)\|_2,$$

and

$$\|(x, y)\|_\infty = \|(\|x\|, \|y\|)\|_\infty.$$

From this we have

$$\|(x, y)\|_\infty \leq \|(x, y)\|_2 \leq \|(x, y)\|_1 \leq \sqrt{2}\|(x, y)\|_2 \leq 2\|(x, y)\|_\infty.$$

(5.3.3) Let $\mathcal{X}$ be a Banach space, and $M \subset \mathcal{X}$ a subspace. Show that the following statements are equivalent:

(i) M is closed;

(ii) if $\{x_n\} \subset M$ is Cauchy, $x = \lim_n x_n$ exists and $x \in M$.

Answer. Suppose first that M is closed. Let $\{x_n\} \subset M$ be Cauchy. As $\mathcal{X}$ is complete, the sequence is convergent; let $x = \lim_n x_n$. Being a limit of element in M, the point x is in the closure of M; but M is closed, so $x \in M$.

Conversely, suppose that M is not closed. Then $\mathcal{X} \setminus M$ is not open; so there exists $x \in \mathcal{X} \setminus M$ that is a limit point for M; that is, there exists $\{x_n\} \subset M$ with $x_n \to x$ and $x \notin M$.

(5.3.4) Let $\mathcal{X}$ be a normed space (not necessarily complete), $M \subset \mathcal{X}$ a subspace. Show that the following statements are equivalent:

(i) M is closed;

(ii) if $\{x_n\} \subset M$ and $x_n \to x$, then $x \in M$.

Answer. Suppose first that M is closed. If $\{x_n\} \subset M$ and $x_n \to x$, then $x \in \overline{M} = M$.

Conversely, if M is not closed, then $\mathcal{X} \setminus M$ is not open; so there exists $x \in \mathcal{X} \setminus M$ that is a limit point for M; that is, there exists $\{x_n\} \subset M$ with $x_n \to x$ and $x \notin M$.

(5.3.5) Let $\mathcal{X}$ be a Banach space and $P \in \mathcal{B}(\mathcal{X})$ a projection. Show that P has closed range.

Answer. Let $\{Px_n\}$ be a Cauchy sequence. Since $\mathcal{X}$ is Banach, there exists $y \in \mathcal{X}$ with $Px_n \to y$. Then

$$Py = P(\lim_n Px_n) = \lim_n P^2 x_n = \lim_n Px_n = y.$$

So $y = Py \in P\mathcal{X}$.

(5.3.6) Show that, given a family $\{\mathcal{X}_j\}_{j\in J}$ of Banach spaces, each of

$$\Big(\bigoplus_{j\in J} \mathcal{X}_j\Big)_{c_0}, \qquad \Big(\bigoplus_{j\in J} \mathcal{X}_j\Big)_{\ell^p},$$

$1 \leq p \leq \infty$, is a Banach space.

Answer. In each case the vector space operations are defined pointwise; namely, $(g + \alpha h)(j) = g(j) + \alpha h(j)$.

For the c_0 and ℓ^∞ direct sums the norm is the same,

$$\|g\|_\infty = \sup\{\|g(j)\| : \ j \in J\}.$$

That this is a norm is basically the same proof as the case that the usual infinity norm on $\ell^\infty(\mathbb{N})$ is a norm. Namely, $\|g+h\|_\infty \leq \|g\|_\infty + \|h\|_\infty$ by the triangle inequality on each $\mathcal{X}_j$ and the fact that the supremum of a sum is at most the sum of the suprema. That $\|\alpha g\| = |\alpha|\,\|g\|$ follows from the corresponding property on each $\mathcal{X}_j$ and that non-negative scalars can be exchanged with the supremum. For the case $p < \infty$, using Minkowski's Inequality (Corollary 2.8.10) for $\ell^p(J)$,

$$\begin{aligned}\|g+h\|_p &= \Big(\sum_j \|g(j)+h(j)\|^p\Big)^{1/p} \leq \Big(\sum_j \big(\|g(j)\|+\|h(j)\|\big)^p\Big)^{1/p} \\ &\leq \Big(\sum_j \|g(j)\|^p\Big)^{1/p} + \Big(\sum_j \|h(j)\|^p\Big)^{1/p} \\ &= \|g\|_p + \|h\|_p\end{aligned}$$

So all that remains is to show that the spaces are complete. Let $\{g_n\} \subset \Big(\bigoplus_{j\in J} \mathcal{X}_j\Big)_{\ell^\infty}$ be Cauchy. Fix $\varepsilon > 0$; then there exists n_0 such that $\|g_n - g_m\|_\infty < \varepsilon$ whenever $m, n \geq n_0$. For any fixed $j \in J$ we have $\|g_n(j) - g_m(j)\|_\infty \leq \|g_m - g_m\|_\infty$ and so the sequence $\{g_n(j)\} \subset \mathcal{X}_j$ is Cauchy. Thus for each j there exists a limit $g(j) = \lim_n g_n(j)$. Now we need to show that $\|g\|_\infty < \infty$ and that $g_n \to g$. We have, for $n, m \geq n_0$,

$$\begin{aligned}\|g_n(j) - g(j)\| &\leq \|g_n(j) - g_m(j)\| + \|g_m(j) - g(j)\| \\ &< \varepsilon + \|g_m(j) - g(j)\|.\end{aligned}$$

As we are free to choose m and $g_m(j) \to g(j)$, we get $\|g_n(j) - g(j)\| \leq \varepsilon$, and this occurs for all j. Thus $\|g_n - g\|_\infty < \varepsilon$, and this shows that $g_n \to g$. Using n big enough

$$\|g\|_\infty \leq \|g - g_n\|_\infty + \|g_n\|_\infty < \infty$$

(since $\{\|g_n\|\}_n$ is Cauchy and hence bounded). This establishes the completeness of $\Big(\bigoplus_{j\in J} \mathcal{X}_j\Big)_{\ell^\infty}$. For the c_0 sum all the above applies, but now we have that $\lim_j g_n(j) \to 0$ for all n and we want to show the same for g. And this follows from

$$\|g(j)\| \leq \|g(j) - g_n(j)\| + \|g_n(j)\| \leq \|g - g_n\|_\infty + \|g_n(j)\|.$$

Then $\limsup_j \|g(j)\| \leq \|g - g_n\|_\infty$; as we are free to choose n and $g_n \to g$, we get $\limsup_j \|g(j)\| = 0$ and therefore the limit exists and is zero. Thus $g \in \Big(\bigoplus_{j\in J} \mathcal{X}_n\Big)_{c_0}$.

And now the case $p < \infty$. Again we have a Cauchy sequence $\{g_n\} \subset \Big(\bigoplus_{j\in J} \mathcal{X}_n\Big)_{\ell^p}$. We still have the inequality $\|g(j)\| \leq \|g\|_p$, so the existence of the limit g is proven exactly as before. This provides g as a poitnwise limit for the net; we need to show that it is a limit in the p-norm; that is, that $\|g_n - g\|_p \to 0$. We have

$$\|g_n - g\|_p = \Big(\sum_j \|g_n(j) - g(j)\|^p\Big)^{1/p}.$$

Fix $\varepsilon > 0$. Since the sequence is Cauchy there exists n_0 such that $\|g_n - g_m\|_p < \varepsilon$ whenever $m, n \geq n_0$. We will mimic the proof of Theorem 2.8.12. Because $\{g_n\}$ is Cauchy, we inductively choose a subsequence $\{g_{n_k}\}$ such that $\|g_{n_{k+1}} - g_{n_k}\| < 2^{-k}$ for all k.

By Minkowsky's Integral Inequality (2.49),

$$\begin{aligned}\Bigg(\sum_j \Big(\sum_{k=1}^{\infty} \|g_{n_{k+1}}(j) - g_{n_k}(j)\|\Big)^p\Bigg)^{1/p} &\leq \sum_{k=1}^{\infty}\Big(\sum_j \|g_{n_{k+1}}(j) - g_{n_k}(j)\|^p\Big)^{1/p} \\ &= \sum_{k=1}^{\infty} \|g_{n_{k+1}} - g_{n_k}\|_p \leq \sum_{k=1}^{\infty} 2^{-k} = 1.\end{aligned}$$

This implies that $\sum_j \Big(\sum_{k=1}^{\infty} \|g_{n_{k+1}}(j) - g_{n_k}(j)\|\Big)^p < \infty$ for all j, and hence the function

$$g(j) = g_{n_1}(j) + \sum_{k=1}^{\infty} g_{n_{k+1}}(j) - g_{n_k}(j)$$

is defined for each j, since the series converges absolutely. We also have

$$\begin{aligned}\|g\|_p &\leq \|g_{n_1}\|_p + \Big\|\sum_{k=1}^{\infty} g_{n_{k+1}} - g_{n_k}\Big\|_p \\ &\leq \|g_{n_1}\|_p + \sum_{k=1}^{\infty} \|g_{n_{k+1}} - g_{n_k}\|_p < \|g_{n_1}\|_p + 1,\end{aligned}$$

so $g \in \Big(\bigoplus_{j\in J} \mathcal{X}_n\Big)_{\ell^p}$. Since the definition of g telescopes,

$$\|g - g_{n_h}\|_p = \Big\|\sum_{k=h}^{\infty} g_{n_{k+1}} - g_{n_k}\Big\|_p \leq \sum_{k=h}^{\infty} 2^{-k} = 2^{-h+1},$$

and so $g_{n_k} \to g$. As this was a subsequence of a Cauchy sequence, $g_n \to g$; thus the space $\Big(\bigoplus_{j\in J} \mathcal{X}_n\Big)_{\ell^p}$ is complete.

(5.3.7) Show that, in the particular case where $\mathcal{X}_n = \mathcal{X}$ for all n, and $\bigoplus_{n\in\mathbb{N}} \mathcal{X}$ denoting any of the three kind of sums in Exercise 5.3.6,

(i) $\bigoplus_{n\in\mathbb{N}} \mathcal{X} \simeq \mathcal{X} \oplus \bigoplus_{n\in\mathbb{N}} \mathcal{X}$;

(ii) $\bigoplus_{n\in\mathbb{N}} \mathcal{X} \simeq \bigoplus_{n\in\mathbb{N}} \mathcal{X} \oplus \bigoplus_{n\in\mathbb{N}} \mathcal{X}$.

In all cases the isomorphisms are isometric.

Answer.

(i) Let $\Gamma : \mathcal{X} \oplus \bigoplus_{n\in\mathbb{N}} \mathcal{X} \to \bigoplus_{n\in\mathbb{N}} \mathcal{X}$ be given by

$$(\Gamma(x,g))(n) = \begin{cases} x, & n = 1 \\ g(n-1), & n > 1 \end{cases}$$

Equivalently we may write $(\Gamma(x,g))(n) = \delta_{1,n}x + (1-\delta_{1,n})\, g(n-1)$. This is linear, for

$$\begin{aligned}(\Gamma(x+\alpha y, g+\alpha h))(n) &= \delta_{1,n}(x+\alpha y) + (1-\delta_{1,n})\,(g(n-1)+\alpha h(n-1)) \\ &= \delta_{1,n}x + (1-\delta_{1,n})\, g(n-1) \\ &\qquad + \alpha\delta_{1,n}y + (1-\delta_{1,n})\, h(n-1).\end{aligned}$$

For the infinity norm,

$$\|\Gamma(x,g)\|_\infty = \sup\{\|\Gamma(x,g)(n)\| :\ n\} = \|(x,g)\|_\infty.$$

And when $p < \infty$

$$\begin{aligned}\|\Gamma(x,g)\|_p^p &= \sum_n \|\delta_{1,n}x + (1-\delta_{1,n})\, g(n-1)\|^p \\ &= \|x\|^p + \sum_n \|g(n)\|^p = \|(x,g)\|_p^p.\end{aligned}$$

In either case, Γ is isometric. We finish by showing that Γ is surjective; indeed, given $g \in \bigoplus_{n\in\mathbb{N}} \mathcal{X}$, we have $g = \Gamma(g(1), g')$, where $g'(n) = g(n-1)$.

(ii) The idea is now very similar, but we take $\Gamma : \bigoplus_{n\in\mathbb{N}} \mathcal{X} \oplus \bigoplus_{n\in\mathbb{N}} \mathcal{X} \to \bigoplus_{n\in\mathbb{N}} \mathcal{X}$ to be

$$(\Gamma(g,h))(n) = \begin{cases} g(n), & n \ \text{ odd} \\ h(n), & n \ \text{ even} \end{cases}$$

The linearity, isometry, and surjectivity are proven in an analog way to that of the previous case.

(5.3.8) In the setting of Exercise 2.3.27, show that

$$\left(\bigoplus_j L^2(K,\mu_j)\right)_{\ell^2} \simeq L^2(X,\Sigma,\mu),$$

Answer. Given $\tilde f = \{f_j\}_j$ with $f_j \in L^2(K,\mu_j)$ for all j, and such that $\sum_j \|f_j\|^2 < \infty$, let $(V\tilde f) : X \to \mathbb{C}$ be given by $(V\tilde f)(x,j) = f_j(x)$. We have that V is linear and

$$\|V\tilde f\|_2^2 = \int_X |Vf|^2\,d\mu = \sum_j \int_K |Vf(x,j)|^2\,d\mu_j(x)$$

$$= \sum_j \int_K |f_j|^2\,d\mu_j = \sum_j \|f_j\|^2 = \|\tilde f\|_2^2,$$

so V is an isometry. Given $f \in L^2(X)$, let $f_j(x) = f(x,j)$. Then f_j is μ_j-measurable and if $\tilde f = \{f_j\}$ we get $V\tilde f = f$.

(5.3.9) If $\mathcal{H}$ is a Hilbert space and $M \subset \mathcal{H}$ is a closed subspace, show that $\mathcal{H}/M$ has a natural Hilbert space structure that makes it a Hilbert space, and that $\mathcal{H}/M$ can be identified with (i.e. it is naturally isomorphic to) $M^\perp$.

Answer. Let P be the orthogonal projection onto $M^\perp$. We have $\xi + M = P\xi + M$: indeed, $I-P$ is the orthogonal projection onto M (Proposition 4.3.8), so $\xi - P\xi = (I-P)\xi \in M$. We define

$$\langle \xi + M, \eta + M\rangle = \langle P\xi, P\eta\rangle. \tag{AB.5.3}$$

This is well-defined since $\xi_1 - \xi_2 \in M$ implies $P\xi_1 - P\xi_2 = 0$. Sesquilinearity is straighforward. If $\langle \xi + M, \xi + M\rangle = 0$, we get $\|P\xi\| = 0$, so $\xi = (I-P)\xi \in M$, and hence $\xi + M = 0$.

We can define the isomorphism $\pi : \mathcal{H}/M \to M^\perp$ by $\pi : \xi + M \longmapsto P\xi$. This is well-defined because if $P\xi_1 = P\xi_2$ then $P(\xi_1 - \xi_2) = 0$, so $\xi_1 - \xi_2 \in M$; this also shows that π is one-to-one. Linearity is clear. And for any $\xi \in M^\perp$, we have $\xi = P\xi = \pi(\xi)$, so π is onto. Finally, π preserves the inner product by (AB.5.3).

(5.3.10) Prove that when $M \subset \mathcal{X}$ is a closed subspace, the quotient norm is a norm *(Hint: think of it as a distance).*

Answer. As M is a subspace,

$$\|v+M\| = \inf\{\|v+m\|:\ m\in M\} = \inf\{\|v-m\|:\ m\in M\} = \operatorname{dist}(m,M).$$

If $\|v+M\|=0$, this means that $\operatorname{dist}(m,M)=0$. Thus $v\in\overline{M}=M$. For any nonzero λ, and using that M is a subspace,

$$\begin{aligned}\|\lambda(v+M)\| &= \inf\{\|\lambda v+m\|:\ m\in M\}\\ &= |\lambda|\,\inf\{\|v+m/\lambda\|:\ m\in M\} = |\lambda|\,\|v+M\|.\end{aligned}$$

For the triangle inequality, fix $\varepsilon>0$. Given $v,w\in\mathcal{X}$ there exist $m_1,m_2\in M$ such that $\|v+m_1\|<\|v+M\|+\varepsilon$, $\|w+m_2\|<\|w+M\|+\varepsilon$. Then

$$\begin{aligned}\|v+w+M\| &\le \|v+w+m_1+m_2\| \le \|v+m_1\|+\|w+m_2\|\\ &\le \|v+M\|+\|w+M\|+2\varepsilon.\end{aligned}$$

As we can do this for all $\varepsilon>0$, we get $\|v+w+M\|\le\|v+M\|+\|w+M\|$.

(5.3.11) Let $\mathcal{X}=C[0,1]$ and $M=\{f\in\mathcal{X}:\ f(1)=0\}$. Show that $\mathcal{X}/M=\{c+M:\ c\in\mathbb{C}\}$ and $\|f+M\|=|f(1)|$.

Answer. We have that $f\sim g$ if and only if $f(1)=g(1)$. Hence $f+M=f(1)+M$ for every f. That is, we can choose a constant function as the representative for each class; which means that $\mathcal{X}/M=\{c+M:\ c\in\mathbb{C}\}$. As for the norm, we only need to calculate the norm for a constant function, since these are representatives. If $g(1)=0$, then $|c+g(1)|=|c|$, which shows that $\|c+g\|\ge|c|$ for all $g\in M$; with $g=0$, we get $\|c+M\|=|c|$.

(5.3.12) Let $\mathcal{X}=C[0,1]$. Fix $t_1,\ldots,t_n\in[0,1]$. Let $M=\{f\in\mathcal{X}:\ f(t_j)=0,\ j=1,\ldots,n\}$. Show that $\mathcal{X}/M\simeq\mathbb{C}^n$ and

$$\|f+M\|=\max\{|f(t_j)|:\ j=1,\ldots,n\}.$$

Answer. We have that $f\sim g$ if and only if $f(t_j)=g(t_j)$ for all j. Hence the map $f+M\longmapsto(f(t_1),\ldots,f(t_n))$ is well-defined $\mathcal{X}/M\to\mathbb{C}^n$. It is surjective, as we can construct a continuous function with n prescribed values (we can make it piece-wise linear, for instance); it is injective by definition of the equivalence relation. And it is linear, hence a vector space isomorphism.

Continuity is a given since we are dealing with finite-dimensional vector spaces (Theorem 5.2.2). As for the norm, since $g(t_j) = 0$ for all j and any $g \in M$, $\|f + g\| \geq \max\{||f(t_1)|, \ldots, |f(t_n)|\}$; and choosing $g = 0$ gives us the reverse inequality.

(5.3.13) Let $\mathcal{X} = C[0,1]$. Let $M = \left\{f \in \mathcal{X} : \int_0^1 f = 0\right\}$. Show that $\mathcal{X}/M \simeq \mathbb{C}$ and that $\|f + M\| = \left|\int_0^1 f\right|$.

Answer. For any f we have $f \sim c$ if $c = \displaystyle\int_0^1 f$. So we have $\mathbb{C}$ as representatives. That is, we can define $\rho : \mathcal{X}/M \to \mathbb{C}$ by

$$\rho(f + M) = \int_0^1 f.$$

If $f + M = g + M$ this means that $\displaystyle\int_0^1 (f - g) = 0$, and so ρ is well-defined and injective. The linearity follows from the linearity of the integral, and the surjectivity from $\rho(c + M) = c$ for all $c \in \mathbb{C}$.

Alternatively we may notice that $M = \ker \psi$, where $\psi(f) = \displaystyle\int_0^1 f$. Then the isomorphism $\mathcal{X}/M \simeq \mathbb{C}$ is the first isomorphism theorem.

As for the norm, if $c = \displaystyle\int_0^1 f$ then

$$\|f + M\| = \|c + M\| = |c|\, \|1 + M\|.$$

So we focus on showing that $\|1 + M\| = 1$. By definition,

$$\|1 + M\| = \inf\left\{\|g\|_\infty : \int_0^1 g = 1\right\}.$$

From $\displaystyle\int_0^1 1 = 1$ we get that $\|1 + M\| \leq 1$. On the other hand, if $\|g\|_\infty < 1 - \delta$ for $\delta > 0$, then

$$\left|\int_0^1 g\right| \leq \int_0^1 |g| \leq 1 - \delta.$$

This shows that if $\int_0^1 g = 1$, then $\|g\|_\infty \geq 1$. Thus $\|1 + M\| \geq 1$ and so $\|1 + M\| = 1$.

(5.3.14) Let $\mathcal{X} = \ell^\infty(\mathbb{N})$ and $M = c_0$. Show that the norm on $\mathcal{X}/M$ is given by

$$\|a + c_0\| = \limsup_n |a_n|, \qquad a \in \ell^\infty(\mathbb{N}).$$

Answer. Let $s = \limsup_n |a_n|$. Fix $\varepsilon > 0$, and $x \in c_0$. There exists n_0 such that $|x_n| < \varepsilon$ for all $n \geq n_0$. By definition of limsup there exists $n \geq n_0$ such that $|a_n| + \varepsilon > s$. Then

$$s < |a_n| + \varepsilon \leq |a_n + x_n| + |x_n| + \varepsilon \leq \|a + x\|_\infty + 2\varepsilon.$$

As $x \in c_0$ as arbitrary, we get that $s \leq \|a + c_0\| + 2\varepsilon$; and ε was arbitrary too, so $s \leq \|a + c_0\|$.

If we take $x = -a\,1_{\{1,\ldots,n_0\}} \in c_0$, then

$$\|a + x\|_\infty = \sup\{|a_n| : \ n \geq n_0\}.$$

It follows that $\|a + c_0\| \leq \sup\{|a_n| : \ n \geq n_0\}$ for all n_0. That is,

$$\|a + c_0\| \leq \lim_{n_0 \to \infty} \{|a_n| : \ n \geq n_0\} = s.$$

5.4. Locally Convex Spaces

(5.4.1) Let $\mathcal{X}$ be a topological vector space and $M \subset \mathcal{X}$ balanced. Show that if $c \in \mathbb{T}$ then $cM = M$.

Answer. The definition of balanced implies that $cM \subset M$ and $c^{-1}M \subset M$. Then

$$M = c(c^{-1}M) \subset cM,$$

and so $cM = M$.

(5.4.2) Let $\mathcal{X}$ be a topological space. Show that the following statements are equivalent:

(i) singletons are closed;

(ii) finite sets are closed;

(iii) $\mathcal{X}$ is T_1: namely, given distinct $x, y \in \mathcal{X}$, there exists an open set V such that $x \in V$, $y \notin V$.

Answer. If singletons are closed, then $\{x_1, \dots, x_n\} = \bigcup_{j=1}^n \{x_j\}$ is a finite union of closed, so closed. The converse is trivial.

Again assuming that singletons are closed, given $x \neq y$, since $\{y\}$ is closed, $V = \mathcal{X} \setminus \{y\}$ is open, and $x \in V$, $y \notin V$. So $\mathcal{X}$ is T_1.

If $\mathcal{X}$ is T_1, given $x \in X$, for each $y \neq x$ there exists V_y, open, with $y \in V_y$ and $x \notin V_y$. Then $V = \bigcup_{y \neq x} V_y$ is open, and it contains all points bar x: that is $V = \mathcal{X} \setminus \{x\}$. As V is open, its complement $\{x\}$ is closed.

(5.4.3) Show that in a topological vector space, all open neighbourhoods around a point x are given by translates by x of neighbourhoods of 0.

Answer. Fix x. The function $f(y) = x + y$ is continuous by definition of TVS. It's inverse $g(y) = -x + y$ is also continuous, so f is a homeomorphism. Given any open set V with $x \in V$, the set $W = f^{-1}(V)$ is open by continuity, and $0 \in W$, since $f(0) = x$. We have $V = f(W) = x + W$.

(5.4.4) Show that any open neighbourhood of 0 in a topological vector space is absorbing, and that any nonzero multiple of an open set is open.

Answer. Fix V open with $0 \in V$. By definition of topological vector space, the map $\gamma : \mathbb{C} \to \mathcal{X}$ given by $\gamma(t) = tx$ is continuous; as $\gamma(0) = 0 \in V$, there exists an open disk $W = B_\delta(0) \subset \mathbb{C}$ with $0 \in W$ and $\gamma(W) \subset V$. This means that if $|t| < \delta$ then $tx \in V$. So V is absorbing.

Let V open and fix $c \in \mathbb{C} \setminus \{0\}$. Also by definition, the map $\alpha : \mathcal{X} \to \mathcal{X}$ given by $\alpha(x) = \frac{1}{c} x$ is continuous. Then $cV = \alpha^{-1}(V)$ is open.

(5.4.5) Let $\mathcal{X}$ be a TVS and $V, W \subset \mathcal{X}$ with V open. Show that $V + W$ is open.

Answer. This was explicitly done in the text! For each $v \in V$, by Exercise 5.4.3 the set $v + W$ is open. Then

$$V + W = \bigcup_{v \in V} (v + W)$$

is open.

(5.4.6) Prove that in a topological vector space the interior and the closure of a convex set are convex.

Answer. Let $A \subset \mathcal{X}$ be convex. Let A_0 be the interior of A. If $x, y \in A_0$ and $t \in (0,1)$, choose open $V_x, V_y \subset A$ with $x \in V_x$, $y \in V_y$. Then $tx \in tV_x$, which is open, and $(1-t)y \in (1-t)V_y$, which is also open. And $tx + (1-t)y \in tV_x + (1-t)V_y$, which is open by Exercise 5.4.5, and a subset of A by convexity. This shows that A_0 is convex.

As for the closure, if $x, y \in \overline{A}$, there exist nets $\{x_j\}$ and $\{y_j\}$ with $x_j \to x$, $y_j \to y$. Then, since addition and multiplication by scalars are continuous,

$$tx + (1-t)y = \lim_j tx_j + (1-t)y_j \in \overline{A},$$

as $tx_j + (1-t)y_j \in A$ by convexity. Hence $\overline{A}$ is convex.

(5.4.7) Prove that, in a topological vector space, the closure of a balanced set is balanced; and if the interior contains 0, then the interior is balanced.

Answer. Assume that $A \subset \mathcal{X}$ is balanced. If $x \in \overline{A}$ and $c \in \mathbb{C}$ with $|c| \leq 1$, there exists a net $\{x_j\}$ with $x_j \to x$. As A is balanced, $cx_j \in A$ for all j. Then, by continuity of the product by scalars, $cx = \lim cx_j \in \overline{A}$.

If $0 \in A$ and $x \in A_0$, the interior of A, by Exercise 5.4.3 there exists an open set V with $0 \in V$ and $x + V \subset A$. Using Lemma 5.4.4, there exists V_0 open and balanced with $0 \in V_0 \subset V$. Then, if $0 < |c| \leq 1$, $cx + cV_0$ is an open neighbourhood of cx, and $cx + cV_0 = c(x + V_0) \subset cA \subset A$, so $cx \in A_0$. When $c = 0$, we have that $cx = 0 \in A_0$ by hypothesis. So A_0 is balanced.

Here is an example of a balanced set such that 0 is not in its interior. Let $\mathcal{X} = \mathbb{C}^2$ and

$$A = \{(z_1, z_2) : \ |z_1| \leq |z_2|\}.$$

Then A is balanced, since multiplying each coordinate by a fixed c with $|c| \leq 1$ will not alter the inequality. And

$$\operatorname{Int} A = \{(z_1, z_2) \in A : \ |z_1| < |z_2|\}; \tag{AB.5.4}$$

indeed, given (z_1, z_2) with $|z_1| < |z_2|$ we can choose $\varepsilon < (|z_2| - |z_1|)/2$. Then if $|w_1 - z_1| < \varepsilon$ and $|w_2 - z_2| < \varepsilon$ we have

$$|w_2| - |w_1| \geq |z_2| - \varepsilon - |z_1| - \varepsilon = |z_2| - |z_1| - 2\varepsilon > 0,$$

showing that the set in (AB.5.4) is open. When $|z_1| = |z_2|$ any ball that contains (z_1, z_2) will contain points (w_1, w_2) with $|w_1| > |w_2|$, so (AB.5.4) does describe Int A. In particular, $0 = (0, 0) \notin \operatorname{Int} A$.

(5.4.8) Give examples in $\mathbb{C}^2$, with the usual topology, of open neighbourhoods of 0 that are:

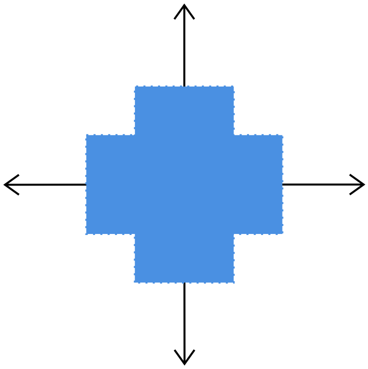
(i) balanced but not convex;

(ii) convex but not balanced.

In each case, does a local basis at 0 for the topology exist where all sets are like that? Are the same examples possible in $\mathbb{C}$?

Answer.

(i) Let $V_0 = \{(z, w) : \ |z| < 1, \ |w| < \frac{1}{2}\}$, $V_1 = \{(z, w) : |z| < \frac{1}{2}, \ |w| < 1\}$, and $V = V_0 \cup V_1$. This is what a real version of V would look like:



The set V is open (union of open), and balanced: both V_0 and V_1 already are balanced, since c with $|c| \leq 1$ will make $|cz| \leq |z|$ and $|cw| \leq |w|$. And it is not convex: for small $\varepsilon > 0$ the points $(\frac{1}{2} - \varepsilon, 1 - \varepsilon)$ and $(1 - \varepsilon, \frac{1}{2} - \varepsilon)$ are in V, but $\frac{1}{2}\,(\frac{1}{2} - \varepsilon, 1 - \varepsilon) + \frac{1}{2}\,(1 - \varepsilon, \frac{1}{2} - \varepsilon) = (\frac{3}{4} - \varepsilon, \frac{3}{4} - \varepsilon)$ is not in V.

Since "crosses" like the above one can be put inside balls, and balls inside them, they induce the same topology as the balls; so there is a basis for $\mathbb{C}^2$ given by sets as above.

In $\mathbb{C}$, an open balanced neighbourhood of 0 is an open disk or all of $\mathbb{C}$, so it is convex. Indeed, if $V \subset \mathbb{C}$ is open and balanced with $0 \in V$, let

$r = \sup\{|z| : z \in V\}$. If $r = \infty$, given any $z \in \mathbb{C}$ there exists $v \in V$ with $|v| > |z|$; then write $v = ae^{it}$, $z = be^{is}$. We have, since $a > b$,

$$z = \left(\frac{b}{a}\right) e^{i(s-t)} \, v \in V.$$

When $r < \infty$, given $z \in \mathbb{C}$ with $|z| < r$ we can find $v \in V$ with $|v| > |z|$ and we can repeat the above argument, so $V = B_r(0)$.

(ii) Let $V = \{z \in \mathbb{C} : -1 < \operatorname{Re} z < 1, \ -\frac{1}{2} < \operatorname{Im} z < \frac{1}{2}\} \times \{w \in \mathbb{C} : |w| < 1\}$. Then V is open, and it is convex: if $t \in [0,1]$ then

$$t(z_1, w_1) + (1-t)(z_2, w_2) = (tz_1 + (1-t)z_2, tw_1 + (1-t)w_2) \in V.$$

But it is not balanced: we have $(\frac{3}{4}, \frac{3}{4}) \in V$, but $i(\frac{3}{4}, \frac{3}{4}) = (\frac{3}{4}i, \frac{3}{4}i) \notin V$.

The same argument with the balls shows that there is indeed a local basis at 0 for the topology of $\mathbb{C}^2$ made of open sets which are convex but not balanced.

In $\mathbb{C}$, let $V = \{z \in \mathbb{C} : 2(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 < 1\}$. Then V is an open neighbourhood of 0, and it is convex (it's an "ellipse"). But it is not balanced: $z = i/\sqrt{2} \in V$, but $iz = -1/\sqrt{2} \notin V$.

(5.4.9) Fill the details in Example 5.4.14, i.e. show that the topology induced by the seminorms agrees with the topology of pointwise-convergence.

Answer. Take the family of seminorms as in Example 5.4.14.

Suppose that $f_j \to f$ pointwise. Given $V(p_1, \ldots, p_n, \varepsilon)$ where $p_k(g) = |g(t_k)|$, for each $k = 1, \ldots, n$ we can choose j_k such that, for $j \geq j_k$, we have $|f_j(t_k) - f(t_k)| < \varepsilon$. Let $j_0 = \max\{j_1, \ldots, j_n\}$. Then, for $j \geq j_0$, $p_k(f_j - f) = |f_j(t_k) - f(t_k)| < \varepsilon$. So $f_j \in f + V(p_1, \ldots, p_n, \varepsilon)$ for all $j \geq j_0$; as the basic neighbourhood was arbitrary, we have shown that $f_j \to f$ in the topology determined by the seminorms.

Conversely, if $f_j \to f$ on the seminorms, given $t \in [0,1]$ and $\varepsilon > 0$, there exists j_0 such that for all $j \geq j_0$ we have $f + V(p_t, \varepsilon)$; that is, $|f_j(t) - f(t)| < \varepsilon$ for all $j \geq j_0$; so $f_j \to f$ pointwise.

(5.4.10) Show that if $\mathcal{H}$ is an infinite-dimensional Hilbert space and $\{\xi_j\}$ is an orthonormal basis, then $\xi_j \xrightarrow{\text{weak}} 0$. *(We defined weak convergence on Example 5.4.13)*

Answer. Fix $\eta \in \mathcal{H}$. We have, since $\sum_j |\langle \eta, \xi_j \rangle|^2 = \|\eta\|^2 < \infty$, that $\langle \eta, \xi_j \rangle \to 0$. That is, $\xi_j \xrightarrow{\text{weak}} 0$.

Note that this only works in infinite dimension, for otherwise the coefficients need not converge to zero.

(5.4.11) Let $V \subset \mathcal{X}$ be an open, balanced, convex, set with $0 \in V$. Let $\varepsilon > 0$. Show that

$$\{x : \ \mu_V(x) < \varepsilon\} = \{x : \ \mu_{\varepsilon V}(x) < 1\} = \varepsilon V.$$

Answer. The set εV is open by continuity of multiplication by scalars, and it is trivial to check that it is convex. By Proposition 5.4.9,

$$\begin{aligned} \varepsilon V &= \varepsilon \{x : \ \mu_V(x) < 1\} = \{\varepsilon x : \ \mu_V(x) < 1\} \\ &= \left\{x : \ \mu_V\left(\frac{1}{\varepsilon} x\right) < 1\right\} = \{x : \ \mu_V(x) < \varepsilon\}. \end{aligned}$$

(5.4.12) Let $\mathcal{X}$ be a TVS and $M \subset \mathcal{X}$ a convex, open, neighbourhood of 0. Show that μ_M is continuous.

Answer. Since μ_M is a real seminorm,

$$|\mu_M(x) - \mu_M(y)| \leq \mu_M(x - y),$$

so it is enough to show that μ_M is continuous at 0. Fix $\varepsilon > 0$. By Exercise 5.4.11 the set εM is open and $\varepsilon M = \{x \in \mu_M(x) < \varepsilon\}$. So we can take εM as the neighbourhood of 0 that guarantees that $|\mu_M(x)| < \varepsilon$ if $x \in \varepsilon M$.

(5.4.13) Let $\mathcal{X}$ be a vector space and $\mathcal{P}$ a family of seminorms that separates points. Show that the sets

$$V_x(p_1, \ldots, p_n, \varepsilon), \qquad x \in \mathcal{X}, \ p_1, \ldots, p_n \in \mathcal{P}, \ \varepsilon > 0,$$

where

$$V_x(p_1, \ldots, p_n, \varepsilon) = \{x' \in \mathcal{X} : \ p_k(x' - x) < \varepsilon, \ k = 1, \ldots, n\}.$$

form a base for a topology.

Answer. The sets clearly cover $\mathcal{X}$, as every x is allowed. So we need to show that given $x, y \in \mathcal{X}$, $p_1, \ldots, p_n, q_1, \ldots, q_m \in \mathcal{P}$ and $\varepsilon > 0$, $\delta > 0$, if

$V_x(p_1,\dots,p_n,\varepsilon)\cap V_y(q_1,\dots,q_m,\delta)\neq\varnothing$ there exist $z\in\mathcal{X}$, $r_1,\dots,r_k\in\mathcal{P}$ and $\gamma>0$ such that

$$V_z(r_1,\dots,r_k,\gamma)\subset V_x(p_1,\dots,p_n,\varepsilon)\cap V_y(q_1,\dots,q_m,\delta).$$

This is achieved by taking $z\in V_x(p_1,\dots,p_n,\varepsilon)\cap V_y(q_1,\dots,q_m,\delta)$,

$$r_1,\dots,r_k=p_1,\dots,p_n,q_1,\dots,q_m,$$

and

$$\gamma=\min_{k,j}\{\varepsilon-p_k(z-x),\delta-q_j(z-y)\}.$$

Then for any $w\in V_z(r_1,\dots,r_k,\gamma)$ we have

$$p_k(w-z)\leq p_k(w-x)+p_k(z-x)\leq\gamma+p_k(z-x)<\varepsilon,$$

and

$$q_j(w-z)\leq q_j(w-y)+q_j(z-y)\leq\gamma+q_j(z-y)<\delta.$$

(5.4.14) Let $\mathcal{X},\mathcal{Y}$ be locally convex spaces. Show that $(x_j,y_j)\to(x,y)$ on $\mathcal{X}\oplus_T\mathcal{Y}$ if and only if $x_j\to x$ and $y_j\to y$.

Answer. Suppose that $(x_j,y_j)\to(x,y)$. By definition this means that $(p\times q)\big((x_j,y_j)-(x,y)\big)\to 0$ for all seminorms p for $\mathcal{X}$ and q for $\mathcal{Y}$, and by definition of the product seminorms this is $p(x_j-x)+q(y_j-y)\to 0$. As $p(x_j-x)\leq p(x_j-x)+q(y_j-y)$ for all j, we get that $p(x_j-x)\to 0$ and similarly $q(y_j-y)\to 0$ for all seminorms. Hence $x_j\to x$ and $y_j\to y$.

Conversely, if $x_j\to x$ and $y_j\to y$ then $p(x_j-x)\to 0$ and $q(y_j-y)\to 0$ for all seminorms. It follows that $p(x_j-x)+q(y_j-y)\to 0$ for all seminorms $p\times q$, showing that $(x_j,y_j)\to(x,y)$.

(5.4.15) Show that the family $\tilde{S}_{\mathcal{X}}$ from Proposition 5.4.21 is indeed a family of seminorms that separates points.

Answer. Fix $p\in S_{\mathcal{X}}$. For $\alpha\in\mathbb{C}$ nonzero,

$$\begin{aligned}\tilde{p}(\alpha x+M)&=\inf\{p(\alpha(x+m/\alpha)):\ m\in M\}=|\alpha|\,\inf\{p(x+m/\alpha):m\in M\}\\&=|\alpha|\,\inf\{p(x+m):\ m\in M\}=|\alpha|\,\tilde{p}(x+M).\end{aligned}$$

For the triangle inequality, let $x, y \in \mathcal{X}$ and fix $\varepsilon > 0$. Choose $m_x, m_y \in M$ such that $p(x + m_x) < \tilde{p}(x + M) + \varepsilon$, $p(y + m_y) < \tilde{p}(x + M) + \varepsilon$. Then

$$\begin{aligned}\tilde{p}(x + M + y + M) &= \inf\{p(x + y + m) : \ m \in M\}\\ &\le p(x + y + m_x + m_y)\\ &\le p(x + m_x) + p(y + m_y)\\ &\le \tilde{p}(x + M) + \tilde{p}(x + M) + 2\varepsilon.\end{aligned}$$

As this works for any $\varepsilon > 0$, the triangle inequality is established. It remains to show that the family separates points. If $x + M \neq 0$, this means that $x \notin M$. As M is closed, there exists a basic neighbourhood $N = \{z : \ p_j(z) < 1, \ j = 1, \dots, r\}$ of 0 such that $M \cap (x + N) = \varnothing$. That is, $m - x \notin N$ for all $m \in M$. Which means that there exists j with $p_j(x + m) \ge 1$ for all $m \in M$, so $\tilde{p}_j(x + M) \ge 1$. This shows that if $\tilde{p}(x + M) = 0$ for all $\tilde{p} \in \tilde{S}_{\mathcal{X}}$ then $x + M = 0$.

5.5. The Dual

(5.5.1) Let $\mathcal{X}$ be a normed space. Show that $\mathcal{X}^*$ is a normed vector space.

Answer. Given $\varphi, \psi \in \mathcal{X}^*$ and $\lambda \in \mathbb{C}$ we can form linear combinations by $(\varphi + \lambda\psi)(x) = \varphi(x) + \lambda\psi(x)$, so $\mathcal{X}^*$ is naturally a vector space if we show that this linear combination is continuous. But this follows from the continuity of φ and ψ and the continuity of the vector space operations: if $x_n \to x$, then

$$(\varphi + \lambda\psi)(x_n) = \varphi(x_n) + \lambda\psi(x_n) \to \varphi(x) + \lambda\psi(x) = (\varphi(x) + \lambda\psi)(x).$$

(5.5.2) Let $\mathcal{X}$ be a TVS. Show that $\mathcal{X}^*$ is a vector space.

Answer. The same argument from Exercise 5.5.1 works.

(5.5.3) Let $\mathcal{X}$ be a finite-dimensional space. Show that $\mathcal{X}^*$ is finite-dimensional and $\dim \mathcal{X}^* = \dim \mathcal{X}$.

Answer. Fix a basis $x_1, \ldots, x_n$. For each $x \in \mathcal{X}$ there are unique numbers $\lambda_j(x) \in \mathbb{C}$ with $x = \sum_j \lambda_j(x)\, x_j$. The uniqueness makes each $\lambda_j : \mathcal{X} \to \mathbb{C}$ linear, since

$$\alpha x + y = \sum_j (\alpha\lambda_j(x) + \lambda_j(y))\, x_j.$$

They are also continuous since $\dim \mathcal{X} < \infty$ (Corollary 5.2.5). Given any $\psi \in \mathcal{X}^*$,

$$\psi(x) = \sum_j \lambda_j(x)\, \psi(x_j) = \Big(\sum_j \psi_j(x_j)\, \lambda_j\Big)(x), \qquad x \in \mathcal{X}.$$

Thus $\mathcal{X}^* = \operatorname{span}\{\lambda_1, \ldots, \lambda_n\}$. And if $\sum_j c_j \lambda_j = 0$, for a fixed x_k we have

$$0 = \sum_j c_j \lambda_j(x_k) = c_k.$$

So $c_1 = \cdots = c_n = 0$, and $\lambda_1, \ldots, \lambda_n$ are linearly independent. Thus $\dim \mathcal{X}^* = \dim \mathcal{X}$.

(5.5.4) Complete the proof of Proposition 5.5.2.

Answer. Assume first that $\dim X/K = 1$. Choose x such that $x + K \neq K$. Then $X/K = \mathbb{C}(x + K)$. Define $\varphi(y) = c_y$, where $c_y \in \mathbb{C}$ is the scalar such that $y + K = c_y\, x + K$. The scalar is unique (because if $ax + K = bx + K$ then $(a - b)x \in K$, and so $a - b = 0$), hence φ is well defined. If $y, z \in \mathcal{X}$ and $\lambda \in \mathbb{C}$, then $(y + \lambda\, z) + K = (y + K) + \lambda\,(z + K)$ by definition of addition in the quotient; so $y + \lambda z + K = c_y\, x + \lambda\, c_z\, x + K = (c_y + \lambda\, c_z)x + K$. Thus $\varphi(y + \lambda z) = \varphi(y) + \lambda\varphi(z)$ by the uniqueness and thus φ is linear.

If $\varphi(y) = 0$, then $y + K = K$, so $y \in K$; thus $\ker \varphi \subset K$. Conversely, if $y \in K$ then $y + K = 0 + K$ so $\varphi(y) = 0$. Then $K = \ker \varphi$.

(5.5.5) Prove Proposition 5.5.4.

Answer. (i) $\Longrightarrow$ (ii): trivial.

(ii) $\Longrightarrow$ (iii): trivial, as we can take $x = 0$.

(iii) $\Longrightarrow$ (iv): assume that φ is continuous at x_0. Then there exists $\delta > 0$ such that $|\varphi(x) - \varphi(x_0)| < 1$ whenever $\|x - x_0\| < \delta$. Now fix $x \in \mathcal{X}$.

Then $x' = \frac{\delta x}{2\|x\|} + x_0$ satisfies $\|x' - x_0\| < \delta$ and therefore

$$|\varphi(x)| = \frac{2\|x\|}{\delta}\, |\varphi(x' - x_0)| < \frac{2\|x\|}{\delta}.$$

(iv) $\implies$ (i): Fix $x_0 \in \mathcal{X}$. Given $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{r}$. Then, if $\|x - x_0\| < \delta$, we have

$$|\varphi(x) - \varphi(x_0)| = |\varphi(x - x_0)| \leq r\|x - x_0\| < \varepsilon.$$

So φ is continuous at x_0.

(5.5.6) Prove Proposition 5.5.6.

Answer. By definition,

$$\|\varphi\| = \inf\{r : |\varphi(x)| \leq r\|x\| \text{ for all } x \in \mathcal{X}\}.$$

Given $\varepsilon > 0$, there exists r as above with $r < \|\varphi\| + \varepsilon$. Then, for any $x \in \mathcal{X}$,

$$|\varphi(x)| \leq r\|x\| \leq (\|\varphi\| + \varepsilon)\|x\|.$$

As this holds for all $\varepsilon > 0$, we get $|\varphi(x)| \leq \|\varphi\|\,\|x\|$. So

$$\|\varphi\| = \min\{r : |\varphi(x)| \leq r\|x\| \text{ for all } x \in \mathcal{X}\}.$$

Now if r is an upper bound for $\{|\varphi(x)| : \|x\| = 1\}$, we have $|\varphi(x/\|x\|)| \leq r$ for all nonzero x, so $|\varphi(x)| \leq r\|x\|$ (which works also for $x = 0$). So, by (i), $\sup\{|\varphi(x)| : \|x\| = 1\} \geq \|\varphi\|$, since $\|\varphi\|$ is below all upper bounds. And since when $\|x\| = 1$ we have $|\varphi(x)| \leq \|\varphi\|\,\|x\| = \|\varphi\|$, we get $\|\varphi\|$ itself is an upper bound, so $\sup\{|\varphi(x)| : \|x\| = 1\} \leq \|\varphi\|$, giving us the equality (ii). For (iii), we simply note that $y = x/\|x\|$ has $\|y\| = 1$, and $|\varphi(x)|/\|x\| = |\varphi(x/\|x\|)|$.

(5.5.7) Use Proposition 5.5.6 to show that (5.10) defines a norm on the space of bounded functionals on $\mathcal{X}$.

Answer. We need to show that the norm as in Proposition 5.5.6 is a norm.

If $\|\varphi\| = 0$, then $|\varphi(x)| = 0$ for all x, and so $\varphi = 0$.

For $\lambda \in \mathbb{C}$,

$$\begin{aligned}\|\lambda\varphi\| &= \sup\{|\lambda\,\varphi(x)| : \|x\| = 1\} = |\lambda|\,\sup\{|\varphi(x)| : \|x\| = 1\}\\ &= |\lambda|\,\|\varphi\|.\end{aligned}$$

And

$$\begin{aligned}\|\varphi+\psi\| &= \sup\{|\varphi(x)+\psi(x)| : \ \|x\|=1\}\\ &\le \sup\{|\varphi(x)|+|\psi(x)| : \ \|x\|=1\}\\ &\le \sup\{\|\varphi\|+|\psi(x)| : \ \|x\|=1\}\\ &= \|\varphi\|+\sup\{|\psi(x)| : \ \|x\|=1\} = \|\varphi\|+\|\psi\|.\end{aligned}$$

(5.5.8) Let $\mathcal{X}$ be a topological vector space, and $\varphi : \mathcal{X} \to \mathbb{C}$ linear. Suppose that there exists an open neighbourhood V of 0 and $c > 0$ such that $|\varphi(v)| \le c$ for all $v \in V$. Prove that φ is continuous

Answer. Fix $\varepsilon > 0$, and let $V_\varepsilon = \frac{\varepsilon}{c} V$. For any $x \in V_\varepsilon$, we can write $x = \frac{\varepsilon}{c} v$, with $v \in V$. Then

$$|\varphi(x)| = \left|\varphi\left(\frac{\varepsilon}{c}\, v\right)\right| = \frac{\varepsilon}{c}|\varphi(v)| \le \frac{\varepsilon}{c}\, c = \varepsilon.$$

So φ is continuous at 0. If now $x_j \to x$, since addition is continuous we have that $x_j - x \to 0$. Then $\varphi(x_j - x) \to 0$, and so $\varphi(x_j) \to \varphi(x)$ by linearity.

(5.5.9) Let $\mathcal{X} = \{f \in C[0,1] : \ f(0) = 0\}$ with the supremum norm, and $\varphi(f) = \int_0^1 f$. Show that $\|\varphi\| = 1$ but $|\varphi(f)| < 1$ for all $f \in \mathcal{X}$ with $\|f\| = 1$.

Answer. If $\|f\| \le 1$, then

$$\left|\int_0^1 f\right| \le \int_0^1 |f| \le \int_0^1 1 = 1,$$

so $\|\varphi\| \le 1$. Now let

$$g_n(t) = \begin{cases} nt, & 0, \le t \le \frac{1}{n} \\ 1, & t \ge \frac{1}{n} \end{cases}$$

Then $g_n \in \mathcal{X}$, $\|g_n\| = 1$ and

$$\int_0^1 g_n = \frac{1}{2n} + 1 - \frac{1}{n} = 1 - \frac{1}{2n}.$$

This shows that $\|\varphi\| > 1 - \frac{1}{2n}$ for all n, so $\|\varphi\| = 1$.

If remains to see that $|\varphi(f)| < 1$ for all $f \in \mathcal{X}$ with $\|f\| = 1$. Given such f, because f is continuous at 0 there exists $\delta > 0$ such that $|f(t)| < \frac{1}{2}$

when $t < \delta$. Then

$$|\varphi(f)| = \left| \int_0^1 f \right| \leq \left| \int_0^\delta f \right| + \int_\delta^1 1 \leq \frac{\delta}{2} + 1 - \delta = 1 - \frac{\delta}{2} < 1.$$

(5.5.10) Let $\mathcal{X}$ be a normed space and $\varphi : \mathcal{X} \to \mathbb{C}$ a linear functional. Show that the following statements are equivalent:

(i) φ is unbounded;

(ii) there exists a sequence $\{y_n\} \subset \mathcal{X}$ such that $\|y_n\| = 1$, $\varphi(y_n) > n$ for all n;

(iii) there exists a sequence $\{x_n\} \subset \mathcal{X}$ such that $x_n \to 0$, and $\varphi(x_n) = 1$ for all n.

Answer. (i) $\implies$ (ii): If φ is unbounded, then for each $n \in \mathbb{N}$ there exists $x_n \in \mathcal{X}$ with $|\varphi(x_n)| > n\|x_n\|$. Take $\lambda_n \in \mathbb{T}$ with $\lambda_n \varphi(x_n) = |\varphi(x_n)|$. If we take $y_n = \lambda_n x_n / \|x_n\|$, then $\|y_n\| = 1$ and

$$\varphi(y_n) = \frac{\lambda_n \varphi(x_n)}{\|x_n\|} = \frac{|\varphi(x_n)|}{\|x_n\|} > n.$$

(ii) $\implies$ (iii): Take the sequence $\{y_n\}$ as in (ii) and define $x_n = y_n/\varphi(y_n)$. Then $\|x_n\| = 1/\varphi(y_n) < 1/n \to 0$, and $\varphi(x_n) = \varphi(y_n)/\varphi(y_n) = 1$.

(iii) $\implies$ (i): Given $\{x_n\}$ as in (iii), let $z_n = x_n/\|x_n\|$. Then $\|z_n\| = 1$ and $|\varphi(z_n)| = |\varphi(x_n)|/\|x_n\| = 1/\|x_n\| \to \infty$, so φ is unbounded.

(5.5.11) Let $\mathcal{X}$ be a Banach space, $\mathcal{X}_0$ a dense subspace and $\varphi : \mathcal{X}_0 \to \mathbb{C}$ a bounded linear functional. Show that φ admits a unique extension $\tilde{\varphi} \in \mathcal{X}^*$.

Answer. Let $x \in \mathcal{X}$. There exists a sequence $\{x_n\} \subset \mathcal{X}_0$ with $x_n \to x$. Since φ is bounded,

$$|\varphi(x_n) - \varphi(x_m)| = |\varphi(x_n - x_m)| \leq \|\varphi\| \, \|x_n - x_m\|,$$

so the sequence $\{\varphi(x_n)\} \subset \mathbb{C}$ is Cauchy. Let $\tilde{\varphi}(x) = \lim_n \varphi(x_n)$, this limit exists since $\mathbb{C}$ is complete. This is well-defined: if $x_n' \to x$, then

$$|\varphi(x_n) - \varphi(x_m')| = |\varphi(x_n - x_m')| \to 0$$

with m, n, so the limit is the same. Linearity of $\tilde{\varphi}$ is straightforward since it is defined as a limit of linear maps. Finally,

$$|\tilde{\varphi}(x)| = \lim_n |\varphi(x_n)| \leq \|\varphi\| \lim_n \|x_n\| = \|\varphi\| \, \|x\|$$

(using Exercise 5.1.2 for the last equality). Thus $\|\tilde{\varphi}\| \leq \|\varphi\|$ and the reverse inequality holds trivially because $\tilde{\varphi}$ is an extension of φ.

Uniqueness: if $\psi \in \mathcal{X}^*$ and $\psi|_{\mathcal{X}_0} = \varphi$, then for any $x \in \mathcal{X}$ there exists $\{x_n\} \subset \mathcal{X}_0$ with $x_n \to x$. Then, using that ψ is continuous,

$$\psi(x) = \lim_n \psi(x_n) = \lim_n \varphi(x_n) = \tilde{\varphi}(x).$$

(5.5.12) Let V be an infinite-dimensional real/complex vector space. Consider linear maps $\varphi_1, \ldots, \varphi_n : V \to \mathbb{C}$ linear. Improve on Proposition 5.5.12 by showing that

$$\dim \bigcap_{j=1}^n \ker \varphi_j = \infty.$$

Answer. Consider as in the proof of Proposition 5.5.12 the linear map $\Gamma : V \to \mathbb{C}^n$ given by $\Gamma(x) = (\varphi_1(x), \ldots, \varphi_n(x))^\top$. Let $Y = \Gamma(V)$, a subspace of $\mathbb{C}^n$. We have

$$V/\ker\Gamma \simeq Y.$$

This forces $\dim\ker\Gamma = \infty$. Indeed, if we had $\dim\ker\Gamma < \infty$, we can choose a basis $y_1, \ldots, y_m$ of Y and a basis $\{z_1, \ldots, z_s\}$ of $\ker\Gamma$. By the isomorphism there is a basis $\{v_1 + \ker\Gamma, \ldots, v_r + \ker\Gamma\}$ of $V/\ker\Gamma$. So any $v \in V$ can be written as a linear combination of $v_1, \ldots, v_r, z_1, \ldots, z_s$ and V would be finite-dimensional. Therefore

$$\dim \bigcap_{j=1}^n \ker \varphi_j = \dim\ker\Gamma = \infty.$$

For a different argument, suppose that

$$\dim V = \infty \qquad \text{and} \qquad \bigcap_{j=1}^n \ker \varphi_j = \operatorname{span}\{z_1, \ldots, z_r\}$$

with $z_1, \ldots, z_r$ linearly independent. Extend $\{z_1, \ldots, z_r\}$ to a basis

$$\{z_1, \ldots, z_r\} \cup \{w_1, w_2, \ldots\}$$

of V. If we let $\{\psi_1, \psi_2, \ldots\}$ be the dual basis of $\{w_1, w_2, \ldots\}$ we get infinitely many linearly independent linear functionals. As $\{z_1, \ldots, z_r\} \subset \ker\psi_j$ for all j, from Lemma 5.5.10 we have $\psi_j \subset \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$, a contradiction since we have infinitely many linearly independent ψ_j.

5.6. Examples of Duals

(5.6.1) Let $q \in [1,\infty)$, $p \in \mathbb{R}$ with $\frac{1}{p}+\frac{1}{q}=1$, $g:\mathbb{N}\to\mathbb{C}$. Put $B_p=\{f\in\ell^p(\mathbb{N}):\ \|f\|_p=1\}$. Show that

$$\max\left\{\,|\langle f,g\rangle| : f\in B_p\right\} = \max\left\{\sum_{k=1}^\infty |g(k)f(k)| : f\in B_p\right\},$$

where

$$\langle f,g\rangle=\sum_{k=1}^\infty g(k)f(k).$$

Answer. Let us denote the left-hand-side by L and the right-hand-side by R. The triangle inequality guarantees that $L\le R$.

Write $g(k)f(k)=e^{i\theta_k}\,|f(k)g(k)|$. Let $f_0:\mathbb{N}\to\mathbb{C}$ be given by $f_0(k)=e^{-i\theta_k}f(k)$. Then $\|f_0\|_p=\|f\|_p=1$, and

$$\left|\sum_{k=1}^\infty g(k)f_0(k)\right| = \left|\sum_{k=1}^\infty e^{-i\theta_k}g(k)f(k)\right| = \sum_{k=1}^\infty |g(k)f(k)|.$$

Then $R\le L$, as any element in the right-hand set appears in the left-hand set. Thus $L=R$.

(5.6.2) Consider the Banach space c_0 (Example 5.1.9). Given $x\in c_0$ show that for any $f\in\ell^1(\mathbb{N})$ the map $x\mapsto\sum_n x_nf_n$ is a continuous linear functional. Use this to prove that there is an isometric embedding of $\ell^1(\mathbb{N})$ into c_0^*. Prove that this embedding is surjective, i.e. $c_0^*=\ell^1(\mathbb{N})$.

Answer. Call the map γ_f. It is linear, since limits and sums are linear:

$$\gamma_f(x+\alpha y)=\sum_n (x_n+\alpha y_n)f_n=\sum_n x_nf_n+\alpha y_nf_n=\gamma_f(x)+\alpha\gamma_f(y).$$

Since $x\in c_0$, we have that $|x_n|\le\|x\|_\infty$ for all n. Then

$$|\gamma_f(x)|=\left|\sum_n x_nf_n\right|\le\sum_n |x_n|\,|f_n|\le\|x\|_\infty\,\|f\|_1.$$

So $\gamma_f \in c_0^*$ and $\|\gamma_f\| \leq \|f\|_1$. Now write $f_n = e^{i\theta_n}|f_n|$ and let x be the sequence with its first m entries consisting of $x_n = e^{-i\theta_n}$, and the rest 0. Then $x \in c_0$, $\|x\|_\infty = 1$ and

$$|\gamma_f(x)| = \sum_{n=1}^m |f_n|.$$

We then get that $\|\gamma_f\| \geq \sum_{n=1}^m |f_n|$; as we can do this for any m we get that $\|\gamma_f\| \geq \|f\|_1$. Thus $\|\gamma_f\| = \|f\|_1$. This shows that $\gamma : f \longmapsto \gamma_f$ is an isometric embedding. It is also linear, since

$$\begin{aligned}\gamma(f+\alpha g)(x) = \gamma_{f+\alpha g}(x) &= \sum_n x_n(f_n + \alpha g_n) = \sum_n x_n f_n + \alpha \sum_n x_n g_n \\ &= \gamma_f(x) + \alpha\gamma_g(x) = [\gamma(f) + \alpha\gamma(g)](x).\end{aligned}$$

For surjectivity, let $\phi \in c_0^*$. Given $x \in c_0$ we have $x = \sum_n x_n e_n$, where e_n are the canonical elements, i.e. $e_n(j) = \delta_{n,j}$. The series converges by Exercise 5.1.9. Thus, as ϕ is continuous and linear,

$$\phi(x) = \sum_n x_n \phi(e_n).$$

So $\phi = \gamma(f)$, where $f_n = \phi(e_n)$, if we are able to show that this f is in $\ell^1(\mathbb{N})$. For this, write $\phi(e_n) = e^{i\theta_n}|\phi(e_n)|$, and let $x \in c_0$ be such that

$$x_n = \begin{cases} e^{-i\theta_n}, & n \leq m \\ 0, & n > m \end{cases}$$

Then, using again that ϕ is continuous and linear,

$$\sum_{n=1}^m |\phi(e_n)| = \sum_{n=1}^m x_n \phi(e_n) = \phi\left(\sum_{n=1}^m x_n e_n\right) = \phi(x) \leq \|\phi\|\,\|x\| = \|\phi\|.$$

As m is arbitrary, this shows that $\|f\|_1 = \sum_n |\phi(e_n)| < \infty$. Thus γ is onto.

(5.6.3) Using the ideas in Exercise 5.6.2, show that the dual of c (the Banach space of convergent sequences with the supremum norm) is $\ell^1(\mathbb{N})$.

Answer. We have that $c = c_0 + \mathbb{C}1$. The proof in Exercise 5.6.2 doesn't apply directly, because we used that $x \in c_0$ to prove surjectivity—and that's essential in some way, since otherwise we would have a "proof" that $\ell^\infty(\mathbb{N})^* = \ell^1(\mathbb{N})$, which is false. The embedding part works fine—as it also does for $\ell^\infty(\mathbb{N})$—so we need to focus in surjectivity, i.e. showing that if $\varphi \in c^*$ then there exists $f \in \ell^1(\mathbb{N})$ with $\varphi = \langle \cdot, f\rangle$. Using Exercise 5.6.2 we have that, on c_0, $\varphi(x) = \langle x, f\rangle$ where $f_n = \varphi(e_n)$. As $c = c_0 + \mathbb{C}1$, the value of φ

changes depending on what $\varphi(1)$ is. The problem, in other words, is that the function $x \longmapsto \lim_n x_n$ cannot possibly come from $\ell^1(\mathbb{N})$ if we try to reuse the embedding from Exercise 5.6.2. We can solve this the following way: we reserve the first coordinate in $\ell^1(\mathbb{N})$ for the value $\varphi(1)$. So we define $\gamma : \ell^1(\mathbb{N}) \to c^*$ by

$$\langle \gamma(f), x + \lambda 1 \rangle = \lambda f_1 + \sum_{n=1}^{\infty} x_n f_{n+1}.$$

Given $\varphi \in c^*$, we have by restriction that $\varphi \in c_0^*$. By Exercise 5.6.2 there exists $f' \in \ell^1(\mathbb{N})$ with $\varphi = \langle \cdot, f' \rangle$. Let

$$f = \varphi(1) e_1 + \sum_{n=2}^{\infty} f'_{n-1} e_n \in \ell^1(\mathbb{N}).$$

This f trivially satisfies $\gamma(f) = \varphi$, so γ is bounded. Finally we check that γ is isometric. For $x = x' + \lambda 1 \in c$ with $x' \in c_0$, we have $|\lambda| \leq \|x\|_\infty$ and $\|x'\|_\infty = \|x - \lambda\|_\infty \leq 2\|x\|_\infty\}$. Then

$$|\gamma(f)(x)| \leq |\lambda|\, f_1 + \sum_{n=1}^{\infty} |x_n|\, |f_{n+1}| \leq \|x\|_\infty \, \|f\|_1.$$

So $\|\gamma(f)\| \leq 2\|f\|_1$. Choosing an appropriate x as in Exercise 5.6.2 we get that $\|x\|_\infty = 1$ and $|\gamma(f)(x)| \geq \sum_{n=1}^{m} |f_n|$, and as we can do this for any m it follows that $\|\gamma\| \geq \|f\|_1$ and thus $\|f\|_1 \leq \|\gamma(f)\| \leq 2\|f\|_1$. In particular, $\|\gamma^{-1}\| \leq 1$ and γ is bicontinuous. It is not clear that c^* is isometrically isomorphic to $\ell^1(\mathbb{N})$.

(5.6.4) Show that c and c_0 are isomorphic as Banach spaces.

Answer. For $x \in c$, write $l_x = \lim_n x_n$. Then define $\gamma : c \to c_0$ by

$$\gamma(x) = (l_x, x_1 - l_x, x_2 - l_x, \ldots).$$

As $l_{x+y} = l_x + l_y$ and $l_{\lambda x} = \lambda\, l_x$, it follows that γ is linear. Also, if $\gamma(x) = 0$, then $l_x = 0$ and $0 = x_1 - l_x = x_1$, etc., so $x = 0$. Finally, given $y \in c_0$,

$$y = \gamma(y_1 + y_2, y_1 + y_3, y_1 + y_4, \ldots).$$

So γ is an linear isomorphism. And γ is bounded, as $|l_x| \leq \|x\|_\infty$ and so

$$\|\gamma(x)\| \leq \max\{|l_x|, \sup\{|x_n - l_x| : \ n \in \mathbb{N}\}\} \leq 2\|x\|.$$

Note that γ is not isometric. For instance if $x = (2, 1, 1, \ldots)$, then $\|\gamma(x)\| = 1$, while $\|x\| = 2$. This failure of γ on being isometric is not a failure of the way γ was chosen, but rather an intrinsic feature (see Exercise 5.6.5). The inverse γ^{-1} is trivially seen to be bounded, again with $\|\gamma^{-1}\| = 2$.

(5.6.5) Show that there is no isometry—linear or not—between c and c_0 *(Hint: consider that 1 is the middle point between 0 and 2).*

Answer. In c we have that $\|2-1\| = \|1-0\| = 1$; that is, the element 1 is at distance 1 from both 0 and 2. More importantly, 1 is the only element with that property. Indeed, suppose that $\|2-z\| = \|z\| = 1$. The second equality gives us that $|z_n| \leq 1$ for all n, while the first equality gives us $|2-z_n| \leq 1$ for all n. Then $2 - |z_n| \leq |2-z_n| \leq 1$, so

$$2 \leq 1 + |z_n| \leq 2.$$

Thus $|z_n| = |2-z_n| = 1$. The latter equality is, after using $|z_n| = 1$,

$$5 - 4\mathrm{Re}\, z_n = 1,$$

which in turn is $\mathrm{Re}\, z_n = 1$. Combined with $|z_n| = 1$, we obtain $z_n = 1$. As n was arbitrary, $z = 1$.

Meanwhile, given $x, y \in c_0$ with $\|y - x\| = 2$, there are uncountably many $z \in c_0$ with $\|x - z\| = \|y - z\| = 1$. Indeed, by translating everything by y we may assume that $y = 0$. That is, $\|x\| = 2$, and we are looking for z with $\|x - z\| = \|z\| = 1$. Since we are in c_0, the norm is actually a maximum. That is, there exists m with $|x_m| = 2$ and $|x_n| \leq 2$ for all n. Since $\lim_n x_n = 0$, there exists n_0 such that $|x_n| < 1/2$ for all $n \geq n_0$. We can define z to have $z_n = x_n/2$ for all $n \leq n_0$, and $z_n = e^{ir_n} x_n/2$ for $n > n_0$; with the exception of $z_m = 1$. Then $|x_n - z_n| \leq 1$ for all n, and $|z_m| = 1$. Thus $\|x - z\| = 1$, $\|z\| = 1$, and we are free to choose the $r_{n_0+1}, r_{n_0+2}, \ldots$ in uncountably many ways.

(5.6.6) Let S be an arbitrary set. Show that $\ell^1(S)^* = \ell^\infty(S)$, where the duality is the same as the one in Exercise 5.6.2.

Answer. We have the natural embedding $\gamma : \ell^\infty(S) \to \ell^1(S)^*$ given by

$$\gamma(x)(y) = \sum_j x_j y_j.$$

Then γ is linear, since for every $y \in \ell^1(S)$

$$\begin{aligned}\gamma(x + \alpha z)(y) &= \sum_j (x_j + \alpha z_j) y_j = \sum_j x_j y_j + \alpha \sum_j z_j y_j \\ &= [\gamma(x) + \alpha\gamma(z)](y).\end{aligned}$$

Also,

$$|\gamma(x)(y)| \leq \sum_j |x_j y_j| \leq \|x\|_\infty \, \|y\|_1, \quad x \in \ell^\infty(S),\; y \in \ell^1(S).$$

Thus $|\gamma(x)| \leq \|x\|_\infty$. Given $\varepsilon > 0$ and k such that $\|x\|_\infty - |x_k| < \varepsilon$, let $y \in \ell^1(\mathbb{N})$ be given by $y_k = \lambda$, where $\lambda x_k = |x_k|$, and $y_j = 0$ if $j \neq k$. Then $\|y\|_1 = 1$, and

$$\gamma(x)(y) = |x_k| > \|x\|_\infty - \varepsilon.$$

As we can do this for any $\varepsilon > 0$, we get that $\gamma(x) = \|x\|_\infty$. So the embedding is isometric. It remains to show that γ is surjective. Let $\varphi \in \ell^1(S)^*$. Write $y \in \ell^1(S)$ as $y = \sum_j y_j e_j$. Then, since φ is linear and bounded,

$$\varphi(y) = \sum_j y_j \varphi(e_j).$$

We have, since $\|e_j\|_1 = 1$, that $|\varphi(e_j)| \leq \|\varphi\|$ for all j. So $x = \{\varphi(e_j)\} \in \ell^\infty(S)$, and $\varphi = \gamma(x)$.

(5.6.7) Let $p \in (1,\infty)$ and $g \in \ell^q(\mathbb{N})$. Show that the map

$$\varphi : f \mapsto \sum_{k=1}^\infty f(k) g(k) \tag{5.21}$$

defines a bounded functional on $\ell^p(\mathbb{N})$ with norm $\|g\|_q$

Answer. Since φ is made up of pointwise evaluations, products by scalars, sums, and limits, all of which are linear, φ is linear itself. Hölder's inequality

$$|\varphi(f)| = \Big|\sum_{k=1}^\infty f(k)g(k)\Big| \leq \|f\|_p \|g\|_q$$

guarantees that φ is well-defined, it is bounded, and $\|\varphi\| \leq \|g\|_q$. Now let $f : \mathbb{N} \to \mathbb{C}$ be given by

$$f(k) = \theta(k)\, |g(k)|^{q-1},$$

where $\theta(k)\, g(k) = |g(k)|$. Then $\theta(k) \in \mathbb{T}$ for all k and

$$\begin{aligned} \|f\|_p^p &= \sum_k |f(k)|^p = \sum_k |g(k)|^{(q-1)p} \\ &= \sum_k |g(k)|^q = \|g\|_q^q < \infty, \end{aligned}$$

so $f \in \ell^p(\mathbb{N})$. And

$$\varphi(f) = \sum_k f(k)g(k) = \sum_k \theta(k)|g(k)|^{q-1}g(k) = \sum_k |g(k)|^q$$
$$= \|g\|_q^q = \|f\|_p \|g\|_q^{q-\frac{q}{p}} = \|f\|_p \|g\|_q,$$

so $\|\varphi\| = \|g\|_q$.

(5.6.8) Let $\varphi \in \ell^\infty(\mathbb{N})^*$. Show that there exist $\varphi_1, \varphi_\infty$ such that $\varphi = \varphi_1 + \varphi_\infty$, where $\varphi_1, \varphi_\infty \in \ell^\infty(\mathbb{N})^*$, $\varphi_\infty|_{c_0} = 0$, and $\varphi_1(x) = \langle x, y\rangle$ for some $y \in \ell^1(\mathbb{N})$ and all $x \in \ell^\infty(\mathbb{N})$.

Answer. Let $\varphi_0 = \varphi|_{c_0}$. By Exercise 5.6.2 there exists $y \in \ell^1(\mathbb{N})$ such that $\varphi_0(x) = \langle x, y\rangle$. Call φ_1 the extension to all of $\ell^\infty(\mathbb{N})$ with the same formula. Let $\varphi_\infty = \varphi - \varphi_1$. Then $\varphi_\infty|_{c_0} = 0$ by construction, φ_1 is bounded by construction, and φ_∞ is bounded because it is a linear combination of bounded functionals.

(5.6.9) Let $\mathcal{M}$ be the σ-algebra of subsets of $[0, 1]$ that are either finite or countable, or alternatively have finite or countable complement. Let μ be the counting measure.

(i) Show that μ is not σ-finite.

(ii) Show that $g(x) = x$ is not measurable.

(iii) Show that $\gamma : f \longmapsto \sum_x x f(x)$ defines a bounded linear functional on $L^1(\mu)$.

(iv) Show that γ is not of the form $\gamma(f) = \int f h\, d\mu$ for $h \in L^\infty(\mu)$.

(v) Conclude that $L^1(\mu)^* \neq L^\infty(\mu)$.

Answer.

(i) If $\mu(E) < \infty$, then E is finite. And a countable union of finite sets is countable, so not all of $[0, 1]$. Hence μ is not σ-finite.

(ii) Take $V = (0, 1/2)$, which is open. Then $g^{-1}(V) = V$, which is not measurable.

(iii) We have

$$\left|\sum_x xf(x)\right| \leq \sum_x x|f(x)| \leq \sum_x |f(x)| = \|f\|_1.$$

So γ is bounded and $\|\gamma\| \leq 1$. Linearity follows from the fact that the series will be absolutely convergent because $f \in L^1$.

(iv) If it were, given $t \in [0,1]$ let $f = 1_{\{t\}}$. Then, as $\mu(\{t\}) = 1$,

$$h(t) = \int fh\,d\mu = \gamma(f) = \sum_x xf(x) = t.$$

So $h(x) = x$ for all $x \in [0,1]$, but then it cannot be measurable.

(v) If they were equal, every γ would be of the form $\gamma(f) = \int fh\,d\mu$, which we showed above is impossible.

(5.6.10) Let $X = \{0,1\}$, and μ the measure given by $\mu(\{0\}) = 1$, $\mu(\{1\}) = \infty$. Show that $L^1(X)^* \neq L^\infty(X)$.

Answer. Here $L^1(X) = \{f : f(1) = 0\}$, so $L^1(X)$ is one-dimensional. On the other hand, $L^\infty(X)$ consists of all $f : \{0,1\} \to \mathbb{C}$, so it is two-dimensional.

(5.6.11) We consider the measure space $[0,1]$ with the counting measure.

(i) Show that $\ell^1[0,1]^* = \ell^\infty[0,1]$, though the measure is not σ-finite.

(ii) Let $\mathcal{X} = \{a : [0,1] \to \mathbb{C} : \operatorname{supp} a$ is countable and $a \in C_0(\operatorname{supp} a)\}$, where the topology is given by the $\|\cdot\|_\infty$ norm . Show that $\mathcal{X}$ is complete and that $\mathcal{X}^* = \ell^1[0,1]$.

Answer.

(i) This is Exercise 5.6.6.

(ii) It is clear that $\mathcal{X}$ is a vector space, so we check for completeness.

Let $\{a_n\} \subset \mathcal{X}$ be Cauchy. Since $|a_n(t) - a_m(t)| \leq \|a_n - a_m\|_\infty$, we get that for each t the number sequence is Cauchy, and so we can define $a(t) = \lim_n a_n(t)$. We get that a is bounded, because $\{a_n\}$ is bounded: if

$c \geq \|a_n\|$ for all n,

$$|a(t)| \leq |a(t) - a_n(t)| + |a_n(t)| \leq |a(t) - a_n(t)| + c.$$

As this works for all n and $a(t) - a_n(t) \to 0$, we get that $|a(t)| \leq c$ for all t. With the same idea we get that a is a norm limit: fix $\varepsilon > 0$ and choose n_0 such that $\|a_n - a_m\|_\infty < \varepsilon$ when $n, m \geq n_0$. Then

$$\begin{aligned} |a(t) - a_n(t)| &\leq |a(t) - a_m(t)| + |a_m(t) - a_n(t)| \\ &\leq |a(t) - a_m(t)| + \|a_m - a_n\|_\infty \\ &\leq |a(t) - a_m(t)| + \varepsilon. \end{aligned}$$

As the right-hand-side does not depend on n, we get that $|a(t)-a_n(t)| \leq \varepsilon$. This shows that $\|a-a_n\|_\infty \leq \varepsilon$, so $a_n \to a$. It remains to show that $\operatorname{supp} a$ is countable and that $a \in C_0(\operatorname{supp} a)$. We have $\operatorname{supp} a \subset \bigcup_n \operatorname{supp} a_n$, so countable. Given $\varepsilon > 0$ choose a_n with $\|a - a_n\|_\infty < \varepsilon/2$. Since $a_n \in C_0(\operatorname{supp} a_n)$ there exists $K \subset \operatorname{supp} a_n$, compact, such that $|a_n(a)| \leq \varepsilon/2$ for $t \notin K$. For such t,

$$|a(t)| \leq |a_n(t)| + |a_n(t) - a(t)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

So $K \cup \operatorname{supp} a$ is a compact subset of $\mathcal{X}$ such that $|a(t)| \leq \varepsilon$ outside of it. Therefore $a \in C_0(\operatorname{supp} a)$. So $\mathcal{X}$ is complete.

Now given $b \in \ell^1[0,1]$, we can define $\langle b, a\rangle = \sum_t b(t)a(t)$, and this is well defined because a is bounded and b is summable. Conversely, if $\varphi \in \mathcal{X}^*$, let $b(t) = \varphi(e_t)$. Let $F \subset [0,1]$ be finite. Let $\beta_t \in \mathbb{T}$ such that $\beta_t\, b(t)\, a(t) = |b(t)\, a(t)|$. Then

$$\begin{aligned} \sum_F |b(t)\, a(t)| &= \sum_F b(t)\, \beta_t a(t) = \sum_F \varphi(e_t)\, \beta_t a(t) \\ &= \varphi\Big(\sum_F \beta_t a(t)\, e_t\Big) \leq \|\varphi\| \, \Big\|\sum_F \beta_t a(t)\, e_t\Big\| = \|\varphi\|\, \|a\|_\infty. \end{aligned}$$

So the series $\sum_t b(t)a(t)$ converges absolutely, and in particular it converges unconditionally. Then, using the continuity of φ,

$$\sum_t b(t)a(t) = \sum_t \varphi(e_t)a(t) = \varphi\Big(\sum_t a(t)e_t\Big) = \varphi(a).$$

(5.6.12) Recall from Exercise 2.3.26 that a measure μ is **semifinite** if for every measurable E with $\mu(E) = \infty$ there exists $F \subset E$ with $0 < \mu(F) < \infty$. Show that for a measure space $(X, \mathcal{A}, \mu)$ the following statements are equivalent:

(i) μ is seminifite;

(ii) the canonical embedding $\Gamma : L^\infty(X) \to L^1(X)^*$ is injective.

When these conditions are satisfied, Γ is actually isometric.

Answer. Suppose first that μ is semifinite. Fix $f \in L^\infty(X)$. We know from Hölder's inequality that $\|\Gamma(f)\| \leq \|f\|_\infty$. Given $\varepsilon > 0$, let $E = \{|f| \geq \|f\|_\infty - \varepsilon\}$. By definition of the infinity norm we have $\mu(E) > 0$. By the semifiniteness there exists $F \subset E$, measurable, with $0 < \mu(F) < \infty$. Now we can define, if we write $f = e^{i\theta_f}|f|$ for an appropriate measurable function θ_f,

$$g = \frac{1}{\mu(F)}\, e^{-i\theta_f}\, 1_F.$$

Then $g \in L^1(X)$ with $\|g\|_1 = 1$, and

$$|\Gamma(f)g| = \Big|\int_X fg\,d\mu\Big| = \frac{1}{\mu(F)}\int_F |f|\,d\mu \geq \|f\|_\infty - \varepsilon.$$

As this can be done for all $\varepsilon > 0$ it follows that $\|\Gamma(f)\| = \|f\|_\infty$ and so Γ is isometric.

Conversely, suppose that μ is not semifinite. This means that there exists measurable E with $\mu(E) = \infty$ and $\mu(F) = 0$ for every measurable $F \subset E$ with $\mu(F) < \infty$. Let $f = 1_E \in L^\infty(X)$. Given $g \in L^1(X)$, put $F = \{g \neq 0\} \cap E$. Then $\mu(F) = 0$, for $\{|g| > 1/n\} \cap E$ is a finite measure subset of E—hence a nullset—and $F = \bigcup_n \{|g| > 1/n\} \cap E$. So

$$\Gamma(f)g = \int_X fg\,d\mu = \int_F g\,d\mu = 0.$$

Then $\Gamma(f) = 0$ and therefore Γ is not injective.

(5.6.13) *(This exercise is non-trivial, and it appears as an exercise by necessity of space; if tackled, it should be considered a project, and some guidance will likely be needed)* A measure μ on $(X, \mathcal{A})$ is **localizable** if it is semifinite and, in addition, given any collection $\mathcal{E}$ of measurable sets, it admits an **essential supremum**: that is a measurable H such that $\mu(E \setminus H) = 0$ for all $E \in \mathcal{E}$ (so $E \subset H$ a.e.) and if H' satisfies the same property then $\mu(H \setminus H') = 0$ (that is, H is the smallest such set up to nullsets).

(i) Show that if μ is semifinite and the canonical map $\Gamma : L^\infty(X) \to L^1(X)^*$ is surjective, then μ is localizable.

(ii) *(This part of the exercise is measure theory, but it is needed for the rest; it allows us to patch measurable functions—notably Radon-Nikodym derivatives—together as long as they agree almost everywhere on the intersection of their domains. It is not a trivial result so the reader might want to skip it and just use it)* Suppose that μ is localizable and that $\{f_j\}$ is a family of measurable real-valued functions, each with domain $D_j \in \mathcal{A}$ and such that $f_j = f_k$ a.e. on $D_j \cap D_k$. For each $q \in \mathbb{Q}$ and each j, let

$$E_{j,q} = \{x \in D_j : \; f_j(x) \geq q\},$$

and let E_q be an essential supremum of $\{E_{j,q} : \; j\}$. Put

$$h'(x) = \sup\{q : \; q \in \mathbb{Q}, \; x \in E_q\},$$

allowing for $\sup \varnothing = -\infty$. Finally, let $h(x) = h'(x)$ if $h'(x) \in \mathbb{R}$ and zero otherwise. Show that h is measurable and that $h|_{D_j} = f_j$ a.e. for all j.

(iii) Show that Γ is an isometric isomorphism if and only if μ is localizable.

Answer.

(i) Fix a collection $\mathcal{E}$ of measurable sets. Let $\mathcal{F}$ be the family of finite unions of elements of $\mathcal{E}$, ordered by inclusion. Note that $E \setminus G$ is a nullset for every $E \in \mathcal{E}$ if and only if it is a nullset for every $E \in \mathcal{F}$. We define $\psi : L^1(\mu) \to \mathbb{C}$ in the following way. Given $f \in L^1(\mu)$ with $f \geq 0$ a.e.,

$$\psi(f) = \lim_{E \in \mathcal{F}} \int_E f \, d\mu. \tag{AB.5.5}$$

The limits exists by monotonicity. Note also that $\psi(f) \leq \|f\|_1$ so the limit is always real. For arbitrary f we write $f = f_1 - f_2 + i(f_3 - f_4)$ with $f_1, f_2, f_3, f_4 \geq 0$ and $f_1 f_2 = f_3 f_4 = 0$ a.e. (so the four functions are unique up to a nullset) and we define $\psi(f) = \psi(f_1) - \psi(f_2) + i\psi(f_3) - i\psi(f_4)$. Linearity of ψ follows with the same idea as in page 133 of the Book. The estimate $|\psi(f)| \leq \|f\|_1$ is direct from (AB.5.5) (that holds for arbitrary $f \in L^1(\mu)$ by linearity of the limit).

Since by hypothesis Γ is surjective, there exists $g \in L^\infty(X)$ such that $\Gamma(g)f = \psi(f)$ for all f. Necessarily, $\|g\|_\infty \leq 1$; so we may assume without loss of generality that $|g| \leq 1$. The function g is necessarily real-valued and non-negative on each $E \in \mathcal{E}$, since by Exercise 2.5.22 we have that $g|_F$ is non-negative for each $F \subset E$ measurable and finite; by the semifiniteness this means that $g|_E$ is non-negative.

Let $H = \{g > 0\}$, measurable since g is. For any $F \in \mathcal{A}$ with $\mu(F) < \infty$ we have $1_F \in L^1(\mu)$ and so

$$\begin{aligned}\int_F g\,d\mu = \int_X 1_F\,g\,d\mu = \psi(1_F) = \lim_{E\in\mathcal{F}} \int_E 1_F\,d\mu \\ = \sup\{\mu(E\cap F):\ E\in\mathcal{F}\}.\end{aligned} \tag{AB.5.6}$$

Given $E \in \mathcal{E}$, if $\mu(E \setminus H) > 0$ then by semifiniteness there exists measurable $F \subset E \setminus H$ with $0 < \mu(F) < \infty$. Then, as $g = 0$ a.e. on H,

$$\mu(F) = \mu(E\cap F) \le \int_F g\,d\mu = 0,$$

a contradiction. Therefore $\mu(E\setminus H) = 0$. If now H' is another measurable set satisfying $\mu(E\setminus H') = 0$ for all $E \in \mathcal{E}$, suppose that $\mu(H\setminus H') > 0$. By the semifiniteness there exists measurable $F \subset H\setminus H'$ with $0 < \mu(F) < \infty$. For every $E \in \mathcal{E}$ we have

$$\begin{aligned}\mu(E\cap F) &= \mu\big([(E\setminus H')\cup(E\cap H')]\cap F\big) = \mu\big((E\cap H')\cap F\big) \\ &\le \mu(H'\cap F) = 0.\end{aligned}$$

Therefore, by (AB.5.6)

$$\int_F g\,d\mu = 0.$$

As $g > 0$ a.e. on F (because $F \subset H$), $\mu(F) = 0$. This is a contradiction, that shows that $\mu(H\setminus H') = 0$. Hence μ is localizable.

(ii) We first note that h' is measurable, for (using Corollary 2.4.5)

$$\{x:\ h'(x) > a\} = \bigcup_{q\in\mathbb{Q}\cap(a,\infty)} E_q \in \mathcal{A}$$

for all $a \in \mathbb{R}$. Given indices k, j and $q \in \mathbb{Q}$,

$$\begin{aligned}E_{j,q}\setminus(X\setminus(D_k\setminus E_{k,q})) &= E_{j,q}\cap\big(D_k\setminus E_{k,q}\big) \\ &\subset \{x\in E_j\cap E_k:\ f_j(x)\ne f_k(x)\}\end{aligned}$$

so it is a nullset. This looks overly complicated with the triple set difference, but it allows us to phrase things in terms such that the definition of essential supremum applies. Then, as E_q is the essential supremum,

$$(E_q\cap D_k)\setminus E_{k,q} = E_q\setminus(X\setminus(D_k\setminus E_{k,q}))$$

is also a nullset. We now form the union of the symmetric differences

$$H_k = \bigcup_{q\in\mathbb{Q}} \big(E_{k,q}\setminus(E_q\cap D_k)\big)\cup\big((E_q\cap D_k)\setminus E_{k,q}\big),$$

which again is a nullset, since $E_{k,q} \subset D_k$ and so $E_{k,q}\setminus(D_k\cap E_q) = E_{k,q}\setminus E_q$ is a nullset for all q. This implies that

$$\text{when } x\in D_k\setminus H_k,\qquad x\in E_q \iff x\in E_{k,q}. \tag{AB.5.7}$$

Indeed, if $x \in (E_q \cap D_k) \setminus H_k$, this means that $x \notin (E_q \cap D_k) \setminus E_{k,q}$, so $x \in E_{k,q}$. Conversely, if $x \in (E_{k,q} \cap D_k) \setminus H_k$, then $x \in X \setminus H_k \subset D_k \cap E_q$. In particular, $x \in E_q$.

Looking at (AB.5.7), if $x \in D_k \setminus H_k$, then $h'(x) \geq q$ if and only if $f_k(x) \geq q$. Doing this for every $q \in \mathbb{Q}$, we have shown that $h' = f_k$ a.e. on D_k.

Finally, we define

$$h(x) = \begin{cases} h'(x), & h'(x) \in \mathbb{R} \\ 0, & h'(x) \in \{-\infty, \infty\} \end{cases}$$

Then h is measurable because it is a modification of h' over a measurable set.

(iii) It remains to show that if μ is localizable, then Γ is surjective. Fix $\psi \in L^1(X)^*$ with $\|\psi\| = 1$. Let $\mathcal{A}_0 = \{F \in \mathcal{A} : \ \mu(F) < \infty\}$. Given $F \in \mathcal{A}_0$ define $\nu_F : \mathcal{A} \to \mathbb{C}$ by

$$\nu_F(E) = \psi(1_{E\cap F}).$$

We have $\nu_F(\varnothing) = 0$. If $E_1, E_2 \in \mathcal{A}$ are disjoint, then

$$\begin{aligned} \nu_F(E_1 \cup E_2) &= \psi(1_{(E_1\cup E_2)\cap F}) = \psi(1_{E_1\cap F} + 1_{E_2\cap F}) \\ &= \psi(1_{E_1\cap F}) + \psi(1_{E_2\cap F}) = \nu_F(E_1) + \nu_F(E_2). \end{aligned}$$

Therefore ν_F is additive. Now let $\{E_k\} \subset \mathcal{A}$ be a countable pairwise disjoint family. Then

$$\sum_k \mu(E_k \cap F) = \mu\Big(\bigcup_k (E_k \cap F)\Big) < \infty.$$

Hence

$$\Big\| \sum_{k>n} 1_{E_k\cap F} \Big\|_1 = \sum_{k>n} \mu(E_k \cap F) \xrightarrow[n\to\infty]{} 0,$$

so the series $\sum_k 1_{E_k\cap F}$ converges in $L^1(\mu)$. Then, as ψ is continuous,

$$\nu_F\Big(\bigcup_k E_k\Big) = \psi\Big(\sum_k 1_{E_k\cap F}\Big) = \sum_k \psi(E_k \cap F) = \sum_k \nu_F(E_k).$$

Thus ν_F is a complex measure. Using again that ψ is bounded,

$$|\nu_F(E)| = |\psi(1_{E\cap F})| \leq \|1_{E\cap F}\|_1 = \mu(E \cap F).$$

So $\nu_F \ll \mu$. By Theorem 2.10.10 there exists a Radon–Nikodym derivative $h_F \in L^1(X)$ with

$$\nu_F(E) = \int_E h_F \, d\mu, \qquad E \in \mathcal{A}.$$

For any E with $\nu_F(E) \neq 0$ we have $\mu(E) > 0$ by the absolute continuity and then

$$\left|\frac{1}{\mu(E)} \int_E h_F \, d\mu\right| = \frac{1}{\mu(E)} |\psi(1_{F\cap E})| \leq \frac{\|1_E\|_1}{\mu(E)} = 1.$$

It follows from Exercise 2.5.22 that $|h_F| \leq 1$ a.e. So $h_F \in L^\infty(\mu)$.

Now let $F_1, F_2 \in \mathcal{A}_0$. We claim that $h_{F_1} = h_{F_2}$ a.e. on $F_1 \cap F_2$. Indeed, given measurable $E \subset F_1 \cap F_2$, since $1_{E\cap F_1} = 1_E = 1_{E\cap F_2}$,

$$\int_E h_{F_1} \, d\mu = \psi(1_{E\cap F_1}) = \psi(1_{E\cap F_2}) = \int_E h_{F_2} \, d\mu.$$

So the function $h_{F_1} - h_{F_2}$ has integral 0 on every measurable subset of the finite measure space $F_1 \cap F_2$. By Exercise 2.5.22, $h_{F_1} = h_{F_2}$ a.e. on $F_1 \cap F_2$. By (ii) there exists a measurable function $h : X \to \mathbb{C}$ such that $h|_F = h_F$ a.e. for every $F \in \mathcal{A}_0$.

Given $F \in \mathcal{A}_0$,

$$\{x \in F : \ |h(x)| > 1\} \subset \{|h_F| > 1\} \cup \{x \in F : \ h(x) \neq h_F(x)\}.$$

The two sets on the right are nullsets, so the set on the left is also a nullset. This forces $\{|h| > 1\}$ to be a nullset; if it were not, by semifiniteness there would exists $F \in \{|h| > 1\}$ with $0 < \mu(F) < \infty$, forcing $\{x \in F : \ |h| > 1\}$ to not be a nullset. Thus $h \in L^\infty(X)$ and $\|h\|_\infty \leq 1$. Now, for any $F \in \mathcal{A}_0$,

$$\Gamma(h)1_F = \int_X 1_F h \, d\mu = \int_F h_F \, d\mu = \nu_F(F) = \psi(1_F).$$

By linearity we get that $\Gamma(h)f = \psi(f)$ for every simple f. As both $\Gamma(h)$ and ψ are continuous, it follows that $\Gamma(h) = \psi$ by Proposition 2.8.17. Finally, if $\|\psi\| \neq 1$, we can scale and apply the proof to the scaled version.

(5.6.14) For $0 < p < 1$, show that $\|f\|_p = \left(\sum_j |f(j)|^p\right)^{1/p}$ is not a norm.

Answer. Fix $k \neq j$ and let $f = \delta_k$ and $g = \delta_j$. Then, as $0 < p < 1$,

$$\|f + g\|_p = 2^{1/p} > 2 = \|f\|_p + \|g\|_p.$$

So the triangle inequality fails and thus the p-norm is not a norm when $p < 1$.

(5.6.15) Show that if $0 < p < 1$ and q satisfies $\frac{1}{p} + \frac{1}{q} = 1$, we have the **Reverse Hölder Inequality**: given $f \in \ell^p(\mathbb{N})$, $g \in \ell^q(\mathbb{N})$,

$$\sum_k |f(k)g(k)| \geq \|f\|_p \|g\|_q.$$

Answer. If $\sum_k |f(k)g(k)| = \infty$, the inequality is trivial. Now assume—by multiplying f by a suitable constant—that $\sum_k |f(k)g(k)| = 1$. Assume also—now multiplying g by a suitable constant—that $\|g\|_q = 1$. Let $h(k) = f(k)g(k)$; our assumptions imply that $h \in \ell^1(\mathbb{N})$, so $h^p \in \ell^{1/p}(\mathbb{N})$. One consequence of $p < 1$ is that $q = p/(p-1) < 0$. Then $g \in \ell^q(\mathbb{N})$ implies that $g(k) \neq 0$ for all k.

We have $1/|g|^p = |g|^{-p} \in \ell^{1/(1-p)}(\mathbb{N})$, since $|g|^{-p/(1-p)} = |g|^q$. Using the usual Hölder inequality for $1/p > 1$ (where the conjugate exponent is $1/(1-p)$),

$$\begin{aligned}\|f\|_p^p \|g\|_q^p = \|f\|_p^p = \sum_k |f(k)|^p = \sum_k |h(k)|^p \left|\frac{1}{g(k)}\right|^p \\ \leq \Big(\sum_k |h(k)|\Big)^p \Big(\sum_k |g(k)|^q\Big)^{1/(1-p)} = \Big(\sum_k |h(k)|\Big)^p.\end{aligned}$$

This is exactly

$$\|f\|_p \|g\|_q \geq \|fg\|_1.$$

(5.6.16) Prove that $\psi : L^q(X) \to L^p(X)^*$, as defined in (5.17), is linear and bounded, and $\|\psi(g)\| \leq \|g\|_q$.

Answer. Linearity follows directly from linearity of integrals. We have, via Hölder,

$$|\psi(g)f| \leq \int_X |g|\,|f|\,d\mu \leq \|g\|_q\,\|f\|_p.$$

This implies $\|\psi(g)\| \leq \|g\|_q$.

(5.6.17) Fix $p \in [1, \infty)$. Let $\mathcal{X} = C[0,1]$ seen as a normed space with the p-norm. Show that the functional $f \longmapsto f(0)$ is unbounded.

Answer. Let $f_n(t) = (n - n^2 t)^{1/p}\, 1_{[0,\frac{1}{n}]}$. Then f_n is continuous, and

$$\|f_n\|_p^p = \int_0^{1/n} (n - n^2 t)\, dt = \frac{1}{2}.$$

Meanwhile, $f_n(0) = n^{1/p}$ becomes arbitrarily large for big enough n.

(5.6.18) For a fixed $0 < p < 1$, consider the vector space

$$\ell^p(\mathbb{N}) = \{x : \mathbb{N} \to \mathbb{C} : \sum_j |x_j|^p < \infty\},$$

with $d_p(x,y) = \sum_j |x_j - y_j|^p$.

(i) Show that d_p is a metric, so $\ell^p(\mathbb{N})$ is a TVS.

(ii) Show that the dual of $\ell^p(\mathbb{N})$ is $\ell^\infty(\mathbb{N})$.

Answer.

(i) We have, for $a, b \geq 0$ and $p' = 1/p > 1$, $a^{p'} + b^{p'} \leq (a+b)^{p'}$ by the binomial series. Thus $a^{1/p} + b^{1/p} \leq (a+b)^{1/p}$. Applying this to $a = |1-r|^p$, $b = |r-t|^p$,

$$\begin{aligned}|1-t|^p = |1-r+r-t|^p &\leq (|1-r| + |r-t|)^p = (a^{1/p} + b^{1/p})^p \\ &\leq a + b = |1-r|^p + |r-t|^p.\end{aligned}$$

Thus

$$\begin{aligned}d_p(x,y) = \sum_n |x_n - y_n|^p &\leq \sum_n |x_n - z_n|^p + |z_n - y_n|^p \\ &= d_p(x,z) + d_p(z,y).\end{aligned}$$

So d_p satisfies the triangle inequality. As $d_p(x,y) = d_p(y,x)$, it is a metric.

(ii) As $0 < p < 1$, we have that $\sum_n |x_n|^p < \infty$ implies that $\sum_n |x_n| < \infty$ for any $x \in \ell^p(\mathbb{N})$ (since $|x_n| < 1$ eventually), so $x = \sum_n x_n e_n$, where $\{e_n\}$ is the canonical basis $e_n = \delta_n$. Moreover, if $d_p(x,0) < 1$, then $|x_n| < 1$ for all j and hence $\sum_n |x_n| \leq \sum_n |x_n|^p$. Thus, for $y \in \ell^\infty(\mathbb{N})$ and $x \in \ell^p(\mathbb{N})$ with $d_p(x,0) < 1$ (i.e., $x \in B_1(0)$),

$$|\langle y, x\rangle| = \Big|\sum_n y_n x_n\Big| \leq \sum_n |y_n x_n| \leq \|y\|_\infty \sum_n |x_n| \leq \|y\|_\infty\, d_p(x,0).$$

It follows that y, as a linear functional, is continuous at 0, and thus continuous.

Let $\phi : \ell^p(\mathbb{N}) \to \mathbb{C}$ be a continuous linear functional. For $x \in \ell^p(\mathbb{N})$, because the p-norm guarantees that the tails of the series converges, we have that $x = \sum_n x_n e_n$ with the series converging in the d_p metric. Therefore

$$\phi(x) = \sum_n x_n \phi(e_n).$$

Suppose that $\{\phi(e_n)\}$ is not bounded. Choose a subsequence such that $|\phi(e_{n_k})| \geq 2^k$. Let $z = \sum_k 2^{-k} e_{n_k}$. We have $z \in \ell^p(\mathbb{N})$, since $\sum_k (2^{-k})^p = (2^p - 1)^{-1}$. And

$$\phi(z) = \sum_k 2^{-k} 2^k = \infty,$$

a contradiction. So $y = \{\phi(e_n)\} \in \ell^\infty(\mathbb{N})$ and $\phi(x) = \langle y, x \rangle$.

(5.6.19) *(another example of a topological vector space with trivial dual)* Let $\mathcal{X} = B[0,1]$, the bounded Borel functions modulo almost everywhere equality. On this complex vector space, define

$$d(f,g) = \int_0^1 \frac{|f-g|}{1+|f-g|}.$$

(i) Show that d is a distance.

(ii) Show that $(\mathcal{X}, d)$ is a topological vector space.

(iii) Show that $\mathcal{X}^* = \{0\}$.

Answer.

(i) It follows readily from the definition that $d(f,g) = d(g,f) \geq 0$ for all $f, g \in \mathcal{X}$. And if $d(f,g) = 0$, then $|f-g| = 0$ a.e. since it is the numerator of an almost everywhere zero function. So it remains to show the triangle inequality. Consider the function $p : [0,\infty) \to [0,\infty)$ given by $p(t) = \frac{t}{1+t}$. This function is differentiable, and

$$p'(t) = \Big(1 - \frac{1}{1+t}\Big)' = \frac{1}{(1+t)^2} > 0.$$

So p is increasing. Given $f, g, h \in \mathcal{X}$, as $|f-g| \leq |f-h| + |h-g|$,

$$\begin{aligned} d(f,g) &= \int_0^1 \frac{|f-g|}{1+|f-g|} \leq \int_0^1 \frac{|f-h|+|h-g|}{1+|f-h|+|h-g|} \\ &= \int_0^1 \frac{|f-h|}{1+|f-h|+|h-g|} + \int_0^1 \frac{|h-g|}{1+|f-h|+|h-g|} \\ &\leq \int_0^1 \frac{|f-h|}{1+|f-h|} + \int_0^1 \frac{|h-g|}{1+|h-g|} = d(f,h) + d(h,g). \end{aligned}$$

(ii) We need to show that points are closed and that the vector space operations are continuous. That points are closed we get for free since $\mathcal{X}$ is a metric space. The continuity of addition follows from the fact that d is translation invariant. Namely, using the translation invariance and the triangle inequality,

$$\begin{aligned} d(f_1+f_2, g_1+g_2) &= d(f_1-g_1, f_2-g_2) \\ &\leq d(f_1-g_1, 0) + f(f_2-g_2, 0), \end{aligned}$$

and so addition is simultaneously continuous in both variables. And if $\alpha_n \to \alpha$ for scalars $\{\alpha_n\}$ and α,

$$\begin{aligned} d(\alpha_n f, \alpha f) = d((\alpha_n-\alpha)f, 0) &= \int_0^1 \frac{|(\alpha_n-\alpha)f|}{1+|(\alpha_n-\alpha)f|} \\ &\leq |\alpha_n-\alpha| \, \|f\|_\infty \xrightarrow[n\to\infty]{} 0. \end{aligned}$$

(iii) Let $V \subset \mathcal{X}$ be nonempty, open, and convex. By translating if needed, we assume that $0 \in V$. So there exists $\delta > 0$ such that $B_\delta(0) \subset V$. Choose $n \in \mathbb{N}$ such that $n > \frac{1}{\delta}$. Fix $f \in \mathcal{X}$. Let $s_0 = 0$ and $s_k = \frac{1}{n}$, for $k = 1, \ldots, n$. Put

$$g_k = n\, f\, 1_{[s_{k-1}, s_k]}$$

Then $g_k \in B[0,1]$ and

$$d(g_k, 0) = \int_{s_{k-1}}^{s_k} \frac{n|f|}{1+n|f|} \leq s_k - s_{k-1} = \frac{1}{n} < \delta.$$

That is, $g_k \in B_\delta(0) \subset V$. Then, as V is convex,

$$f = \sum_{k=1}^{n} \frac{1}{n} g_k \in V.$$

Thus $V = \mathcal{X}$. Now if $\varphi \in \mathcal{X}^*$ the continuity and linearity of φ imply that $\varphi^{-1}(B_1^{\mathbb{C}}(0))$ is nonempty, open, and convex. Then $\varphi^{-1}(B_1^{\mathbb{C}}(0)) = \mathcal{X}$. This means that $|\varphi(x)| < 1$ for all $x \in \mathcal{X}$, which by linearity can only happen if $\varphi = 0$.

5.7. The Hahn–Banach Theorem

(5.7.1) Write a complete proof of Corollary 5.7.6.

Answer. If we define $q(x) = \|\varphi\| \, \|x\|$, then q is a seminorm and for all $x \in W$ we have $|\varphi(x)| \leq \|\varphi\| \, \|x\| = q(x)$, so Theorem 5.7.5 applies. We get $\tilde{\varphi} : V \to \mathbb{C}$, linear, with $\tilde{\varphi}|_W = \varphi$ and $|\tilde{\varphi}(x)| \leq q(x)$ for all $x \in V$.

Writing this as $|\tilde{\varphi}(x)| \leq \|\varphi\| \, \|x\|$ tells us that $\tilde{\varphi}$ is bounded and that $\|\tilde{\varphi}\| \leq \|\varphi\|$. As $\tilde{\varphi}|_W = \varphi$ we also have $\|\varphi\| \leq \|\tilde{\varphi}\|$, and therefore $\|\tilde{\varphi}\| = \|\varphi\|$.

(5.7.2) Let $\mathcal{X}$ be a locally convex space and $\mathcal{Y} \subset \mathcal{X}$ a subspace with $\dim \mathcal{Y} < \infty$. Show that $\mathcal{Y}$ is (topologically) complemented.

Answer. Let $e_1, \ldots, e_n$ be a basis of $\mathcal{Y}$. Define maps, for $j = 1, \ldots, n$, $\varphi_j : \mathcal{Y} \to \mathbb{C}$ by $\varphi_j(\sum_{j=1}^n \alpha_j e_j) = \alpha_j$. This is well-defined, since $e_1, \ldots, e_n$ form a basis. We also have that φ_j is linear, and continuous since $\dim \mathcal{Y} < \infty$ (Theorem 5.4.16). By Corollary 5.7.24 there exist $\psi_1, \ldots, \psi_n \in \mathcal{X}^*$ with $\psi_j|_\mathcal{Y} = \varphi_j$.

Let $P : \mathcal{X} \to \mathcal{X}$ be given by

$$Px = \sum_{j=1}^n \psi_j(x) \, e_n.$$

This map is linear and continuous, since each ψ_j is. We also have $Py \in \mathcal{Y}$ for all $x \in \mathcal{X}$, and $Py = y$ for all $y \in \mathcal{Y}$. So P is a continuous projection onto $\mathcal{Y}$ and by Proposition 5.4.19 the subspace $\mathcal{Y}$ is topologically complemented.

(5.7.3) Let $\mathcal{X}$ be a locally convex space and $V \subset \mathcal{X}$ a closed subspace. Show that if $m = \dim \mathcal{X}/V < \infty$, then V is topologically complemented with complement of dimension m. Give an example to show that V need not be complemented if it is not closed.

Answer. Let $y_1 + V, \ldots, y_m + V$ be a basis of $\mathcal{X}/V$. Take $\varphi_1, \ldots, \varphi_m \in (\mathcal{X}/V)^*$ to be the dual basis, i.e., $\varphi_j(y_k + V) = \delta_{kj}$ (Proposition 1.7.11), and define

$\psi_1, \ldots, \psi_m \in \mathcal{X}^*$ as $\psi_k = \varphi_k \circ q$, where $q : \mathcal{X} \to \mathcal{X}/V$ is the quotient map; the ψ_k are continuous since q and φ_k are (see Proposition 5.3.13).

Let $W = \operatorname{span}\{y_1, \ldots, y_m\}$ and $P : \mathcal{X} \to \mathcal{X}$ given by

$$Px = \sum_{k=1}^{m} \psi_k(x)\, y_k.$$

We have $\dim W = m$, since $y_1, \ldots, y_m$ inherit the linear independence from $y_1 + V, \ldots, y_m + V$. Then P is a bounded projection onto W. The continuity of P follows from that of $\psi_1, \ldots, \psi_m$. The range of P is in W by construction, and if $x \in W$ then $x = \sum_{k=1}^m x_k y_k$ and then

$$Px = \sum_{k=1}^{m} \sum_{j=1}^{m} x_k \psi_j(y_k) y_j = \sum_{k=1}^{m} x_k y_k = x;$$

so P is a projection onto W.

The proof will be complete if we show that $V = (I - P)\mathcal{X}$. If $x = (I - P)x$, then $Px = 0$, so $\psi_k(x) = 0$ for all k. That is, $\varphi_k(x + V) = 0$ for all k. The maps $\varphi_1, \ldots, \varphi_m$ separate points, since they are a basis of $(\mathcal{X}/V)^*$; hence, $x + V = 0$ and thus $x \in V$. Conversely, by definition of q, $P|_V = 0$, so $V \subset (I - P)\mathcal{X}$.

For an example when V is not closed, let $\mathcal{X}$ be any infinite-dimensional Banach space and let $\varphi : \mathcal{X} \to \mathbb{C}$ be any unbounded linear functional. Then $V = \ker \varphi$ is a subspace of $\mathcal{X}$ with $\dim \mathcal{X}/V = 1$, and V is dense so it cannot be topologically complemented.

(5.7.4) Show that $\ell^\infty(\mathbb{N})^* \neq \ell^1(\mathbb{N})$, by proving that the equality would imply that $\ell^\infty(\mathbb{N})/c_0$ has trivial dual, in contradiction with Corollary 5.7.7.

Answer. Suppose that every bounded functional on $\ell^\infty(\mathbb{N})$ comes from an element of $\ell^1(\mathbb{N})$. Denote by $\pi : \ell^\infty(\mathbb{N}) \to c_0$ the quotient map. Let $\varphi \in (\ell^\infty(\mathbb{N})/c_0)^*$. Then $\varphi \circ \pi \in \ell^\infty(\mathbb{N})^*$ (note that the quotient map is bounded by definition of the quotient norm). By hypothesis there exists $y \in \ell^1(\mathbb{N})$ such that $\varphi(\pi(x)) = \langle x, y \rangle$ for all $x \in \ell^\infty(\mathbb{N})$. Since $\pi|_{c_0} = 0$, we get that $\langle x, y \rangle = 0$ for all $x \in c_0$. This implies that $y = 0$, as $\langle e_k, y \rangle = y_k$ for each $k \in \mathbb{N}$. So $\varphi = 0$.

(5.7.5) Let $\mathcal{X}$ be a TVS and $\varphi : \mathcal{X} \to \mathbb{C}$ linear. Show that if $\varphi|_V = 0$ on some neighbourhood V of 0, then $\varphi = 0$.

Answer. Let $x \in \mathcal{X}$. By continuity of the product by scalars we have $\frac{1}{n}\,x \to 0$. As V is an open neighbourhood of 0, there exists n such that $\frac{1}{n}x \in V$. Then $0 = \varphi(\frac{1}{n}x) = \frac{1}{n}\,\varphi(x)$, so $\varphi(x) = 0$.

(5.7.6) *(This is part of the proof of Proposition 5.7.12)* Show that if $\mathcal{X}$ is a TVS and $\varphi : \mathcal{X} \to \mathbb{C}$ is linear and open, then $\operatorname{Re}\varphi : \mathcal{X} \to \mathbb{R}$ is real linear and open.

Answer. Let $V \subset \mathcal{X}$ be open. Then $\varphi(V) \subset \mathbb{C}$ is open. If $t \in \operatorname{Re}\varphi(V)$, there exists $s \in \mathbb{R}$ with $t + is \in \varphi(V)$. From $\varphi(V)$ open we get that there exists $\delta > 0$ such that $B_\delta(t+is) \subset \varphi(V)$. Now if $r \in \mathbb{R}$ with $|r - t| < \delta$, then $|r + is - (t + is)| = |r - t| < \delta$, so $r + is \in \varphi(V)$ and then $r \in \operatorname{Re}\varphi(V)$; thus $(t - \delta, t + \delta) \subset \operatorname{Re}\varphi(V)$, showing that $\operatorname{Re}\varphi(V)$ is open in $\mathbb{R}$.

(5.7.7) *(This is part of the proof of Proposition 5.7.12)* Prove that if $\mathcal{X}$ is a TVS and $\varphi : \mathcal{X} \to \mathbb{C}$ is a linear functional, then φ is an open map if and only if for every neighbourhood Z of $0 \in \mathcal{X}$, $\varphi(Z)$ contains $0 \in \mathbb{C}$ as an interior point.

Answer. If φ is an open map, then $\varphi(Z)$ is open, and so 0 is interior. Conversely, let $V \subset \mathcal{X}$ be open. Fix $v \in V$; then $-v+V$ is an open neighbourhood of 0. By hypothesis, 0 is an interior point of $\varphi(-v+V) = -\varphi(v)+\varphi(V)$. Thus $\varphi(v)$ is an interior point of $\varphi(V)$; indeed, there exists $\delta > 0$ such that $|t| < \delta$ implies $t \in -\varphi(v) + \varphi(V)$, so $\varphi(v) + t \in \varphi(V)$ and hence $B_\delta(\varphi(v)) \in \varphi(V)$.

(5.7.8) Prove that the set Z used at the beginning of the proof of Theorem 5.7.13 is nonempty, open, and convex.

Answer. Since V is open, we have

$$Z_0 = \bigcup_{w \in W} (V - w + z_0),$$

a union of open sets, so open. And as both V, W are convex, if $v_1, v_2 \in V$, $w_1, w_2 \in W$ and $t \in [0, 1]$,

$$\begin{aligned} &t(v_1 - w_1 + x_0) + (1-t)(v_2 - w_2 + x_0) \\ &\qquad\qquad = tv_1 + (1-t)v_2 - [tw_1 + (1-t)w_2] + x_0 \in Z_0, \end{aligned}$$

so Z_0 is convex. We also have that $-Z_0$ is open an convex, and hence $Z = Z_0 \cap (-Z_0)$ is open and convex.

Finally, $Z \neq \varnothing$ since $0 \in Z$.

(5.7.9) Is the condition "V open" necessary for Theorem 5.7.13?

Answer. Yes. The result can fail when neither V nor W is open, already in $\mathbb{R}^2$. For instance, let $\mathcal{X} = \mathbb{R}^2$ and

$$V = \{(0,0)\} \cup \{(x,y): \ x > 0\}, \qquad W = \{(0,1)\}.$$

Both V, W are convex, none is open. Let ψ be a real linear functional and $c \in \mathbb{R}$ with $\psi(x,y) < c \leq \psi(0,1)$ for all $x > 0$. Linear functionals in $\mathbb{R}^2$ are of the form $\psi(x,y) = ax + by$ for some $a, b \in \mathbb{R}$. So we need

$$ax + by < c \leq b, \qquad x > 0, \ y \in \mathbb{R}.$$

Taking first the limit as $x \searrow 0$, we have $by \leq c \leq b$ for all $y \in \mathbb{R}$. This can only happen if $b = c = 0$. So now our inequality is $ax < 0$ for all $x > 0$. This works with any $a < 0$. But we also have $(0,0) \in V$, and this requires the inequality $0 < 0$, which is impossible.

(5.7.10) Show that if $\mathcal{X}$ is a locally convex space and $M \subset \mathcal{X}$ is a subspace, then the closure $\overline{M}$ of M is

$$\overline{M} = \bigcap_{f \in K_M} \ker f,$$

where $K_M = \{f \in \mathcal{X}^* : \ M \subset \ker f\}$.

Answer. Let $f \in K_M$. Then $M \subset \ker f$. As f is bounded, $\ker f$ is closed; so $\overline{M} \subset \ker f$. As this works for all $f \in K_M$, we have that $\overline{M} \subset \bigcap_{f \in K_M} \ker f$. Now let $y \in \bigcap_{f \in K_M} \ker f \setminus \overline{M}$. If such y exists, by Hahn–Banach (Corollary 5.7.19) there exists $g \in \mathcal{X}^*$ such that $g|_{\overline{M}} = 0$ and $g(y) = 1$. But this is impossible since $g \in K_M$, which requires $y \in \ker g$. So no such y can exist, showing that $\overline{M} = \bigcap_{f \in K_M} \ker f$.

(5.7.11) Use Exercise 5.7.10 to show that if $\mathcal{X}$ is locally convex and $M \subset \mathcal{X}$ is a subspace, then M is dense in $\mathcal{X}$ if and only $\{f \in \mathcal{X}^* : \ f = 0 \text{ on } M\} = \{0\}$.

Answer. Assume first that M is dense. Then by Exercise 5.7.10 we get that $\mathcal{X} = \bigcap_{f \in K_M} \ker f$. If $f = 0$ on M, then $f = 0$ on $\overline{M}$ by continuity. Thus $K_M = \{0\}$. Conversely, if $K_M = \{0\}$, then by Exercise 5.7.10 we have $\overline{M} = \ker 0 = \mathcal{X}$.

(5.7.12) In Remark 5.7.26 it is shown that the map $f \longmapsto \int_0^1 f$ is not continuous on $\mathcal{X} = L^p[0,1]$, $0 < p < 1$. Prove this explicitly by finding a sequence $\{f_n\} \subset \mathcal{X}$ such that $d_p(f_n, 0) \to 0$ while $\int_0^1 |f_n| \nearrow \infty$.

Answer. Let $f_n = \frac{1}{\sqrt{\log n}}\, t^{-1}\, 1_{[\frac{1}{n},1]}$. Then

$$\begin{aligned} d_p(f_n, 0) &= \frac{1}{(\log n)^{p/2}} \int_{1/n}^1 t^{-p}\, dt \\ &= \frac{1}{(\log n)^{p/2}} \left(\frac{1}{1-p} - \frac{1}{(1-p)n^{1-p}} \right) \xrightarrow[n\to\infty]{} 0. \end{aligned}$$

Meanwhile,

$$\int_0^1 f_n = \frac{1}{(\log n)^{1/2}} \int_{1/n}^1 t^{-1}\, dt = (\log n)^{1/2} \xrightarrow[n\to\infty]{} \infty.$$

(5.7.13) Let $\mathcal{X}$ be a locally convex space, and $x, y \in \mathcal{X}$. Show that if $\operatorname{Re}\varphi(x) = \operatorname{Re}\varphi(y)$ for all $\varphi \in \mathcal{X}^*$, then $x = y$.

Answer. Given $\varphi \in \mathcal{X}^*$, consider $\psi = i\varphi \in \mathcal{X}^*$. Then

$$\operatorname{Im}\varphi(x) = -\operatorname{Re} i\psi(x) = -\operatorname{Re} i\psi(y) = \operatorname{Im}\varphi(y).$$

Thus $\varphi(x) = \varphi(y)$ for all $\varphi \in \mathcal{X}^*$. By Corollary 5.7.7,

$$\|x - y\| = \sup\{|\varphi(x-y)| : \ \varphi \in \mathcal{X}^*, \ \|\varphi\| = 1\} = 0,$$

so $x = y$.

Here is an alternative argument. If $x \neq y$, we apply Theorem 5.7.18 to the compact sets $\{x\}$ and $\{y\}$, so there exists $\varphi \in \mathcal{X}^*$ with $\operatorname{Re}\varphi(x) < \operatorname{Re}\varphi(y)$. This contradicts our hypothesis, so it follows that $x = y$.

(5.7.14) Let $\mathcal{Y}$ be a balanced, convex, closed subset of a locally convex space $\mathcal{X}$, and let $x \in \mathcal{X} \setminus \mathcal{Y}$. Show that there exists $\varphi \in \mathcal{X}^*$ with $\varphi(x) > 1$ and $|\varphi(y)| \leq 1$ for all $y \in \mathcal{Y}$.

Answer. Apply Theorem 5.7.18 to the sets $\{x\}$ and $\mathcal{Y}$ to obtain $\varphi' \in \mathcal{X}^*$ and $c \in \mathbb{R}$ with

$$\operatorname{Re}\varphi'(x) < c < \operatorname{Re}\varphi'(y), \quad y \in \mathcal{Y}.$$

As $\mathcal{Y}$ is convex and balanced, $y \in \mathcal{Y}$ if and only if $-y \in \mathcal{Y}$, so we also have

$$\operatorname{Re}\varphi'(y) < c' < \operatorname{Re}\varphi'(-x), \quad y \in \mathcal{Y},$$

where $c' = -c$. Also from $\mathcal{Y}$ being balanced and convex we get $0 \in \mathcal{Y}$, so $c' > 0$. Now let $\varphi'(x) = |\varphi'(x)|\, e^{i\theta}$ be the polar form. Let $\varphi = (e^{-i\theta}/c')\,\varphi'$. Then

$$\varphi(x) = \frac{e^{-i\theta}}{c'}\,\varphi'(x) = \frac{|\varphi'(x)|}{c'} \geq \frac{\operatorname{Re}\varphi'(-x)}{c'} > 1.$$

For any $y \in \mathcal{Y}$, write $\varphi'(y) = |\varphi'(y)|e^{i\gamma}$ to get $e^{-i\gamma}\varphi'(y) \geq 0$, and

$$|\varphi(y)| = \frac{|\varphi'(y)|}{c'} = \frac{e^{-i\gamma}\varphi'(y)}{c'} = \frac{\varphi'(e^{-i\gamma}y)}{c'} = \frac{\operatorname{Re}\varphi'(e^{-i\gamma}\,y)}{c'} \leq \frac{c'}{c'} = 1.$$

CHAPTER 6

Huge consequences of Baire's Category Theorem

6.1. Bounded linear operators

(6.1.1) Prove the First Isomorphism Theorem for linear operators: given vector spaces $\mathcal{X}, \mathcal{Y}$ and $T : \mathcal{X} \to \mathcal{Y}$ linear, then

$$\mathcal{X} / \ker T \simeq \operatorname{ran} T$$

canonically.

Answer. Define $\tilde{T} : \mathcal{X} / \ker T \to \operatorname{ran} T$ by $\tilde{T}(x + \ker T) = Tx$. This is well-defined since T is zero on $\ker T$. Linearity is straightforward, as is surjectivity: given $Tx \in \operatorname{ran} T$, $Tx = \tilde{T}(x + \ker T)$. As for injectivity, if $\tilde{T}(x + \ker T) = 0$, this means that $Tx = 0$ and so $x \in \ker T$, which is the same as saying that $x + \ker T = 0$.

(6.1.2) Prove Proposition 6.1.2.

Answer. If T is continuous, it is continuous at 0. If T is continuous at 0 and $x_j \to x$, then $x_j - x \to 0$, so $Tx_j - Tx = T(x_j - x) \to 0$, and then T is continuous at x. If T is continuous at x, with the same idea we can show that T is continuous at 0; this means that for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\|x\| < \delta$ implies $\|Tx\| < \varepsilon$. Then for $x \in \mathcal{X}$ we get that $\|\delta x/(2\|x\|)\| < \delta$, so $\|T(\delta x/(2\|x\|)\| < \varepsilon$; and this gives us

$$\|Tx\| \leq \frac{2\varepsilon}{\delta}\|x\|.$$

(6.1.3) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, and $T : \mathcal{X} \to \mathcal{Y}$ be linear. Show that the following statements are equivalent:

(i) T is unbounded;

(ii) there exists a sequence $\{x_n\} \subset \mathcal{X}$ such that $\|x_n\| = 1$ and $\|Tx_n\| > n$ for all n;

(iii) there exists a sequence $\{x_n\} \subset \mathcal{X}$ such that $x_n \to 0$ and $\|Tx_n\| = 1$ for all n.

Answer. This is a direct generalization of Exercise 5.5.10, and the argument is entirely similar.

If T is unbounded, for each n there exists $z_n \in \mathcal{X}$ with $\frac{\|Tz_n\|}{\|z_n\|} > n$. Then $x_n = \frac{z_n}{\|z_n\|}$ satisfies $\|x_n\| = 1$ and $\|Tx_n\| > n$ for all n.

If now we assume that we have a sequence $\{x_n\}$ with $\|x_n\| = 1$ and $\|Tx_n\| > n$ for all n, let $z_n = \frac{x_n}{\|Tx_n\|}$. Then $\|z_n\| = \frac{1}{\|Tx_n\|} < \frac{1}{n} \to 0$, and $\|Tz_n\| = 1$ by construction.

Finally, if $\{x_n\} \subset \mathcal{X}$ such that $x_n \to 0$ and $\|Tx_n\| = 1$ for all n, then $\frac{\|Tx_n\|}{\|x_n\|} = \frac{1}{\|x_n\|} \to \infty$ and T is unbounded.

(6.1.4) Prove that (6.1) defines a norm in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$, and that

$$\|Tx\| \leq \|T\|\,\|x\|, \qquad T \in \mathcal{B}(\mathcal{X}, \mathcal{Y}), \quad x \in \mathcal{X}.$$

Answer. Since $\|T\| = \inf\{r : \|Tx\| \leq r\|x\|,\ x \in \mathcal{X}\}$, given $\varepsilon > 0$ there exists $r > 0$ such that $\|T\| > r - \varepsilon$ and $\|Tx\| \leq r\|x\|$ for all x. Then, for each x,

$$\|Tx\| \leq (\|T\| + \varepsilon)\|x\| = \|T\|\,\|x\| + \varepsilon\,\|x\|.$$

As this happens for all $\varepsilon > 0$, we get that $\|Tx\| \leq \|T\|\,\|x\|$.

To show that T is a norm, we have that $\|T\| \geq 0$ by definition. If $\|T\| = 0$, Then $\|Tx\| \leq 0$, and so $Tx = 0$ for all x, which is $T = 0$.

Given $\lambda \in \mathbb{C}$,

$$\begin{aligned}\|\lambda T\| &= \inf\{r > 0 : \ \|\lambda Tx\| \le r\|x\|, \ x \in \mathcal{X}\} \\ &= \inf\left\{r > 0 : \ \|Tx\| \le \frac{r}{|\lambda|}\,\|x\|, \ x \in \mathcal{X}\right\} \\ &= |\lambda| \ \inf\left\{\frac{r}{|\lambda|} > 0 : \ \|Tx\| \le \frac{r}{|\lambda|}\,\|x\|, \ x \in \mathcal{X}\right\} \\ &= |\lambda| \ \inf\{r > 0 : \ \|Tx\| \le r\,\|x\|, \ x \in \mathcal{X}\} \\ &= |\lambda|\,\|T\|.\end{aligned}$$

For the triangle inequality,

$$\begin{aligned}\|(T+S)x\| &= \|Tx + Sx\| \le \|Tx\| + \|Sx\| \\ &\le \|T\|\,\|x\| + \|S\|\,\|x\| = (\|T\| + \|S\|)\,\|x\|.\end{aligned}$$

As this occurs for all $x \in \mathcal{X}$, we have shown that $\|T+S\| \le \|T\| + \|S\|$.

Alternatively, if we already have Proposition 6.1.4, we can use the equality

$$\|T\| = \sup\{\|Tx\| : \ \|x\| = 1\}$$

for a simpler proof. Indeed, as

$$\|(T+S)x\| = \|Tx + Sx\| \le \|Tx\| + \|Sx\|$$

and the supremum is subadditive, we get that $\|T+S\| \le \|T\| + \|S\|$. If $\|T\| = 0$, then $\|Tx\| = 0$ for all x, so $T = 0$. And, since $\|\lambda Tx\| = |\lambda|\,\|Tx\|$, we get that $\|\lambda T\| = |\lambda|\,\|T\|$.

(6.1.5) Prove Proposition 6.1.4.

Answer. We know from Exercise 6.1.4 that

$$\|Tx\| \le \|T\|\,\|x\| \tag{AB.6.1}$$

for all $x \in \mathcal{X}$. This shows that

$$\|T\| = \min\{r : \ \|Tx\| \le r\|x\|, \ x \in \mathcal{X}\}.$$

When $\|x\| = 1$, by the above, $\|Tx\| \le \|T\|\,\|x\| = \|T\|$. So $\|T\|$ is an upper bound for the set $\{\|Tx\| : \ \|x\| = 1\}$. And if $s = \sup\{\|Tx\| : \ \|x\| = 1\}$, Then for any x we have

$$\left\|\frac{Tx}{\|x\|}\right\| \le s,$$

which implies that $\|Tx\| \le s\|x\|$ for all s. Therefor $s \ge \|T\|$, and so $s = \|T\|$.

For the third equality we just note that

$$\{\|Tx\| : \ \|x\| = 1\} = \{\|Tx\|/\|x\| : \ x \ne 0\}.$$

Finally, take $x \in \mathcal{X}$. Using (AB.6.1) twice,

$$\|TSx\| \leq \|T\| \, \|Sx\| \leq \|T\| \, \|S\| \, \|x\|.$$

As $\|TS\|$ is the infimum of the constants that may appear in the above inequality, we get that $\|TS\| \leq \|T\| \, \|S\|$.

(6.1.6) Prove Proposition 6.1.5 (*Hint: Proposition 5.5.8 is a particular case*).

Answer. Let $\{T_n\} \subset \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be a Cauchy sequence. From the reverse triangle inequality,

$$| \, \|T_n\| - \|T_m\| \, | \leq \|T_n - T_m\|,$$

which shows that the number sequence $\{\|T_n\|\}$ is Cauchy. In particular there exists $c > 0$ with $\|T_n\| \leq c$ for all n. For any fixed $x \in \mathcal{X}$,

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\| \, \|x\|,$$

so the sequence $\{T_n x\} \subset \mathcal{Y}$ is Cauchy. As $\mathcal{Y}$ is complete, the sequence is convergent and we may define $Tx = \lim_n T_n x$. Since the T_n and limits are linear, it follows that T is a linear function. If we fix $x \in \mathcal{X}$ and let $\varepsilon > 0$, there exists n with $\|Tx - T_n x\| < \varepsilon$. Then

$$\|Tx\| \leq \|Tx - T_n x\| + \|T_n x\| \leq \varepsilon + \|T_n\| \, \|x\| \leq \varepsilon + c\|x\|.$$

As we can do this for every $\varepsilon > 0$, we get that $\|Tx\| \leq c\|x\|$. This works for all $x \in \mathcal{X}$ with the same c, so we have shown that T is bounded.

Finally, we need to show that T is a (norm) limit of the T_n. Fix $\varepsilon > 0$. There exists n_0 such that $\|T_n - T_m\| < \varepsilon$ if $n, m \geq n_0$. Then, if $n, m \geq n_0$,

$$\|Tx - T_n x\| \leq \|Tx - T_m x\| + \|(T_m - T_n)x\| \leq \|Tx - T_m x\| + \varepsilon \, \|x\|.$$

Taking the limit as $m \to \infty$, we get

$$\|Tx - T_n x\| \leq \varepsilon \, \|x\|$$

for all $n \geq n_0$. Hence $\|T - T_n\| < \varepsilon$ and therefore $\|T - T_n\| \to 0$.

(6.1.7) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. Show that if $\dim \mathcal{X} = n$, $\dim \mathcal{Y} = m$, then $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ can be identified with $M_{m,n}(\mathbb{C})$.

Answer. We note first that any linear map $\mathcal{X} \to \mathcal{Y}$ is bounded (Corollary 5.2.5). The identification will depend on the choice of bases on both $\mathcal{X}$ and $\mathcal{Y}$. Fix bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ of $\mathcal{X}$ and $\mathcal{Y}$ respectively. Given $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ there exist coefficients t_{kj}, $k = 1, \ldots, m$, $j = 1, \ldots, n$, such that

$$Te_j = \sum_{k=1}^{m} t_{kj} f_k.$$

Given $x \in \mathcal{X}$, we can write $x = \sum_j x_j e_j$, and so

$$Tx = \sum_{j=1}^{n} \sum_{k=1}^{m} t_{kj} x_j f_k = \sum_{k=1}^{m} \Big(\sum_{j=1}^{n} t_{kj} x_j \Big) f_k. \tag{AB.6.2}$$

Let $\gamma : \mathcal{B}(\mathcal{X}, \mathcal{Y}) \to M_{m,n}(\mathbb{C})$ be given by $\gamma(T) = [t_{kj}]$. Because the representation of a vector in a basis is unique, $\gamma(T) = \gamma(S)$ implies $T = S$, so γ is injective; it also implies that is linear. Given $[t_{kj}] \in M_{m,n}(\mathbb{C})$ we can use (AB.6.2) to define $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ such that $\gamma(T) = [t_{kj}]$; so γ is surjective.

(6.1.8) Let $\mathcal{Y}$ be an infinite-dimensional normed space. Show that the space $\mathcal{B}(\mathbb{C}, \mathcal{Y})$ is infinite-dimensional.

Answer. Any linear $T : \mathbb{C} \to \mathcal{Y}$ is determined by its value at 1. Given $n \in \mathbb{N}$, since $\mathcal{Y}$ is infinite-dimensional we can find $y_1, \ldots, y_n \in \mathcal{Y}$, linearly independent. Define $T_j\lambda = \lambda y_j$. As $T_j(\alpha\lambda + \mu) = (\alpha\lambda + \mu)y_j = \alpha T_j\lambda + T_j\mu$, the operator T_j is linear for all j. Boundedness is automatic since $\mathbb{C}$ is finite-dimensional. Now if $\sum_j \alpha_j T_j = 0$, then

$$0 = \sum_j \alpha_j T_j 1 = \sum_j \alpha_j y_j.$$

As $y_1, \ldots, y_n$ are linearly independent, we get that $\alpha_1 = \cdots = \alpha_n = 0$. So $T_1, \ldots, T_n$ are linearly independent. This works for any n, showing that $\dim \mathcal{B}(\mathbb{C}, \mathcal{Y}) = \infty$.

(6.1.9) Let $\mathcal{X}, \mathcal{Y}$ be a Banach spaces and $T : \mathcal{X} \to \mathcal{Y}$ linear and isometric, that is $\|Tx\| = \|x\|$ for all $x \in \mathcal{X}$. Show that $\operatorname{ran} T$ is closed.

Answer. Let $\{Tx_n\}$ be Cauchy. As $\|x_n - x_m\| = \|T(x_n - x_m)\| = \|Tx_n - Tx_m\|$, the sequence $\{x_n\}$ is Cauchy in $\mathcal{X}$. Let $x = \lim x_n$. The operator T is bounded, so $\lim_n Tx_n = Tx$ and hence $\{Tx_n\}$ converges to $Tx \in \operatorname{ran} T$.

(6.1.10) Let $\mathcal{X},\mathcal{Y}$ be normed spaces.

(i) Show that if $\varphi\in\mathcal{X}^*$, $y_0\in\mathcal{Y}$, and $T:x\longmapsto\varphi(x)y_0$, then $\|T\|=\|\varphi\|\,\|y_0\|$.

(ii) Show that if $\varphi_1,\ldots,\varphi_n\in\mathcal{X}^*$ and $y_1,\ldots,y_n\in\mathcal{Y}$, then $T:x\longmapsto\sum_j\varphi_j(x)y_j$ is bounded.

(iii) Show that if $\mathcal{Y}$ is Banach, $\varphi_1,\varphi_2,\ldots\in\mathcal{X}^*$ and $y_1,y_2,\ldots\in\mathcal{Y}$ with $\sum_j\|\varphi_j\|\,\|y_j\|<\infty$, then $T:x\longmapsto\sum_j\varphi_j(x)y_j$ is bounded.

Answer.

(i) If $Tx=\varphi(x)y_0$, then

$$\|T\|=\sup\{\|Tx\|:\ \|x\|=1\}=\sup\{\|\varphi(x)\,y_0\|:\ \|x\|=1\}$$
$$=\|y_0\|\,\sup\{|\varphi(x)|:\ \|x\|=1\}=\|y_0\|\,\|\varphi\|.$$

A slightly more convoluted approach:

$$\|Tx\|=\|\varphi(x)\,y_0\|=|\varphi(x)|\,\|y_0\|\leq\|\varphi\|\,\|y_0\|\,\|x\|,$$

so $\|T\|\leq\|\varphi\|\,\|y_0\|$. Fix $\varepsilon>0$ and let $x\in\mathcal{X}$ such that $\|x\|=1$ and $|\varphi(x)|\geq(\|\varphi\|-\varepsilon)$. Then

$$\|Tx\|=|\varphi(x)|\,\|y_0\|\geq(\|\varphi\|-\varepsilon)\,\|y_0\|.$$

As ε was arbitrary, $\|T\|\geq\|\varphi\|\,\|y_0\|$.

(ii) We prove (iii), as it has (ii) as a particular case. When

$$Tx=\sum_{j=1}^\infty\varphi_j(x)\,y_j$$

with $\sum_j\|\varphi_j\|\,\|y_j\|<\infty$ (the series for T exists because of this last condition—which forces the tails to be small—and the fact that $\mathcal{Y}$ is Banach) we have, using that the norm is continuous,

$$\|Tx\|=\lim_N\Big\|\sum_{j=1}^N\varphi(x)\,y_j\Big\|$$
$$\leq\limsup_N\sum_{j=1}^N\|\varphi\|\,\|y_j\|\,\|x\|$$
$$=\Big(\sum_{j=1}^\infty\|\varphi\|\,\|y_j\|\Big)\,\|x\|,$$

so T is bounded and $\|T\| \leq \sum_{j=1}^{\infty} \|\varphi\| \, \|y_j\|$. The case with a finite sum is a particular case of this, as we can take $\varphi_j = 0$ for $j > n$.

(6.1.11) Let $T : c_{00} \to c_{00}$ be given by

$$T(x_1, x_2, \cdots) = \Big(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots\Big).$$

Show that T is bounded and bijective.

Answer. The proof of $\|T\| = 1$ and injectivity are exactly the same as for Exercise 6.1.12. In fact, there is no need to even re-do the proof, as $c_{00} \subset c_0$, the restriction of an injective map is injective, and the canonical basis is in c_{00}.

Surjectivity: if $x \in c_{00}$, then $x = T(x_1, 2x_2, 3x_3, \ldots)$. The sequence $\{nx_n\}$ is in c_{00} for all $x \in c_{00}$ because of the finite support.

(6.1.12) Let $T : c_0 \to c_0$ be given by

$$T(x_1, x_2, \cdots) = \Big(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots\Big).$$

Show that T is bounded (with $\|T\| = 1$) and injective, but not surjective.

Answer. We have

$$\|Tx\| = \sup\{|x_k/k| : \; k\} \leq \sup\{|x_k| : \; k\} = \|x\|.$$

So T is bounded and $\|T\| \leq 1$. As $Te_1 = e_1$, this gives us $\|Te_1\| = 1$ with $\|e_1\| = 1$, so $\|T\| = 1$. It $Tx = 0$, then $\frac{1}{n} x_n = 0$ for all n, so $x_n = 0$ and $x = 0$; hence T is injective.

Consider $b \in c_0$ where $b_n = 1/n$ for all n. If $b = Ta$, then $b_n = a_n/n$ for all n, so $a_n = 1$ for all n. But then $a \notin c_0$. It follows that T is not surjective.

6.2. Invertibility in $\mathcal{B}(\mathcal{X})$

(6.2.1) Show that if R is a ring with unit, and both a and ab are invertible, then b is invertible. Using the algebra $\mathcal{B}(\ell^2(\mathbb{N}))$, show that it is possible to have ab invertible with neither a nor b invertible.

Answer. Since ab is invertible, there exists $c \in R$ with $cab = abc = 1$. As both a and c are invertible, we get $a^{-1}c^{-1} = a^{-1}(abc)c^{-1} = b$. So b is a product of invertible elements and thus invertible with inverse ca.

In $\mathcal{B}(\ell^2(\mathbb{N}))$ we may take the left and right unilateral shifts T and S as in page 424 of the Book. Then $TS = I$ even though neither S nor T are invertible.

(6.2.2) Let $S \in \mathcal{B}(\mathcal{X})$ with $\|I - S\| = 1$. Decide (and justify) whether such an S is always invertible, sometimes invertible and sometimes not, or never invertible.

Answer. The operator S can fail to be invertible; easiest example is $S = 0$. But there are also examples where S is invertible. For a trivial example of this situation, let $S = 2I$. Another example is $\gamma = \frac{1}{2} + i\frac{\sqrt{3}}{2}$, and put $S = \gamma I$. Then S is invertible with inverse $\gamma^{-1} I$, and $\|I - S\| = |1 - \gamma| = 1$. For a slightly less trivial example let $\mathcal{X} = \mathbb{C}^2$ with the 2-norm (that is, the usual Euclidean norm) and let

$$S = \begin{bmatrix} \gamma & 0 \\ 0 & \bar{\gamma} \end{bmatrix},$$

with the same γ as above. It is not hard to check that because S is diagonal its norm is $|\gamma| = 1$, and $\|I - S\| = \max\{|1 - \gamma|, |1 - \bar{\gamma}|\} = 1$.

(6.2.3) Let $S \in \mathcal{B}(\mathcal{X})$ such that $\|I - S\| = \frac{1}{2}$. Decide (and justify) whether such an S is always invertible, sometimes invertible and sometimes not, or never invertible.

Answer. Here we can apply Lemma 6.2.1 to $T = I - S$ and we get that S is invertible with inverse

$$S^{-1} = \sum_{k=0}^{\infty}(I-S)^k.$$

(6.2.4) Let $S \in \mathcal{B}(\mathcal{X})$ such that $S^2 = S$. Decide (and justify) whether such an S is always invertible, sometimes invertible and sometimes not, or never invertible.

Answer. There is the possibility that $S = I$, in which case it is invertible. It could also be 0, in which case it would not be invertible. There are always many non-trivial idempotents on $\mathcal{X}$ if $\dim \mathcal{X} > 1$ (example: fix $y \in \mathcal{X}$ and $\varphi \in \mathcal{X}^*$ with $\varphi(y) = 1$, and define $Sx = \varphi(x)y$). A non-trivial idempotent S cannot be invertible (if $I = ST$ then $S = S^2T = ST = I$), and as $I - S$ is also a non-trivial idempotent, it cannot be invertible either. To see this, note that $\ker S = \operatorname{ran}(I - S)$. So if $\ker S$ is trivial, then $S = I$.

6.3. Baire's Theorem and its Corollaries

(6.3.1) Let $\mathcal{X}$ be a metric space. Show that $A \subset \mathcal{X}$ is nowhere dense if and only if $\mathcal{X} \setminus \overline{A}$ is open and dense.

Answer. We always have $\mathcal{X} \setminus \overline{A}$ open, so all that matters is whether it is dense.

Suppose that $\mathcal{X} \setminus \overline{A}$ is not dense. Then there exists $x \in \mathcal{X}$ and $r > 0$ such that $B_r(x) \cap (\mathcal{X} \setminus \overline{A}) = \varnothing$. Thus $B_r(x) \subset \overline{A}$ and A is not nowhere dense.

Conversely, suppose that $\mathcal{X} \setminus \overline{A}$ is dense. For any $x \in \mathcal{X}$ and $r > 0$, we have $B_r(x) \cap (\mathcal{X} \setminus \overline{A}) \neq \varnothing$. So it is not possible for $B_r(x)$ to be inside $\overline{A}$; thus the interior of $\overline{A}$ is empty, and A is nowhere dense.

(6.3.2) Prove that Baire's Category Theorem 6.3.1 holds for locally compact Hausdorff topological spaces (*Hint: use the finite intersection property instead of completeness*).

Answer. Fix $W \subset \mathcal{X}$ open. Since V_1 is dense, $W \cap V_1 \neq \varnothing$. As V_1 is also open $W \cap V_1$ is open. So its interior is nonempty: and with $\mathcal{X}$ being locally compact, there exists a nonempty open $K_1 \subset W \cap V_1$, with $\overline{K_1}$ compact. Now $K_1 \cap V_2$ is open and nonempty, and we can repeat the process. Inductively: we now assume that we have nonempty K_n open with $\overline{K_n}$ compact and $\overline{K_n} \subset K_{n-1} \cap V_n$. Reasoning as above we obtain a nonempty open set K_{n+1} with $\overline{K_{n+1}}$ compact and $\overline{K_{n+1}} \subset K_n \cap V_n$. This way we obtain a family $\{\overline{K_n}\}$ of compact sets with the finite intersection property, inside the compact set $\overline{K_1}$; thus $\cap_n \overline{K_n} \neq \varnothing$. Let $x \in \cap_n \overline{K_n}$. For each n,

$$x \in K_n \subset K_{n-1} \cap V_n \subset V_n.$$

So $x \in V_n$ for all n, and then $x \in \bigcap_n V_n$. Therefore $W \cap \bigcap_n V_n \neq \varnothing$. As W was any open set, it follows that $\bigcap_n V_n$ is dense.

(6.3.3) Show that any Hamel basis of an infinite-dimensional Banach space is uncountable (*Hint: show that finite-dimensional subspaces are nowhere dense*).

Answer. If a subspace contains a ball, then it has to be the whole space. Indeed, suppose that $B_r(y) \subset \mathcal{X}_0 \subset \mathcal{X}$, where $\mathcal{X}_0$ is a subspace. Let $x \in \mathcal{X}$. Then, as $y \in \mathcal{X}_0$,

$$x = \frac{2\|x\|}{r}\left[\overbrace{\left(\frac{r}{2\|x\|}\,x + y\right)}^{\in B_r(y)} - y\right] \in \mathcal{X}_0$$

As finite-dimensional subspaces are closed, being proper subspaces the interior of their closure is empty, so they are nowhere dense. If $\mathcal{X}$ has a countable Hamel basis $\{x_n\}$, we can write

$$\mathcal{X} = \bigcup_n \operatorname{span}\{x_1, \ldots, x_n\},$$

contradicting Theorem 6.3.1.

(6.3.4) Show an example of an infinite-dimensional normed space that is a countable union of nowhere dense subsets.

Answer. As per Exercise 6.3.3, any normed space with a countable Hamel basis will do. For instance $\mathcal{X} = \mathbb{C}[x]$, with the norm $\|p\| = \max\{|p(t)| : t \in [0,1]\}$. In this case—and in the case of any other normed space with a countable Hamel basis—we can write the nowhere dense sets explicitly, as in Exercise 6.3.3. Namely, $\mathcal{X} = \bigcup_n \{p : \deg p \leq n\}$.

(6.3.5) Let $\mathcal{H}$ be a Hilbert space. Prove that $T \in \mathcal{B}(\mathcal{H})$ is surjective if and only if it admits a right inverse $S \in \mathcal{B}(\mathcal{H})$. The same assertion is not true for Banach spaces (Remark 6.2.5).

Answer. If $TS = I$, then $\xi = TS\xi$, so T is surjective. Conversely, if T is surjective, consider the restriction T_0 of T to $(\ker T)^\perp$. Then T_0 is bijective, so by the Inverse Mapping Theorem 6.3.6 there exists $S : \mathcal{H} \to (\ker T)^\perp$, linear and bounded, with $T_0 S = I$. Then $TS = T_0 S = I$, showing that S is a right inverse for T.

(6.3.6) Let $\mathcal{X}$ be a Banach space. Show that $T \in \mathcal{B}(\mathcal{X})$ admits a right inverse if and only if T is surjective and $\ker T$ is complemented.

Answer. Let $P : \mathcal{X} \to \ker T$ be a continuous projection onto $\ker T$. Let $M = (I - P)\mathcal{X}$. If $x \in M$ and $Tx = 0$, then $x \in M \cap \ker T = \{0\}$, so T is injective on M. Thus $T : M \to \mathcal{X}$ is a bounded bijective operator. By the Inverse Mapping Theorem there exists $S : \mathcal{X} \to M$, bounded, with $TS = I_{\mathcal{X}}$.

Conversely, if $TS = I_{\mathcal{X}}$ and $x \in \mathcal{X}$, then $x = T(Sx)$, so T is surjective. It remains to show that the existence of S guarantees that $\ker T$ is complemented. From $TS = I$ we see that S is injective. So given any $x \in \mathcal{X}$ there exists a unique $x_1 \in S\mathcal{X}$ such that $Tx = Tx_1$. As $x - x_1 \in \ker T$, we have $x = (x - x_1) + x_1 \in \ker T + S\mathcal{X}$. If $x \in S\mathcal{X} \cap \ker T$, then $x = Sy$ for some y, and $y = TSy = Tx = 0$, so $x = 0$. Thus $\mathcal{X} = \ker T + S\mathcal{X}$ is a direct sum. If $Sx_n \to 0$, then $x_n = TSx_n \to 0$, so $S\mathcal{X}$ is closed. Finally, consider the projection $P : \mathcal{X} \to S\mathcal{X}$. Given $x \in \mathcal{X}$, we have $Px = x_1$, with $x_1 \in S\mathcal{X}$ and $Tx = Tx_1$. Let $y \in \mathcal{X}$ with $x_1 = Sy$. Using that S is bounded,

$$\begin{aligned}\|Px\| &= \|x_1\| = \|Sy\| \leq \|S\|\,\|y\| = \|S\|\,\|TSy\| \\ &= \|S\|\,\|Tx_1\| = \|S\|\,\|Tx\| \leq \|S\|\,\|T\|\,\|x\|\end{aligned}$$

and so P is bounded. This shows that $S\mathcal{X}$ is topologically complemented, and thus so is $\ker T$ (since P is bounded, so is $I - P$).

Alternatively, we can use Proposition 6.3.9 and the fact that $\mathcal{X} = \ker T \oplus S\mathcal{X}$ and both subspaces are closed.

(6.3.7) Show that T^{-1} in Example 6.3.8 is unbounded.

Answer. For each n, $T^{-1}e_n = n\,e_n$. Thus $\|e_n\| = 1$ and $\|T^{-1}e_n\| = n$.

(6.3.8) Prove that the Closed Graph Theorem 6.3.12 implies the Open Mapping Theorem 6.3.5.

Answer. Assume first that T is bijective and bounded, and that its graph is closed. Since the flip is a homemorphism $\mathcal{X}\times\mathcal{Y}\to\mathcal{Y}\times\mathcal{X}$, we get that $\mathcal{G}(T^{-1}) = \{(Tx, x):\ x\in\mathcal{X}\}$ is closed. By the Closed Graph Theorem, T^{-1} is bounded. Then, for any V open, $TV = (T^{-1})^{-1}(V)$ is open by continuity of T^{-1}.

Now if we assume that T is bounded but only surjective, the linear map $\tilde{T}:\mathcal{X}/\ker T\to\mathcal{Y}$ is bijective and bounded:

$$\begin{aligned}\|\tilde{T}(x+\ker T)\| = \|Tx\| &= \inf_{z\in\ker T}\|Tx+Tz\|\\ &\leq \|T\|\inf_{z\in\ker T}\|x+z\| = \|T\|\,\|x+\ker T\|.\end{aligned}$$

By the above, $\tilde{T}$ is open. We also proved in Proposition 5.3.13 that the canonical quotient map $\pi:\mathcal{X}\to\mathcal{X}/\ker T$ is open. Then $T = \tilde{T}\circ\pi$ is open, being a composition of open maps.

(6.3.9) Complete the proof of Theorem 6.3.12 by showing that for a linear map $T:\mathcal{X}\to\mathcal{Y}$ with $\mathcal{X},\mathcal{Y}$ Banach, the graph $\mathcal{G}(T)$ is closed if and only if for every sequence $\{x_k\}\subset\mathcal{X}$ and $y\in\mathcal{Y}$, if $x_k\to 0$ and $Tx_k\to y$ then $y=0$.

Answer. If $\mathcal{G}(T)$ is closed, $x_k\to 0$, and $Tx_k\to y$, then $\{(x_k, Tx_k)\}$ is Cauchy in $\mathcal{G}(T)$. As $\mathcal{G}(T)$ is complete—closed subset of the complete space $\mathcal{X}\oplus\mathcal{Y}$—the limit of the sequence is in $\mathcal{G}(T)$. So $(0,y)\in\mathcal{G}(T)$, which implies $y=0$ as $(0,0)\in\mathcal{G}(T)$ and $\mathcal{G}(T)$ is a graph.

Conversely, suppose that $\{(x_k, Tx_k)\}$ is a Cauchy sequence in $\mathcal{X}\oplus\mathcal{Y}$. As $\max\{\|x\|,\|y\|\}\leq\|(x,y)\|$, we have that both $\{x_k\}$ and $\{Tx_k\}$ are Cauchy in $\mathcal{X}$ and $\mathcal{Y}$ respectively. Both $\mathcal{X}$ and $\mathcal{Y}$ are complete, so there exist $x\in\mathcal{X}$, $y\in\mathcal{Y}$ with $x_k\to x$ and $Tx_k\to y$. Then $x_k - x\to 0$ and $T(x_k-x)\to y-Tx$. The hypothesis then gives us that $y-Tx=0$, that is $y=Tx$. Thus $(x_k, Tx_k)\to(x,Tx)$, showing that $\mathcal{G}(T)$ is complete, and thus closed.

(6.3.10) Let $\mathcal{X}$ be a Banach space and $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ closed subspaces such that $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. Let $T \in \mathcal{B}(\mathcal{X})$ such that $T\mathcal{X}_1 \subset \mathcal{X}_1$ and $T\mathcal{X}_2 \subset \mathcal{X}_2$. Show that T is invertible if and only if $T|_{\mathcal{X}_1} \in \mathcal{B}(\mathcal{X}_1)$ and $T|_{\mathcal{X}_2} \subset \mathcal{B}(\mathcal{X}_2)$ are invertible.

Answer. Suppose that T is invertible. Let P be the projection only $\mathcal{X}_1$ induced by the decomposition $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. We have, for $x \in \mathcal{X}_1$, $TPx_1 = Tx_1 = PTx_1$. Similarly, $TPx_2 = PTx_2$ for any $x_2 \in \mathcal{X}_2$; it follows that $TPx = TP(x_1+x_2) = PT(x_1+x_2) = PTx$. That is, $PT = TP$. Multiplying on the left and right by T^{-1} we get that $T^{-1}P = PT^{-1}$; this shows that $T^{-1}\mathcal{X}_1 \subset \mathcal{X}_1$, so $S = T^{-1}|_{\mathcal{X}_1} \in \mathcal{B}(\mathcal{X}_1)$. It follows that $ST|_{\mathcal{X}_1} = TS|_{\mathcal{X}_1} = I_{\mathcal{X}_1}$. Similarly, $T|_{\mathcal{X}_2}$ is invertible.

Conversely, suppose that there exist $S_1 \in \mathcal{B}(\mathcal{X}_1)$ and $S_2 \in \mathcal{B}(\mathcal{X}_2)$ such that $S_1T|_{\mathcal{X}_1} = T|_{\mathcal{X}_1}S = I_{\mathcal{X}_1}$ and $S_2T|_{\mathcal{X}_2} = T|_{\mathcal{X}_2}S = I_{\mathcal{X}_2}$. Let $S = S_1 \oplus S_2$, that is $S(x_1 + x_2) = S_1x_1 + S_2x_2$. Then

$$ST(x_1 + x_2) = S_1Tx_1 + S_2Tx_2 = I_{\mathcal{X}_1}x_1 + I_{\mathcal{X}_2}x_2 = x_1 + x_2,$$

and similarly $TS = I_{\mathcal{X}}$. So T is invertible.

(6.3.11) Let $\mathcal{X}$ be a Banach space and $T : \mathcal{X} \to \mathcal{X}^*$ be a linear map. Use the Closed Graph Theorem to show that if $(Tx)y = (Ty)x$ for all $x, y \in \mathcal{X}$, then T is bounded.

Answer. Suppose that $x_n \to x$ and $Tx_n \to \phi$. If we show that $\phi = Tx$, then the Closed Graph Theorem implies that T is bounded. For any $y \in \mathcal{X}$,

$$\phi(y) = \lim_n (Tx_n)y = \lim_n (Ty)x_n = (Ty)x = (Tx)y.$$

As y was arbitrary, we conclude that $\phi = Tx$ and hence T is bounded.

(6.3.12) Let $\mathcal{X}$ be a Banach space and $T : \mathcal{X} \to \mathcal{X}^*$ be a linear map. Use the Uniform Boundedness Principle to show that if $(Tx)y = (Ty)x$ for all $x, y \in \mathcal{X}$, then T is bounded.

Answer. Consider the family $\{Tx\}_{\|x\|\leq 1}$ of bounded linear functionals. We have

$$|(Tx)y| = |(Ty)x| \leq \|Ty\|\,\|x\| = \|Ty\|.$$

Hence $\sup\{|(Tx)y| : \ \|x\| \leq 1\} < \infty$ for each y. By the Uniform Boundedness Principle, $\sup\{\|Tx\| : \ \|x\| \leq 1\} < \infty$, and this number is $\|T\|$.

(6.3.13) *(Compare with Exercise 6.3.14)* Show an example of a normed space $\mathcal{X}$ and $T \in \mathcal{B}(\mathcal{X})$ such that for every $x \in \mathcal{X}$ there exists $n \in \mathbb{N}$ with $T^n x = 0$, but such that $T^m \neq 0$ for all $m \in \mathbb{N}$.

Answer. Let $\mathcal{X} = c_{00}$ and T the right-shift, that is

$$T(a_1, a_2, \ldots) = (a_2, a_3, \ldots).$$

Then $T^n a = 0$ if $a_n = a_{n+1} = \cdots = 0$. But $T^m e_{m+1} = e_1$, so $T^m \neq 0$ for all m.

(6.3.14) *(Compare with Exercise 6.3.13)* Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$. Assume that, for each $x \in \mathcal{X}$, there exists $n \in \mathbb{N}$ such that $T^n x = 0$. Show that there exists $m \in \mathbb{N}$ such that $T^m = 0$.

Answer. The hypothesis gives us that $\mathcal{X} = \bigcup_n \ker T^n$, a union of closed subspaces (closed, because T^n is bounded and so $\ker T^n = (T^n)^{-1}(\{0\})$ is closed). By Baire's Category Theorem—as in Remark 6.3.3—there exists m such that $\ker T^m$ is not nowhere dense: so $\ker T^m$ contains a ball. A subspace that contains a ball is necessarily the whole space (see Exercise 6.3.3 for a proof), so $\ker T^m = \mathcal{X}$.

CHAPTER 7

Weak topologies

7.1. The weak topology

(7.1.1) Prove that if $\mathcal{T}_1 \subset \mathcal{T}_2$ are topologies on a set $\mathcal{X}$, with $\mathcal{T}_1$ Hausdorff and $\mathcal{T}_2$ compact, then $\mathcal{T}_1 = \mathcal{T}_2$ (*Hint: consider the identity map* $(\mathcal{X}, \mathcal{T}_2) \to (\mathcal{X}, \mathcal{T}_1)$).

Answer. Let $i : (\mathcal{X}, \mathcal{T}_2) \to (\mathcal{X}, \mathcal{T}_1)$ be the identity map, $i(x) = x$. Given $V \in \mathcal{T}_1$, since $V \in \mathcal{T}_2$ we have $i^{-1}(V) = V \in \mathcal{T}_2$. So i is continuous. Now let $K \subset \mathcal{X}$ be closed. Since $\mathcal{T}_2$ is compact, K is compact in $(\mathcal{X}, \mathcal{T}_2)$. The image of a compact set under a continuous function is compact; so $K = i(K)$ is compact in $(\mathcal{X}, \mathcal{T}_1)$. From $\mathcal{T}_1$ Hausdorff, compact sets are closed. So i maps closed sets to closed sets. Thus, if $V \in \mathcal{T}_2$, then $\mathcal{X} \setminus V$ is closed in $(\mathcal{X}, \mathcal{T}_2)$, and so $\mathcal{X} \setminus V = i(\mathcal{X} \setminus V)$ is closed in $(\mathcal{X}, \mathcal{T}_1)$. So $V \in \mathcal{T}_1$, and we have shown that $\mathcal{T}_2 \subset \mathcal{T}_1$.

(7.1.2) Let $\mathcal{X}$ be a TVS. Show that the weak topology $\sigma(\mathcal{X}, \mathcal{X}^*)$ is the weakest topology such that every $\varphi \in \mathcal{X}^*$ is continuous.

Answer. Let $\mathcal{T}$ be a topology on $\mathcal{X}$ such that φ is continuous for each $\varphi \in \mathcal{X}^*$. Then $\varphi^{-1}(V) \in \mathcal{T}$ for all $V \subset \mathbb{C}$ open. By definition of $\sigma(\mathcal{X}, \mathcal{X}^*)$ we have $\sigma(\mathcal{X}, \mathcal{X}^*) \subset \mathcal{T}$.

(7.1.3) Prove that

$$\sigma(\mathcal{X}, \mathcal{X}^*) = \operatorname{Top}\{\varphi^{-1}(B_r(\lambda)) : \ \varphi \in \mathcal{X}^*, \ r > 0, \ \lambda \in \mathbb{C}\}.$$

Answer. Since $B_r(\lambda)$ is open for all $r > 0$ and all $\lambda \in \mathbb{C}$, $\varphi^{-1}(B_r(\lambda)) \in \sigma(\mathcal{X}, \mathcal{X}^*)$. This shows that

$$\sigma(\mathcal{X}, \mathcal{X}^*) \supset \operatorname{Top}\{\varphi^{-1}(B_r(\lambda)) : \ \varphi \in \mathcal{X}^*, \ r > 0, \ \lambda \in \mathbb{C}\}.$$

Now given any $V \subset \mathbb{C}$ open, for each $v \in V$ there exists $r_v > 0$ with $B_{r_v}(v) \subset V$. Then

$$V = \bigcup_{v \in V} B_{r_v}(v).$$

Thus

$$\begin{aligned}\varphi^{-1}(V) = \varphi^{-1}\Big(\bigcup_{v \in V} B_{r_v}(v)\Big) &= \bigcup_{v \in V} \varphi^{-1}(B_{r_v}(v)) \\ &\in \operatorname{Top}\{\varphi^{-1}(B_r(\lambda)) : \ \varphi \in \mathcal{X}^*, \ r > 0, \ \lambda \in \mathbb{C}\}.\end{aligned}$$

So

$$\sigma(\mathcal{X}, \mathcal{X}^*) \subset \operatorname{Top}\{\varphi^{-1}(B_r(\lambda)) : \ \varphi \in \mathcal{X}^*, \ r > 0, \ \lambda \in \mathbb{C}\}.$$

(7.1.4) Show that the sets $N(\varphi_1, \ldots, \varphi_k; \varepsilon)$, indexed by a positive integer $k \in \mathbb{N}$ and $\varphi_1, \ldots, \varphi_k \in \mathcal{X}^*$, $\varepsilon > 0$, form a local base for $\sigma(\mathcal{X}, \mathcal{X}^*)$ at 0.

Answer. The intersection of two of the sets has the same form, so all we need to show is that any open set contains an element of the purported local base.

Let V be a weak-open set with $0 \in V$. By Lemma 7.1.3 there exist $\varepsilon > 0$ and $\varphi_1, \ldots, \varphi_k \in \mathcal{X}^*$ with $\bigcap_j \varphi_j^{-1}(B_\varepsilon(0)) \subset V$. As $\bigcap_j \varphi_j^{-1}(B_\varepsilon(0)) = N(\varphi_1, \ldots, \varphi_k; \varepsilon)$, we are done.

(7.1.5) Show that the sets $N(\varphi_1,\ldots,\varphi_k;\varepsilon)$, where $\varphi_1,\ldots,\varphi_k \in \mathcal{X}^*$ are linearly independent and $\varepsilon > 0$, form a local base for $\sigma(\mathcal{X},\mathcal{X}^*)$ at 0.

Answer.
We will show that these neighbourhoods generate the same open sets than the 'not-necessarily-linearly independent neighbourhoods. If $\varphi_1,\ldots,\varphi_n$ are linearly dependent, after reordering we may assume that $\varphi_1,\ldots,\varphi_r$ are linearly independent, and that $\varphi_{r+1},\ldots,\varphi_n$ are linear combinations of those, say

$$\varphi_{r+j} = \sum_{k=1}^{r} c_{jk}\varphi_k.$$

Then

$$|\varphi_{r+j}(x)| \leq \sum_{k=1}^{r} |c_{jk}|\,|\varphi_k(x)| \leq \left(\sum_{k=1}^{r} |c_{jk}|\right) \max\{|\varphi_k(x)| : \ k=1,\ldots,r\}.$$

If we put $c = \max_j\{(\sum_{k=1}^r |c_{jk}|) : \ j=1,\ldots,n-r\}$, then

$$N(\varphi_1,\ldots,\varphi_r;\varepsilon/c\} \subset N(\varphi_1,\ldots,\varphi_n;\varepsilon\} \subset N(\varphi_1,\ldots,\varphi_r;\varepsilon\}.$$

The mutual containment guarantees that both families generate the same topology.

(7.1.6) Let $\mathcal{X}$ be a normed space and $\{x_n\}$ a weakly convergent net with $x_n \xrightarrow{\text{weak}} x$. Show that $\|x\| \leq \liminf_n \|x_n\|$. Find an example where the inequality is strict.

Answer. We have, by Corollary 5.7.7

$$\|x\| = \max\{|\varphi(x)| : \ \varphi \in \mathcal{X}^*, \ \|\varphi\| \leq 1\}.$$

For any $\varphi \in \mathcal{X}^*$ with $\|\varphi\| \leq 1$, since $|\varphi(x_j)| \leq \|x_j\|$, and choosing a subnet $\{x_{j_k}\}$ such that $\lim_k \|x_{j_k}\| = \liminf_j \|x_j\|$,

$$|\varphi(x)| = \lim_j |\varphi(x_j)| = \lim_k |\varphi(x_{j_k})| \leq \lim_k \|x_{j_k}\| = \liminf_j \|x_j\|.$$

For an example where the inequality is strict, it was shown in Example 7.1.6 that an orthonormal basis on an infinite-dimensional Hilbert space is a sequence $\{\xi_n\}$ with $\|\xi_n\| = 1$ for all n and $\xi_n \xrightarrow{\text{weak}} 0$.

(7.1.7) Use Proposition 7.1.4 to show that if $\varphi \in \mathcal{X}^*$, $c \in \mathbb{R}$, then $\{x \in \mathcal{X} : \operatorname{Re}\varphi(x) \geq c\}$ is weakly closed.

Answer. Because $\mathcal{X}^*$ is complete in the weak topology (Proposition 7.1.5, and Proposition 5.5.8 and Theorem 7.1.16), it is enough to show that if $\{x_j\} \subset \{x \in \mathcal{X} : \operatorname{Re}\varphi(x) \geq c\}$ and $x_j \to x$ weakly, then $x \in \{x \in \mathcal{X} : \operatorname{Re}\varphi(x) \geq c\}$. And this is trivial: since $\operatorname{Re}\varphi(x_j) \geq c$ for all j, we have $\operatorname{Re}\varphi(x) = \lim_j \operatorname{Re}\varphi(x_j) \geq c$.

(7.1.8) Prove that the weak topology is a locally convex topology given by the seminorms $p_\varphi(x) = |\varphi(x)|$, $\varphi \in \mathcal{X}^*$.

Answer. Using Proposition 7.1.4,

$$x_j \xrightarrow{\text{weak}} x \iff \forall \varphi \in \mathcal{X}^*,\ \varphi(x_j) \to \varphi(x) \iff \forall \varphi \in \mathcal{X}^*,\ p_\varphi(x_j - x) \to 0.$$

The only observation one needs is that $\varphi(x_j) \to \varphi(x) \iff \varphi(x_j - x) \to 0$ by linearity.

(7.1.9) Prove Proposition 7.1.12.

Answer. Let $K \subset \mathcal{X}$ be weakly compact. For each $\varphi \in \mathcal{X}^*$,

$$\sup\{|\hat{x}(\varphi)| : x \in K\} = \sup\{|\varphi(x)| : x \in K\} < \infty,$$

since φ is weakly continuous and a continuous function maps compact sets to compact sets. By the Uniform Boundedness Principle (Theorem 6.3.16), applied to $\hat{K} \subset \mathcal{B}(\mathcal{X}^*, \mathbb{C}) = \mathcal{X}^{**}$, we get that $\sup\{\|x\| : x \in K\} = \sup\{\|\hat{x}\| : x \in K\} < \infty$.

(7.1.10) Prove Corollary 7.1.17.

Answer. The set $K = \operatorname{conv}\{x_n : n \in \mathbb{N}\}$ is convex, so by Theorem 7.1.16 we have $x \in \overline{K}^{\sigma(\mathcal{X}, \mathcal{X}^*)} = \overline{K}$. This gives the existence of the net $\{x'_m\}$. When $\mathcal{X}$ is metrizable, by taking a countable local base around x, we can extract a convergent subsequence out of $\{x'_m\}$.

(7.1.11) Show that the unit ball of c_0 is not weakly compact.

Answer. Consider the sequence

$$g_n = (\overbrace{1, \ldots, 1}^{n \text{ times}}, 0, \ldots).$$

If the unit ball of c_0 is weakly compact, then $\{g_n\}$ has a cluster point z. That would mean that for every $y \in \ell^1(\mathbb{N})$, $\langle y, g_n\rangle \to \langle y, z\rangle$. With $y = e_k$, we get

$$z(k) = \langle e_k, z\rangle = \lim_n \langle e_k, g_n\rangle = 1.$$

But then $z = 1 \notin c_0$.

(7.1.12) Let $\mathcal{X}$ be a normed space and $\{x_n\} \subset \mathcal{X}$ such that $\|x_n\| \geq \delta > 0$ for all n, and such that $x_n \xrightarrow{\text{weak}} 0$. Show that $x_n/\|x_n\| \xrightarrow{\text{weak}} 0$.

Answer. Given $\varphi \in \mathcal{X}^*$,

$$\left|\varphi\left(\frac{x_n}{\|x_n\|}\right)\right| = \frac{1}{\|x_n\|}\,|\varphi(x_n)| \leq \frac{1}{\delta}\,|\varphi(x_n)| \to 0.$$

(7.1.13) Show that the boundedness requirement in Lemma 7.1.18 cannot be dispensed with. That is find an example, for $p \in (1, \infty]$, of a sequence $\{f_n\} \subset \ell^p(\mathbb{N})$ such that $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$ and $h \in \ell^q(\mathbb{N})$ such that $\langle f_n, h\rangle$ does not converge to 0.

Answer. Since $p > 1$, we have that $1 \leq q < \infty$. Let $f_n = n^2 e_n$, $n \in \mathbb{N}$. Then $f_n \in \ell^p(\mathbb{N})$ for all n, and $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$. Indeed, if $g = \sum_{k=1}^m a_k e_k$, then $\langle f_n, g\rangle = 0$ for all $n > m$. Meanwhile, if we put $h = \sum_k k^{-3/(2q)}\, e_k$ then $h \in \ell^q(\mathbb{N})$ while $\langle f_n, h\rangle = n^{2-3/(2q)} \geq n^{1/2} \to \infty$.

(7.1.14) Show that the $p > 1$ requirement in Lemma 7.1.18 cannot be dispensed with. That is, find an example of a bounded sequence $\{f_n\} \subset \ell^1(\mathbb{N})$ such that $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$ and $h \in \ell^\infty(\mathbb{N})$ such that $\langle f_n, h\rangle$ does not converge to 0.

Answer. Let $f_n = e_n$; then $\|f_n\|_1 = 1$ for all n and $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$. But $\langle f_n, 1\rangle = 1$ for all n.

(7.1.15) Let $\mathcal{X}$ be a normed space, with $\mathcal{X}^*$ separable. Show that the weak topology is metrizable on the unit ball $B_1^{\mathcal{X}}(0)$. *(Hint: if in need of inspiration, look at the proof of Corollary 7.2.20)*

Answer. Let $\{\varphi_n\} \subset B_1^{\mathcal{X}^*}(0)$ be a dense sequence. Define

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n} |\varphi_n(x-y)|, \qquad x, y \in B_1^{\mathcal{X}}(0).$$

The series converges, since $|\varphi_n(x-y)| \leq \|\varphi_n\| \, (\|x\| + \|y\|) \leq 2$. It is also translation invariant, symmetric and satisfies the triangle inequality, so it is a translation invariant metric. If $x_j \xrightarrow{\text{weak}} 0$, fix $\varepsilon > 0$ and choose n_0 such that $\sum_{n>n_0} 2^{-n} < \varepsilon/4$. Choose also j_0 such that $|\varphi_n(x_j)| < \varepsilon/2$, if $j > j_0$, $n = 1, \ldots, n_0$. Then, for $j > j_0$,

$$d(x_j, 0) = \sum_{n=1}^{n_0} 2^{-n} |\varphi_n(x-y)| + \sum_{n>n_0} 2^{-n} |\varphi_n(x-y)| < \sum_{n=1}^{n_0} 2^{-n} \frac{\varepsilon}{2} + 2\,\frac{\varepsilon}{4} < \varepsilon.$$

It follows that $d(x_j, 0) \to 0$. Conversely, suppose that $d(x_j, 0) \to 0$. Fix $\varphi \in \mathcal{X}^*$ and $\varepsilon > 0$; assume initially that $\|\varphi\| \leq 1$. There exists n such that $\|\varphi - \varphi_n\| < \varepsilon$. Then

$$\begin{aligned} |\varphi(x_j)| &\leq |\varphi(x_j) - \varphi_n(x_j)| + |\varphi_n(x_j)| \\ &\leq \|\varphi - \varphi_n\| + |\varphi_n(x_j)| \leq \varepsilon + 2^n \, d(\varphi_j, 0). \end{aligned}$$

Thus

$$\limsup_j |\varphi(x_j)| \leq \varepsilon.$$

By the Limsup Routine, $\lim_j |\varphi(x_j)| = 0$. When φ is arbitrary, we apply the above to $\varphi/\|\varphi\|$.

(7.1.16) Let $\mathcal{X}$ be a Banach space and $K \subset \mathcal{X}$ convex. Show that K is weakly closed if and only if $K \cap \overline{B_r(0)}$ is weakly closed for all $r > 0$.

Answer. We know that $\overline{B_r(0)}$ is weakly closed by Theorem 7.1.16. So when K is weakly closed, $K \cap \overline{B_r(0)}$ is weakly closed, being an intersection of closed sets.

For the converse, let $\{x_n\} \subset K$ be a Cauchy sequence. Then $x_n \to x \in \mathcal{X}$. As a norm-convergent sequence is bounded, there exists $r > 0$ such that $\|x_n\| < r$ for all n. Then $\{x_n\} \subset K \cap \overline{B_r(0)}$; this set, being weakly closed, is also norm-closed by Theorem 7.1.16, so $x \in K \cap \overline{B_r(0)}$. In particular $x \in K$, which shows that K is norm closed. As K is convex, Theorem 7.1.16 gives us that K is weakly closed.

(7.1.17) Let $\mathcal{H}$ be a separable Hilbert space and fix an orthonormal basis $\{\xi_n\}$. Let $S \subset \mathcal{H}$ be

$$S = \{\xi_m + m\xi_n : \ n, m \in \mathbb{N}\}.$$

Let $\overline{S}^{\text{w-seq}}$ denote the weak sequential closure of S. Prove that the inclusion $\overline{S}^{\text{w-seq}} \subset \overline{\overline{S}^{\text{w-seq}}}^{\text{w-seq}}$ is proper by showing that

$$0 \in \overline{\overline{S}^{\text{w-seq}}}^{\text{w-seq}} \setminus \overline{S}^{\text{w-seq}}.$$

Answer. Let $\{\xi_{m_k} + m_k\xi_{n_k}\}$ be a weakly convergent subsequence of S. By Proposition 7.1.11, the sequence is bounded. Since $\|\xi_{m_k} + m_k\xi_{n_k}\| = (1 + m_k^2)^{1/2}$ (or $1 + m_k$ if $m_k = n_k$) we deduce that m_k takes only finitely many values $r_1, \ldots, r_p$. As $\xi_n \xrightarrow{\text{weak}} 0$, the only possible limit of the sequence is some ξ_{r_k}. Thus

$$\overline{S^{\text{w-seq}}} = S \cup \{\xi_n\}_n.$$

In particular, $0 \notin \overline{S}^{\text{w-seq}}$. On the other hand, since $\xi_n \xrightarrow{\text{weak}} 0$ as already mentioned, $0 \in \overline{\overline{S}^{\text{w-seq}}}^{\text{w-seq}}$.

(7.1.18) With S as in Exercise 7.1.17, show that $0 \in \overline{S}^{\sigma(\mathcal{H},\mathcal{H}^*)}$.

Answer. Let N be a basic neighbourhood of 0. We have $N = \{\eta \in \mathcal{H} : |\langle \eta, \nu_j\rangle| < 1, \ j = 1, \ldots, s\}$ for some $\nu_1, \ldots, \nu_s \in \mathcal{H}$. Since $\{\xi_n\}$ is orthonormal, we can choose m such that $|\langle \xi_m, \nu_j\rangle| < \frac{1}{2}$, $j = 1, \ldots, s$. And then we can choose n such that $|\langle \xi_n, \nu_j\rangle| < \frac{1}{2m}$, $j = 1, \ldots, s$. Then $|\langle \xi_m + m\xi_n, \nu_j\rangle| < \frac{1}{2} + \frac{1}{2} = 1$, and so $\xi_m + m\xi_n \in N$. If we denote these chosen numbers by m_N, n_N, we get a net $\{\xi_{m_N} + m_N\xi_{n_N}\}$ in S that converges weakly to 0.

(7.1.19) Let $p \in (1, \infty]$. Consider the sequence $\{g_n\} \subset L^p[0, 1]$ given by

$$g_n(t) = \operatorname{sgn}(\sin(\pi n t)).$$

Show that $\{g_n\}$ converges weakly to 0, but $\{g_n\}$ does not converge pointwise to 0.

Answer. For the weak convergence, fix $0 < a < b < 1$. The function g_n will have integral equal to zero on any interval $\left[\frac{2k-2}{n}, \frac{2k}{n}\right]$. So all that survives on the integral are the integrals from a to the closest number $\frac{2k-2}{n}$ and from the last $\frac{2j}{n}$ to b. This gives us

$$\left|\int_a^b g_n(t)\,dt\right| \leq \frac{2}{n}.$$

In other words, $\langle g_n, 1_{[a,b]}\rangle \to 0$ for all $a < b$. This immediately extends to linear combinations, so $\langle g_n, f\rangle \to 0$ for all step functions f. As step functions are uniformly dense in the continuous functions (in the supremum norm, hence also in the q-norm as we are in a finite-measure context), and the continuous functions are dense in $L^q[0,1]$, we deduce via Hölder that the step functions are dense in $L^q[0,1]$. Then, given $\varepsilon > 0$, for any $f \in L^q[0,1]$ there exists a step function f_0 such that $\|f - f_0\|_q < \varepsilon$. This implies that (since $\|g_n\|_p \leq 1$ for all n)

$$\limsup_n |\langle g_n, f\rangle| \leq \limsup_n |\langle g_n, f_0\rangle| + \limsup_n \|g_n\|_q \, \|f_0 - f\|_q \leq \varepsilon.$$

By the Limsup Routine, $\langle g_n, f\rangle \to 0$.

The pointwise limit already fails at $t = 1/2$, for the sequence $\{g_n(1/2)\}$ takes the values $0, 1, -1$ infinitely often. Also, when t is irrational the sequence $\{\sin n\pi t\}_n$ is dense in $[-1,1]$, and so the limit does not exist a.e.

(7.1.20) Generalize the idea in Remark 7.1.13 to construct an example of a weakly-convergent unbounded net in $\ell^p(\mathbb{N})$ for $p \in [1, \infty)$.

Answer. We consider the set $R = \{n^{1/q} e_n : \ n \in \mathbb{N}\}$, where q is the conjugate exponent to p and $\{e_n\}$ is the canonical basis. Suppose that $0 \notin \overline{R}^{\sigma(\ell^p(\mathbb{N}), \ell^q(\mathbb{N}))}$. Then there is a weak-open neighbourhood of 0

$$W = \{f : \ |\langle g_j, f\rangle| < 1 : \ j = 1, \ldots, N\}$$

with $W \cap \overline{R}^{\sigma(\ell^p(\mathbb{N}), \ell^q(\mathbb{N}))} = \varnothing$, where we are using Proposition 5.6.3. So for each $n \in \mathbb{N}$ we can find an index $j_n \in \{1, \ldots, N\}$ with

$$1 \leq |\langle g_{j_n}, n^{1/q} e_n\rangle|.$$

Then

$$\sum_{j=1}^N \|g_j\|_q^q = \sum_{j=1}^N \sum_{n=1}^\infty |\langle g_j, e_n\rangle|^q = \sum_{n=1}^\infty \sum_{j=1}^N |\langle g_j, e_n\rangle|^q \geq \sum_{n=1}^\infty \frac{1}{n} = \infty.$$

It follows that $0 \in \overline{R}^{\sigma(\ell^p(\mathbb{N}), \ell^q(\mathbb{N}))}$.

7.2. The weak* topology

(7.2.1) Show that the natural embedding $\iota : \mathcal{X} \to \mathcal{X}^{**}$ given by $\iota(x) = \hat{x}$ is a linear isometry.

Answer. Linearity follows from

$$\begin{aligned}\iota(\alpha x + y)(g) &= g(\alpha x + y) = \alpha g(x) + g(y)\\ &= \alpha\iota(x)(g) + \iota(y)(g), \qquad g \in \mathcal{X}^*.\end{aligned}$$

As this works for all g, we we $\iota(\alpha x + y) = \alpha\iota(x) + \iota(y)$.

Isometry: this was already done in (7.2). We have

$$\begin{aligned}\|\hat{x}\| &= \sup\{|\hat{x}(g)| : \ g \in \mathcal{X}^*, \ \|g\| = 1\}\\ &= \sup\{|g(x)| : \ g \in \mathcal{X}^*, \ \|g\| = 1\} = \|x\|\end{aligned}$$

by Corollary 5.7.7.

(7.2.2) Show that the sets $N(x_1, \dots, x_k; \varepsilon)$, where $x_1, \dots, x_k \in \mathcal{X}$ are linearly independent and $\varepsilon > 0$, form a base for $\sigma(\mathcal{X}^*, \mathcal{X})$.

Answer. We will show that these neighbourhoods generate the same open sets than the "non-linearly independent" neighbourhoods. If $x_1, \dots, x_n$ are linearly dependent, after reordering and removing duplicates we may assume that the first $x_1, \dots, x_r$ are linearly independent, and that $x_{r+1}, \dots, x_n$ are linear combinations of those, say

$$x_{r+j} = \sum_{k=1}^{n} c_{jk} x_k.$$

Then, for any $\varphi \in \mathcal{X}^*$,

$$|\varphi(x_{r+j})| \leq \sum_{k=1}^{r} |c_{jk}|\, |\varphi(x_k)| \leq \left(\sum_{k=1}^{r} |c_{jk}|\right) \max\{|\varphi(x_k)| : \ k = 1, \dots, r\}.$$

If we put $c = \max\{(\sum_{k=1}^{r} |c_{jk}|) : \ j = 1, \dots, n - r\}$, then

$$N(x_1, \dots, x_r; \varepsilon/c\} \subset N(x_1, \dots, x_n; \varepsilon\} \subset N(x_1, \dots, x_r; \varepsilon).$$

So these "linearly independent" neighbourhoods generate the same topology.

(7.2.3) Prove Lemma 7.2.2.

Answer. Let $\varphi, \psi \in \mathcal{X}^*$ with $\varphi \neq \psi$. By definition, this means that there exists $x \in \mathcal{X}$ with $\varphi(x) \neq \psi(x)$. Let $\delta = |\varphi(x) - \psi(x)|/2$. Then the weak* open sets

$$V_\varphi = \hat{x}^{-1}(B_\delta(\varphi(x)), \quad V_\psi = \hat{x}^{-1}(B_\delta(\psi(x)))$$

are disjoint, $\varphi \in V_\varphi$, $\psi \in V_\psi$.

(7.2.4) Prove Proposition 7.2.3.

Answer. Assume first that $\varphi_j \xrightarrow{\text{weak}^*} \varphi$. Fix $x \in \mathcal{X}$, and let $\varepsilon > 0$. Since $\hat{x}^{-1}(B_\varepsilon(\varphi(x)))$ is a weak*-open neighbourhood of φ, there exists j_0 such that, for any $j > j_0$, $\varphi_j \in \hat{x}^{-1}(B_\varepsilon(\varphi(x)))$, which means exactly that $|\varphi_j(x) - \varphi(x)| < \varepsilon$. So $\varphi_j(x) \to \varphi(x)$.

Conversely, if $\varphi_j(x) \to \varphi(x)$ for all $x \in \mathcal{X}$, let V be a weak*-open neighbourhood of φ. Then there exist $x_0 \in \mathcal{X}$, $\varepsilon > 0$, $c \in \mathbb{C}$, such that $\varphi \in \hat{x}_0^{-1}(B_\varepsilon(c)) \subset V$. Since $\varphi(x_0) \in B_\varepsilon(c)$, there exists $\varepsilon' < \varepsilon$ with $B_{\varepsilon'}(\varphi(x_0)) \subset B_\varepsilon(c)$. So $\varphi \in \hat{x}_0^{-1}(B_{\varepsilon'}(\varphi(x_0))$. Since $\varphi_j(x_0) \to \varphi(x_0)$, there exists j_0 such that, for all $j > j_0$, $|\varphi_j(x_0) - \varphi(x_0)| < \varepsilon'$, which means that for $j > j_0$ we have $\varphi_j \in \hat{x}_0^{-1}(B_{\varepsilon'}(\varphi(x_0))) \subset V$. As V was arbitrary, $\varphi_j \xrightarrow{\text{weak}^*} \varphi$.

(7.2.5) Prove Proposition 7.2.6 *(Hint: take a good look at the proof of Proposition 7.1.11)*. Show also that the completeness of $\mathcal{X}$ is crucial and cannot be dispensed with.

Answer. Let $\{\varphi_n\} \subset \mathcal{X}$ with $\varphi_n \xrightarrow{\text{weak}^*} \varphi$. For any $x \in \mathcal{X}$ we have by definition that $\varphi_n(x) \to \varphi(x)$. The numeric sequence $\{\varphi_n(x)\}$ is convergent, and thus bounded. Then

$$\sup\{|\varphi_n(x)| : \ n \in \mathbb{N}\} < \infty$$

for all $x \in \mathcal{X}$. By the Uniform Boundedness Principle (Theorem 6.3.16), which applies since $\mathcal{X}$ is complete ,we get

$$\sup\{\|\varphi_n\| : \ n \in \mathbb{N}\} < \infty.$$

For the case where $\mathcal{X}$ fails to be complete, consider $\mathcal{X} = c_{00}$. Then $\mathcal{X}^* = \ell^1(\mathbb{N})$. For each $n \in \mathbb{N}$, let $\varphi_n = n\,\delta_n$, which can be seen as the element

$n\,e_n \in \ell^1(\mathbb{N})$. Then $\|\varphi_n\| = n$, but $\lim_n \varphi_n(x) = 0$ for all $x \in c_{00}$ since x has only finitely many nonzero entries.

(7.2.6) Let $\mathcal{X}$ be a normed space and $\{\varphi_n\}$ a weak*-convergent net in $\mathcal{X}^*$ with $\varphi_n \xrightarrow{\text{weak}^*} \varphi$. Show that $\|\varphi\| \leq \liminf_n \|\varphi_n\|$. Find an example where the inequality is strict.

Answer. We have, by definition of the norm in $\mathcal{X}^*$,

$$\|\varphi\| = \sup\{|\varphi(x)| : \ x \in \mathcal{X}, \ \|\mathcal{X}\| \leq 1\}.$$

For any $x \in \mathcal{X}$ with $\|x\| \leq 1$, since $|\varphi_n(x)| \leq \|\varphi_n\|$, and choosing a subnet $\{\varphi_{n_k}\}$ such that $\lim_k \|\varphi_{n_k}\| = \liminf_n \|\varphi_n\|$,

$$|\varphi(x)| = \lim_n |\varphi_n(x)| = \lim_k |\varphi_{n_k}(x)| \leq \lim_k \|\varphi_{n_k}\| = \liminf_n \|\varphi_n\|.$$

For an example, it was shown in Example 7.1.6 that an orthonomal basis on an infinite-dimensional Hilbert space is a sequence $\{\xi_n\}$ with $\|\xi_n\| = 1$ for all n and $\xi_n \xrightarrow{\text{weak}} 0$. We can consider $\{\xi_n\} \subset \mathcal{H}^*$ via $\xi_n(\eta) = \langle \eta, \xi_n \rangle$. In this context we have $\xi_n \xrightarrow{\text{weak}^*} 0$.

(7.2.7) Prove the assertions in Examples 7.2.7.

Answer.

(i) For the weak convergence $e_j \xrightarrow{\text{weak}} 0$ in c_0, take $x \in \ell^1(\mathbb{N})$. Then

$$\langle e_j, x \rangle = \sum_k e_j(k) x_k = x_j \to 0,$$

since $\sum_k |x_k| < \infty$.

(ii) Now when we consider $e_j \in \ell^1(\mathbb{N})$, the weak*-topology is given by the elements of c_0. For $x \in c_0$,

$$\langle e_j, x \rangle = \sum_k e_j(k) x_k = x_j \to 0.$$

For weak convergence, we can now use any $x \in \ell^\infty(\mathbb{N})$. If we take $x = 1_{2\mathbb{N}}$, then

$$\langle e_j, x \rangle = \begin{cases} 1, & j \in 2\mathbb{N} \\ 0, & j \notin 2\mathbb{N} \end{cases}$$

so the limit doesn't exist.

(iii) Consider, for each $n \in \mathbb{N}$, the element $y_n = \frac{1}{n}\sum_{j=1}^{n} e_j \in K$. For any $x \in c_0$,

$$\langle y, x\rangle = \frac{1}{n}\sum_{j=1}^{n}\langle e_j, x\rangle = \frac{1}{n}\sum_{j=1}^{n} x(j).$$

Given $\varepsilon > 0$, there exists j_0 such that $|x(j)| < \varepsilon/2$ for all $j \geq j_0$. Then

$$|\langle y_n, x\rangle| \leq \frac{1}{n}\sum_{j=1}^{n_0}|x(j)| + \frac{1}{n}\sum_{n_0+1}^{n}|x(j)| \leq \frac{1}{n}\sum_{j=1}^{n_0}|x(j)| + \frac{\varepsilon}{2}.$$

It now follows that, for n big enough, $|\langle y_n, x\rangle| < \varepsilon$. So $\langle y_n, x\rangle \to 0$. Thus $0 \in \overline{K}^{w^*}$.

When we consider the weak topology, the dual is $\ell^\infty(\mathbb{N})$ and now we can take $x = 1$. For any $y \in K$ we have

$$0 = \langle 0, x\rangle < \frac{1}{2} < 1 = \frac{1}{n}\sum_{j=1}^{n}\langle e_j, x\rangle.$$

The above says that the functional 1 strictly separates $\overline{K}^{w}$ and $\{0\}$. Thus $0 \notin \overline{K}^{w}$.

(7.2.8) Let $\mathcal{X} = \ell^1[0,1]$, so that its dual is $\mathcal{X}^* = \ell^\infty[0,1]$ (see Exercise 5.6.11), and put

$$M = \{x \in \mathcal{X}^* : \operatorname{supp} x \ \text{ is countable}\}.$$

Show that M is closed in norm and in the weak topology, and that it is dense in the weak*-topology.

Answer. Since $\mathcal{X}^*$ with the norm topology is a metric space, we can deal with sequences instead of nets. So let $\{x_n\} \subset M$ be a Cauchy sequence. As $\mathcal{X}^*$ is complete, there exists $x = \lim x_n \in \mathcal{X}^*$. If S_n denotes the support of x_n, put $S = \bigcup_n S_n$. Then, as $x(k) = \lim_n x_n(k)$ for each k (from $|x(k) - x_n(k)| \leq \|x - x_n\|_\infty$), we have that $\operatorname{supp} x \subset S$, which is countable; thus $x \in M$. Then M is closed in norm. As M is convex (it is a subspace), its weak closure agrees with its norm closure (Theorem 7.1.16), so

$$\overline{M}^{\sigma(\mathcal{X}^*,\mathcal{X}^{**})} = \overline{M} = M.$$

In the weak* topology, though, M is dense. To see this, fix $z \in \ell^\infty[0,1]$. A weak*-neighbourhood of z is of the form

$$W = \{y \in \ell^\infty[0,1] : \ |\langle (y - z), w_j\rangle| < 1, \ j = 1, \ldots, m\}.$$

where $w_1, \ldots, w_m \in \ell^1[0,1]$. Let $S = \bigcup_{j=1}^m \operatorname{supp} w_j$, and let $x = z|_S$ (that is, $x(k) = z(k)$ if $k \in S$, and 0 otherwise). Then $x \in M$ and $x \in W$ (the latter, because $\langle (x-z), w_j \rangle = 0$ for all j). This shows that there exists a net $\{x_\alpha\} \subset M$ such that $x_\alpha \xrightarrow{\text{weak}^*} z$.

(7.2.9) Let $\mathcal{X}$ be a finite-dimensional normed space. Show that $\mathcal{X}$ is reflexive.

Answer. Since $\dim \mathcal{X} < \infty$, the closed unit ball is compact (Corollary 5.2.4). Then the closed unit ball is weakly compact, and $\mathcal{X}$ is reflexive by Proposition 7.2.21.

(7.2.10) Complete the proof of Lemma 7.2.17 by showing that if (i) is false, then $\alpha \notin \overline{\gamma(B_1^{\mathcal{X}}(0))}$.

Answer. The negation of (i) means that there exists $\varepsilon > 0$ such that for all $x \in B_1^{\mathcal{X}}(0)$ there exists $k \in \{1, \ldots, n\}$ with $|\varphi_k(x) - \alpha_k| \geq \varepsilon$. Let

$$d = \operatorname{dist}\left(\alpha, \overline{\gamma(B_1^{\mathcal{X}}(0))}\right) = \operatorname{dist}\left(\alpha, \gamma(B_1^{\mathcal{X}}(0))\right).$$

Since d is an infimum, there exists $x \in B_1^{\mathcal{X}}(0)$ such that $\|\alpha - \gamma(x)\| \leq d + \frac{\varepsilon}{2}$. By the above there exists k such that $|\varphi_k(x) - \alpha_k| \geq \varepsilon$. Then

$$\varepsilon \leq |\varphi_k(x) - \alpha_k\| \leq \|\gamma(x) - \alpha\| \leq d + \frac{\varepsilon}{2},$$

and we obtain that $d \geq \frac{\varepsilon}{2}$.

(7.2.11) By Goldstine's Theorem the unit ball of c_0 is weak*-dense in the unit ball of $\ell^\infty(\mathbb{N})$. Given $f \in \ell^\infty(\mathbb{N})$ with $\|f\| \leq 1$, find a net $\{g_j\} \subset c_0$ such that $\|g_j\| = 1$ for all j, and $g_j \to f$ in the weak*-topology.

Answer. Define $g_m = e_{m+1} + \sum_{n=1}^m f(n)\, e_n$. Then $g_m \in c_0$ (it is in c_{00}, actually) and $\|g_m\|_\infty = 1$. For any $x \in \ell^1(\mathbb{N})$,

$$|\langle f - g_m, x\rangle| = \left|(f(m) - 1)x(m) + \sum_{n > m+1} f(n)x(n)\right|$$

$$\leq 2 \sum_{n>m} |x(n)| \xrightarrow[m\to\infty]{} 0$$

since $x \in \ell^1(\mathbb{N})$.

(7.2.12) Let $\mathcal{X}$ be a normed space, and $\varphi : \mathcal{X}^* \to \mathbb{C}$ a weak*-continuous linear map. Use the Closed Graph Theorem to show that φ is bounded *(This is not hard, though it is not the easiest way to prove this)*.

Answer. Let $\{f_n\} \subset \mathcal{X}^*$ such that $f_n \to f$ and $\varphi(f_n) \to c$. By the weak*-continuity,

$$c = \lim_n \varphi(f_n) = \varphi(f).$$

So the graph of φ is closed, and φ is bounded by the Closed Graph Theorem (Theorem 6.3.12).

(7.2.13) Let $\mathcal{X}$ be a normed space and $\psi \in \mathcal{X}^{**}$ nonzero. Show that there exists $\{x_j\} \subset \mathcal{X}$ with $\hat{x}_j \xrightarrow{\text{weak}^*} \psi$ and $\|x_j\| = \|\psi\|$ for all j.

Answer. Assume without loss of generality that $\|\psi\| = 1$. By Theorem 7.2.18 there exists $\{z_j\} \subset \mathcal{X}$ with $\|z_j\| \leq 1$ and $\hat{z}_j \xrightarrow{\text{weak}^*} \psi$. Fix $\varepsilon > 0$. By definition of the norm, there exists $g \in \mathcal{X}^*$ such that $\|g\| = 1$ and $\psi(g) \geq 1 - \varepsilon$. Since

$$1 \geq \|z_j\| \geq g(z_j) \to \psi(g) \geq 1 - \varepsilon,$$

there exists $j(\varepsilon)$ such that $\|z_j\| \geq 1 - \varepsilon/2$ for all $j \geq j(\varepsilon)$. That is, $\|z_j\| \to 1$. Then $x_j = z_j / \|z_j\|$ satisfies $\|x_j\| = 1$ and for any $g \in \mathcal{X}^*$

$$\hat{x}_j(g) - \psi(g) = \frac{g(z_j)}{\|z_j\|} - g(z_j) + g(z_j) - \psi(g) \to 0.$$

7.3. Polars and Prepolars

(7.3.1) Show that if $X \subset \mathcal{X}$ is a subspace, then

$$X^o = \{\varphi \in \mathcal{X}^* : \ \varphi(x) = 0 \ \text{ for all } x \in X\}.$$

Similarly, show that if $Y \subset \mathcal{X}^*$ is a subspace, then

$$Y_o = \{x \in \mathcal{X} : \ \varphi(x) = 0 \ \text{ for all } \varphi \in Y\}.$$

Answer. Let $\varphi \in X^o$. Since X is a subspace, given any $x \in X$ we have $nx \in X$ for all $n \in \mathbb{N}$. Then

$$1 \geq |\varphi(nx)| = n\,|\varphi(x)|.$$

As n is arbitrary, $\varphi(x) = 0$. The other inclusion is trivial.

Similarly, if $x \in Y_o$, for any $\varphi \in Y$ we have $|\varphi(x)| \leq 1$. As Y is a subspace, we get $n|\varphi(x)| \leq 1$ for all n, and so $\varphi(x) = 0$.

(7.3.2) Show that

$$\big[B_1^{\mathcal{X}}(0)\big]^o = \overline{B_1(0)^{\mathcal{X}^*}}, \qquad \big[B_1(0)^{\mathcal{X}^*}\big]_o = \overline{B_1(0)^{\mathcal{X}}}.$$

Answer. We have

$$\begin{aligned}\big[B_1^{\mathcal{X}}(0)\big]^o &= \{\varphi \in \mathcal{X}^* : \ |\varphi(x)| \leq 1 \ \text{ for all } \|x\| \leq 1\} \\ &= \{\varphi \in \mathcal{X}^* : \ \|\varphi\| \leq 1\} = \overline{B_1(0)^{\mathcal{X}^*}}.\end{aligned}$$

Similarly,

$$\begin{aligned}\big[B_1(0)^{\mathcal{X}^*}\big]_o &= \{x \in \mathcal{X} : \ |\varphi(x)| \leq 1 \ \text{ for all } \|\varphi\| \leq 1\} \\ &= \{x \in \mathcal{X} : \ \|x\| \leq 1\} = \overline{B_1(0)^{\mathcal{X}}}.\end{aligned}$$

(7.3.3) Prove Proposition 7.3.2.

Answer.

(i) The inequality $|\varphi(x)| \leq 1$ survives convex combinations, multiplication by scalars of absolute value at most 1, and pointwise limits.

(ii) The inequality $|\varphi(x)| \leq 1$ survives convex combinations, multiplication by scalars of absolute value at most 1, and pointwise limits.

(iii) If $\varphi \in X_2^o$, then $|\varphi(x)| \leq 1$ for all $x \in X_2$; in particular, for all $x \in X_1$, so $\varphi \in (X_1)^o$.

(iv) If $x \in (Y_2)_o$, then $|\varphi(x)| \leq 1$ for all $\varphi \in Y_2$; in particular, for all $\varphi \in Y_1$. Thus $x \in (Y_1)_o$.

(v) Let $x \in \Big(\bigcup_j Y_j\Big)_o$. If $\varphi = \sum_k t_k\varphi_k$ with each φ_k in some Y_j, $t_k \geq 0$ for all k, and $\sum_k t_k = 1$, then

$$\Big|\sum_k t_k\varphi_k(x)\Big| \leq \sum_k t_k|\varphi_k(x)| \leq \sum_k t_k = 1.$$

So $x \in \Big(\operatorname{conv}\bigcup_j Y_j\Big)_o$. The reverse inclusion is automatic by (iv).

(vi) Let $\varphi \in \bigcup_j X_j^o$. Then $\varphi \in X_{j_0}^o$ for some j_0. If $x \in \bigcap_j X_j$, then $x \in X_{j_0}$ and so $|\varphi(x)| \leq 1$.

For the failure of the reverse inclusion, let $\mathcal{X} = \mathcal{X}^* = \mathbb{C}$. Take $X_1 = \{0,1\}$, $X_2 = \{0,2\}$. Then $X_1 \cap X_2 = \{0\}$, so $\Big(\bigcap_j X_j\Big)^o = \mathcal{X}^* = \mathbb{C}$, while $X_1^o = \overline{\mathbb{D}}$ and $X_2^o = \frac{1}{2}\,\overline{\mathbb{D}}$.

(vii) Let $\varphi \in \bigcap_j X_j^o$. If $x \in \bigcup_j X_j$, then $x \in X_{j_0}$ for some j_0, and so $|\varphi(x)| \leq 1$ since $\varphi \in X_{j_0}^o$. Conversely, if $\varphi \in \Big(\bigcup_j X_j\Big)^o$, then $|\varphi(x)| \leq 1$ for all $x \in X_j$, for all j. Then $\varphi \in X_j^o$ for all j.

(viii) The inequality $|\varphi(x)| \leq 1$ is weakly continuous on x (and the weak closure of the ball agrees with the norm closure by convexity) and weak*-continuous on φ.

(7.3.4) Prove Corollary 7.3.5.

Answer. Since X, Y are subspaces, $\operatorname{cb} X = X$, $\operatorname{cb} Y = Y$. If X is dense in $\mathcal{X}$, by Proposition 7.3.4 $(X^o)_o = \overline{X} = \mathcal{X}$. This says that each functional in X^o is zero on all of $\mathcal{X}$, so $X^o = \{0\}$. Conversely, if $X^o = \{0\}$ then again by Proposition 7.3.4 $\overline{X} = (X^o)_o = \mathcal{X}$.

If Y is weak*-dense in $\mathcal{X}^*$, then by Proposition 7.3.4

$$\mathcal{X}^* = \overline{Y}^{w^*} = (Y_o)^o.$$

This means that every $f \in \mathcal{X}^*$ is zero on Y_o, so $Y_o = \{0\}$ by Hahn–Banach (Corollary 5.7.19). Conversely, if $Y_o = \{0\}$, then by Proposition 7.3.4

$$\overline{Y}^{w^*} = (Y_o)^o = \{0\}^o = \mathcal{X}^*.$$

(7.3.5) Let $\mathcal{X}$ be a normed space, and $K \subset \mathcal{X}$ a subspace. Show that $K^{oo} = \overline{J_{\mathcal{X}} K}^{w^*}$, where $J_{\mathcal{X}}$ is the usual embedding of $\mathcal{X}$ in $\mathcal{X}^{**}$.

Answer. We first verify that $J_K K \subset K^{oo}$. Given $x \in K$ and $\varphi \in K^o$,

$$(J_K x)\varphi = \varphi(x) = 0,$$

so $J_K x \in K^{oo}$ since K is a subspace.

We have that K^{oo} is weak*-closed by Proposition 7.3.2, so $\overline{J_{\mathcal{X}} K}^{w^*} \subset K^{oo}$. Conversely, let $\Phi \in K^{oo} \setminus \overline{J_{\mathcal{X}} K}^{w^*}$ be nonzero. By geometric Hahn–Banach (Theorem 5.7.18) and Proposition 7.2.10 there exists $\varphi \in \mathcal{X}^*$ such that $\Phi(\varphi) = 1$ and $\varphi|_K = 0$. But then $\varphi \in \mathcal{K}^o$ for all $n \in \mathbb{N}$ and $\Phi(2\varphi) = 2 > 1$, contradicting that $\Phi \in K^{oo}$. Thus $K^{oo} = \overline{J_{\mathcal{X}} K}^{w^*}$.

7.4. Spaces of Continuous Functions

(7.4.1) Show that $C_0(T)$ is complete.

Answer. This is the usual argument that uniform convergence of continuous functions is continuous, together with the vanishing at infinity part.

Let $\{f_n\} \subset C_0(T)$ be Cauchy. From $|f_n(t) - f_m(t)| \leq \|f_n - f_m\|_\infty$ we deduce that $\{f_n(t)\}$ is Cauchy for every $t \in T$; and as $\mathbb{C}$ is complete, the pointwise limit $f(t) = \lim_n f_n(t)$ exists. And the convergence is uniform: if $\varepsilon > 0$ is given and $\|f_n - f_m\| < \varepsilon$ when n, m are big enough,

$$|f(t) - f_m(t)| = \lim_n |f_n(t) - f_m(t)| \leq \limsup_n \|f_n - f_m\| < \varepsilon.$$

Fix $\varepsilon > 0$ and $s \in T$. There exists n such that $\|f - f_n\| < \varepsilon/3$. As f_n is continuous, there exists an open neighbourhood V of s such that $|f_n(t) - f_n(s)| < \varepsilon/3$ if $t \in V$. In such case,

$$\begin{aligned}|f(t) - f(s)| &\leq |f(t) - f_n(t)| + |f_n(t) - f_n(s)| + |f_n(s) - f(t)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

So f is continuous. Similarly, there exists $K \subset T$ compact with $|f_n| < \varepsilon/2$ on $T \setminus K$. Then, for $t \in T \setminus K$ and n big enough,

$$|f(t)| \leq |f(t) - f_n(t)| + |f_n(t)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{2} < \varepsilon.$$

So f vanishes at infinity, which shows that $f \in C_0(T)$ and, as a consequence, that $C_0(T)$ is complete.

(7.4.2) Show that if T is locally compact Hausdorff, then $C_0(T)$ separates points: that is, given $s, t \in T$ with $s \neq t$, there exists $f \in C_0(T)$ with $f(s) \neq f(t)$.

Answer. As T is Hausdorff, there exists V open with $s \in V$, $t \notin V$. Applying Urysohn's Lemma with $K = \{s\}$ and V, there exists $f \in C_c(T)$ with $f(s) = 1$ and $f(t) = 0$.

(7.4.3) Show that if X, Y are topological vector spaces, the following statements are equivalent:

(i) $C_{\mathbb{R}}(X)$ and $C_{\mathbb{R}}(Y)$ are isomorphic as rings;

(ii) $C_{\mathbb{R}}(X)$ and $C_{\mathbb{R}}(Y)$ are isomorphic as real algebras.

Answer. Only one implication needs to be proven, since algebras are rings. So suppose that $\Gamma : C_{\mathbb{R}}(X) \to C_{\mathbb{R}}(Y)$ is a ring isomorphism. For any $n \in \mathbb{N}$ we have $\Gamma(nf) = \Gamma(f + \cdots + f) = n\Gamma f$. We also have, for $n \in \mathbb{N}$, $\Gamma((-n)f) + \Gamma(nf) = \Gamma 0 = 0$, so $\Gamma((-n)f) = -\Gamma(nf) = (-n)\Gamma f$. For nonzero $m \in \mathbb{Z}$, $m\Gamma(f/m) = \Gamma(mf/m) = \Gamma f$. Hence for any $q \in \mathbb{Q}$, written as $q = n/m$,

$$\Gamma(qf) = \Gamma\left(\frac{m}{n} f\right) = \frac{m}{n} \Gamma f = q \Gamma f.$$

If $f \geq 0$, we can write $f = g^2$ with $g = \sqrt{f}$. Then $\Gamma f = \Gamma(g^2) = (\Gamma g)^2 \geq 0$. So Γ preserves order. Fix $r \in \mathbb{R}$ and sequences $\{p_n\}, \{q_n\} \subset \mathbb{Q}$ with $p_n \nearrow r$, $q_n \searrow r$. For any $f \geq 0$ we have $p_n f \leq rf \leq q_n f$. Then

$$p_n \Gamma f = \Gamma(p_n f) \leq \Gamma(rf) \leq \Gamma(q_n f) = q_n \Gamma f.$$

Taking limit we get $r\Gamma f \leq \Gamma(rf) \leq r\Gamma f$ and so $\Gamma(rf) = r\Gamma f$. For arbitrary f we can write $f = f^+ - f^-$ and the additivity of Γ gives $\Gamma(rf) = \Gamma(rf^+) - \Gamma(rf^-) = r\Gamma f$; so Γ is real linear.

(7.4.4) Show that the function d defined on (7.22) is indeed a metric, and that it induces the topology of T.

Answer. If $d(t,s) = 0$, then $f_n(t) = f_n(s)$ for all n; using the density, $f(t) = f(s)$ for all $f \in C(T)$. Then by Urysohn's Lemma (Theorem 2.6.5), $s = t$. The triangle inequality follows directly:

$$\begin{aligned} d(t,s) &= \sum_{n=1}^{\infty} 2^{-n} \left|f_n(t) - f_n(s)\right| \\ &\leq \sum_{n=1}^{\infty} 2^{-n} \left|f_n(t) - f_n(r)\right| + \left|f_n(r) - f_n(s)\right| \\ &= \sum_{n=1}^{\infty} 2^{-n} \left|f_n(t) - f_n(r)\right| + \sum_{n=1}^{\infty} \left|f_n(r) - f_n(s)\right| \\ &= d(t,r) + d(r,s). \end{aligned}$$

There is no issue manipulating the series as the condition $\|f_n\| \leq 1$ guarantees that they converge absolutely and uniformly. That $d(t,s) = d(s,t)$ follows from $|f_n(t) - f_n(s)| = |f_n(s) - f_n(t)|$. So d is a distance.

If $d(t_j, t) \to 0$, then for each n we have $f_n(t_j) \to f_n(t)$. Fix $\varepsilon > 0$; as $\{f_n\}$ is dense in the unit ball, if $f \in C(T)$ and $\|f\| \leq 1$ then there exists n such that $\|f - f_n\| < \varepsilon$. Hence

$$\begin{aligned} |f(t_j) - f(t)| &\leq |f(t_j) - f_n(t_j)| + |f_n(t_j) - f_n(t)| + |f_n(t) - f(t)| \\ &\leq 2\|f - f_n\| + |f_n(t_j) - f_n(t)|. \end{aligned}$$

Then $\limsup_j |f(t_j) - f(t)| \leq 2\varepsilon$. By the Limsup Routine, $\lim_j f(t_j) = f(t)$. This also works for arbitrary $f \in C(T)$ as we can scale it into the unit ball. Then $t_j \to t$ by Corollary 2.6.7.

Conversely, suppose that $t_j \to t$. Fix $\varepsilon > 0$ and choose n_0 such that we have $\sum_{n>n_0} 2^{-n} < \varepsilon/4$. Choose j_0 such that, for $j \geq j_0$, $|f_n(t_j) - f_n(t)| < \varepsilon/2$, $n = 1, \ldots, n_0$. Then, when $j \geq j_0$,

$$\begin{aligned} d(t_j, t) &= \sum_{n=1}^{n_0} 2^{-n} \left|f_n(t_j) - f(t)\right| + \sum_{n>n_0} 2^{-n} \left|f_n(t_j) - f(t)\right| \\ &\leq \sum_{n=1}^{n_0} 2^{-n} \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

Thus $d(t_j, t) \to 0$.

(7.4.5) Prove that a compact metric space is separable.

Answer. For each n, we have $T \subset \bigcup_{t\in T} B_{1/n}(t)$. By compactness, there exist $t_{n,1}, \ldots, t_{n,k(n)}$ such that for each $t \in T$ there exists j with $|t - t_{n,j}| < 1/n$. Then $\bigcup_n \{t_{n,1}, \ldots, t_{n,k(n)}\}$ is countable and dense.

(7.4.6) Let $\{k_s\}_{s\in\mathbb{R}} \subset C_0(\mathbb{R})$ be the family of functions $k_s(t) = st(1 + s^2t^2)^{-1}$, and $h(t) = (1+t)^{-1}$. Show that the algebra generated by $\{h\} \cup \{k_s\}$ is dense in $C_0(\mathbb{R})$.

Answer. We want to apply Stone–Weierstrass. Let $\mathcal{A}$ be the complex algebra generated by the $\{k_s\}$ and h. The functions are real-valued, so $\mathcal{A}$ is selfadjoint. They separate points, for if $x \neq y$, we can assume without loss of generality that $x \neq 0$ (otherwise, switch roles) and consider $k_1(x) \neq 0$; since $k_s(t) \xrightarrow[s\to\infty]{} 0$ we may choose s so that $|k_s(y)| < |k_1(x)|$. The algebra $\mathcal{A}$ also vanishes nowhere, for $h(0) = 1$ and for $t \neq 0$ we have $k_s(t) \neq 0$ for all s. So the algebra $\mathcal{A}$ is selfadjoint, it separates points, and it vanishes nowhere. By Corollary 7.4.23, $\mathcal{A}$ is dense in $C_0(\mathbb{R})$.

(7.4.7) Let S, T be locally compact Hausdorff spaces. Show that

$$\overline{\operatorname{span} C_0(S)C_0(T)} = C_0(S \times T),$$

where each product fg is identified with the function $(s,t) \longmapsto f(s)g(t)$.

Answer. If we write $\alpha(f,g)$ for the function $(s,t) \longmapsto f(s)g(t)$ as above, let

$$\mathcal{A} = \operatorname{span}\{\alpha(f,g) : \ f \in C_0(S), \ g \in C_0(T)\}.$$

It is clear that $\mathcal{A}$ is an algebra, for $\alpha(f_1,g_1)\alpha(f_2,g_2) = \alpha(f_1f_2, g_1g_2)$. Suppose first that S, T are compact. Then $S \times T$ is compact. We have $1 = \alpha(1,1) \in \mathcal{A}$. Also, $\mathcal{A}$ is selfadjoint for $\alpha(f,g)^* = \alpha(f^*, g^*)$. And, given $(s_1,t_1),(s_2,t_2) \in S \times T$ by Urysohn's Lemma there exists $f \in C_0(S)$ with $f(s_1) = 1$, $f(s_2) = 0$. The function $\alpha(f,1)$ takes the value 1 at (s_1,t_1), and 0 at (s_2,t_2). Then Stone–Weierstrass guarantees that $\mathcal{A}$ is dense in $C(S \times T)$.

When at least one of S and T is not compact, given $(s,t) \in S \times T$ we can still get $f \in C_0(S)$ and $g \in C_0(T)$ with $f(s) = g(t) = 1$. Then

$\alpha(f,g)$ is nonzero at (s,t) and so $\mathcal{A}$ vanishes nowhere. The argument from the previous paragraph still applies to show that $\mathcal{A}$ still separates points. By Corollary 7.4.23, $\mathcal{A}$ is dense.

(7.4.8) Show that $c_0 = C_0(\mathbb{N})$.

Answer. We know that $c_0 = \{g : \mathbb{N} \to \mathbb{C} : \lim_n g(n) = 0\}$. The compact sets in $\mathbb{N}$ are the finite sets, so $g \in C_0(\mathbb{N})$ if and only if $\lim_n g(n) = 0$. Indeed, if $x\ C_0(\mathbb{N})$ and $\varepsilon > 0$, there exists $K \subset \mathbb{N}$ compact with $|x(n)| < \varepsilon$ for all $n \notin K$. With $n_0 = \max K$, we have that if $n > n_0$ then $|x(n)| < \varepsilon$, showing that $\lim_n x(n) = 0$. Conversely, if $\lim_n x(n) = 0$, by definition of limit given $\varepsilon > 0$ there exists n_0 with $|x(n)|\varepsilon$ whenever $n > n_0$. Then $K = \{1, \ldots, n_0\}$ is compact and for $n \notin K$ we have $|x(n)| < \varepsilon$.

(7.4.9) Show that the bounded set $\mathcal{R} = \{f \in C([0,1] : 0 \leq f \leq 1,\ f|_{[\frac12,1]} = 0\}$ in $C[0,1]$ admits no supremum.

Answer. Let g be an upper bound for $\mathcal{R}$. Then $g \geq 1$ on $[0,1/2]$ and $g \geq 0$ on $[1/2,1]$. As $g(1/2) \geq 1$ and g is continuous, there exists $\delta > 0$ such that $g(t) \geq \frac12$ for all $t \in (\frac12, \frac12 + \delta)$. Let

$$h(t) = \begin{cases} \frac14 - \frac{1}{2\delta}\left|t - \frac12 - \frac{\delta}{2}\right|, & t \in \left(\frac12, \frac12 + \delta\right) \\ 0, & t \in [0,1] \setminus \left(\frac12, \frac12 + \delta\right) \end{cases}$$

Then $h \in C[0,1]$ with $0 \leq h \leq \frac14$, $h = 0$ outside of $(\frac12, \frac12 + \delta)$, and $h(1/2 + \delta/2) = 1/4$. Then $g - h \leq g$, $g - h \geq 1$ on $[0, \frac12]$ and $g - h \geq 0$ on $[1/2, 1]$. So $g - h$ is an upper bound for $\mathcal{R}$, showing that g cannot be a least upper bound.

(7.4.10) Let T be an extremally disconnected topological space, and $U, V \subset T$ disjoint open subsets. Show that $\overline{U} \cap \overline{V} = \varnothing$.

Answer. Because T is extremally disconnected, both $\overline{U}$ and $\overline{V}$ are open. From Exercise 1.8.32 we get $\overline{U} \cap V = \varnothing$. And as $\overline{U}$ is open, we can apply Exercise 1.8.32 again to get $\overline{U} \cap \overline{V} = \varnothing$.

(7.4.11) Prove Lemma 7.4.29 by modifying the proof of Lemma 5.7.3 appropriately.

Answer. Since v_0 and Z are linearly independent, for each $x \in Z + \mathbb{R}v_0$ there exist unique $c \in \mathbb{R}$ and $z \in Z$ with $x = cv_0 + z$. Fix $z_1, z_2 \in Z$; then

$$S(z_1 + z_2) \leq q(z_1 + z_2) \leq q(z_1 + v_0) + q(z_2 - v_0),$$

and we deduce that

$$S(z_2) - q(z_2 - v_0) \leq -S(z_1) + q(z_1 + v_0) \qquad \text{(AB.7.1)}$$

This can be done for all $z_1, z_2 \in Z$, so by the order-completeness of $C(T)$ (via Proposition 7.4.27) we conclude that there exists $d \in C_{\mathbb{R}}(T)$ with

$$d = \sup\{S(z_2) - q(v_0 - z_2) : z_2 \in Z\}.$$

By (AB.7.1) we also have

$$d \leq \inf\{-S(z_1) + q(v_0 + z_1) : \; z_1 \in Z\}.$$

Now define the linear map $\tilde{S}(cv_0 + z) = cd + Sz$; this is well-defined by the uniqueness of c and z. Then $\operatorname{ran} \tilde{S} \subset C_{\mathbb{R}}(T)$ and $\tilde{S}|_Z = S$. Fix $c \in \mathbb{R} \setminus \{0\}$, $z \in Z$. Suppose first that $c > 0$. Then

$$\begin{aligned} \tilde{S}(cv_0 + z) &= cd + Sz = c\left(d + S\left(\frac{z}{c}\right)\right) \\ &\leq c\left(-S\left(\frac{z}{c}\right) + q\left(v_0 + \frac{z}{c}\right) + S\left(\frac{z}{c}\right)\right) \\ &= q(cv_0 + z). \end{aligned}$$

Similarly, if $c < 0$,

$$\begin{aligned} \tilde{S}(cv_0 + z) &= cd + Sz = (-c)\left(-d + S\left(\frac{z}{-c}\right)\right) \\ &\leq (-c)\left(-S\left(\frac{z}{-c}\right) + q\left(\frac{z}{-c} - v_0\right) + S\left(\frac{z}{-c}\right)\right) \\ &= q(cv_0 + z). \end{aligned}$$

(7.4.12) Prove Proposition 7.4.30 by modifying the proof of Theorem 5.7.4 appropriately.

Answer. Let $\mathcal{F}$ be the family of all (W', S'), where W' is a subspace of V with $W \subset W'$, $S' : W' \to C_{\mathbb{R}}(T)$ extends S and $S'(x) \leq q(x)$ for all $x \in W'$; the family $\mathcal{F}$ is trivially nonempty as $(W, S) \in \mathcal{F}$. In $\mathcal{F}$, we consider the partial order $(W_1, S_1) \leq (W_2, S_2)$ if $W_1 \subset W_2$ and $S_2|_{W_1} = S_1$. Let $\{(W_j, S_j)\}$ be a chain in $\mathcal{F}$; put $W' = \bigcup_j W_j$, which being an increasing union of subspaces is a subspace of V, and let $S' : W' \to C_{\mathbb{R}}(T)$ be given by $S'x = S_j x$ if $x \in W_j$. The compatibility given by the order guarantees that S' is well-defined. It is clear that $(W', S') \in \mathcal{F}$ and it is an upper bound for the chain. By Zorn's Lemma there exists a maximal element $(Z, \tilde{S})$ in $\mathcal{F}$. If $Z \subsetneq V$, we can use Lemma 7.4.29 to contradict the maximality of $(Z, \tilde{S})$. So $Z = V$ and $\tilde{S}$ is the desired extension. The condition $\tilde{S} \leq q$ comes for free since every element in $\mathcal{F}$ satisfies it.

(7.4.13) Prove Theorem 7.4.31 by adapting the proof of Theorem 5.7.5 appropriately.

Answer. In the real-valued case, we use Proposition 7.4.30 to get $\tilde{S} : V \to C_{\mathbb{R}}(T)$ with $\tilde{S}|_W = S$ and $\tilde{S}(x) \leq q(x)$ for all $x \in V$. Because q is a seminorm and thus $q(-x) = q(x)$, for any $x \in V$ we have the inequality $\tilde{S}(-x) \leq q(-x) = q(x)$; this we may write as $-q(x) \leq \tilde{S}(x)$. Together with $\tilde{S}(x) \leq q(x)$, this gives us $|\tilde{S}x(t)| \leq q(x)$ for all $t \in T$ and hence $\|\tilde{S}x\|_\infty \leq q(x)$.

Now consider the complex case. Let S_1 be the real-valued, real-linear map given by $(S_1 x)(t) = \operatorname{Re}(Sx)(t)$, $t \in T$. Then $|(S_1 x)(t)| \leq |(Sx)(t)| \leq q(x)$ for all $x \in W$ and $t \in T$, so we can apply the previous part of the proof to obtain $\tilde{S}_1 : V \to C_{\mathbb{R}}(T)$ with $\tilde{S}_1|_W = \operatorname{Re} S$ and $|\tilde{S}_1(x)| \leq q(x)$ for all $x \in W$. Using Lemma 5.7.1 (note that pointwise evaluation of a function is a linear functional), define a new map

$$(\tilde{S}x)(t) = (\tilde{S}_1 x)(t) - i\,(\tilde{S}_1(ix))(t), \qquad t \in T.$$

Then, if $x \in W$,

$$\begin{aligned}(\tilde{S}x)(t) &= \operatorname{Re}(Sx)(t) - i \operatorname{Re}(S(ix))(t) \\ &= \operatorname{Re}(Sx)(t) - i \operatorname{Re} i(Sx)(t) = (Sx)(t).\end{aligned}$$

Now fix $x \in V$. Let $|(\tilde{S}x)(t)|\, e^{i\theta} = (\tilde{S}x)(t)$ be the polar form of the complex number $(\tilde{S}x)(t)$. Then $|(\tilde{S}x)(t)| = e^{-i\theta}(\tilde{S}x)(t) = (\tilde{S}(e^{-i\theta}x))(t)$. Hence

$$\begin{aligned}|(\tilde{S}x)(t)| &= (\tilde{S}(e^{-i\theta}x))(t) = (\operatorname{Re}\tilde{S}(e^{-i\theta}x))(t) \\ &\leq q(e^{-i\theta}x) = q(x).\end{aligned}$$

(7.4.14) Let T be a discrete topological space. Show that βT is extremally disconnected without using Proposition 7.4.34. To show that an open set $V \subset \beta T$ has open closure, consider the function $f = 1_{V \cap \delta(T)} \in C(\delta(T))$.

Answer. Let $V \subset \beta T$ be open and let $f : T \to \mathbb{C}$ be $f = 1_{\delta^{-1}(V \cap \delta(T))} = 1_{V\cap\delta(T)} \circ \delta$. Then f is continuous since T is discrete and hence every function is continuous. By the universal property (7.18) of the Stone–Čech compactification there exists $\tilde{f} \in C(\beta T)$ with $\tilde{f} \circ \delta = f$. If $v \in V$, there exists a net $\{v_j\} \subset V \cap \delta(T)$ with $v_j \to v$. Then

$$\tilde{f}(v) = \lim_j f(\delta^{-1}(v_j)) = 1_{V\cap\delta(T)}(v_j) = 1.$$

By the continuity, $\tilde{f}(v) = 1$ for all $v \in \overline{V}$. Conversely, if $w \in \beta T \setminus \overline{V}$ then there exists a net $\{w_j\} \subset \delta(T)$ with $w_j \notin V$ for all j. Then

$$\tilde{f}(w) = \lim_j f(\delta^{-1}(w_j)) = 1_{V\cap\delta(T)}(w_j) = 0.$$

Thus $\tilde{f} = 1_{\overline{V}}$. Then $\overline{V} = (\tilde{f})^{-1}(B_{1/2}(1))$ is open and βT is extremally disconnected.

7.5. Convexity

(7.5.1) Show that $\phi : \mathcal{Y} \to \mathbb{R}$ is convex if and only if

$$\phi\Big(\sum_{j=1}^{n} t_j y_j\Big) \leq \sum_{j=1}^{n} t_j\, \phi(y_j) \qquad (7.29)$$

for all $y_1, \ldots, y_n \in \mathcal{Y}$ and all $t_1, \ldots, t_n \in [0,1]$ with $\sum_j t_j = 1$.

Answer. The converse is just the case $n = 2$. So suppose that ϕ is convex, $y_1, \ldots, y_n \in \mathcal{Y}$, $t_1, \ldots, t_n \in [0,1]$, and $\sum_j t_j = 1$. We may assume without loss of generality that $t_j > 0$ for all j. The proof goes by induction. The case $n = 2$ is the hypothesis. So assume that (7.29) holds for $n - 1$. Then, with

$c = \sum_{k=1}^{n-1} t_k$, the numbers $t_1/c, \ldots, t_{n-1}/c$ are convex coefficients and thus

$$\begin{aligned}\phi\Big(\sum_{j=1}^n t_j y_j\Big) &= \phi\Big(t_n y_n + \sum_{j=1}^{n-1} t_j y_j\Big) = \phi\Big((1-c)y_n + c\sum_{j=1}^{n-1}(t_j/c)y_j\Big)\\ &\le (1-c)\phi(y_n) + c\,\phi\Big(\sum_{j=1}^{n-1}(t_j/c)y_j\Big)\\ &\le (1-c)\phi(y_n) + c\sum_{j=1}^{n-1}(t_j/c)\phi(y_j) = \sum_{j=1}^n t_j\phi(y_j).\end{aligned}$$

(7.5.2) Show that if $\mathcal{X}$ locally is convex, then any extreme point of the convex set $K \subset \mathcal{X}$ is a boundary point of K.

Answer. It is enough to show that interior points are not extreme. Suppose that $x \in K$ is interior. Then there exists an open, convex, balanced, neighbourhood of 0 such that $x + V \subset K$. Let $y \in V$, nonzero. We have $x \pm y \in x + V \subset K$, and

$$x = \tfrac{1}{2}(x+y) + \tfrac{1}{2}(x-y),$$

so x is not extreme.

(7.5.3) Let $\mathcal{X}$ be a real or complex vector space and $Y \subset \mathcal{X}$. Show that $\operatorname{conv} Y$ is the smallest convex subset of $\mathcal{X}$ that contains Y.

Answer. Let $Z \subset X$ be convex with $Y \subset Z$. The convexity of Z guarantees that $\operatorname{conv} Y \subset Z$. Then $\operatorname{conv} Y \subset \bigcap\{Z : \text{ convex}, Y \subset Z\}$. As $\operatorname{conv} Y$ is itself one of the Z, we get

$$\operatorname{conv} Y = \bigcap\{Z : \text{ convex}, Y \subset Z\}.$$

(7.5.4) Prove Proposition 7.5.8.

Answer. It is not obvious how to go directly from (ii) to (iii) without passing through (i). So we take the less direct approach.

(i) $\Longrightarrow$ (ii) This is the definition of extreme point.

(ii) $\Longrightarrow$ (i) Suppose that $x = ty' + (1-t)z'$ with $t \in (0,1)$ and $y', z' \in \mathcal{X}$. Assume, without loss of generality, that $t > \frac{1}{2}$. Let $y = y'$ and $z = 2x - y'$. We have, by convexity,

$$z = 2x - y' = (2t-1)y' + 2(1-t)z' \in K$$

(note that the condition $t > \frac{1}{2}$ guarantees that $2t-1$ and $2(1-t)$ are convex coefficients). By definition of y and z,

$$x = \tfrac{1}{2}\,(y+z),$$

so the hypothesis implies that $y = z = x$. As $y' = y$, we have

$$z' = \frac{x - ty}{1-t} = \frac{(1-t)y}{1-t} = y = x.$$

Thus x is extreme.

(i) $\Longrightarrow$ (iii) If $y, z \in K \setminus \{x\}$ and $t \in [0,1]$, then $ty + (1-t)z \neq x$ since x is extreme; so $ty + (1-t)z \in K \setminus \{x\}$.

(iii) $\Longrightarrow$ (i) Suppose that x is not extreme. The there exist $t \in (0,1)$ and $y, z \in K \setminus \{x\}$ with $x = ty + (1-t)z$. Hence $K \setminus \{x\}$ is not convex.

(7.5.5) Let $C \subset \mathbb{R}^n$ be convex, and $w \in \mathbb{R}^n \setminus C$. Show that there exists $v \in \mathbb{R}^n$ with $\|v\| = 1$ and $\alpha \in \mathbb{R}$ such that $\langle w, v\rangle \leq \alpha$ and $\langle z, v\rangle \geq \alpha$ for all $z \in C$. This can be done via Hahn–Banach, but an elementary proof is desired.

Answer. The closure $\overline{C}$ of C is convex and the result follows if we prove it for $\overline{C}$. So we may assume without loss of generality that C is closed. Because C is closed the distance between w and C is attained (proof: there should be elements that are at almost the distance from w; take a closed ball centered on w that contains these points, and by cutting C with this ball we may assume that C is compact, and a sequence of points approaching the distance will have a limit; or use Lemma 4.3.4).

Let $c \in C$ with $\|w - c\| = \operatorname{dist}(w, C)$. Put $v_0 = w - c$. For any $z \in C$ we have $\|w - c\| \leq \|w - z\|$. In particular, if $z \in C$ and $t \in (0,1)$ we have

$$\|w - c\| \leq \|w - (tc + (1-t)z)\|.$$

Using that $w - (tc + (1-t)z) = w - c + (1-t)(c - z)$ we get

$$\begin{aligned}\|w - c\|^2 &\leq \|w - c + (1-t)(c-z)\|^2 \\ &= \|w-c\|^2 + (1-t)^2\|c - z\|^2 + 2\langle w - c, (1-t)(c-z)\rangle.\end{aligned}$$

This simplifies, since $t < 1$, to

$$\langle w - c, c - z\rangle \geq -\frac{1}{2}\,(1-t)\|c - z\|^2.$$

As t can be chosen arbitrarily close to 1, this gives $\langle w - c, c - z\rangle \geq 0$ for all $z \in C$. This gives us $\langle v_0, z\rangle \leq \langle v_0, c\rangle$ for all $z \in C$. Put $v = -v_0$ and $\alpha = \langle v, c\rangle$. Then $\langle v, z\rangle \geq \alpha$ and

$$\langle v, w\rangle = \langle v, w - c\rangle + \langle v, c\rangle = -\|w - c\|^2 + \alpha \leq \alpha.$$

Finally, we can replace v with $v/\|v\|$ and α with $\alpha/\|v\|$.

(7.5.6) Let $K \subset \mathbb{R}^n$ be compact and convex. Show that K has an extreme point. *(Hint: choose two points that realize the diameter, and show that they are extreme).*

Answer. We use the Euclidean norm, so $\mathbb{R}^n$ is a real Hilbert space. Let $D = \sup\{\|x - y\| : x, y \in K\}$. Since K is compact this has to be a maximum. So there exist $x, y \in K$ such that $D = \|x - y\|$. Now suppose that $x = \frac{1}{2}\,x_1 + \frac{1}{2}\,x_2$. Then $\|x_1 - y\| \leq D = \|x - y\|$ and similarly $\|x_2 - y\| \leq \|x - y\|$. Using the Parallelogram Identity (4.2), that still works in a real inner product space,

$$\begin{aligned} 2\|x - y\|^2 &\geq \|x_1 - y\|^2 + \|x_2 - y\|^2 \\ &= \frac{1}{2}\,\|x_1 + x_2 - 2y\|^2 + \frac{1}{2}\,\|x_1 - x_2\|^2 \\ &= 2\|x - y\|^2 + \frac{1}{2}\,\|x_1 - x_2\|^2. \end{aligned}$$

It follows that $\|x_1 - x_2\| = 0$, so $x_1 = x_2$ and x is extreme.

(7.5.7) A **hyperplane** is a subset $H \subset \mathbb{R}$ of the form $z + V$, where $V \subset \mathbb{R}^n$ is subspace of dimension $n - 1$. Show that $H \subset \mathbb{R}^n$ is a hyperplane if and only if there exist $\alpha \in \mathbb{R}$ and $v \in \mathbb{R}^n$ such that $H = \{x \in \mathbb{R}^n : \langle x, v\rangle = \alpha\}$.

Answer. Suppose that $H = z + V$. As $\dim V = n - 1$, we have that $\dim V^\perp = 1$. So $V^\perp = \mathbb{R}\,v$ for some unit vector v. For any $w \in H$ we have $w - z \in V$, so $\langle w - z, v\rangle = 0$. Let $\alpha = \langle z, v\rangle$. Then $\langle w, v\rangle = \alpha$. And if $\langle w, v\rangle = \alpha$, then $\langle w - z, v\rangle = 0$, which implies that $w - z \in V$.

Conversely, suppose that $H = \{x \in \mathbb{R}^n : \langle x, v\rangle = \alpha\}$. Let

$$V = \{x \in \mathbb{R}^n : \langle x, v\rangle = 0\} = \{v\}^\perp.$$

This is a subspace of dimension $n - 1$. Fix $h_0 \in H$ and $v_0 \in V$, and let $z = h_0 - v_0$. For any $x \in H$,

$$\langle x - z, v\rangle = \langle x, v\rangle - \langle h_0, v\rangle + \langle v_0, v\rangle = \alpha - \alpha + 0 = 0.$$

So $x - z \in V$, and therefore $x \in z + V$. And this also works the other way, if $x = z + w$ with $w \in V$, then $\langle x, v\rangle = \alpha$ so $x \in H$. That is, $H = z + V$.

(7.5.8) Let $C \subset \mathbb{R}^n$ be convex. A **supporting hyperplane** for C is a hyperplane $H \subset \mathbb{R}^n$ such that

(i) C is in one of the two half-spaces determined by H; namely, there exists $\alpha \in \mathbb{R}$ such that $\langle x, y\rangle \geq \alpha$ for all $x \in C$ and $y \in H$;

(ii) $H \cap \partial C \neq \varnothing$.

Show that for each $x \in \partial C$ there exists a supporting hyperplane H for C such that $x \in H$.

Answer. If we provide H for $\overline{C}$, it works for C as well, as they both have the same boundary. For each $n \in \mathbb{N}$ there exists $v \in B_1(0)$ such that $x_n = x + \frac{1}{n}v \notin C$. Apply Exercise 7.5.5 to x_n and C; so there exists $\alpha_n \in \mathbb{R}$ and $v_n \in \mathbb{R}^n$ such that

$$\langle x_n, v_n\rangle \leq \alpha_n, \qquad \langle z, v_n\rangle \geq \alpha_n, \qquad z \in C.$$

From $\|v_n\| = 1$ we get (since $x \in \overline{C}$) that $\alpha_n \leq \langle x, v_n\rangle \leq \|x\|$. And

$$\alpha_n \geq \langle x_n, v_n\rangle \geq -\|x_n\|\,\|v_n\| \geq (-\|x\| - \frac{1}{n}) \geq (-\|x\| - 1).$$

If follows that $\{\alpha_n\}$ is a bounded sequence, and so it admits a convergent subsequence to an element α. Also, as $\|v_n\| = 1$ for all n, the sequence $\{v_n\}$ admits a convergent subsequence to an element v. Then, as $x_n \to x$,

$$\langle x, v\rangle \leq \alpha, \qquad \langle z, v\rangle \geq \alpha, \qquad z \in C.$$

Now let $\beta = \langle x, v\rangle$ and let H be the hyperplane $H = \{y \in \mathbb{R}^n : \langle y, v\rangle = \beta\}$. Then $x \in H$ and for any $z \in C$ we have $\langle z, v\rangle \geq \alpha \geq \beta$.

(7.5.9) Let $C \subset \mathbb{R}^n$ be convex, $x \in \partial C$ and H a supporting hyperplane for C at x. Show that H is convex and that $\operatorname{Ext}(H \cap C) \subset \operatorname{Ext} C$.

Answer. If $H = \{x : \langle x, v\rangle = \alpha\}$ and $h_1, h_2 \in H$, $t \in [0,1]$,

$$\langle th_1 + (1-t)h_2, v\rangle = t\langle h_1, v\rangle + (1-t)\langle h_2, v\rangle = t\alpha + (1-t)\alpha = \alpha,$$

and H is convex. Then $H \cap C$ is convex. Let $x \in \operatorname{Ext}(H \cap C)$. If $x = ty + (1-t)z$ with $y, z \in C$ and $t \in [0,1]$, we have

$$\alpha = \langle x, v\rangle = t\langle y, v\rangle + (1-t)\langle z, v\rangle \geq t\alpha + (1-t)\alpha = \alpha.$$

This forces $\langle y, v\rangle = \langle z, v\rangle = \alpha$, and so $y, z \in H \cap C$. But then $y = z = x$, and $x \in \operatorname{Ext} C$.

(7.5.10) In Example (7.5.1), show that K is convex and establish $\operatorname{Ext} K$.

Answer. The proof depends on how one defines what a regular polygon is. So we will stay with an intuitive argument. The convexity is clear, as any points joined by a segment will have the segment inside the closure of the polygon.

Regarding extreme points, for any $x \in \operatorname{int} K$, any line through it will touch the boundary at two points: then x is a convex combination of those two points, and thus not extreme.

Any point in the middle of an edge in the boundary is a convex combination of the corresponding two vertices, so not extreme.

Any line through a vertex either goes into K, or along an edge, or doesn't touch K other than at the vertex. In all three cases the vertex cannot be in between two points of K. So the vertices are the extreme points.

(7.5.11) In Example (7.5.2), show that K is convex and establish $\operatorname{Ext} K$.

Answer. For the convexity, suppose that $t \in (0, 1)$ and $v^2+w^2 \leq 1$, $u^2+z^2 \leq 1$. With $x = (v, w)$, $y = (u, z)$, we have $\|x\| \leq 1$, $\|y\| \leq 1$. Then

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\| \leq t + 1 - t = 1,$$

so $tx + (1-t)y$ is in the unit disk.

For the extreme points, if $\|x\| < 1$, then $x = t\,x' + (1-t)\,0$, where $t = \|x\|$ and $x' = x/t$. So only the circle can contain extreme points. And if $\|x\| = 1$ and $x = ty + (1-t)z$ with $\|y\| \leq 1$, $\|z\| \leq 1$, then

$$1 = \|x\| \leq t\|y\| + (1-t)\|z\| \leq 1,$$

which forces $\|y\| = 1$ and $\|z\| = 1$. Now we have

$$\begin{aligned} 1 &= \|ty + (1-t)z\|^2 = t^2\,\|y\|^2 + (1-t)^2\,\|z\|^2 + 2t(1-t)\,\langle y, z\rangle \\ &\leq t^2 + (1-t)^2 + 2t(1-t) = (t + 1 - t)^2 = 1. \end{aligned}$$

In particular we get $\langle y, z\rangle = \|y\|\,\|z\|$. By Theorem 4.2.2, there exists $c \in \mathbb{R}$ with $y = cz$. Then $|c| = 1$. And $x = ty + (1-t)cy$ forces $t + (1-t)c = 1$, and so $c = 1$ as above. Then $y = z = x$ and x is extreme.

(7.5.12) In Example (7.5.3), show that K is closed, convex and establish $\operatorname{Ext} K$.

Answer. K is closed, since its complement is $\{(x,y):\ y<0\}$. Indeed, given $(x,y)\in K^c$ the ball $B_{-y/2}(x,y)$ is entirely contained in $\{(x,y):\ y<0\}$, which is then open.

Convexity: if $y\geq 0$ and $w\geq 0$, then $ty+(1-t)w\geq 0$ for all $t\in[0,1]$. So $t(x,y)+(1-t)(v,w)=(tx+(1-t)v,ty+(1-t)w)\in K$.

Extreme points: given any $(x,y)\in K$ we have

$$(x,y)=\tfrac{1}{2}\,(x-1,y)+\tfrac{1}{2}\,(x+1,y),$$

so (x,y) is not extreme.

(7.5.13) In Example (7.5.4), show that K is convex and establish $\operatorname{Ext} K$.

Answer. Convexity: if $y>0$ and $w>0$, then $ty+(1-t)w>0$ for all $t\in[0,1]$. So $t(x,y)+(1-t)(v,w)=(tx+(1-t)v,ty+(1-t)w)\in K$. If $y>0$, then

$$t(x,y)+(1-t)(x_1,0)=(tx+(1-t)x_1,ty).$$

For $t=0$, we get $(x_1,0)\in K$. For $t>0$, the second coordinate is $ty>0$, so $t(x,y)+(1-t)(x_1,0)\in K$, and K is convex.

Extreme points: given any $(x,y)\in K$ with $y>0$ we have

$$(x,y)=\tfrac{1}{2}\,(x-1,y)+\tfrac{1}{2}\,(x+1,y),$$

so (x,y) is not extreme. Suppose that $(x_1,0)=t(x,y)+(1-t)(v,w)$ for some $t\in(0,1)$; since $y\geq 0$ and $w\geq 0$, from $ty+(1-t)w=0$ we conclude that $y=w=0$. The only point in K with second coordinate zero is $(x_1,0)$, so $v=x_1$. Thus $(x_1,0)$ is extreme.

(7.5.14) Show that, in the complex plane $\operatorname{Ext}\overline{\mathbb{D}}=\mathbb{T}$.

Answer. This is of course the same as Exercise 7.5.11, but we include here an argument using the language of complex numbers.

Write $\lambda\in\overline{\mathbb{D}}$ as $\lambda=re^{i\theta}$. If $r<1$, let $\delta=\frac{1-r}{2}$. Then $(r\pm\delta)e^{i\theta}\in\mathbb{D}$ and

$$\lambda=\tfrac{1}{2}\,(r+\delta)e^{i\theta}+\tfrac{1}{2}\,(r-\delta)e^{i\theta},$$

so λ is not extreme.

When $r = 1$, if $e^{i\theta} = \frac{1}{2}\,\alpha + \frac{1}{2}\,\beta$ with $|\alpha| \leq 1$ and $|\beta| \leq 1$, then

$$1 = |e^{i\theta}| \leq \tfrac{1}{2}\,|\alpha| + \tfrac{1}{2}\,|\beta| \leq 1,$$

so $|\alpha| = |\beta| = 1$. We would have $e^{i\theta} = \frac{1}{2}\,e^{i\eta} + \frac{1}{2}\,e^{i\nu}$, which we may write as

$$1 = \tfrac{1}{2}\,e^{i(\eta-\theta)} + \tfrac{1}{2}\,e^{i(\nu-\theta)}.$$

The real part of this equality is $1 = \frac{1}{2}\,\cos(\eta-\theta) + \frac{1}{2}\,\cos(\nu-\theta)$. As both cosines are at most 1, we obtain $1 = \cos(\eta-\theta) = \cos(\nu-\theta)$, which gives $\eta = \theta + 2k\pi$, $\nu = \theta + 2j\pi$, and thus $\alpha = \beta = \lambda$. By Proposition 7.5.8, $\lambda \in \operatorname{Ext}\overline{\mathbb{D}}$.

(7.5.15) In Example (7.5.9), show that K is convex and establish $\operatorname{Ext} K$.

Answer. Any unit ball with respect to a norm will be convex: if $\|f\|, \|g\| \leq 1$, then

$$\|tf + (1-t)g\| \leq t\|f\| + (1-t)\|g\| = 1 + 1 - t = 1.$$

As for the extreme points, suppose $f \in C_0(\mathbb{R})$ and $\|f\| \leq 1$. Given $\varepsilon = \frac{1}{2}$, there exists n such that $|f(x)| < \frac{1}{2}$ for all $|x| \geq n$. Let $g \in C_0(\mathbb{R})$ be continuous, with $\operatorname{supp} g \subset [n, n+1]$ and $\|g\| = \frac{1}{2}$. Then $|f \pm g| \leq 1$, and so $f = \frac{1}{2}\,(f+g) + \frac{1}{2}\,(f-g)$ with $f \pm g \in C_0(\mathbb{R})$, and therefore f is not extreme.

(7.5.16) In Example (7.5.10), show that K is convex and establish $\operatorname{Ext} K$.

Answer. Suppose first that there exist $\mu_1, \mu_2 \in \delta_t = \frac{1}{2}\,\mu_1 + \frac{1}{2}\,\mu_2$ (it is enough to use $\frac{1}{2}$ by Proposition 7.5.8). We have

$$1 = \|\delta_t\| = \frac{1}{2}\,\|\mu_1 + \mu_2\| \leq \frac{1}{2}\,\big(\|\mu_1\| + \|\mu_2\|\big) \leq 1.$$

We see that they are equalities, so in particular $\|\mu_1\| = \|\mu_2\| = 1$. Let V be an open neighbourhood of t. By Urysohn's Lemma (Theorem 2.6.5) there exists $f_V \in C_0(T)$ with $f(t) = 1$, $0 \leq f_V \leq 1$, and $f = 0$ on V^c. From

$$1 = \delta_t(f_V) = \frac{1}{2}\,\big(\mu_1(f_V) + \mu_2(f_V)\big) \leq 1$$

we deduce that $\mu_1(f_V) = \mu_2(f_V) = 1$ for all V (details at the end). Let $K \subset T \setminus \{t\}$ be compact. As $T \setminus K$ is an open neighbourhood of $\{t\}$, we have $\mu_1(f_{T\setminus K}) = 1$. Also, since $0 \leq f_{T\setminus K} \leq 1$,

$$1 = \int_{T\setminus K} f_{T\setminus K}\,d\mu_1 \leq \int_{T\setminus K} f_{T\setminus K}\,d|\mu_1| \leq |\mu_1|(T \setminus K) \leq 1.$$

Thus $|\mu_1|(T \setminus K) = 1$, and therefore $|\mu_1|(K) = 0$. As $|\mu_1|$ is Radon and $|\mu_1|(K) = 0$ for any compact $K \subset T \setminus \{t\}$, it follows that $|\mu_1|(T \setminus \{t\}) = 0$, and so $|\mu_1|(\{t\}) = 1$, which implies that $|\mu_1| = \delta_t$. The same argument applies to μ_2. At this stage we have $\mu_1 = \alpha_1 \, \delta_t$ and $\mu_2 = \alpha_2 \, \delta_t$ with $\alpha_1, \alpha_2 \in \mathbb{T}$. But again from $1 = \delta_t(\{t\})$ we get $2 = \alpha_1 + \alpha_2$, which implies $\alpha_1 = \alpha_2 = 1$, and therefore δ_t is extreme.

Conversely, suppose that $\operatorname{supp} |\mu|$ contains at least two points t_1 and t_2. Because T is Hausdorff there exist disjoint open sets $V_1, V_2 \subset T$ with $t_1 \in V_1$ and $t_2 \in V_2$. By definition of support we have $|\mu|(V_1) > 0$, $|\mu|(V_2) > 0$. Let $t = |\mu|(V_1) \in (0, 1)$. The fact that $V_1 \cap V_2 = \varnothing$ guarantees that $|\mu|(V_2) \leq 1 - t$. Define measures

$$\eta_1 = t^{-1} \, \mu|_{V_1}, \qquad \eta_2 = (1-t)^{-1} \, \mu|_{V_2}.$$

Then $\|\eta_j\| = |\eta_j|(T) \leq 1$, $j = 1, 2$, and

$$\mu = t \, \eta_1 + (1-t) \, \eta_2.$$

Therefore μ is not extreme.

It remains to explain why if $\frac{1}{2}(z_1 + z_2) = 1$ for two complex numbers with $|z_1| \leq 1$ and $|z_2| \leq 1$, then $z_1 = z_2 = 1$. The equality is satisfied by the real parts. This gives

$$\big(1 - \operatorname{Re} z_1\big) + \big(1 - \operatorname{Re} z_2\big) = 0.$$

As both real parts are real numbers with absolutely value at most 1, it follows that $\operatorname{Re} z_1 = \operatorname{Re} z_2 = 1$. Then $|\operatorname{Im} z_1|^2 = 1 - \big(\operatorname{Re} z_1\big)^2 = 0$ and thus $z_1 = 1$; similarly we obtain $z_2 = 1$.

(7.5.17) Prove Proposition 7.5.6.

Answer. A simple observation is that if $\|x\| \leq 1$, $\|y\| \leq 1$, and $\|tx + (1-t)y\| = 1$ for some $t \in (0, 1)$, then $\|x\| = \|y\| = 1$. This follows from

$$1 = \|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\| \leq t + 1 - t = 1.$$

Then $t(1 - \|x\|) + (1-t)(1 - \|y\|) = 0$, and as this is sum of nonnegative terms, they both have to be zero and hence $\|x\| = \|y\| = 1$.

One can show that (i) (in the form "$\|(u+v)/2\| = 1$ for $\|u\| = \|v\| = 1$ implies $u = v$") is equivalent to

$$\|x\| = \|y\| = 1, \ \|tx + (1-t)y\| = 1 \implies x = y \tag{AB.7.2}$$

(proof at the end).

$(i) \implies (ii)$: if we have $\|x + y\| = \|x\| + \|y\|$, we can rewrite this as

$$\left\| \frac{\|x\|}{\|x\| + \|y\|} \, \frac{x}{\|x\|} + \frac{\|y\|}{\|x\| + \|y\|} \, \frac{y}{\|y\|} \right\| = 1.$$

Now (AB.7.2) applies and we get that $x/\|x\| = y/\|y\|$.

$(ii) \implies (iii)$: let $\phi \in X^*$ and $x, y \in X$ with $\|\varphi\| = 1$, $\phi(x) = \phi(y) = 1$. Then

$$2 = \|x\| + \|y\| = \phi(x) + \phi(y) = \phi(x+y) \leq \|x+y\|.$$

This gives us equality in the triangle inequality, and so by (ii) we have that $y = \lambda x$ with $\lambda > 0$. Then

$$\lambda = \lambda\phi(y) = \phi(\lambda y) = \phi(x) = 1$$

and so $y = x$.

$(iii) \implies (i)$: Suppose that u, v are unit vectors with $\|\frac{u+v}{2}\| = 1$. Use Hahn–Banach (Corollary 5.7.7) to construct $\phi \in X^*$ with $\|\phi\| = 1$ and $\phi(u+v) = \|u+v\|$. Then

$$2 = \phi(u+v) = \phi(u) + \phi(v) \leq \|\phi\|\,\|u\| + \|\phi\|\,\|v\| = 2.$$

As in the observation at the beginning, this forces $\phi(u) = \phi(v) = 1$. By (iii) we have $u = v$.

We finish by proving (AB.7.2). We only need to show that $(i) \implies$ (AB.7.2), as the converse is trivial. Suppose that $\|x\| = \|y\| = 1$ and $\|tx + (1-t)y\| = 1$ for some $t \in (0,1)$. We can rewrite the last equality as

$$\left\|2t\Big(\frac{x+y}{2}\Big) + (1-2t)y\right\| = 1$$

(this only works if $t \leq 1/2$, but if it isn't we can do the same argument with $1-t$ instead). Then

$$1 = \left\|2t\Big(\frac{x+y}{2}\Big) + (1-2t)y\right\| \leq 2t\left\|\frac{x+y}{2}\right\| + (1-2t).$$

This forces $\|(x+y)/2\| = 1$, and then by hypothesis we conclude that $x = y$.

(7.5.18) Show that $\|\cdot\|_p$ is strictly convex for $1 < p < \infty$, while $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are not strictly convex.

Answer. We write an argument for $\ell^p(\mathbb{N})$ but the same idea works in any $L^p(X)$ for X with at least two points.

For $1 < p < \infty$, if $x, y \in \ell^p(\mathbb{N})$ and $\|x+y\|_p = \|x\|_p + \|y\|_p$, we have equality in Minkowski's Inequality (2.47). Looking at the proof of Corollary 2.8.10, this means that the inequalities in the proof are equalities. In particular we get

$$\begin{aligned}\|x+y\|_p^p &= \sum_k |x_k+y_k|\,|x_k+y_k|^{p-1}\\ &= \sum_k |x_k|\,|x_k+y_k|^{p-1} + \sum_k |y_k|\,|x_k+y_k|^{p-1}\\ &= \|x\|_p \Big(\sum_k |x_k+y_k|^{q(p-1)}\Big)^{1/p} + \|y\|_p \Big(\sum_k |x_k+y_k|^{q(p-1)}\Big)^{1/p}.\end{aligned}$$

The second equality forces $|x_k+y_k| = |x_k|+|y_k|$ for all k, and so $y_k = \lambda_k x_k$ for $\lambda_k \geq 0$. The equality in the third and last equality is equality in Hölder's inequality for the pairs of functions $x, |x+y|^{p-1}$ and $y, |x+y|^{p-1}$. But equality in Hölder implies, looking at the proof of Theorem 2.8.8, pointwise equality on Young's inequality for the normalized versions of the functions. As the logarithm is strictly convex, we get

$$\frac{|x_k|}{\|x\|_p} = \frac{|x_k+y_k|^{p-1}}{\||x+y|^{p-1}\|_q} = \frac{|y_k|}{\|y\|_p}.$$

Then

$$\lambda_k|x_k| = |y_k| = \frac{\|y\|_p}{\|x\|_p}\,|x_k|, \qquad\qquad k\in\mathbb{N},$$

so $\lambda_k = \frac{\|y\|_p}{\|x\|_p}$ for all k. That is, $y=\lambda x$ for a certain $\lambda>0$. Then $\|\cdot\|_p$ is strictly convex by Proposition 7.5.6.

For $\|\cdot\|_1$, we have $\|e_1\|_1 = \|e_2\|_1 = 1$, and $\|e_1+e_2\|_1 = 2$, so not strictly convex.

For $\|\cdot\|_\infty$, $\|e_1\|_\infty = \|e_1+e_2\|_\infty = 1$ and $\|e_1+(e_1+e_2)\|_\infty = 2$, so not strictly convex.

(7.5.19) Let $\mathcal{X},\mathcal{Y}$ be normed spaces and $V\in\mathcal{B}(\mathcal{X},\mathcal{Y})$ an isometry.

(i) Show that if V is surjective, then V maps the set $\operatorname{Ext}\overline{B_1^{\mathcal{X}}(0)}$ onto $\operatorname{Ext}\overline{B_1^{\mathcal{Y}}(0)}$.

(ii) Show that if V is not surjective, the above may fail, i.e. construct an example of a linear isometry that maps an extreme point to a non-extreme point.

(iii) *(This one may be a little harder because we want the same domain and codomain; but examples exist that are not convoluted)* Find a Banach space $\mathcal{X}$ with $\operatorname{Ext}\overline{B_1^{\mathcal{X}}(0)} \neq \varnothing$, an isometry $V\in\mathcal{B}(\mathcal{X})$, and $e\in\operatorname{Ext}\overline{B_1^{\mathcal{X}}(0)}$ such that $Ve\notin\operatorname{Ext}\overline{B_1^{\mathcal{X}}(0)}$.

Answer.

(i) Suppose that V is surjective, and $x \in \overline{B_1^{\mathcal{X}}(0)}$ is extreme. If $Vx = \frac{1}{2}\,y' + \frac{1}{2}\,z'$ with $\|y'\| \leq 1$, $\|z'\| \leq 1$, by the surjectivity there exist $y, z \in \mathcal{X}$ with $y' = Vy$, $z' = Vz$. Because V is isometric, $y, z \in \overline{B_1^{\mathcal{X}}(0)}$. Then

$$Vx = \frac{1}{2}\,Vy + \frac{1}{2}\,Vz = V\big(\frac{1}{2}\,y + \frac{1}{2}\,z\big).$$

By the injectivity of V, we get that $x = \frac{1}{2}\,y + \frac{1}{2}\,z$. As x is extreme, $y = z = x$, and hence $y' = Vy = Vx$, $z' = Vz = Vx$.

If $y \in \overline{B_1^{\mathcal{Y}}(0)}$ is extreme, let $x = V^{-1}y$. If $x = \frac{1}{2}\,y + \frac{1}{2}\,z$ for $y, z \in \mathcal{X}$, then $y = \frac{1}{2}\,Vy + \frac{1}{2}\,Vz$, and $Vy, Vz \in \overline{B_1^{\mathcal{Y}}(0)}$. As y is extreme, $Vy = Vz = y = Vx$. Then the injectivity of V gives us $z = y = x$, and so x is extreme. So V is surjective from the extreme points to the extreme points.

(ii) Let $\mathcal{X}$ be any Banach space such that $\operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)} \neq \varnothing$. Let $V : \mathcal{X} \to \mathcal{X} \oplus c_0$ be given by $Vx = x \oplus 0$, where on $\mathcal{X} \oplus c_0$ we consider the norm $\|x \oplus y\| = \max\{\|x\|, \|y\|\}$. If $e \in \operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)}$, then $Ve = e \oplus 0$ is not extreme, since $e \oplus 0 = \frac{1}{2}\,(e \oplus 1) + \frac{1}{2}\,(e \oplus (-1))$ and $\|e \oplus 1\| = \|e \oplus (-1)\| = 1$.

(iii) Let $\mathcal{X} = \ell^\infty(\mathbb{N})$. Let $V \in \mathcal{B}(\mathcal{X})$ be the linear map induced by $Ve_k = e_{2k}$ on the canonical basis. Then V is an isometry. Consider $1 \in \overline{B_1^{\mathcal{X}}(0)}$; then $1 \in \operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)}$ and $V1 = \sum_k e_{2k}$ is not extreme, since $V1 = \frac{1}{2}\,(e_1 + V1) + \frac{1}{2}\,(-e_1 + V1)$, and $\|e_1 + V1\| = \| - e_1 + V1\| = 1$.

(7.5.20) Prove Exercise 5.6.5 using convexity ideas. That is, show that c and c_0 are not isometrically isomorphic as Banach spaces (*Hint: show that the unit ball of c_0 has no extreme points, while the unit ball of c does, and use Exercise 7.5.19*)

Answer. Suppose that $\gamma : c \to c_0$ is an isometric isomorphism. By Exercise 7.5.19, it preserves extreme point.

We claim that $x \in c$ is extreme in the unit ball if and only if $|x_n| = 1$ for all n. Indeed, if $|x_n| = 1$ for all n, and $x = tz + (1-t)w$ with t, w in the unit ball, then for each n we have $x_n = tz_n + (1-t)w_n$, which forces $z_n = w_n$ since x_n is extreme in the unit disk; so $z = w = x$ and x is extreme. Conversely, if $|x_m| < 1$ for some m, let $\delta = (1 - |x_m|)/2$ and let z be equal to x with the exception of $z_m = x_n + \delta$, and w equal to x with the exception of $w_m = x_n - \delta$. Then $x = (z + w)/2$, and x is not extreme.

In c_0, for any x there exists m (many, actually), with $|x_m| < 1$. Now we can repeat the argument from the previous paragrpah to show that x cannot be extreme.

(7.5.21) Is convexity of K_j used in the proof of Lemma 7.5.14? Where?

Answer. Yes, it is used. The last line in the proof uses the computation that $\operatorname{conv}(\bigcup_{r=1}^m K_r) = \alpha(T \times K_1 \times \cdots \times K_m)$. And that is where convexity is used. An element of $\operatorname{conv}(\bigcup_{r=1}^m K_r)$ is a priori not necessarily of the form $\sum_j t_j x_j$ with $x_j \in K_j$; it actually is, but that's because the K_j are convex. We refer to Example 7.5.15

(7.5.22) Let $\mathcal{X}$ be a TVS, $K \subset \mathcal{X}$, and $x_1, x_2, \ldots \in K$. Show that
$$\sum_{k=1}^\infty 2^{-k} x_k \in \overline{\operatorname{conv}} K.$$

Answer. Since $\displaystyle\sum_{k=1}^m 2^{-k} = 1 - 2^{-m}$, we have that
$$2^{-m} x_1 + \sum_{k=1}^m 2^{-k} x_k \in \operatorname{conv} K,$$
and it converges to $\displaystyle\sum_{k=1}^\infty 2^{-k} x_k$.

(7.5.23) Prove Lemma 7.5.14 in a straightforward way, that is without using the fact that a continuous image of compact is compact.

Answer. Consider a net $\{x_j\}_j$, with $x_j \in \operatorname{conv}(\bigcup_{r=1}^m K_r)$. Using that the sets $K_1, \ldots, K_m$ are convex, as explained in the proof of the lemma we may write
$$x_j = \sum_{r=1}^m t_{jr} x_{jr}, \quad \text{where } t_{jr} \in [0,1],\ \sum_{r=1}^m t_{jr} = 1, \text{ and } x_{jr} \in K_j.$$
By compactness of $[0,1]$ and $K_1, \ldots, K_m$, we may successively choose convergent subnets $\{t_{j_k r}\}_k \subset [0,1]$ and $\{x_{j_k r}\}_k \in K_j$, $r = 1, \ldots, m$. Let $t_r = \lim_k t_{j_k r}$, $x_r = \lim_k x_{j_k r}$, $r = 1, \ldots, m$. Then $\sum_{r=1}^m t_r = \lim_k \sum_{r=1}^m t_{j_k r} = 1$, and
$$\sum_{r=1}^m t_r x_r = \lim_k \sum_{r=1}^m t_{j_k r} x_{j_k r} = \lim_k x_{j_k}.$$
So the net converges in $\operatorname{conv}(\bigcup_{r=1}^m K_r)$, which is then compact.

(7.5.24) For Hilbert spaces, we proved in Lemma 4.3.4 that given a closed convex set K and $x \notin K$, the distance between x and K is achieved, and it is achieved at a unique point. The same is not true for an arbitrary Banach space. Let $\mathcal{X} = C[0,1]$, with the supremum norm, and let

$$K = \left\{ g \in C[0,2] : \int_0^1 g - \int_1^2 g = 1 \right\}$$

Show that K is closed and convex, that $\operatorname{dist}(0,K) = \frac{1}{2}$, and that $\|g\| > \frac{1}{2}$ for all $g \in K$.

Answer. We have that K is closed and convex, since it is the preimage of a point by a bounded linear functional. Clearly $0 \notin K$. For $g \in K$, we have

$$1 = \int_0^1 g - \int_1^2 g \leq \int_0^2 |g| \leq 2\|g\|_\infty. \tag{AB.7.3}$$

So $\|g\|_\infty \geq \frac{1}{2}$ for all $g \in K$.

Now suppose that $\|g\|_\infty = \frac{1}{2}$. Then both inequalities above are equalities. From the last inequality now turned equality,

$$0 = \int_0^2 \|g\|_\infty - |g|.$$

As the integrand is non-negative, $|g| = \|g\|_\infty = \frac{1}{2}$. The first inequality in (AB.7.3), now turned equality, is now

$$\int_0^1 \left(\frac{1}{2} - g\right) + \int_1^2 \left(\frac{1}{2} + g\right) = 0.$$

Looking at the real parts,

$$\int_0^1 \left(\frac{1}{2} - \operatorname{Re} g\right) + \int_1^2 \left(\frac{1}{2} + \operatorname{Re} g\right) = 0.$$

As $-\frac{1}{2} \leq \operatorname{Re} g \leq \frac{1}{2}$ both integrands are non-negative, which forces $\operatorname{Re} g = \frac{1}{2}$ on $[0,1]$ and $\operatorname{Re} g = -1/2$ on $[1,2]$. But g is continuous and so is $\operatorname{Re} g$, making this is impossible. Thus $\|g\|_\infty > \frac{1}{2}$ for all $g \in K$.

It remains to show that the distance from 0 to K is actually $1/2$, that is, that we can find $g \in K$ with $\|g\|_\infty$ as close to $\frac{1}{2}$ as desired. The above reasoning showed that the distance would be achieved by a function that is $1/2$ on $[0,1]$, and $-1/2$ on $[1,2]$; of course this would not be continuous, which is the point. But we can get arbitrarily close to $1/2$ as follows. Let

$$g_n(x) = \begin{cases} \frac{1}{2} + \frac{1}{n} + \frac{1}{n-1}\Big(x - 1 + \frac{1}{n}\Big), & 0 \le x \le 1 - \frac{1}{n} \\ -\frac{1}{2} - \frac{n+1}{2}\Big(x - 1 - \frac{1}{n}\Big), & 1 - \frac{1}{n} \le x \le 1 + \frac{1}{n} \\ -\frac{1}{2}, & 1 + \frac{1}{n} \le x \le 2 \end{cases}$$

This g_n was constructed as consisting of three segments: namely the lines joining

$$(0, \frac{1}{2}),\ (1 - \frac{1}{n}, \frac{1}{2} + \frac{1}{n}),\ (1 + \frac{1}{n}, -\frac{1}{2}),\ (2, -\frac{1}{2}).$$

Then $g_n \in K$ and $\|g_n\|_\infty = \frac{1}{2} + \frac{1}{n}$.

7.6. The Cantor Space

(7.6.1) Show that η, as in the proof of Proposition 7.6.4, is a continuous bijection.

Answer. If $\eta(\{g_m\}) = \eta(\{h_m\})$ this means that $g_k(r) = h_k(r)$ for all k, r. This says that $g_k = h_k$ for all k, and so $\{g_m\} = \{h_m\}$. So η is injective. Given $h \in 2^{\mathbb{N}}$ define, for each k, $g_k : \mathbb{N} \to \{0,1\}$ by $g_k(r) = h(n_{k,r})$. Then $\eta(\{g_m\}) = h$ and so η is surjective.

For continuity, if $\{p_\ell\} \subset (2^{\mathbb{N}})^{\mathbb{N}}$ and $p_\ell \xrightarrow[\ell\to\infty]{} 0$, this means that

$$p_\ell(n)(m) \xrightarrow[\ell\to\infty]{} 0 \qquad \text{for all } n, m.$$

Then

$$\eta(p_\ell)(n_{k,r}) = p_\ell(k)(r) \xrightarrow[\ell\to\infty]{} 0$$

for all k, r, so $\eta(p_\ell) \to 0$. Hence η is continuous.

(7.6.2) Let X, Y be topological spaces and $\alpha : X \to Y$ a homeomorphism. Define $\alpha^{\mathbb{N}} : X^{\mathbb{N}} \to Y^{\mathbb{N}}$ by $\alpha^{\mathbb{N}}(\{x_n\}) = \{\alpha(x_n)\}$. Show that $\alpha^{\mathbb{N}}$ is a homeomorphism.

Answer. We define $(\alpha^{-1})^{\mathbb{N}} : Y^{\mathbb{N}} \to X^{\mathbb{N}}$ by $(\alpha^{-1})^{\mathbb{N}}(\{x_n\}) = (\alpha^{-1}(x_n)\}$. Then $\alpha^{\mathbb{N}} \circ (\alpha^{-1})^{\mathbb{N}} = \mathrm{id}_{Y^{\mathbb{N}}}$ and $(\alpha^{-1})^{\mathbb{N}} \circ \alpha^{\mathbb{N}} = \mathrm{id}_{X^{\mathbb{N}}}$, so $\alpha^{\mathbb{N}}$ is bijective.

If a net $\{p_j\} \subset X^{\mathbb{N}}$ converges to $p \in X^{\mathbb{N}}$, we have that $p_j(n) \to p(n)$ for all n. Then

$$\alpha^{\mathbb{N}}(p_j)(n) = \alpha(p_j(n)) \xrightarrow[j]{} \alpha(p(n)) = \alpha^{\mathbb{N}}(p)(n).$$

by the continuity of α. An analog computation shows that $(\alpha^{-1})^{\mathbb{N}}$ is continuous, so $\alpha^{\mathbb{N}}$ is bicontinuous.

(7.6.3) In the proof of Theorem 7.6.5, it is claimed that using $\nu : T \to \overline{\mathbb{D}}^{\mathbb{N}}$ and the homeomorphic embedding of $\overline{\mathbb{D}}$ into $[0,1]^2$ one gets a continuous injective $\tilde{\nu} : T \to [0,1]^{\mathbb{N}}$, homeomorphic onto its image. Write the details to justify these assertions.

Answer. Let $\eta : \overline{\mathbb{D}} \to [0,1]^2$ be given by

$$\eta(a+ib) = \Big(\frac{a+1}{2}, \frac{b+1}{2}\Big).$$

This map is continuous by since convergence in $\overline{\mathbb{D}}$ is equivalent to convergence of the real and imaginary parts. As η is injective and continuous with compact domain, it is a homeomorphism onto its image by Exercise 1.8.38. So, using Exercise 7.6.2, $\eta^{\mathbb{N}} \circ \nu : T \to \big([0,1]^2\big)^{\mathbb{N}}$ is a homeomorphism onto its image.

Let $\rho : \big([0,1]^2\big)^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$ be given by $\rho(\{(s_n,t_n)\} = (s_1,t_1,s_2,t_2,\ldots)$. Since we consider pointwise convergence in both spaces, ρ is a continuous injection, and by the compactness of $\big([0,1]^2\big)^{\mathbb{N}}$ (by Tychonoff, Theorem 1.8.24) we get from Exercise 1.8.38 that ρ is a homeomorphism onto its image. Then

$$\tilde{\nu} = \rho \circ \eta \circ \nu : T \to [0,1]^{\mathbb{N}}$$

is a continuous homeomorphism onto its image.

CHAPTER

8

Fourier Series and Hilbert Function Spaces

8.1. Fourier Series

(8.1.1) Let $f(t) = t$, $t \in [-\pi, \pi]$. Show that

$$S_{f,n}(t) = \sum_{k=1}^{n} \frac{(-1)^{k+1} \sin kt}{k}.$$

Answer. As f is odd, $\hat{f}(0) = 0$. For $k \neq 0$,

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} t\, e^{-ikt}\, dt = \frac{1}{2\pi} \left. \frac{t\, e^{-ikt}}{-ik} \right|_{-\pi}^{\pi} = \frac{2\pi(-1)^k}{-2\pi ik} = \frac{(-1)^{k+1}}{ik}.$$

To form $S_{f,n}$, if we put together the terms with k and $-k$ we get (note that $(-1)^k = (-1)^{-k}$)

$$\frac{(-1)^{k+1} e^{-ikt}}{i(-k)} + \frac{(-1)^{k+1} e^{ikt}}{ik} = \frac{(-1)^{k+1}}{ik} (e^{ikt} - e^{-ikt})$$

$$= \frac{(-1)^{k+1}}{ik}\, 2i \sin kt = \frac{2(-1)^{k+1}}{k} \sin kt.$$

Thus

$$S_{f,n}(t) = \sum_{k=1}^{n} \frac{(-1)^{k+1}\sin kt}{k}.$$

(8.1.2) Let $f(t) = t^2$, $t \in [-\pi, \pi]$. Show that

$$S_{f,n}(t) = \frac{\pi^2}{3} + \sum_{k=1}^{n} \frac{4(-1)^k \cos kt}{k^2}.$$

Answer. We have

$$\hat{f}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} t^2\,dt = \frac{\pi^2}{3}.$$

For $k \neq 0$,

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} t^2\,e^{-ikt}\,dt = -\frac{1}{2\pi}\,\frac{2t\,e^{-ikt}}{-ik}\Bigg|_{-\pi}^{\pi} = \frac{2(-1)^k}{k^2}.$$

To form $S_{f,n}$, if we put together the terms with k and $-k$ we get

$$\frac{2(-1)^k e^{-ikt}}{k^2} + \frac{2(-1)^k e^{ikt}}{k^2} = \frac{4(-1)^k}{k^2}\,\cos kt.$$

Thus

$$S_{f,n}(t) = \frac{\pi^2}{3} + \sum_{k=1}^{n} \frac{4(-1)^k \cos kt}{k^2}.$$

(8.1.3) Let $f(t) = e^t$, $t \in [-\pi, \pi]$. Show that

$$S_{f,n}(t) = \frac{\sinh\pi}{2\pi} + \sum_{k=1}^{n} \frac{2(-1)^k \sinh\pi\,\big(\cos kt - k\,\sin kt\big)}{\pi(1+k^2)}.$$

Answer. We have

$$\hat{f}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^t\,dt = \frac{\sinh\pi}{2\pi}.$$

For $k \neq 0$,

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^t\,e^{-ikt}\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{(1-ik)t}\,dt = \frac{e^{(1-ik)t}}{2\pi(1-ik)}\Bigg|_{-\pi}^{\pi}$$

$$= \frac{(-1)^k(e^{\pi}-e^{-\pi})}{2\pi(1-ik)} = \frac{(-1)^k \sinh\pi}{\pi(1-ik)} = \frac{(-1)^k \sinh\pi\,(1+ik)}{\pi(1+k^2)}.$$

To form $S_{f,n}$, if we put together the terms with k and $-k$ we get

$$\begin{aligned}\hat{f}(-k)e^{-ikt} + \hat{f}(k)e^{ikt} &= \frac{(-1)^k \sinh\pi\,(1-ik)e^{-ikt}}{\pi(1+k^2)} + \frac{(-1)^k \sinh\pi\,(1+ik)e^{ikt}}{\pi(1+k^2)} \\ &= \frac{(-1)^k \sinh\pi}{\pi(1+k^2)} \big((1-ik)e^{-ikt} + (1+ik)e^{ikt}\big) \\ &= \frac{(-1)^k \sinh\pi}{\pi(1+k^2)} \big(2\cos kt - 2k\sin kt\big).\end{aligned}$$

Thus

$$S_{f,n}(t) = \frac{\sinh\pi}{2\pi} + \sum_{k=1}^{n} \frac{2(-1)^k \sinh\pi\,\big(\cos kt - k\,\sin kt\big)}{\pi(1+k^2)}.$$

(8.1.4) Let $f : [-\pi,\pi] \to \mathbb{C}$ be given by $f(t) = \operatorname{sgn}(t)$. Show that

$$S_f(t) = \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\sin(2k-1)t}{2k-1}.$$

If available, use graphing software to display $S_{f,n}(t)$ for $n = 5$, $n = 50$, and $n = 500$.

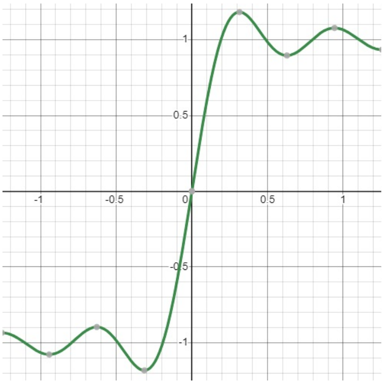
Answer. We have

$$\begin{aligned}\hat{f}(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{sgn}(t)\, e^{-ikt}\,dt = -\frac{1}{2\pi}\int_{-\pi}^{0} e^{-ikt}\,dt + \frac{1}{2\pi}\int_{0}^{\pi} e^{-ikt}\,dt \\ &= -\frac{e^{-ikt}}{-2\pi ik}\bigg|_{-\pi}^{0} + \frac{e^{-ikt}}{-2\pi ik}\bigg|_{0}^{\pi} = \frac{1}{2\pi ik}\big(1 - (-1)^k - (-1)^k + 1\big) \\ &= \frac{1}{\pi ik}\big(1-(-1)^k\big) = \begin{cases} \frac{2}{\pi ik}, & k \text{ odd} \\ 0, & k \text{ even}\end{cases}\end{aligned}$$

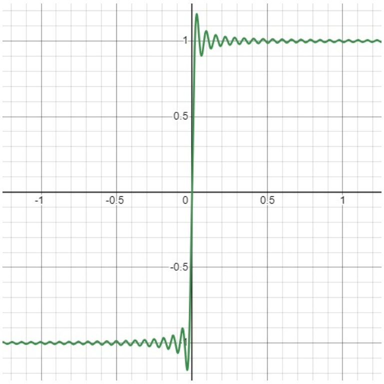
Then, as only the odd coefficients are nonzero,

$$\begin{aligned}S_f(t) &= \sum_{k=1}^{\infty} \hat{f}(2k-1)e^{i(2k-1)t} + \hat{f}(-2k+1)e^{-i(2k-1)t} \\ &= \sum_{k=1}^{\infty} \frac{2}{\pi i(2k-1)}\big[e^{i(2k-1)t} - e^{-i(2k-1)t}\big] \\ &= \sum_{k=1}^{\infty} \frac{2}{\pi i(2k-1)} 2i\sin(2k-1)t = \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\sin(2k-1)t}{2k-1}.\end{aligned}$$

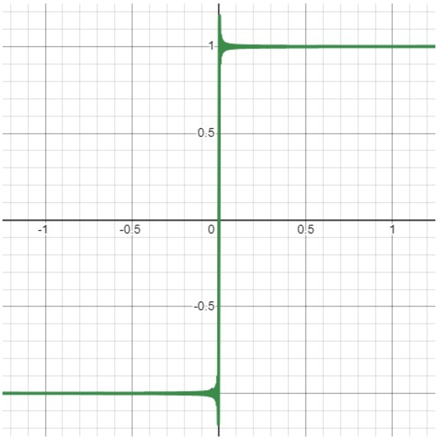
Now, the pictures.



$S_{f,n}(t)$ when $f(t) = \text{sgn}(t)$, $n = 5$

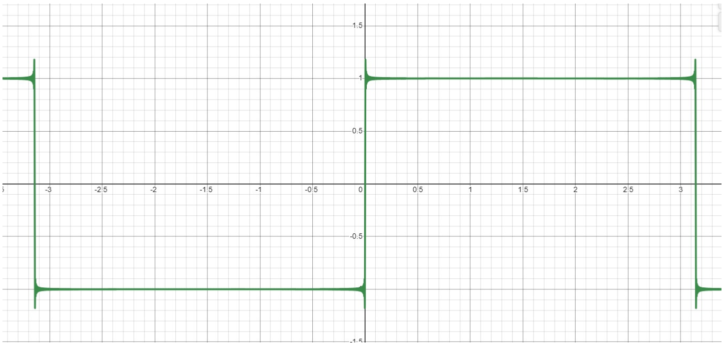


$S_{f,n}(t)$ when $f(t) = \text{sgn}(t)$, $n = 50$



$S_{f,n}(t)$ when $f(t) = \text{sgn}(t)$, $n = 500$

Another picture, showing $S_{f,500}(t)$ more globally:



$S_{f,n}(t)$ when $f(t) = \operatorname{sgn}(t)$, $\qquad n = 500$

(8.1.5) Mimicking Example 8.1.7, show that

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}.$$

Answer. We use $f(t) = t^2$. We have, for $k \neq 0$,

$$\begin{aligned}
\hat{f}(k) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} t^2\, e^{-ikt}\, dt = \frac{1}{-2\pi ik}\, t^2 e^{-ikt}\Big|_{-\pi}^{\pi} + \frac{1}{2\pi ik} \int_{-\pi}^{\pi} 2te^{-ikt}\, dt \\
&= -\frac{1}{2\pi ik}\, t^2 e^{-ikt}\Big|_{-\pi}^{\pi} - \frac{1}{\pi k^2}\, te^{-ikt}\Big|_{-\pi}^{\pi} - \frac{1}{\pi k^2} \int_{-\pi}^{\pi} e^{-ikt}\, dt \\
&= -\frac{1}{\pi k^2}\, te^{-ikt}\Big|_{-\pi}^{\pi} = \frac{2(-1)^{k+2}}{k^2}.
\end{aligned}$$

Also

$$\hat{f}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} t^2\, dt = \frac{\pi^3}{3}$$

and

$$\|f\|_2^2 = \int_{-\pi}^{\pi} t^4\, dt = \frac{2\pi^5}{5}.$$

Now (8.3) gives us (putting together the terms corresponding to k and $-k$, as they are equal),

$$\frac{2\pi^5}{5} = 2\pi \left(\frac{\pi^4}{9} + \sum_{k=1}^{\infty} \frac{4}{k^4} \right).$$

Solving for the series we get

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}.$$

(8.1.6) Let $m, n \in \mathbb{N}$. Show that, as functions,

$$\cos^m t \, \sin^n t \in \operatorname{span}\{a \cos kt + b \sin kt : \; a, b \in \mathbb{R}, \; k \in \mathbb{N}\}.$$

Answer. By Lemma 8.1.2 we have that $\cos^m t \, \sin^n t \in \operatorname{span}\{a \cos kt + b \sin kt : \; a, b \in \mathbb{C}, \; k \in \mathbb{N}\}$. Now if

$$\cos^m t \, \sin^n t = \sum_{k=0}^{\ell} a_k \cos kt + b_k \sin kt,$$

as the left-hand-side is real, we may replace a_k and b_k with their reals parts and the equality still holds.

(8.1.7) Let

$$f(t) = \begin{cases} \frac{\pi - t}{2}, & 0 \leq t \leq \pi \\ -\frac{\pi + t}{2}, & -\pi \leq t < 0 \end{cases}$$

Use f and Corollary 8.1.13 to show that there exists $c > 0$ such that $\left| \sum_{k=1}^{n} \frac{\sin kx}{k} \right| \leq c$ for all n and all x.

Answer. If we show that the series is the Fourier series of a function of bounded variation, the result follows from Corollary 8.1.13. The function f is of bounded variation, as it is piecewise continuous on a compact set.

The Fourier coefficients are $\hat{f}(0) = 0$ and for $n \neq 0$

$$\begin{aligned} \hat{f}(k) &= -\frac{1}{2\pi} \int_{-\pi}^{0} \frac{\pi + t}{2} \, e^{-ikt} \, dt + \frac{1}{2\pi} \int_{0}^{\pi} \frac{\pi - t}{2} \, e^{-ikt} \, dt \\ &= -\frac{1}{4\pi} \int_{-\pi}^{\pi} t \, e^{-ikt} \, dt - \frac{1}{4} \int_{-\pi}^{0} e^{-ikt} \, dt + \frac{1}{4} \int_{0}^{\pi} e^{-ikt} \, dt \\ &= -\frac{1}{4\pi} \left. \frac{te^{-ikt}}{-ik} \right|_{-\pi}^{\pi} - \frac{1}{4} \left. \frac{e^{-ikt}}{-ik} \right|_{-\pi}^{0} + \frac{1}{4} \left. \frac{e^{-ikt}}{-ik} \right|_{0}^{\pi} \\ &= \frac{(-1)^k}{2ik} + \frac{1}{4ik} - \frac{(-1)^k}{4ik} - \frac{(-1)^k}{4ik} + \frac{1}{4ik} = \frac{1}{2ik}. \end{aligned}$$

When we write the Fourier series, the corresponding positive and negative terms match like

$$\frac{e^{ikt}}{2ik} + \frac{e^{-ikt}}{2i(-k)} = \frac{1}{k}\,\frac{e^{ikt} - e^{-ikt}}{2i} = \frac{\sin kt}{k}.$$

Hence, as $\hat{f}(0) = 0$,

$$\sum_{k=1}^{n} \frac{\sin kt}{k} = S_{f,n}(t).$$

(8.1.8) Show that $\lim_n \|D_n\|_1 = \infty$.

Answer. We have, since $\sin t < t$ for $t \geq 0$,

$$\begin{aligned}
\|D_n\|_1 &\geq \int_0^{\pi} |D_n(t)|\,dt = \int_0^{\pi} \left|\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{t}{2}}\right| dt \\
&\geq \int_0^{\pi} \left|\frac{\sin\left(n+\frac{1}{2}\right)t}{t}\right| dt = \int_0^{\left(n+\frac{1}{2}\right)\pi} \frac{|\sin t|}{t}\,dt \\
&\geq \int_0^{n\pi} \frac{|\sin t|}{t}\,dt \geq \sum_{k=0}^{n-1} \int_{k\pi+\frac{\pi}{4}}^{k\pi+\frac{3\pi}{4}} \frac{|\sin t|}{t}\,dt \\
&\geq \frac{1}{\sqrt{2}} \sum_{k=0}^{n-1} \int_{k\pi+\frac{\pi}{4}}^{k\pi+\frac{3\pi}{4}} \frac{1}{t}\,dt \geq \frac{1}{\sqrt{2}} \sum_{k=0}^{n-1} \frac{1}{k\pi + \frac{3\pi}{4}} \xrightarrow[n\to\infty]{} \infty.
\end{aligned}$$

(8.1.9) Let $f \in L^1[-\pi,\pi]$ such that f is even. Show that $S_f(t)$ is of the form

$$S_f(t) = \sum_{k=0}^{\infty} a_k \cos kt,$$

where

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt, \qquad a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\cos kt\,dt, \quad k \geq 1.$$

Answer. Using that f is even and the substitution $t \to -t$,

$$\hat{f}(-k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,e^{-i(-k)t}\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(-t)\,e^{-ikt}\,dt = \hat{f}(k).$$

Then

$$\hat{f}(k)e^{ikt} + \hat{f}(-k)e^{-ikt} = \hat{f}(k)\big(e^{ikt} + e^{-ikt}\big) = 2\hat{f}(k)\,\cos kt.$$

Thus

$$S_f(t) = \hat{f}(0) + \sum_{k=1}^{\infty} \hat{f}(k)e^{ikt} + \hat{f}(-k)e^{-ikt} = \sum_{k=1}^{\infty} 2\hat{f}(k)\,\cos kt.$$

Since $\hat{f}(k) = \hat{f}(-k)$,

$$\hat{f}(k) = \frac{\hat{f}(k) + \hat{f}(-k)}{2} = \frac{1}{4\pi}\int_{-\pi}^{\pi} f(t)\left(e^{ikt} + e^{-ikt}\right) dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,\cos kt\,dt.$$

Thus

$$a_k = 2\hat{f}(k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\cos kt\,dt.$$

(8.1.10) Let $f \in L^1[-\pi,\pi]$ such that f is odd. Show that $S_f(t)$ is of the form

$$S_f(t) = \sum_{k=1}^{\infty} b_k \sin kt,$$

where

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\sin kt\,dt, \qquad k \geq 1.$$

Answer. Using that f is odd and the substitution $t \to -t$,

$$\hat{f}(-k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,e^{-i(-k)t}\,dt = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f(-t)\,e^{-ikt}\,dt = -\hat{f}(k).$$

Then

$$\hat{f}(k)e^{ikt} + \hat{f}(-k)e^{-ikt} = \hat{f}(k)\big(e^{ikt} - e^{-ikt}\big) = 2i\hat{f}(k)\,\sin kt.$$

Thus

$$S_f(t) = \sum_{k=1}^{\infty} \hat{f}(k)e^{ikt} + \hat{f}(-k)e^{-ikt} = \sum_{k=1}^{\infty} 2i\hat{f}(k)\,\sin kt.$$

Since $\hat{f}(k) = -\hat{f}(-k)$,

$$\begin{aligned}\hat{f}(k) &= \frac{\hat{f}(k) - \hat{f}(-k)}{2} = \frac{1}{4\pi}\int_{-\pi}^{\pi} f(t)\left(e^{-ikt} - e^{ikt}\right) dt \\ &= -i\,\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,\sin kt\,dt.\end{aligned}$$

Thus

$$a_k = 2i\hat{f}(k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\sin kt\,dt.$$

(8.1.11) Given $f, g \in L^1[-\pi, \pi]$, consider the **convolution**

$$(f * g)(t) = \int_{-\pi}^{\pi} f(x)\, g(t-x)\, dx,$$

where f, g are considered extended periodically to $[-2\pi, 2\pi]$. Show that $\widehat{f * g}(n) = 2\pi\, \hat{f}(n)\, \hat{g}(n)$.

Answer. Below, Fubini (Theorem 2.7.12) applies because $f, g \in L^1[-\pi, \pi]$. We have

$$\begin{aligned}
\widehat{f * g}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f * g)(t)\, e^{-int}\, dt \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x)\, g(t-x)\, e^{-inx}\, e^{-in(t-x)}\, dx\, dt \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx} \int_{-\pi}^{\pi} g(t-x)\, e^{-in(t-x)}\, dt\, dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx \int_{-\pi}^{\pi} g(t)\, e^{-int}\, dt \\
&= 2\pi\, \hat{f}(n)\, \hat{g}(n).
\end{aligned}$$

(8.1.12) Use the following steps to show that the function

$$f(t) = \sum_{k=1}^{\infty} \frac{1}{k^2} \sin\left[\left(2^{k^3} + 1\right) \frac{|t|}{2}\right], \qquad t \in [0, \pi]$$

is in $C[-\pi, \pi]$ and $\{S_{f,n}(0)\}$ diverges.

(i) Show that f is continuous.

(ii) Use Exercise 8.1.9 to write $S_f(t) = \sum_{k=1}^{\infty} a_k \cos kt$, where

$$a_k = \frac{2}{\pi} \sum_{j=1}^{\infty} \frac{1}{j^2} \lambda_{k, 2^{j^3}-1},$$

and

$$\lambda_{k,h} = \int_0^{\pi} \sin\left[(2h+1)\frac{t}{2}\right] \cos kt\, dt.$$

(iii) Show that $\sum_{k=0}^{h} \lambda_{k,h} \geq \frac{1}{2} \log h$ and $\sum_{k=0}^{n} \lambda_{k,h} \geq 0$ for all n, h.

(iv) Conclude that there exists a sequence $\{n_j\}$ such that

$$S_{f,n_j}(0) \to \infty,$$

by showing that

$$S_{f,2^{j^3}-1}(0) \geq \frac{j^3-1}{2j^2} \log 2.$$

Answer.

(i) The series for f converges uniformly by comparison with $\sum_k \frac{1}{k^2}$. The extension to $[-\pi, \pi]$ satisfies the same bound, and the extension is continuous at 0 since $f(0) = 0$.

(ii) We have, using either Dominated Convergence or Fubini,

$$\begin{aligned} a_k &= \frac{2}{\pi} \int_0^{\pi} \sum_{j=1}^{\infty} \frac{1}{j^2} \sin\left[\left(2^{j^3}+1\right)\frac{t}{2}\right] \cos kt\, dt \\ &= \frac{2}{\pi} \sum_{j=1}^{\infty} \frac{1}{j^2} \int_0^{\pi} \sin\left[\left(2\, 2^{j^3-1}+1\right)\frac{t}{2}\right] \cos kt\, dt \\ &= \frac{2}{\pi} \sum_{j=1}^{\infty} \frac{1}{j^2} \lambda_{k,2^{j^3-1}}. \end{aligned}$$

(iii) We have

$$\begin{aligned} \lambda_{k,h} &= \int_0^{\pi} \sin\left[(2h+1)\frac{t}{2}\right] \cos kt\, dt \\ &= \frac{1}{2} \int_0^{\pi} \left[\sin\left(\frac{2h+1}{2}+k\right)t + \sin\left(\frac{2h+1}{2}-k\right)t\right] dt \\ &= \frac{1}{2h+2k+1} + \frac{1}{2h-2k+1} = \frac{2h+1}{(2h+1)^2-4k^2}. \end{aligned}$$

When $k \leq h$ this gives us $\lambda_{k,h} \geq 0$ and hence $\sum_{k=0}^{n} \lambda_{k,h} \geq 0$ if $n \leq h$. When $n \geq h$,

$$2\sum_{k=0}^{n} \lambda_{k,h} = \sum_{k=0}^{n} \frac{1}{2h+2k+1} + \sum_{k=0}^{n} \frac{1}{2h-2k+1}$$
$$= \sum_{k=h}^{n+h} \frac{1}{2k+1} + \sum_{k=n-h}^{h} \frac{1}{2k+1}$$
$$= \frac{1}{2h+1} + \sum_{k=n-h}^{n+h} \frac{1}{2k+1} \geq 0.$$

This last estimate also allows us to do

$$2\sum_{k=0}^{h} \lambda_{k,h} = \frac{1}{2h+1} + \sum_{k=0}^{2h} \frac{1}{2k+1} \geq \sum_{k=0}^{2h} \int_{2k+1}^{2k+3} \frac{1}{t}\,dt$$
$$= \int_{1}^{4h+3} \frac{1}{t}\,dt = \log(4h+3) \geq \log h.$$

(iv) Now we have

$$S_{f,2^{j^3-1}}(0) = \sum_{k=0}^{2^{j^3-1}} a_k = \frac{2}{\pi} \sum_{k=0}^{2^{j^3-1}} \sum_{h=1}^{\infty} \frac{1}{h^2}\lambda_{k,2^{h^3-1}} \geq \frac{2}{\pi}\frac{1}{j^2} \sum_{k=0}^{2^{j^3-1}} \lambda_{k,2^{j^3-1}}$$
$$\geq \frac{2}{\pi}\frac{1}{j^2}\log(2^{j^3-1}) = \frac{2}{\pi}\frac{j^3-1}{j^2}\log 2 \xrightarrow[j\to\infty]{} \infty.$$

8.2. The Fourier Transform

(8.2.1) Let $f \in L^1(\mathbb{R}^n)$ and $c \in \mathbb{R}$. Show that the Fourier transform is linear, and that if $g(x) = f(cx)$, then $\hat{g}(\xi) = \frac{1}{|c|^n}\hat{f}\left(\frac{\xi}{c}\right)$.

Answer. The linearity follows directly from the linearity of the integral. We have, with the change of variables $x \longmapsto x/c$,

$$\hat{g}(\xi) = \int_{\mathbb{R}^n} f(cx)\,e^{-2\pi i\langle \xi,x\rangle}\,dx = \frac{1}{|c|^n}\hat{f}\left(\frac{\xi}{c}\right).$$

(8.2.2) Let $f \in L^1(\mathbb{R})$ and such that $g : x \longmapsto xf(x)$ is also in L^1. Show that $\hat{f}$ is differentiable and $(\hat{f})'(\xi) = -2\pi i\hat{g}(\xi)$.

Answer. The two conditions on f allow us to apply Dominated Convergence and differentiate under the integral symbol (properly, instead of taking limits as $h \to 0$ we need to work for sequences $\{h_n\}$, but the result works for all sequences and hence allows us to take the limit as $h \to 0$). So

$$(\hat{f})'(\xi) = -2\pi i \int_{\mathbb{R}} xf(x)\, e^{-2\pi i\xi x}\, dx = -2\pi i\hat{g}(\xi).$$

(8.2.3) Let $f \in L^1(\mathbb{R})$ differentiable with $f' \in L^1(\mathbb{R})$. Show that $\widehat{f'}(\xi) = 2\pi i\xi\, \hat{f}(\xi)$.

Answer. Integrating by parts,

$$\widehat{f'}(\xi) = \int_{\mathbb{R}} f'(x)\, e^{-2\pi i\xi x}\, dx = 2\pi i\xi \int_{\mathbb{R}} f(x)\, e^{-2\pi i\xi x}\, dx = 2\pi i\xi\, \hat{f}(\xi).$$

(8.2.4) Let $f, g \in L^1(\mathbb{R}^n)$. Show that

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\, \hat{g}(\xi).$$

Answer. The fact that both f, g are in $L^1(\mathbb{R}^n)$ allows us to use Fubini's Theorem. Then

$$\begin{aligned}
\widehat{f * g}(\xi) &= \int_{\mathbb{R}^n} (f * g)(x)\, e^{-2\pi i \langle \xi, x\rangle}\, dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)\, g(y)\, e^{-2\pi i \langle \xi, x\rangle}\, dy\, dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)\, g(y)\, e^{-2\pi i \langle \xi, x\rangle}\, dx\, dy \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)\, e^{-2\pi i \langle \xi, x\rangle}\, dx\, g(y)\, dy \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \langle \xi, x+y\rangle}\, dx\, g(y)\, dy \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \langle \xi, x\rangle}\, dx\, g(y)\, e^{-2\pi i \langle \xi, y\rangle}\, dy \\
&= \int_{\mathbb{R}^n} \hat{f}(\xi)\, g(y)\, e^{-2\pi i \langle \xi, y\rangle}\, dy \\
&= \hat{f}(\xi)\, \hat{g}(\xi).
\end{aligned}$$

8.3. Hilbert Function Spaces

(8.3.1) Show that if $a \in \ell^2(\mathbb{N})$ then $f = \sum_n a_n z^n$ is analytic on $\mathbb{D}$.

Answer. For each $z \in \mathbb{D}$ the series converges, since

$$\Big| \sum_{n=m}^{\infty} a_n z^n \Big| \leq \Big(\sum_{n=m}^{\infty} |a_n|^2 \Big)^{1/2} \Big(\sum_{n=m}^{\infty} |z|^{2n} \Big)^{1/2} \leq \|a\|_2 \sqrt{\frac{|z|^m}{1-|z|^2}}.$$

The estimate shows that the convergence is uniform in any proper subdisk of $\mathbb{D}$. This guarantees that we can differentiate term by term, and hence f is holomorphic/analytic.

(8.3.2) Show that if $f, g \in B_a(\mathbb{D})$ and $f = \sum_n a_n z^n$, $g = \sum_n b_n z^n$, then

$$\langle f, g\rangle_B = \sum_{n=0}^{\infty} \frac{a_n \overline{b_n}}{n+1}.$$

Answer. Since $|z| < 1$ for all z, the uniform convergence allows us to write

$$\langle f, g\rangle = \sum_{n,m} a_n \overline{b_m}\, \frac{1}{\pi} \int_{\mathbb{D}} z^n \bar{z}^m \, dz.$$

Assume without loss of generality that $n \geq m$. When $n > m$, $z^n \bar{z}^m = z^{n-m}\, |z|^m$ and so if $z = re^{it}$ then $z^n \bar{z}^m = r^{n+m}\, e^{i(n-m)t}$. Hence

$$\frac{1}{\pi} \int_{\mathbb{D}} z^n \bar{z}^m \, dz = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 r^{n+m+1}\, e^{i(n-m)t} \, dr\, dt = 0,$$

since the integral of the exponential is zero. When $n = m$ we get

$$\frac{1}{\pi} \int_{\mathbb{D}} |z|^{2n} \, dz = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 r^{2n+1} \, dr\, dt = \frac{1}{n+1}.$$

(8.3.3) Show that the Bergman space is a Hilbert function space (that is, complete and point evaluations are bounded), and that the set $\{z^n\}_{n=0}^{\infty}$ is orthogonal and total.

Answer.

We have

$$\begin{aligned}\langle z^n, z^m\rangle_B &= \frac{1}{\pi} \int_{\mathbb{D}} z^n \overline{z^m} \, dm(z) \\ &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 r^{n+m+1}\, e^{it(n-m)} \, dr\, dt = \frac{2}{n+m+1}\, \delta_{n,m}.\end{aligned}$$

If $\langle f, z^n\rangle_B = 0$ for all n, then $f = 0$, and so $\{z^n\}$ is orthogonal and total.

Given $z \in \mathbb{D}$ and $f \in L_a^2(\mathbb{D})$, for any $r \in (0,1)$ Cauchy's Formula gives us

$$\begin{aligned}|f(z)| &= \frac{1}{2\pi} \Big| \int_{r\mathbb{T}} \frac{f(w)}{w - z} \, dw \Big| = \frac{1}{2\pi} \Big| \int_0^{2\pi} \frac{f(re^{it}\, ie^{it}}{re^{it}} \, dt \Big| \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})| \, dt \leq \sqrt{2\pi} \Big(\int_0^{2\pi} |f(re^{it})|^2 \, dt \Big)^{1/2}.\end{aligned}$$

Squaring both sides, multiplying by r, and integrating from 0 to 1, we get

$$|f(z)|^2 \leq 4\pi \int_{\mathbb{D}} |f|^2 = 4\pi \|f\|_B.$$

So $z \longmapsto f(z)$ is bounded for each $z \in \mathbb{D}$.

It remains to check that $L^2_a(\mathbb{D})$ is complete. If $\{f_n\}$ is a Cauchy sequence in $L^2_a(\mathbb{D})$, then by the above estimate

$$|f_n(z) - f_m(z)| \leq 4\pi \, \|f_n - f_m\|_B.$$

So the sequence is uniformly Cauchy, which guarantees that the limit exists and it is analytic and in $L^2(\mathbb{D})$. The only non-obvious part of the last sentence is that the limit f is square integrable. We have

$$|f(z) - f_m(z)| = \lim_n |f_n(z) - f_m(z)| \leq \limsup_n \|f_n - f_m\|_B.$$

Choosing m sufficiently big, we can make the left-hand-side as small as we want. Now we can do, using Fatou's Lemma,

$$\begin{aligned} \|f - f_m\|_B^2 &= \frac{1}{\pi} \int_{\mathbb{D}} |f(z) - f_m(z)|^2 \, dm \\ &= \frac{1}{\pi} \int_{\mathbb{D}} \liminf_n |f(z)_n - f_m(z)|^2 \, dm \\ &\leq \liminf_n \|f_n - f_m\|^2 \end{aligned}$$

and this goes to zero with m.

(8.3.4) Let $p \in [1, \infty]$, and consider the corresponding Hardy space $\mathbb{H}^p(\mathbb{D})$, that is the space of analytic functions $f(z) = \sum_n a_n z^n$ on the disk, such that $a \in \ell^p(\mathbb{N})$ and $\sup_{0<r<1} \|f_r\|_p < \infty$, where $f_r(t) = f(re^{it})$, $t \in [-\pi, \pi]$. Show that $\|f\| = \lim_{r\to 1} \|f_r\|_p$ is a norm, and that $\mathbb{H}^p(\mathbb{D})$ is complete with that norm.

Answer.

Let $0 < r < s < 1$. By Cauchy's Integral Formula applied on the curve $\gamma(\theta) = s\, e^{i\theta}$,

$$f(re^{it}) = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{f(se^{i\theta})}{se^{i\theta} - re^{it}} \, si\, e^{i\theta} \, d\theta.$$

We have

$$\begin{aligned}\operatorname{Re}\frac{se^{i\theta}}{se^{i\theta}-re^{it}} &= \operatorname{Re}\frac{1}{1-\frac{r}{s}\,e^{i(t-\theta)}} = \operatorname{Re}\frac{1}{1-\frac{r}{s}\cos(t-\theta)-i\,\frac{r}{s}\,\sin(t-\theta)} \\ &= \frac{1-\frac{r}{s}\cos(t-\theta)}{(1-\frac{r}{s}\cos(t-\theta))^2+\frac{r^2}{s^2}\sin^2(t-\theta)} \\ &= \frac{1-\frac{r}{s}\cos(t-\theta)}{1+\frac{r^2}{s^2}-2\frac{r}{s}\,\cos(t-\theta)}\end{aligned}$$

This allows us to write, applying the above to αf for $\alpha \in \mathbb{T}$ with $\alpha\, f(re^{it}) = |f(r^{it})|$,

$$\alpha\, f(re^{it}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \alpha\,(se^{i\theta})\, P_{r/s}(t-\theta)\, d\theta,$$

where

$$P_\rho(\phi) = \frac{1-\rho\cos\phi}{1+\rho^2-2\rho\cos\phi}.$$

The function $P_\rho(\phi)$ satisfies, for $0<\rho<1$, $P_\rho(\phi)\geq 0$ and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_\rho(\phi)\, d\phi = 1.$$

Because $x \longmapsto |x|^p$ is convex (as $p \geq 1$), Jensen's Inequality (Theorem 7.5.4), applied to the functional of integration against $P_{r/s}(t-\theta)$, gives

$$|f(re^{it})|^p \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(se^{i\theta})|^p\, P_{r/s}(t-\theta)\, d\theta.$$

It follows that

$$\begin{aligned}\int_{-\pi}^{\pi} |f(re^{it})|^p\, dt &\leq \int_{-\pi}^{\pi} \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(se^{i\theta})|^p\, P_{r/s}(t-\theta)\, d\theta\, dt \\ &= \int_{-\pi}^{\pi} |f(se^{i\theta})|^p \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{r/s}(t-\theta)\, dt\, d\theta \\ &= \int_{-\pi}^{\pi} |f(se^{i\theta})|^p\, d\theta.\end{aligned}$$

That is, $\|f_r\|_p \leq \|f_s\|_p$. So the numbers $\|f_r\|_p$ are monotone on r and bounded, so the limit at 1 exists.

Next we check that $\|\cdot\|$ is a norm. Suppose that $\|f\| = 0$. This means, by the monotonicity of $\|f_r\|_p$, that $\|f_r\|_p = 0$ for all $r \in (0,1)$. Let $z \in \mathbb{D}$.

Choose any $r \geq r_0$ with $|z| < r < 1$. By Cauchy's Integral Formula,

$$|f(z)| = \Big|\frac{1}{2\pi i}\oint_{r\mathbb{T}} \frac{f(w)}{w-z}\,dw\Big| = \frac{1}{2\pi}\Big|\int_{-\pi}^{\pi} \frac{f(e^{it})}{re^{it}-z}\,dt\Big|$$

$$\leq \frac{1}{2\pi\,(r-|z|)}\int_{-\pi}^{\pi} |f(re^{it})|\,dt$$

$$\leq \frac{(2\pi)^{1/q}}{2\pi\,(r-|z|)}\,\|f_r\|_p = 0.$$

So $f = 0$. We have $\|\lambda f\| = \lim_{r\to 1}\|\lambda f_r\| = |\lambda|\,\|f\|$. And

$$\|f+g\| = \lim_{r\to 1}\|f_r + g_r\|_p \leq \lim_{r\to 1}\|f_r\|_p + \lim_{r\to 1}\|g_r\|_p = \|f\| + \|g\|.$$

It remains to show that $\mathbb{H}^p$ is complete. Suppose that $\{f_n\} \subset \mathbb{H}^p$ is Cauchy. The same estimates we used above show that there is $c > 0$ such that $|f(z)| \leq c\|f_r\|_p$ as long as $r > |z|$. In that situation,

$$|f_n(z) - f_m(z)| \leq c\,\|(f_n - f_m)_r\|_p,$$

and so $\{f_n(z)\}$ is uniformly Cauchy in $r\mathbb{D}$ and thus convergent to an analytic function f. As we can do this for any $r < 1$, we get that $f \in A(\mathbb{D})$. And because the convergence is uniform in $r\mathbb{D}$ for any r, we have that $\|f_r\|_p = \lim_n \|(f_n)_r\|_p$. From $|\,\|f_n\| - \|f_m\|\,| \leq \|f_n - f_m\|$, we get that the number sequence $\{\|f_n\|\}$ is Cauchy, and thus convergent, say $\|f_n\| \to c'$. Fix $\varepsilon > 0$. Choose n such that $\|f_r - (f_n)_r\|_p < \varepsilon$ and $|\,\|f_n\| - c'| < \varepsilon$. Then

$$|\,\|f_r\|_p - c'\,| \leq |\,\|f_r\|_p - \|(f_n)_r\|_p\,| + |\,\|(f_n)_r\|_p - \|f_n\|\,| + |\,\|f_n\| - c'|$$

$$\leq 2\varepsilon + |\,\|(f_n)_r\|_p - \|f_n\|\,|.$$

It follows that

$$\limsup_{r\to 1}|\,\|f_r\|_p - c'\,| \leq 2\varepsilon,$$

for any $\varepsilon > 0$, showing that $\lim_{r\to 1}\|f_r\|_p = c'$, and $f \in \mathbb{H}^p$.

(8.3.5) Show that $\mathcal{D}(\mathbb{D})$ is a Hilbert function space, and find its reproducing kernel and its feature map.

Answer. Let us first determine the inner product. Given $f(z) = \sum_n a_n z^n$ and $g(z) = \sum_n b_n z^n$,

$$\begin{aligned}\langle f, g\rangle &= a_0\overline{b_0} + \sum_{n,m=1}^{\infty} a_n\overline{b_m}\,\frac{nm}{\pi}\int_{\mathbb{D}} z^{n-1}\,\overline{z^{m-1}}\,dz\\ &= a_0\overline{b_0} + \sum_{n=1}^{\infty} a_n\overline{b_n}\,\frac{n^2}{\pi}\int_{\mathbb{D}} |z|^{2n-2}\,dz\\ &= a_0\overline{b_0} + \sum_{n=1}^{\infty} a_n\overline{b_n}\,\frac{n^2}{\pi}\int_0^{2\pi}\int_0^1 r^{2n-2}\,r\,dr\,d\theta\\ &= a_0\overline{b_0} + \sum_{n=1}^{\infty} n\,a_n\overline{b_n}.\end{aligned}$$

We have $f'(z) = \sum_{n=1}^{\infty} na_n z^{n-1}$, so

$$f(z) = \sum_n a_n z^n = f(0) + \sum_{n=1}^{\infty}\frac{na_n\,z\,\overline{z}^{n-1}}{n} = \langle f, k_z\rangle,$$

where

$$k_z(w) = 1 + \sum_{n=1}^{\infty}\frac{\bar{z}^{n-1}w^n}{n} = 1 - \frac{1}{\bar{z}}\log(1-\bar{z}w).$$

(8.3.6) Let $\mathcal{H}$ be a Hilbert Function Space over X with reproducing kernel k, and $\mathcal{H}_0 \subset \mathcal{H}$ a closed subspace. Show that $\mathcal{H}_0$ is a Hilbert Function Space over X and that its reproducing kernel is given by $k_0(x,y) = \langle P(k_x), P(k_y)\rangle$, where P is the orthogonal projection onto $\mathcal{H}_0$.

Answer. The elements of $\mathcal{H}_0$ are in $\mathcal{H}$, so points evaluations are bounded. Given $x \in X$, the Riesz Representation Theorem guarantees that there exists $k_x' \in \mathcal{H}_0$ such that $g(x) = \langle g, k_x'\rangle$ for all $g \in \mathcal{H}_0$. We also have that $f(x) = \langle f, k_x\rangle$ for all $f \in \mathcal{H}$. As $Pg = g$ for all $g \in \mathcal{H}_0$,

$$\langle g, k_x'\rangle = g(x) = (Pg)(x) = \langle Pg, k_x\rangle = \langle g, Pk_x\rangle.$$

This occurs for all $g \in \mathcal{H}_0$, hence $k_x' = Pk_x$. And then

$$k_0(x,y) = \langle k_x', k_y'\rangle = \langle Pk_x, Pk_y\rangle.$$

CHAPTER

9

Bounded operators on a Banach space

9.1. Linear operators and continuity

(9.1.1) Show that if $T, S \in \mathcal{B}(\mathcal{X})$ and both TS and ST are invertible, then both T and S are invertible. Find an example where TS is invertible, but neither T nor S are.

Answer. By hypothesis there exist $X, Y \in \mathcal{B}(\mathcal{X})$ with $XTS = I_{\mathcal{X}} = STY$. We can see this as $(XT)S = S(TY) = I_{\mathcal{X}}$. Also,

$$XT = XTI_{\mathcal{X}} = XTSTY = I_{\mathcal{X}}TY = TY.$$

Then the element $XT = TY$ is the inverse of S. We can similarly use the equalities $TSX = YST = I_{\mathcal{X}}$ to conclude that T is invertible.

Such an example is discussed at the beginning of Section 6.2. Namely, let $\mathcal{X} = \ell^2(\mathbb{N})$, and

$$Tx = (x_2, x_3, \ldots,), \qquad Sx = (0, x_1, x_2, \ldots).$$

Since $\|Tx\| \leq \|x\|$ and $\|Sx\| = \|x\|$ for all x, we have $T, S \in \mathcal{B}(\mathcal{X})$. Also, $TS = I$, but $ST \neq I$ (T is not injective!). So neither T nor S are invertible, even though T admits a right inverse and S a left inverse.

(9.1.2) Show that if $\dim \mathcal{X} < \infty$, then any linear $T : \mathcal{X} \to \mathcal{X}$ is bounded.

Answer. Because $\dim \mathcal{X} < \infty$, all norms are equivalent (Theorem 5.2.2). Consider the norm $p(x) = \|x\| + \|Tx\|$. Then there exists $c > 0$ with $\|x\| + \|Tx\| \leq c\|x\|$. Thus $\|Tx\| \leq (c-1)\|x\|$.

(9.1.3) Show that if $\dim \operatorname{ran} T < \infty$, it is not necessarily true that T is bounded.

Answer. Using the ideas at the end of Section 4.5, we may construct an unbounded linear functional $f : \mathcal{X} \to \mathbb{C}$. Then $Tx = f(x)\, x_0$ is a rank-one operator that is not bounded.

(9.1.4) Let $\mathcal{X}$ be a normed space and $A \in \mathcal{B}(\mathcal{X})$ invertible. Show that

$$\|A\|\,\|A^{-1}\| = \sup\left\{\frac{\|Ax\|}{\|Ay\|} :\ x, y \in \mathcal{X},\ \|x\| = \|y\|\right\}.$$

Answer. For any x, y with $\|x\| = \|y\|$ we have, with $z = Ay$,

$$\frac{\|Ax\|}{\|Ay\|} = \frac{\|Ax\|}{\|x\|}\,\frac{\|y\|}{\|Ay\|} = \frac{\|Ax\|}{\|x\|}\,\frac{\|A^{-1}z\|}{\|z\|} \leq \|A\|\,\|A^{-1}\|.$$

Now fix $\varepsilon > 0$. Then there exist x, z such that

$$\frac{\|Ax\|}{\|x\|} > \|A\| - \varepsilon \qquad\qquad \frac{\|A^{-1}z\|}{\|z\|} > \|A^{-1}\| - \varepsilon.$$

By rescaling z if needed, we may assume that $\|A^{-1}z\| = \|x\|$. With $y = A^{-1}z$,

$$\frac{\|Ax\|}{\|Ay\|} = \frac{\|Ax\|}{\|x\|}\,\frac{\|A^{-1}z\|}{\|z\|} > (\|A\| - \varepsilon)(\|A^{-1}\| - \varepsilon).$$

As $\|x\| = \|y\|$ and this can be done for each $\varepsilon > 0$, this shows that

$$\|A\|\,\|A^{-1}\| = \sup\left\{\frac{\|Ax\|}{\|Ay\|} :\ \|x\| = \|y\|\right\}.$$

(9.1.5) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{X} \to \mathcal{Y}$ linear, injective, and such that $T\big(\overline{B_1^{\mathcal{X}}(0)}\big) = \overline{B_1^{\mathcal{Y}}(0)}$. Show that T is isometric.

Answer. If $x \in \mathcal{X}$ with $\|x\| \leq 1$, then $Tx \in \overline{B_1^{\mathcal{Y}}(0)}$, so $\|Tx\| \leq 1$. This shows that T is bounded and $\|T\| \leq 1$. Fix $x \in \mathcal{X}$ with $\|x\| = 1$. If $\|Tx\| < 1$, let $c = 1/\|Tx\|$ (note that $Tx \neq 0$ by the injectivity of T). Then $c > 1$ and $\|T(cx)\| = 1$. But $\|cx\| = c > 1$, so $cx \notin \overline{B_1^{\mathcal{X}}(0)}$; and by hypothesis there exists $x' \in \overline{B_1^{\mathcal{X}}(0)}$ with $Tx' = T(cx)$, contradicting the injectivity. It follows that $\|Tx\| = 1$. Now for arbitrary nonzero x,

$$\|Tx\| = \|x\| \left\| T\left(\frac{x}{\|x\|}\right) \right\| = \|x\|.$$

(9.1.6) *(this exercise does not require sophisticated ideas but it is far from trivial; we include it here because it is where it corresponds topic-wise, but it likely requires more experience with operators)*

Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$. We define numbers

$$a(T) = \min\{k \in \mathbb{N} : \ \ker T^k = \ker T^{k+1}\}$$

and

$$d(T) = \min\{k \in \mathbb{N} : \ \operatorname{ran} T^k = \operatorname{ran} T^{k+1}\}.$$

(i) Show that if both $a(T)$ and $d(T)$ are finite, then they are equal.

(ii) If $n = a(T) = d(T)$, show that $\operatorname{ran} T^n$ is closed and that $\mathcal{X} = \operatorname{ran} T^n \oplus \ker T^n$.

Answer. Let $m = a(T)$, $n = d(T)$. We always have the inclusions

$$\ker T \subset \ker T^2 \subset \ker T^3 \subset \cdots \subset \ker T^m$$

and

$$\operatorname{ran} T \supset \operatorname{ran} T^2 \supset \operatorname{ran} T^3 \supset \cdots \supset \operatorname{ran} T^n.$$

(i) Note that $\ker T^p = \ker T^m$ for all $p \geq m$. Indeed, if $x \in \ker T^{m+2} \setminus \ker T^{m+1}$, this means that $T^{m+2}x = 0$ while $T^{m+1}x \neq 0$. Then $Tx \in \ker T^{m+1} \setminus \ker T^m$, a contradiction. Similarly, $\operatorname{ran} T^p = \operatorname{ran} T^n$ for all $p \geq n$; indeed, if $y \in \operatorname{ran} T^{n+1}$then $y = T^{n+1}z$; and $T^n z \in \operatorname{ran} T^n = \operatorname{ran} T^{n+1}$ so $T^n z = T^{n+1}w$ for some w and then $y = T^{n+2}w \in \operatorname{ran} T^{n+2}$.

Suppose that $m > n$. We have

$$\operatorname{ran} T^n = \operatorname{ran} T^{n+1} \qquad \text{and} \qquad \ker T^m = \ker T^{m+1}.$$

Let $x \in \ker T^m$. We have $T^{m-1}x \in \operatorname{ran} T^{m-1} = \operatorname{ran} T^m$, so there exists w with $T^{m-1}x = T^m w$. Then $0 = T^m x = T^{m+1}w$. Thus $w \in \ker T^{m+1} = \ker T^m$, which means that $T^{m-1}x = T^m w = 0$. This shows that $\ker T^m \subset$

$\ker T^{m-1}$ and hence they are equal, a contradiction. We have shown then that $m \leq n$.

Now suppose that $m < n$. Let $x \in \operatorname{ran} T^m \setminus \operatorname{ran} T^{m+1}$. Let $w \in \mathcal{X}$ such that $x = T^m w$. We have $T^{m+1}y - T^m w \neq 0$ for all $y \in \mathcal{X}$; that is, $Ty - w \notin \ker T^m$ for all $y \in \mathcal{X}$. As $\ker T^m = \ker T^n$, we get that $0 \neq T^n(Ty - w) = T^{n+1}y - T^n w$ for all y. But $T^n w \in \operatorname{ran} T^n = \operatorname{ran} T^{n+1}$, a contradiction. Thus $m = n$.

(ii) We note first that $\operatorname{ran} T^n \cap \ker T^n = \{0\}$. If $x = T^n w$ and $T^n x = 0$, then $T^{2n} w = 0$. As $\ker T^{2n} = \ker T^n$ we get that $0 = T^n w = x$.

Fix $x \in \mathcal{X}$. Since $\operatorname{ran} T^n = \operatorname{ran} T^{2n}$ there exists $z \in \mathcal{X}$ with $T^n x = T^{2n} z$. Let $w = x - T^n z$. Then $T^n w = T^n x - T^{2n} z = 0$ and so $w \in \ker T^n$. So $x = T^n z + w$ with $T^n z \in \operatorname{ran} T^n$ and $w \in \ker T^n$. Thus

$$\mathcal{X} = \operatorname{ran} T^n \oplus \ker T^n.$$

Finally we show that $\operatorname{ran} T^n$ is closed. Since $\ker T^n$ is closed, the quotient $\mathcal{X}/\ker T^n$ is Banach (Proposition 5.3.13). Consider the map $\tilde{T} : \mathcal{X}/\ker T^n \to \operatorname{ran} T^k$ given by $\tilde{T}(x + \ker T^n) = T^n x$. The map $\tilde{T}$ is linear and injective by construction. The condition $\operatorname{ran} T^n = \operatorname{ran} T^{n+1}$ makes it surjective, for given $T^n x$ there exists $y \in \mathcal{X}$ with $T^n x = T^{n+1} y$; then $\tilde{T}(Ty + \ker T^n) = T^{n+1} y = T^n x$. Now consider the map

$$\tilde{S} : \mathcal{X}/\ker T^n \to \mathcal{X}/\ker T^n, \qquad \tilde{S}(x + \ker T^n) = Tx + \ker T^n.$$

This is linear and well-defined: if $x - x' \in \ker T^n$ then $x - x' \in \ker T^{n+1}$ and so $Tx - Tx' \in \ker T^n$. It is also surjective. Indeed, given $x \in \mathcal{X}$ the equality $\operatorname{ran} T^n = \operatorname{ran} T^{n+1}$ implies that there exists $y \in \mathcal{X}$ with $T^n x = T^{n+1} y$. Then $x - Ty \in \ker T^n$, so $x + \ker T^n = \tilde{S}(y + \ker T^n)$. And it is injective; if $\tilde{S}(x + \ker T^n) = \ker T^n$ this means that $Tx \in \ker T^n$ and this is $x \in \ker T^{n+1} = \ker T^n$, so $x + \ker T^n = 0$. It follows that $\tilde{S}$ is a bounded bijection of $\mathcal{X}/\ker T^n$. By the Inverse Mapping Theorem $\tilde{S}$ is invertible.

We have, with q the quotient map

$$\tilde{T}\,\tilde{S}^{-n}\,q(T^n x) = \tilde{T}(\tilde{S}^{-n}(T^n x + \ker T^n)) = \tilde{T}(x + \ker T^n) = T^n x$$

and

$$\tilde{S}^{-n}\,q\,\tilde{T}(x + \ker T^n) = \tilde{S}^{-n}(T^n x + \ker T^n) = x + \ker T^n.$$

So $\tilde{T}$ has a bounded inverse. Consider $\{T^n x_k\}$ Cauchy. Since $\mathcal{X}$ is Banach there exists $z \in \mathcal{X}$ with $T^n x_k \to z$. As $\tilde{T}$ is bicontinuous we have that $\{x_k + \ker T^n\}$ is Cauchy in $\mathcal{X}/\ker T^n$. This latter space is also Banach, so there exists a limit $y + \ker T^n = \lim_k x_k + \ker T^n$. Applying $\tilde{T}$, which is bounded,

$$z = \lim_k T^n x_k = \lim_k \tilde{T}(x_k + \ker T^n) = \tilde{T}(y + \ker T^n) = T^n y.$$

Therefore $\operatorname{ran} T^n$ is closed.

(9.1.7) Show that (9.2) does indeed correspond to the operator-block matrix form of ST.

Answer. We have, since $\sum_h P_h = I_{\mathcal{X}}$ and $P_h P_\ell = 0$ whenever $h \neq \ell$,

$$(TS)_{kj} = P_k TSP_j = \Big(\sum_{h=1}^n P_k TP_h\Big)\Big(\sum_{\ell=1}^n P_\ell SP_j\Big)$$
$$= \sum_{h=1}^n P_k TP_h SP_j = \sum_{h=1}^n T_{kh} S_{hj}.$$

(9.1.8) Prove Proposition 9.1.3.

Answer. If $T\mathcal{X}_1 \subset \mathcal{X}_1$, then $T_{21} = P_2 TP_1 = 0$. Conversely, if $T_{21} = 0$, then $TP_1 x = T_{11} P_1 x + T_{21} P_1 x = T_{11} P_1 x \subset \mathcal{X}_1$.

9.2. Banach Algebras and the Spectrum

(9.2.1) Prove that for every $a \in \mathcal{A}$, the resolvent $\rho(a)$ is open.

Answer. By definition of spectrum, if $\mathcal{A}$ is not unital we need to work on the unitization of $\mathcal{A}$. So we assume without loss of generality that $\mathcal{A}$ is unital. Let $\lambda \in \rho(a)$. Then $a-\lambda I \in GL(\mathcal{A})$. For any $\mu \in \mathbb{C}$ with $|\mu-\lambda| < \|(a-\lambda I)^{-1}\|^{-1}$,

$$\|a - \lambda I - (a - \mu I)\| = |\mu - \lambda| < \|(a - \lambda I)^{-1}\|^{-1}.$$

By Proposition 9.2.2, $a - \mu I \in GL(\mathcal{A})$, so $\mu \in \rho(a)$. This shows that every $\lambda \in \rho(a)$ is interior, so $\rho(a)$ is open.

(9.2.2) If $\mathcal{A} = M_n(\mathbb{C})$ and $a \in M_n(\mathbb{C})$, show that the spectrum $\sigma(a)$ consists of the eigenvalues of a.

Answer. For $a - \lambda I$ to be non-invertible, an equivalent condition is that there exists nonzero $v \in \mathbb{C}^n$ with $(A - \lambda I)v = 0$; that is, $Av = \lambda v$. So λ is an eigenvalue precisely when $A - \lambda I$ is not invertible.

(9.2.3) Let $a \in \ell^\infty(\mathbb{N})$. Show that $\sigma(a) = \overline{\{a(n) : \ n \in \mathbb{N}\}}$.

Answer. If $\lambda = a(m)$, then $(a - \lambda)b$ has zero as its m^{th} entry for any $b \in \ell^\infty(\mathbb{N})$, so $a - \lambda$ cannot be invertible. Thus $\{a(n) : \ n \in \mathbb{N}\} \subset \sigma(a)$, and as $\sigma(a)$ is closed, $\overline{\{a(n) : \ n \in \mathbb{N}\}} \subset \sigma(a)$. Conversely, if $\lambda \notin \overline{\{a(n) : \ n \in \mathbb{N}\}}$, then there exists $\delta > 0$ with $|a(n) - \lambda| > \delta$ for all n. Then $1/|a(n) - \lambda| < 1/\delta$ for all n, which shows that $b(n) = 1/(a(n) - \lambda)$ gives $b \in \ell^\infty(\mathbb{N})$ with $b(a - \lambda) = (a - \lambda b) = 1$, and so $a - \lambda$ is invertible.

(9.2.4) Let $f \in L^\infty(X, \Sigma, \mu)$. Show that $\sigma(f) = \operatorname{ess\,ran} f$.

Answer. We can write the essential range E of f as

$$E = \{z \in \mathbb{C} : \ \mu(\{x : \ |f(x) - z| < \varepsilon\}) > 0 \ \text{ for all } \varepsilon > 0\}.$$

Suppose first that $\lambda \notin E$. Then there exists $\varepsilon > 0$ such that $|f(x) - \lambda| \geq \varepsilon$ a.e. Thus $g = \frac{1}{f - \lambda} \in L^\infty(x)$, as $|g| \leq \frac{1}{\varepsilon}$ a.e. It follows that $f - \lambda$ has inverse g, and so $\lambda \notin \sigma(f)$.

Conversely, suppose that $\lambda \in E$. Then $\mu(\{|f - \lambda| < 1/n\}) > 0$ for all $n \in \mathbb{N}$. Let $h_n = 1_{\{|f - \lambda| < 1/n\}}$. We have $h_n \in L^\infty(X)$, $\|h_n\| = 1$ for all n, and $\|(f - \lambda)h_n\| \leq 1/n \to 0$. This shows that $f - \lambda$ is not invertible: if we had $g \in L^\infty(X)$ with $g(f - \lambda) = 1$, then

$$1 = \|h_n\| = \|g(f - \lambda)h_n\| \leq \|g\| \, \|(f - \lambda)h_n\| \leq \frac{\|g\|}{n}$$

for all $n \in \mathbb{N}$, a contradiction. Thus $\lambda \in \sigma(f)$.

(9.2.5) If T is compact Hausdorff and $\mathcal{A} = C(T)$, show that for any $f \in \mathcal{A}$ we have

$$\sigma(f) = f(T).$$

Answer. If $\lambda \notin \sigma(f)$, then there exists $g \in C(T)$ with $g(f-\lambda) = 1$. For any $t \in T$,

$$|f(t) - \lambda| = \frac{1}{|g(t)|} \geq \frac{1}{\|g\|}.$$

So $\lambda \notin f(T)$.

Conversely, suppose that $\lambda \notin f(T)$. As $f(T)$ is compact, so closed, there exists $\delta > 0$ such that $f(t) - \lambda \geq \delta$ for all $t \in T$. Then $g = 1/(f-\lambda) \in C(T)$ is an inverse for $f - \lambda$, and so $\lambda \notin \sigma(f)$.

(9.2.6) Let T be a locally compact Hausdorff space, and $\mathcal{A} = C_0(T)$. Show that the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$ is $C(T_\infty)$, the continuous functions on the one-point compactification of T.

Answer. By the uniqueness of the minimal unitization (Proposition 9.2.21) it is enough to show that $C(T_\infty)$ satisfies that there exists an isometric monomorphism $\pi : C_0(T) \to C(T_\infty)$ such that $\pi(C_0(T))$ is a maximal ideal.

Let $\pi : C_0(T) \to C(T_\infty)$ be given by $\pi(f)(t) = f(t)$ for $t \in T$ and $\pi(f)(\infty) = 0$. The fact that $f \in C_0(T)$ makes $\pi(f)$ continuous on $C(T_\infty)$. Because π is defined in terms of pointwise evaluation, it is a homomorphism. And

$$\|\pi(f)\| = \max\{0, \{|f(t)| : t \in T\}\} = \max\{|f(t)| : t \in T\} = \|f\|,$$

so π is isometric. Since $\pi(\mathcal{A}) = \{g \in C(T_\infty) : g(\infty) = 0\}$, $\pi(C_0(T))$ is an ideal. And for any $g \in C(T_\infty)$ we have $g - g(\infty)1 \in \pi(C_0(T))$, so $g \in g(\infty)1 + \pi(C_0(T))$; hence the codimension of $\pi(C_0(T))$ is 1.

(9.2.7) If T is locally compact Hausdorff and $f \in C_0(T)$, show that

$$\sigma(f) = \overline{f(T)} = \{0\} \cup f(T).$$

Do this is two ways: directly, and by using characters.

Answer. Fix $\varepsilon > 0$. There exists $K \subset T$, compact, such that $|f| < \varepsilon$ on $T \setminus K$. As $f(K)$ is compact, we have that $\overline{f(T)} = \overline{f(K)} \cup \overline{f(T \setminus K)} \subset f(K) \cup \overline{B_\varepsilon(0)} \subset f(T) \cup \overline{B_\varepsilon(0)}$. Then

$$\overline{f(T)} \subset \bigcap_{\varepsilon>0} f(T) \cup \overline{B_\varepsilon(0)} = f(T) \cup \bigcap_{\varepsilon>0} \overline{B_\varepsilon(0)} = f(T) \cup \{0\}.$$

We also have $0 \in \overline{f(T)}$, so $\overline{f(T)} = f(T) \cup \{0\}$.

Now for the spectrum. By definition of the spectrum we need to work on the unitization of $C_0(T)$, which is $C(T_\infty)$ by Exercise 9.2.6. So we consider $f \in C(T_\infty)$ with $f(\infty) = 0$.

Using characters: by Proposition 7.4.7 and Proposition 9.2.24 we have

$$\sigma(f) = \{f(t) : \ t \in T\} \cup \{0\} = f(T) \cup \{0\} = \overline{f(T)}.$$

Directly: Exercise 9.2.5 gives us

$$\sigma(f) = f(T_\infty) = f(\infty) \cup f(T) = \{0\} \cup f(T).$$

This also gives the equality $\{0\} \cup f(T) = \overline{f(T)}$.

(9.2.8) Let R be a unital ring, and $a_1, \dots, a_m \in R$. Show that if $a_k a_j = a_j a_k$ for all j, k then $a_1 \cdots a_m$ is invertible if and only if each a_j is invertible.

Answer. Suppose that $a_1 \cdots a_m$ is invertible. Because we can commute the elements, it is enough to show that a_1 is invertible. Let b be an inverse to $a_1 \dots a_m$. So $ba_1 \dots a_m = 1$, which we may rewrite as $(ba_2 \dots a_m)a_1 = 1$, so a_1 has a left inverse. Also $a_1 \dots a_m b = 1$, which shows that a_1 has a right inverse. Then a_1 is invertible, as the existence of a left inverse c and a right inverse d guarantee that they are equal and thus an inverse to a_1: if $ca_1 = 1$, $a_1 d = 1$, then

$$c = c1 = ca_1 d = 1d = d.$$

For the converse, if $a_1, \dots, a_m$ are all invertible, then $a_m^{-1} \cdots a_1^{-1}$ is an inverse for $a_1 \cdots a_m$.

(9.2.9) Prove Proposition 9.2.2.

Answer. Assume first that $a = 1$. As $\|1 - b\| < 1$, the series $\sum_{k=0}^\infty (1-b)^k$ converges: indeed, the tails satisfy

$$\Big\| \sum_{k=m}^{n} (1-b)^k] \Big\| \leq \sum_{k=m}^{n} \|1-b\|^k < \frac{\|1-b\|^m}{1 - \|1-b\|}.$$

Then, with $c = \sum_{k=0}^\infty (1-b)^k$,

$$c(1-b) = (1-b)c = \sum_{k=1}^{\infty} (1-b)^k = c - 1,$$

showing that $cb = bc = 1$, and so b is invertible with $b^{-1} = c$.

For arbitrary a,
$$\|1 - a^{-1}b\| = \|a^{-1}(a-b)\| \leq \|a^{-1}\|\,\|a-b\| < 1.$$
By the first part of the answer, $a^{-1}b$ is invertible, so $b = a(a^{-1}b)$ is invertible.

(9.2.10) Complete the proof of Proposition 9.2.23 by showing that if $J \subset \mathcal{A}$ is maximal and $\tau(a)$ is the unique scalar such that $a + J = \tau(a)I_{\mathcal{A}} + J$, then τ is a character.

Answer. From $a+b+J = \tau(a+b)I_{\mathcal{A}} + J$ and $a+b+J = a+J+b+J = (\tau(a)+\tau(b))I_{\mathcal{A}} + J$, the uniqueness gives $\tau(a+b) = \tau(a)+\tau(b)$. Similarly,
$$\begin{aligned}\tau(ab)I_{\mathcal{A}} + J = ab + J &= (a+J)(b+J)\\ &= (\tau(a)I_{\mathcal{A}} + J)(\tau(b)I_{\mathcal{A}} + J)\\ &= \tau(a)\tau(b)I_{\mathcal{A}} + J,\end{aligned}$$
and then the uniqueness gives $\tau(ab) = \tau(a)\tau(b)$.

(9.2.11) Use Lemma 9.2.20 and Proposition 7.4.13 to give an alternative proof of Proposition 7.4.6.

Answer. Let $\varphi : C(T) \to \mathbb{C}$ be a character. We know from Lemma 9.2.20 that $\ker\varphi$ is maximal. In particular, it is closed. By Proposition 7.4.13 there exists a closed $T_0 \subset T$ such that $\ker\varphi = \{f : \ f|_{T_0} = 0\}$. Because $\ker\varphi$ is maximal, necessarily T_0 is a singleton, for otherwise we can remove a point a get a larger ideal. Thus $T_0 = \{t_0\}$ for some $t_0 \in T$. As $f - \varphi(f) \in \ker\varphi$, we get that $f = \varphi(f) + h$, where $h \in \ker\varphi$. This means that $h(t_0) = 0$, and so $f(t_0) = \varphi(f) + h(t_0) = \varphi(f)$.

(9.2.12) Consider the Banach algebra $\mathcal{A} = \ell^\infty(\mathbb{N})$ and $K \subset \mathbb{C}$ compact. Show that there exists $a \in \mathcal{A}$ with $\sigma(a) = K$.

Answer. Let $Q = \{q_n\}$ be dense in K (use Exercise 7.4.5, or the fact that $\mathbb{C}$ is separable and metric together with Proposition 1.8.5). Let $a \in A$ be given by $a_n = q_n$. By Exercise 9.2.3,
$$\sigma(a) = \overline{Q} = K.$$

(9.2.13) Show that $\Sigma(M_n(\mathbb{C})) = \varnothing$.

Answer. Let $\tau : M_n(\mathbb{C}) \to \mathbb{C}$ be linear and multiplicative. Consider the canonical matrix units $\{E_{kj}\}$. We have, for $k \neq j$,

$$\tau(E_{kj}) = \tau(E_{kk}E_{kj}E_{jj}) = \tau(E_{jj}E_{kk}E_{kj}) = \tau(0) = 0.$$

And

$$\tau(E_{kk}) = \tau(E_{kj}E_{jk}) = \tau(E_{kj})\tau(E_{jk}) = \tau(0)\tau(0) = 0.$$

So the only linear and multiplicative map $M_n(\mathbb{C}) \to \mathbb{C}$ is the zero map.

(9.2.14) Let $\mathcal{A}, \mathcal{B}$ be unital Banach algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a unital homomorphism. Show that, for any $a \in \mathcal{A}$, $\sigma(\pi(a)) \subset \sigma(a)$.

Answer. Let $\lambda \in \mathbb{C} \setminus \sigma(a)$; then there exists $b \in \mathcal{A}$ with $b(a - \lambda I_{\mathcal{A}}) = I_{\mathcal{A}}$. Thus

$$\pi(b)(\pi(a) - \lambda I_{\mathcal{A}}) = \pi(b(a - \lambda I_{\mathcal{A}})) = I_{\mathcal{A}},$$

so $\lambda \in \mathbb{C} \setminus \sigma(\pi(a))$. Thus $\mathbb{C} \setminus \sigma(a) \subset \mathbb{C} \setminus \sigma(\pi(a))$, which is the same as saying that $\sigma(\pi(a)) \subset \sigma(a)$.

(9.2.15) Consider the Banach algebra $\mathcal{A} = M_n(\mathbb{C})$. Fix $\varepsilon > 0$ and define

$$A_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad B_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \varepsilon & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Show that $\sigma(A_n) = \{0\}$, $\sigma(B_n) = \{\omega : \ \omega^n = \varepsilon\}$, and $\|A_n - B_n\| = \varepsilon$.

Answer. The matrix A_n is upper triangular, so $\det(A_n - \lambda I_n) = (-1)^n \lambda^n$ and hence $\sigma(A_n) = \{0\}$. For B_n, calculating $\det(B_n - \lambda I_n)$ along the first column we get

$$\det B_n = (-1)^n \lambda^n - (-1)^n \varepsilon.$$

Thus the characteristic polynomial of B_n is $\lambda^n - \varepsilon$, and so $\sigma(B) = \{\omega : \ \omega^n = \varepsilon\}$. As for the norm, $A_n - B_n = \varepsilon\, E_{n1}$, and so $\|A_n - B_n\| = \varepsilon\, \|E_{n1}\| = \varepsilon$.

An alternative way to find the spectrum of B_n is to look at the eigenvectors. If $B_n x = \lambda x$, this gives the equalities

$$x_2 = \lambda x_1, \quad x_3 = \lambda x_2, \quad \cdots \quad x_n = \lambda x_{n-1}, \quad \varepsilon x_1 = \lambda x_n.$$

We cannot have $\lambda = 0$, for it would force $x_k = 0$ for all k. Similarly, we cannot have $x_1 = 0$, for it would propagate to $x_k = 0$ for all k. So we may assume without loss of generality that $x_1 = 1$, and then $x_k = \lambda^{k-1}$ and $\lambda^n = \varepsilon$.

(9.2.16) Let $\varepsilon > 0$. Use Exercise 9.2.15 to construct operators

$$A, B \in \mathcal{B}\Big(\bigoplus_{n=2}^{\infty} \mathbb{C}^n\Big)$$

such that

$$\|A - B\| = \varepsilon, \qquad \sigma(A) = \{0\}, \qquad \sigma(B) = \mathbb{T}.$$

Answer.
We construct, acting on the Hilbert space $H = \bigoplus_{n=1}^{\infty} \mathbb{C}^n$, the operators

$$A = \bigoplus_{n=2}^{m} A_n, \qquad B = \bigoplus_{n=2}^{m} B_n.$$

Then

$$\|A - B\| = \sup_n \|A_n - B_n\| = \varepsilon,$$

$$\sigma(A) = \bigcup_n \sigma(A_n) = \{0\},$$

and

$$\sigma(B) = \bigcup_n \sigma(B_n) = \overline{\bigcup_n \{w : \ w^n = \varepsilon, \ n = 1, \ldots, m\}} = \mathbb{T}.$$

9.3. The Riesz Functional Calculus

(9.3.1) Let $\mathcal{X}$ be a vector space, $T : \mathcal{X} \to \mathcal{X}$ linear and $\lambda \in \mathbb{C}$, $x \in \mathcal{X}$ such that $Tx = \lambda x$. If $(zI - T)$ is invertible, show that $(zI - T)^{-1}x = (z - \lambda)^{-1}x$.

Answer. Note that $zI - T$ invertible implies that $z \neq \lambda$; for $\ker(\lambda I - T) \neq \{0\}$. We have $(zI - T)x = zx - Tx = zx - \lambda x = (z - \lambda)x$. Applying $(zI - T)^{-1}$

to both sides and multiplying both sides by $(z-\lambda)^{-1}$ we get by linearity

$$(zI-T)^{-1}x=(z-\lambda)^{-1}x.$$

(9.3.2) Let $\mathcal{A}$ be a Banach algebra and $a\in\mathcal{A}$. One can define the exponential $\exp(a)$ by functional calculus,

$$\exp(a)=\frac{1}{2\pi i}\int_\gamma e^z\,(zI-a)^{-1}\,dz$$

for some curve γ that contains $\sigma(a)$. One can also define the exponential via the usual series

$$e^a=\sum_{k=0}^\infty\frac{a^k}{k!}.$$

Show that the series makes sense in $\mathcal{A}$, and that $\exp(a)=e^a$.

Answer. We have

$$\left\|\sum_{k=n}^m\frac{a^k}{k!}\right\|\leq\sum_{k=n}^m\frac{\|a^k\|}{k!}\leq\sum_{k=n}^m\frac{\|a\|^k}{k!}.$$

This last sum is a tail for the series of the usual exponential, so the partial sums for e^a are Cauchy in $\mathcal{A}$. As $\mathcal{A}$ is complete, the limit e^a exists.

Let $p_n(z)=\sum_{k=0}^n\frac{z^k}{k!}$. On any bounded set $p_n(z)\to e^z$ uniformly. As the functional calculus is a continuous homomorphism,

$$\begin{aligned}\exp(a)&=\frac{1}{2\pi i}\int_\gamma\lim_{n\to\infty}p_n(z)\,(zI-a)^{-1}\,dz\\&=\lim_{n\to\infty}\frac{1}{2\pi i}\int_\gamma p_n(z)\,(zI-a)^{-1}\,dz\\&=\lim_{n\to\infty}p_n(a)=e^a.\end{aligned}$$

(9.3.3) Let $\mathcal{A}$ be a Banach algebra and $a,b\in\mathcal{A}$.

(i) Show that if a,b commute (that is, $ab=ba$) then $e^{a+b}=e^ae^b$.

(ii) Show an example where $e^{a+b}\neq e^ae^b$.

Answer.

(i) When $ab = ba$, the proof of $e^{a+b} = e^a e^b$ runs exactly like the numerical case (Exercise 1.5.6).

(ii) Let $\mathcal{A} = M_2(\mathbb{C})$, and

$$a = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

From $a^2 = a$,

$$e^a = I_2 + \sum_{k=1}^{\infty} \frac{a}{k!} = I_2 + (e-1)a = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}.$$

Similarly $(a+b)^2 = a+b$, so

$$e^{a+b} = I_2 + \sum_{k=1}^{\infty} \frac{a+b}{k!} = I_2 + (e-1)(a+b) = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}.$$

And $b^2 = 0$, so

$$e^b = I_2 + b = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Thus

$$e^{a+b} = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix} \neq \begin{bmatrix} e & e \\ 0 & 0 \end{bmatrix} = e^a e^b.$$

(9.3.4) Prove Theorem 9.3.5.

Answer. By construction, $p_T(\lambda) = 0$ for all $\lambda \in \sigma(T)$. Then, for a simple curve γ that surrounds $\sigma(T)$,

$$p_T(T) = \frac{1}{2\pi i} \int_\gamma p_T(z)\,(zI_{\mathcal{A}} - T)^{-1}\,dz = 0$$

by Proposition 9.3.6, since p_T agrees with 0 on $\sigma(T)$.

(9.3.5) Let $\mathcal{X}$ be a Banach space, $S, T \in \mathcal{B}(\mathcal{X})$, and $f \in H(T)$. Show that if $ST = TS$, then $Sf(T) = f(T)S$.

Answer. From $(zI - T)S = S(zI - T)$, by multiplying both left and right with $(zI-T)^{-1}$ we get that $S(zI-T)^{-1} = (zI-T)^{-1}S$. Then, since multiplication by an operator is continuous, we have

$$Sf(T) = \frac{1}{2\pi i}\int_\gamma f(z)\,S(zI-T)^{-1}\,dz$$
$$= \frac{1}{2\pi i}\int_\gamma f(z)\,(zI-T)^{-1}S\,dz = f(T)S.$$

Another way to prove the assertion is to note that as $f(z) = \lim_n p_n(z)$ uniformly for polynomials p_n, we have that $f(T) = \lim_n p_n(T)$, and then

$$Sf(T) = \lim_n Sp_n(T) = \lim_n p_n(T)S = f(T)S,$$

where the commutation $Sp_n(T) = p_n(T)S$ is straightforward from commuting S with T repeatedly.

(9.3.6) Let $T = \begin{bmatrix}1&0&3\\2&1&2\\0&0&2\end{bmatrix}$. Find matrices A and B such that $A^4 = T^3$, and $T = e^B$. Are they unique? *(This exercise is not really related to Riesz Functional Calculus)*

Answer. The characteristic polynomial is $p_T(\lambda) = (\lambda-1)^2(\lambda-2)$. We have that, seeing it in block form,

$$T = \begin{bmatrix}X & y\\0 & 2\end{bmatrix}.$$

If we put

$$V = \begin{bmatrix}I & z\\0 & 1\end{bmatrix}, \qquad W = \begin{bmatrix}I & w\\0 & 1\end{bmatrix}$$

then $VW = \begin{bmatrix}I & z+w\\0 & 1\end{bmatrix}$. So

$$VTV^{-1} = \begin{bmatrix}X & y+(2-X)z\\0 & 2\end{bmatrix}$$

We want $y+(2-X)z = 0$. As $(2-X)^{-1} = \begin{bmatrix}1&0\\2&1\end{bmatrix}$, we get that

$$z = -(2-X)^{-1}y = \begin{bmatrix}-3\\-8\end{bmatrix}.$$

Then

$$VTV^{-1} = \begin{bmatrix}X & 0\\0 & 2\end{bmatrix}.$$

This allows us to answer the question by answering the questions for X and for 2, by assuming—before conjugating back with V^{-1} and V—that

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & a_2 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 & 0 \\ 0 & b_2 \end{bmatrix},$$

so

$$A^4 = \begin{bmatrix} A_1^4 & 0 \\ 0 & a_2^4 \end{bmatrix}, \qquad e^B = \begin{bmatrix} e^{B_1} & 0 \\ 0 & e^{b_2} \end{bmatrix},$$

We need a_2 to be a fourth root of 8, and $b_2 = \log 2$. The characteristic polynomial of X is $p_X(\lambda) = (\lambda - 1)^2$. So if $e^{B_1} = X$, this means that for each eigenvalue μ_1, μ_2 of B_1 we have by the Spectral Mapping Theorem that $e^{\mu_j} = 1$. So both eigenvalues of B_1 have to be zero. This means (thinking of the Jordan form) that $B_1 = SE_{12}S^{-1}$ for some invertible matrix S. Then $X = e^{B_1} = Se^{E_{12}}S^{-1}$. And since $E_{12}^2 = 0$,

$$X = e^{B_1} = Se^{E_{12}}S^{-1} = S(I + E_{12})S^{-1} = I + SE_{12}S^{-1} = I + B_1.$$

So we need $B_1 = X - I = 2E_{21}$. That is,

$$B_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}.$$

This means that the only possibility for B_0 such that $e^{B_0} = VTV^{-1}$ is

$$B_0 = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & \log 2 \end{bmatrix}$$

and so the only possibility for B such that $e^B = T$ is

$$\begin{aligned} B = V^{-1}B_0V &= \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & \log 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -8 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 3\log 2 \\ 2 & 0 & 8\log 2 - 6 \\ 0 & 0 & \log 2 \end{bmatrix}. \end{aligned}$$

For the equation $A_1^4 = X^3$ if $A_1^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then the equality $(A_1^2)^2 = X^3$ is

$$\begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 6 & 1 \end{bmatrix}.$$

From $c(a + d) = 6$ we know that $a + d \neq 0$; then $b(a + d) = 0$ implies $b = 0$. Thus $a^2 = d^2 = 1$, and as $a + d \neq 0$, $a = d = \pm 1$. Then $c = \pm 3$, depending on the sign of a. So we get two possibilities for A_1^2, namely

$$A_1^2 = \pm \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}.$$

Repeating the argument for the square root of this we get

$$A_1 = \pm \begin{bmatrix} 1 & 0 \\ 3/2 & 1 \end{bmatrix}.$$

Then A_0 such that $A_0^4 = (VTV^{-1})^3 = VT^3V^{-1}$ is

$$A_0 = \begin{bmatrix} a & 0 & 0 \\ 3a/2 & a & 0 \\ 0 & 0 & 2^{1/4}\omega \end{bmatrix}, \qquad a \in \{-1, 1\},\ \omega \in \{1, -1, i, -i\}.$$

And then

$$\begin{aligned} A = V^{-1}A_0V &= \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 3a/2 & a & 0 \\ 0 & 0 & 2^{1/4}\omega \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -8 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 2^{1/4} \times 3\omega - 3a \\ 3a/2 & a & -27a/2 + 2^{33/4}\omega \\ 0 & 0 & 2^{1/4}\omega \end{bmatrix}, \end{aligned}$$

for $a \in \{-1, 1\}$ and $\omega \in \{1, -1, i, -i\}$. So there are eight possibilities for A.

(9.3.7) Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$ such that $\sigma(T)$ is not connected.

(i) Show that $\sigma(T) = K_1 \cup K_2$, with K_1, K_2 compact and disjoint.

(ii) Show that there exist closed subspaces $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ such that $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$ and $S\mathcal{X}_1 \subset \mathcal{X}_1$ and $S\mathcal{X}_2 \subset \mathcal{X}_2$ for all $S \in \mathcal{B}(\mathcal{X})$ such that $ST = TS$.

(iii) Denote $T_1 = T|_{\mathcal{X}_1} \in \mathcal{B}(\mathcal{X}_1)$ and $T_2 = T|_{\mathcal{X}_2} \in \mathcal{B}(\mathcal{X}_2)$. Show that $\sigma(T_1) = K_1$ and $\sigma(T_2) = K_2$.

(iv) Prove that there exists $S : \mathcal{X} \to \mathcal{X}_1 \oplus \mathcal{X}_2$, invertible, with $T = S^{-1}(T_1 \oplus T_2)S$.

Answer.

(i) By assumption there exist disjoint open sets $V, W \subset \mathbb{C}$ such that $\sigma(T) = (V \cap \sigma(T)) \cup (W \cap \sigma(T))$ with both components nonempty. Let $K_1 = V \cap \sigma(T)$, $K_2 = W \cap \sigma(T)$. Let $\{V_j\}$ be an open cover for K_1. Then W and the $\{V_j\}$ form an open cover for $\sigma(T)$. So there exist $j_1, \ldots, j_m$ such that $K_1 \subset V_{j_1} \cup \cdots \cup V_{j_m}$; it follows that K_1 is compact. Similarly, K_2 is compact.

(ii) The function 1_V is holomorphic on the open set $V \cup W$. As both components are nonempty, 1_V takes both values 0 and 1 on $\sigma(a)$. Then $1_V(T)$ is a projection with $\sigma(1_V(T)) = 1_V(\sigma(T)) = \{0,1\}$, so it is proper. Similarly, $1_W(T)$ is a proper projection. As $1_V + 1_W = 1$ on $V \cup W$, we have that $1_V(T) + 1_W(T) = I_{\mathcal{X}}$ (this is contained in (ii) in Theorem 9.3.2). Now let $\mathcal{X}_1 = 1_V(T)\mathcal{X}$, $\mathcal{X}_2 = 1_W(T)\mathcal{X}$ and we have $\mathcal{X}_1 + \mathcal{X}_2 = \mathcal{X}$, as any $x \in \mathcal{X}$ can be written as $x = 1_V(T)x + 1_W(T)x$. The sum is direct, for if $x \in \mathcal{X}_1 \cap \mathcal{X}_2$, then

$$x = 1_V(T)1_W(T)x = (1_V 1_W)(T)x = (0)(T)x = 0.$$

We have that $\mathcal{X}_1$ is closed, for if $\{x_n\} \subset \mathcal{X}_1$ is Cauchy, then $x_n \to x$ for some $x \in \mathcal{X}$ since $\mathcal{X}$ is Banach; and since $1_V(T) \in \mathcal{B}(\mathcal{X})$ is bounded, $1_V(T)x = \lim_n 1_V(T)x_n = \lim_n x_n = x$ and hence $x \in \mathcal{X}_1$. Similarly, $\mathcal{X}_2$ is closed. From Proposition 6.3.9 we get that this sum is topological. If $ST = TS$, then $S1_V(T) = 1_V(T)S$ by Exercise 9.3.5. For any $x \in \mathcal{X}_1$,

$$Sx = S1_V(T)x = 1_V(T)Sx \in \mathcal{X}_1.$$

So $S\mathcal{X}_1 \subset \mathcal{X}_1$. Similarly, $S\mathcal{X}_2 \subset \mathcal{X}_2$.

(iii) Let $w \in \mathbb{C} \setminus K_1$. Then the function $g(z) = w - z$ is nowhere zero on K_1, and by compactness there exists $c > 0$ with $|g(z)| \geq c$ on K_1. Then $g(z) \neq 0$ on an open set that contains K_1, which guarantees that $1/g \in H(T_1)$. Then $(1/g)(T) \in \mathcal{B}(\mathcal{X}_1)$ is the inverse of $g(T) = wI - T$ and so $w \notin \sigma(T_1)$. Conversely, suppose that $w \in K_1$. As $w \notin K_2$, the operator $wI_{\mathcal{X}_2} - T|_{\mathcal{X}_2}$ is invertible by the argument we just did. If we had that $wI_{\mathcal{X}_1} - T|_{\mathcal{X}_1}$ is invertible, then $wI - T$ would be invertible by Exercise 6.3.10, a contradiction since $w \in \sigma(T)$. Thus $w \in \sigma(T_1)$. An analog argument shows that $\sigma(T_2) = K_2$.

(iv) Let $S : \mathcal{X} \to \mathcal{X}_1 \oplus_T \mathcal{X}_2$ be given by $Sx = (1_V(T)x, 1_W(T)x)$. The operator S is bounded, for

$$\|Sx\| = \|1_V(T)x\| + \|1_W(T)x\| \leq (\|1_V(T)\| + \|1_W(T)\|)\,\|x\|.$$

It is injective, for $Sx = 0$ implies that $1_V(T)x = 1_W(T)x = 0$, and then $x = 1_V(T)x + 1_W(T)x = 0$. And it is surjective, for if $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$, then $(x_1, x_2) = S(x_1 + x_2)$. By Theorem 6.3.6, S is invertible. We have

$$\begin{aligned} STx &= (1_V(T)Tx, 1_W(T)Tx) = (T1_V(T)x, T1_W(T)x) \\ &= (T_1 1_V(T)x, T_2 1_W(T)x) = (T_1 \oplus T_2)Sx. \end{aligned}$$

Thus $ST = (T_1 \oplus T_2)S$. As S is invertible, $T = S^{-1}(T_1 \oplus T_2)S$.

(9.3.8) Let $\mathcal{A}$ be the Banach algebra $C(X)$ for compact Hausdorff X. Let $f \in \mathcal{A}$ and $g \in H(f)$. Show that $g(f) = g \circ f$.

Answer. Since g is holomorphic, it is a uniform limit of polynomials. As the holomorphic functional calculus is continuous, it is enough to show that $p(f) = p \circ f$ for any $p \in \mathbb{C}[z]$. And by linearity on both sides it is enough to assume that $p(z) = z^k$ for some k. Now $(p \circ f)(x) = f(x)^k = p(f)(x)$. Thus $p \circ f = p(f)$.

(9.3.9) Prove Proposition 9.3.8.

Answer. The matrices T_j are exactly what comes out of applying an r-fold version of Exercise 9.3.7. The direct sum $T_1 \oplus \cdots \oplus T_r$ can be seen as blocks in the diagonal of T. We can form $S = S_1 \oplus \cdots \oplus S_r$.

For each j, $N_j = T_j - \lambda_j I_{n_j}$ has spectrum $\{0\}$. This means that the characteristic polynomial is $p_{n_j}(z) = z^{n_j}$. Then $N_j^{n_j} = p_{N_j}(N_j) = 0$ by Cayley–Hamilton (Theorem 9.3.5).

9.4. Adjoints

(9.4.1) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that if T is bounded below, then $\operatorname{ran} T$ is closed.

Answer. Let $\{y_n\}$ be a Cauchy sequence in $\operatorname{ran} T$, with $y_n \to y$ (this y exists since $\mathcal{Y}$ is Banach). For each n there exists $x_n \in \mathcal{X}$ with $y_n = Tx_n$. Since

$$\|x_n - x_m\| \leq c\|Tx_n - Tx_m\| = c\|y_n - y_m\|$$

and the sequence $\{y_n\}$ is Cauchy, it follows that the sequence $\{x_n\}$ is Cauchy. As $\mathcal{X}$ is Banach, there exists $x \in \mathcal{X}$ with $x_n \to x$. Since T is bounded, $Tx = \lim Tx_n = \lim y_n = y$. This shows that $\operatorname{ran} T$ is closed.

(9.4.2) Let $T \in \mathcal{B}(\mathcal{X})$. Show that the map $T \longmapsto T^*$ is linear and anti-multiplicative.

Answer. We have, for $g \in \mathcal{Y}^*$ and $x \in \mathcal{X}$,

$$[(T + \alpha S)^* g](x) = g((T + \alpha S)x) = g(Tx) + \alpha g(Sx) = ((T^* + \alpha S^*)g)(x).$$

As this holds for all x and all g, we get $(T + \alpha S)^* = T^* + \alpha S^*$.

If $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, then for $g \in \mathcal{Z}^*$

$$[(ST)^* g](x) = g(STx) = (S^* g)(Tx) = (T^* S^* g)(x).$$

So $(ST)^* = T^* S^*$.

(9.4.3) Given $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, show that its adjoint T^* is bounded, and that $\|T^*\| = \|T\|$.

Answer. Given $g \in \mathcal{Y}^*$, we have

$$\begin{aligned} \|T^* g\| &= \sup\{|T^* g(x)| : \ x \in \mathcal{X}, \ \|x\| = 1\} \\ &= \sup\{|g(Tx)| : \ x \in \mathcal{X}, \ \|x\| = 1\} \\ &\leq \|T\| \, \|g\|, \end{aligned}$$

so T^* is bounded and $\|T^*\| \leq \|T\|$.

Given $x \in \mathcal{X}$, we have (using Corollary 5.7.7)

$$\begin{aligned} \|Tx\| &= \max\{|g(Tx)| : \ g \in \mathcal{Y}^*, \ \|g\| = 1\} \\ &= \max\{|T^* g(x)| : \ g \in \mathcal{Y}^*, \ \|g\| = 1\} \\ &\leq \|T^*\| \, \|x\|, \end{aligned}$$

so $\|T\| \leq \|T^*\|$.

(9.4.4) Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be an isometric isomorphism. Show that T^* is an isometric isomorphism.

Answer. The isomorphism part follows directly from Corollary 9.4.9. Also, for $g \in \mathcal{Y}^*$,

$$\begin{aligned}\|T^*g\| &= \sup\{|T^*g(x)| : x \in \mathcal{X}, \|x\| = 1\}\\ &= \sup\{|g(Tx)| : x \in \mathcal{X}, \|x\| = 1\}\\ &= \sup\left\{\frac{|g(Tx)|}{\|Tx\|} : Tx \neq 0\right\}\\ &= \sup\{|g(y)| : y \in \mathcal{Y}, \|y\| = 1\} = \|g\|\end{aligned}$$

and T^* is isometric.

(9.4.5) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that

$$\|T^{**}\hat{x}\| = \|Tx\|, \qquad x \in \mathcal{X}.$$

Answer. Using just definitions,

$$\begin{aligned}\|T^{**}\hat{x}\| &= \sup\{|(T^{**}\hat{x})g| : g \in \mathcal{Y}^*, \|g\| = 1\}\\ &= \sup\{|\hat{x}(T^*g)| : g \in \mathcal{Y}^*, \|g\| = 1\}\\ &= \sup\{|(T^*g)x| : g \in \mathcal{Y}^*, \|g\| = 1\}\\ &= \sup\{|(g(Tx)| : g \in \mathcal{Y}^*, \|g\| = 1\}\\ &= \|Tx\|.\end{aligned}$$

(9.4.6) Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that if $\mathcal{X}$ is reflexive, then $T^{**} = T$.

Answer. Since $\mathcal{X}$ is reflexive, any element of $\mathcal{X}^{**}$ is of the form $\hat{x}$, with $x \in \mathcal{X}$. Then

$$(T^{**}\hat{x})g = \hat{x}(T^*g) = T^*g(x) = g(Tx) = \hat{Tx}\, g.$$

Thus $T^{**} = T$.

(9.4.7) Use Proposition 9.4.2 to give an alternative proof of Proposition 7.2.10.

Answer. If $\Gamma \in (\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))$, this means that $\Gamma : \mathcal{X}^* \to \mathbb{C}$ is $\sigma(\mathcal{X}^*, \mathcal{X})$-continuous. By Proposition 9.4.2, there exists $S : \mathbb{C}^* \to \mathcal{X}$ such that $\Gamma = S^*$.

We have $\mathbb{C}^* = \mathbb{C}$. Let $x = S1$. Then

$$\Gamma(\psi)\lambda = S^*(\psi)\lambda = \psi(\lambda x) = \psi(x)\lambda, \qquad \psi \in \mathcal{X}^*,\ \lambda \in \mathbb{C},$$

and so $\Gamma = \hat{x}$.

(9.4.8) In the context of the proof of Proposition 9.4.10, show that

$$\|g_{x,y}\| = \|x\|\,\|y\|.$$

Answer. We have

$$|g_{x,y}(T)| = |(Tx)y| \leq \|Tx\|\,\|y\| \leq \|T\|\,\|x\|\,\|y\|,$$

so $\|g_{x,y}\| \leq \|x\|\,\|y\|$. Conversely, fix $\varepsilon > 0$ and choose $f \in \mathcal{Y}^*$ with $\|f\| = 1$ such that $\|y\| \leq |f(y)| + \varepsilon$; and let T be the rank-one operator that maps x to $f \in \mathcal{Y}^*$. Then $\|T\| = \|x\|$ and

$$|g_{x,y}(T)| = |(Tx)y| = |f(y)| \geq \|y\| - \varepsilon.$$

(9.4.9) Prove (v) and (vi) in Proposition 9.4.6 without using polars nor prepolars.

Answer. (v) Suppose first that T is injective. If $\operatorname{ran} T^*$ is not weak*-dense, take $g \in \mathcal{X}^* \setminus \overline{\operatorname{ran} T^*}^{w^*}$; by Proposition 7.2.10 and Corollary 5.7.19 there exists $x_0 \in \mathcal{Y}$ such that $\hat{x_0}(g) = 1$ and $\hat{x_0}(T^*f) = 0$ for all $f \in \mathcal{Y}^*$. Then

$$0 = \hat{x_0}(T^*f) = T^*f(x_0) = f(Tx_0)$$

for all $f \in \mathcal{Y}^*$. By Corollary 5.7.7 we obtain $Tx_0 = 0$, and then $x_0 = 0$ by injectivity; a contradiction. So $\operatorname{ran} T^*$ is weak*-dense in $\mathcal{X}^*$.

Conversely, if $\operatorname{ran} T^*$ is weak*-dense in $\mathcal{X}^*$ and $Tx = 0$, then for any $f \in \mathcal{X}^*$ we have $f = \lim T^*g_j$ for some net $\{g_j\} \subset \mathcal{Y}^*$ and the limit in the weak*-topology. Then

$$f(x) = \lim_j T^*g_j(x) = \lim_j g_j(Tx) = 0.$$

As f was arbitrary, we conclude by Corollary 5.7.7 that $x = 0$, and T is injective.

(vi) Assume that T^* is injective. If $\operatorname{ran} T$ is not dense, there exists $y \in \mathcal{Y} \setminus \overline{\operatorname{ran} T}$. By Hahn–Banach (Corollary 5.7.19) there exists $g \in \mathcal{Y}^*$ with $g(y) = 1$ and $g(Tx) = 0$ for all $x \in \mathcal{X}$. But then $0 = g(Tx) = T^*g(x)$ for all x, so $T^*g = 0$. As T^* is injective, $g = 0$, a contradiction. So $\operatorname{ran} T$ is dense.

Conversely, if $\operatorname{ran} T$ is dense in $\mathcal{Y}$ and $T^*g = 0$, then $g(Tx) = 0$ for all x; as $\operatorname{ran} T$ is dense, $g(y) = 0$ for all $y \in \mathcal{Y}$, so $g = 0$ and T^* is injective.

(9.4.10) The following is a "counterexample" to Proposition 9.4.2. Find the mistake.

> *Take $\mathcal{Y} = c_{00} \subset \mathcal{X} = \ell^1(\mathbb{N})$; so we consider the 1-norm on $\mathcal{Y}$. Because $\mathcal{Y}$ is dense in $\ell^1(\mathbb{N})$, we have $\mathcal{X}^* = \mathcal{Y}^* = \ell^\infty(\mathbb{N})$.*
>
> *Define $S : \mathcal{Y}^* \to \mathcal{X}^*$, that is $S : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$, by*
> $$Sw = \Big(\sum_n \frac{w(n)}{n^2}, 0, 0, \ldots\Big).$$
> *If $w_j \to 0$ weak*, this means that $\sum_n w_j(n)x(n) \to 0$ for all $x \in \mathcal{X}$. In particular $\sum_n \frac{w_j(n)}{n^2} \to 0$, and it follows S is weak*-weak* continuous. If we had $S = T^*$, with $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ this would mean that, for each $w \in \ell^\infty(\mathbb{N})$ and $x \in \mathcal{X}$,*
> $$(Sw)x = w(Tx).$$
> *This translates to*
> $$\sum_n \frac{w(n)x(1)}{n^2} = \sum_n w(n)\,(Tx)(n).$$
> *As this should work for all $w \in \ell^\infty$, it follows that we need*
> $$Tx = \left(\frac{x(1)}{n^2}\right)_n.$$
> *But then $Tx \notin \mathcal{Y}$ for any nonzero x, and so $T \notin \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

Answer. Inspired by Remark 7.2.8.

The problem lies in the assertion that S is weak*-continuous. It is not. The sequence $\{\frac{1}{n^2}\}$ is **not** in $\mathcal{Y}$! For $k \in \mathbb{N}$, let $w_k \in \ell^\infty(\mathbb{N})$ be given by $w_k = \sum_{j \geq k} k e_k$. Then $w_k \to 0$ weak*, since any sequence in $\mathcal{Y}$ is eventually zero, but
$$\sum_n \frac{w_k(n)}{n^2} = \sum_{n \geq k} \frac{k}{n^2} \geq k \int_k^\infty \frac{1}{x^2}\,dx = 1.$$
So Sw_k does not converge weak* to zero.

(9.4.11) Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be Banach spaces, $S : \mathcal{Y} \to \mathcal{Z}$ linear, and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range. Use Lemma 9.4.5 to show that if $ST \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$ then S is bounded.

Answer. Suppose that $y_n \to 0$ and $Sy_n \to z$. Since T has closed range, applying Lemma 9.4.5 we get $c > 0$ and elements $x_n \in \mathcal{X}$ such that $Tx_n = y_n$ for each $n \in \mathbb{N}$ and

$$0 \le \|x_n\| \le c\|Tx_n\| = c\,\|y_n\|.$$

As $y_n \to 0$, this shows that $x_n \to 0$. Knowing this,

$$z = \lim_n Sy_n = \lim_n STx_n = 0,$$

the last equality by the continuity of ST. Now the Closed Graph Theorem implies that S is bounded.

(9.4.12) We use notation from Section 9.3. Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$, and $f \in H(T)$. Show that $f(T^*) = f(T)^*$.

Answer. Since the Riesz functional calculus is continuous and f is a uniform limit of polynomials, it is enough to show the equality for polynomials. As polynomials are linear combinations of monomials and taking adjoints is linear, it is enough to show the equality for a monomial. When $f(z) = z^n$, we have for any $\varphi \in \mathcal{X}^*$ and any $x \in \mathcal{X}$

$$(f(T^*)\varphi)x = ((T^*)^n\varphi)x = ((T^*)^{n-1}\varphi)(Tx)$$
$$= \cdots = \varphi(T^n x) = (f(T)^*\varphi)x.$$

Thus $f(T^*) = f(T)^*$.

9.5. The Spectrum of a Linear Operator

(9.5.1) Let $T \in \mathcal{B}(\mathcal{X})$, $\alpha, \beta \in \mathbb{C}$. Show that $\sigma(\alpha T + \beta I) = \alpha\sigma(T) + \beta$.

Answer. If $\alpha = 0$, both sides of the equality are $\{\beta\}$; so we assume $\alpha \neq 0$. As

$$\alpha T + \beta I - (\alpha\lambda + \beta)I = \alpha(T - \lambda I),$$

$T - \lambda I$ is invertible if and only if $\alpha T + \beta I - (\alpha\lambda + \beta)I$ is invertible. So $\sigma(\alpha T + \beta) = \alpha\sigma(T) + \beta$.

(9.5.2) Let $\mathcal{X}$ be a Banach space with dimension (finite or infinite) at least 2. Given $a, b \in \mathcal{X}$ linearly independent, let $\mathcal{X}_0 = \operatorname{span}\{a, b\}$. Show that

(i) there exists $\varphi \in \mathcal{X}^*$ with $\varphi(\alpha a + \beta b) = \beta$;

(ii) there exists a bounded surjective projection $P : \mathcal{X} \to \mathcal{X}_0$;

(iii) the linear operator $T : \mathcal{X} \to \mathcal{X}$ given by $Tx = \varphi(Px)a$ is bounded;

(iv) $\sigma(T) = \{0\}$;

(v) there exists $T \in \mathcal{B}(\mathcal{X})$ with $\|T\| = 1$ and $\operatorname{spr}(T) = 0$.

Answer.

(i) Define a linear functional $\varphi : \mathcal{X}_0 \to \mathbb{C}$ by $\varphi_0(\alpha a + \beta b) = \beta$. Since $\mathcal{X}_0$ is finite-dimensional, φ_0 is bounded. By Hahn–Banach, there exists $\varphi : \mathcal{X} \to \mathbb{C}$, bounded, with $\varphi|_{\mathcal{X}_0} = \varphi_0$.

(ii) Similarly, there is a bounded linear functional $\psi \in \mathcal{X}^*$ such that $\psi(\alpha a + \beta b) = \alpha$. Now define $P : \mathcal{X} \to \mathcal{X}$ by $Px = \psi(x)a + \varphi(x)b$. Then P is a bounded operator and $P|_{\mathcal{X}_0} = \operatorname{id}_{\mathcal{X}_0}$.

(iii) This is just the estimate

$$\|\varphi(Px)a\| \leq |\varphi(Px)|\,\|a\| \leq \|\varphi\|\,\|P\|\,\|a\|\,\|x\|.$$

(iv) Let $\mathcal{X}_1 = \ker P$. It is a closed subspace, since P is bounded. For any $x \in \mathcal{X}$, we have

$$x = Px + (x - Px);$$

as $x - Px \in \ker P$, $\mathcal{X} = \mathcal{X}_0 + \mathcal{X}_1$. If $x \in \mathcal{X}_0 \cap \mathcal{X}_1$, then $x = Px = 0$, so $\mathcal{X}_0 \cap \mathcal{X}_1 = \{0\}$.

(v) We use the $T \in \mathcal{B}(\mathcal{X})$ given by $Tx = \varphi(Px)a$. We have $\|Tx\| = |\varphi(Px)\,\|a\| \leq \|\varphi\|\,\|P\|\,\|a\|\,\|x\|$, so T is indeed bounded. This T does not necessarily satisfy $\|T\| = 1$, but that we can solve by dividing by $\|T\|$. So we concentrate on the spectrum.

Suppose that $\lambda \in \mathbb{C} \setminus \{0\}$. We proved that $x = \alpha a + \beta b + c$, with $c \in \ker P$. Then

$$(T - \lambda I)x = Tx - \lambda x = (\beta - \lambda\alpha)a - \lambda\beta b - \lambda c.$$

The map

$$S(ra + sb + c') = -\tfrac{1}{\lambda}(r + \tfrac{s}{\lambda})\, a - \tfrac{1}{\lambda} s\, b - \tfrac{1}{\lambda}\, c'$$

is easily seen to be an algebraic inverse for $T - \lambda I$. Since $T - \lambda I$ is bounded, by the Inverse Mapping Theorem 6.3.6 the linear map S is bounded and $T - \lambda I$ is invertible. So $\sigma(T) \subset \{0\}$; as the spectrum is always nonempty, $\sigma(T) = \{0\}$.

A more direct argument can be done with more knowledge about compact operators. Since T is compact, by Theorem 9.6.13 every nonzero element of the spectrum is an eigenvalue. Actually, since T is finite-rank, it is not hard to see it even without Theorem 9.6.13. So if $\lambda \in \sigma(T) \setminus \{0\}$, we have $\varphi(Px)a = \lambda x$ for some nonzero x. In particular $x = \alpha a$ for some $\alpha \in \mathbb{C}$. Then the equality becomes

$$\lambda\alpha a = \varphi(Px)a = \alpha\varphi(a)a = 0.$$

Thus $\alpha = 0$ and $x = 0$, showing that λ cannot be an eigenvalue. Thus $\sigma(T) = \{0\}$.

(9.5.3) Let $\mathcal{X}$ be a normed space and $T \in \mathcal{B}(\mathcal{X})$. Show that if there exists nonzero $p \in \mathbb{C}[x]$ such that $p(T) = 0$ then T admits an eigenvalue.

Answer. We may assume that $T \neq 0$, for if $T = 0$ then 0 is an eigenvalue for T.

Let $q \in \mathbb{C}$ with least degree such that $q(T) = 0$ (the existence of q is guaranteed by the existence of p). Then all roots of q are eigenvalues for T. Indeed, given λ with $q(\lambda) = 0$, we may write $q(t) = (t - \lambda)r(t)$ with $\deg r < \deg q$. The minimality of q then guarantees that $r(T) \neq 0$. We may write $q(T) = 0$ as $r(T)T = \lambda r(T)$. As $r(T) \neq 0$, choose v such that $w = r(T)v \neq 0$. Then $Tw = \lambda w$.

(9.5.4) Let $\mathcal{X}$ be a normed space and $T \in \mathcal{B}(\mathcal{X})$. Suppose that $p \in \mathbb{C}[x]$ satisfies $p(T) = 0$. Show that there exists a root of p that is an eigenvalue for T. Are all roots of p eigenvalues of T? Provide roof/counterexample.

Answer. Exercise 9.5.3 shows that some root of p is an eigenvalue of T, because the minimality of q guarantees that q divides p. Indeed, by the division algorithm we have $p(t) = q(t)s(t) + r(t)$, where $\deg r < \deg q$. Since $q(T) = p(T) = 0$ we get that $r(T) = 0$, and then the minimality of q forces $r = 0$. Then q divides p and so the root of q that was found in Exercise 9.5.3 is also a root of p.

It is possible for p to have roots that are not eigenvalues of T. For a trivial example, let $T = I_{\mathcal{X}}$. Then the polynomial $p(t) = (t-2)(t-1)$ satisfies $p(T) = 0$, but 2 is not an eigenvalue for T.

(9.5.5) Let $\mathcal{X}$ be an infinite-dimensional normed space. Show that there exists $T \in \mathcal{B}(\mathcal{X})$ such that $p(T) \neq 0$ for all $p \in \mathbb{C}[x]$.

Answer. By Exercise 9.5.3, if $p(T) = 0$ for some polynomial then T has an eigenvalue. Thus if T has no eigenvalues, it cannot be zero under any polynomial. We saw in the text that the unilateral shift S has no eigenvalues (Example 9.5.10); then $p(S) \neq 0$ for all $p \in \mathbb{C}[x]$. Another common example is $T \in \mathcal{B}(C[0,1])$ given by $(Tf)(t) = tf(t)$; again this operator has no eigenvalues (Example 9.5.7).

(9.5.6) Let $\mathcal{X}$ be a normed space and $T \in \mathcal{B}(\mathcal{X})$. Suppose that $p(T) \neq 0$ for all $p \in \mathbb{C}[x]$. Does this imply that T has no eigenvalues?

Answer. No. Consider the unilateral shift S as above, say acting on $\ell^1(\mathbb{N})$, and form $T = 1 \oplus S$, acting on $\mathbb{C} \oplus \ell^1(\mathbb{N})$. Then $p(T) = p(1) \oplus p(S)$, so it is nonzero for all p. But 1 is an eigenvalue for T.

A more drastic example is S^*. We have $p(S^*) = 0$ for all $p \in \mathbb{C}[x]$, for otherwise we would have $\bar{p}(S) = 0$. But $\sigma_p(S^*) = \mathbb{D}$.

(9.5.7) Let $\mathcal{X}$ be a Banach space and $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ subspaces with $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. Let $T \in \mathcal{B}(\mathcal{X})$ such that $T\mathcal{X}_1 \subset \mathcal{X}_1$ and $T\mathcal{X}_2 \subset \mathcal{X}_2$. Show that

$$\sigma(T) = \sigma(T|_{\mathcal{X}_1}) \cup \sigma(T|_{\mathcal{X}_2}).$$

Answer. It is enough to show that T is invertible if and only if $T_1 = T|_{\mathcal{X}_1}$ and $T_2 = T|_{\mathcal{X}_2}$ are invertible. Suppose that T is invertible; then there exists $S \in \mathcal{B}(\mathcal{X})$ with $ST = TS = I_{\mathcal{X}}$. As $T\mathcal{X}_1 + T\mathcal{X}_2 = T\mathcal{X} = \mathcal{X}$ and $T\mathcal{X}_j \subset \mathcal{X}_j$ for $j = 1, 2$, it follows from T invertible that $T\mathcal{X}_1 = \mathcal{X}_1$ and $T\mathcal{X}_2 = \mathcal{X}_2$. Indeed, if $x \in \mathcal{X}_1$ then $x = Tz_1 + Tz_2$ for $z_1 \in \mathcal{X}_1$ and $z_2 \in \mathcal{X}_2$; as $Tz_1 \in \mathcal{X}_1$ and $Tz_2 \in \mathcal{X}_2$, we get $Tz_2 = 0$ and therefore $x \in T\mathcal{X}_1$. Applying S we obtain $\mathcal{X}_1 = S\mathcal{X}_1$ and $\mathcal{X}_2 = S\mathcal{X}_2$. Then $S|_{\mathcal{X}_1}$ is an inverse for T_1 and $S|_{\mathcal{X}_2}$ is an inverse for T_2.

Conversely, if S_1, S_2 are inverses for T_1 and T_2, respectively, then $S_1 \oplus S_2$ is an inverse for T.

(9.5.8) Write the details of Example 9.5.6.

Answer. When $p < \infty$,

$$\|M_b x\|_p^p = \sum_{n=1}^{\infty} |b(n)x(n)|^p \leq \|b\|_\infty^p \sum_{n=1}^{\infty} |x(n)|^p = \|b\|_\infty^p \, \|x\|_p^p.$$

So $\|M_b\| \leq \|b\|_\infty$. Given $\varepsilon > 0$ there exists n such that $|b(n)| > \|b\|_\infty - \varepsilon$. Then

$$\|b\|_\infty - \varepsilon < |b(n)| = |(M_b e_n)(n)| \leq \|M_b e_n\|_p \leq \|M_b\|.$$

As this can be done for all $\varepsilon > 0$, we get $\|M_b\| = \|b\|_\infty$. A similar idea works when $p = \infty$.

Now we work on the spectrum. Since $M_b e_n = b(n)e_n$, $\{b(n) : \; n\} \subset \sigma_p(M_b)$ (we have two different uses of p, here p is for "point" and not the number p). If $\operatorname{dist}(\lambda, \{b(n) : \; n\}) \geq \delta > 0$, then we can form $a \in \ell^\infty(\mathbb{N})$ where

$$a(n) = \frac{1}{b(n) - \lambda}.$$

Let A be the multiplication operator induced by a. Then

$$[A(M_b - \lambda I)x](n) = \frac{1}{b(n) - \lambda}\,(b(n) - \lambda)x(n) = x(n).$$

Thus $A(M_b - \lambda I) = I$. Similarly, $(M_b - \lambda I)A = I$, showing that $\lambda \notin \sigma(M_b)$. Thus $\sigma(M_b) \subset \overline{\{b(n) : \; n\}}$, and then $\sigma(M_b) = \overline{\{b(n) : \; n\}}$ since it is closed and it contains $b(n)$ for all n. Suppose that $\lambda \in \sigma_p(M_b)$. Then there exists nonzero $x \in \ell^p(\mathbb{N})$ with $M_b x = \lambda x$. At the level of entries this looks like

$$b(n)x(n) = \lambda x(n).$$

For any n such that $x(n) \neq 0$, we have $\lambda = b(n)$, showing that $\sigma_p(M_b) \subset \{b(n) : \; n\}$ and giving us $\sigma_p(M_b) = \{b(n) : \; n\}$.

When $\lambda \in \sigma(T) \setminus \{\{b(n) : \; n\}$, we have that $\ker(M_b - \lambda I) = \{0\}$. By being in the closure of $\{b(n) : \; n\}$, there exists a subsequence $\{b_{n_k}\}$ with $b_{n_k} \xrightarrow[k]{} \lambda$. For each k we have $\|e_{n_k}\| = 1$, and

$$\|M_b T - \lambda I)e_{n_k}\| = |b_{n_k} - \lambda| \xrightarrow[k\to\infty]{} 0.$$

Thus $M_b - \lambda I$ is not bounded below, showing that $\lambda \in \sigma_{ap}(M_b)$ (so $\lambda_{ap}(M_b) = \sigma(M_b)$, as it also contains the eigenvalues), and also that $\operatorname{ran}(M_b - \lambda I)$ is not closed. For any $x = \sum_{j=1}^n c_j e_j \in c_{00}$, we have

$$x = \sum_{j=1}^n c_j e_j = (M_b - \lambda)\Big(\sum_{j=1}^n \frac{c_j}{b(n) - \lambda}\, e_j\Big).$$

Then $c_{00} \in \operatorname{ran}(M_b - \lambda I)$ and $\operatorname{ran}(M_b - \lambda I)$ is dense, implying that $\lambda \in \sigma_c(M_b)$.

(9.5.9) Show that $\sigma_p(T)$, $\sigma_r(T)$, and $\sigma_c(T)$ are mutually disjoint.

Answer. By definition, $\sigma_p(T) \cap \sigma_r(T) = \sigma_p(T) \cap \sigma_c(T) = \varnothing$. And $\sigma_r(T) \cap \sigma_c(T) = \varnothing$, since in one case $T - \lambda I$ is required to have dense range and in the other it is require not to have it.

(9.5.10) Prove the relations (9.15), (9.16), and (9.17).

Answer. If $\lambda \in \sigma_p(T)$, then there exists nonzero $x \in \ker(T - \lambda I)$. Normalizing, and since the kernel is a subspace, we may assume that $\|x\| = 1$. Now take $x_n = x$ for all n, and the definition of λ in the approximate point spectrum is satisfied. This establishes (9.15).

For (9.16), suppose that $\lambda \in \sigma_c(T)$. By definition, $\lambda \notin \sigma_p(T)$. If $\lambda \notin \sigma_{ap}(T)$, it means that $T - \lambda I$ is bounded below: there exists $c > 0$ with $\|(T - \lambda I)x\| \geq c\|x\|$ for all x. Then $T - \lambda I$ has closed range by Exercise 9.4.1; with dense and closed range then T would be surjective, a contradiction. So $\lambda \in \sigma_{ap}(T)$. By definition, $\lambda \notin \sigma_r(T)$, so $\sigma_c(T) \subset \sigma_{ap}(T) \setminus (\sigma_r(T) \cup \sigma_p(T))$.

Conversely, if $\lambda \in \sigma_{ap}(T) \setminus (\sigma_r(T) \cup \sigma_p(T))$, then $T - \lambda I$ is injective (because $\lambda \notin \sigma_p(T)$), it has dense range (because $\lambda \notin \sigma_r(T)$), and it is not bounded below (because $\lambda \in \sigma_{ap}(T)$). So $T - \lambda I$ cannot be surjective (as it would be invertible by the Inverse Mapping Theorem). So $T - \lambda I$ has dense range but it is not surjective, proving that $\lambda \in \sigma_c(T)$.

For (9.17), we have $\sigma_p(T) \cap (\sigma_r(T) \cup \sigma_c(T)) = \varnothing$ by Exercise 9.5.9. And also $\sigma_r(T) \cap \sigma_c(T) = \varnothing$ by definition, since either the range is dense or it is not. Given any $\lambda \in \sigma(T)$, either $\lambda \in \sigma_p(T)$, or $T - \lambda I$ is injective. In the latter case, as $T - \lambda I$ is not invertible, it cannot be surjective; so either the range of $T - \lambda I$ is not dense, giving $\lambda \in \sigma_r(T)$, or $T - \lambda I$ is dense but not closed, giving $\lambda \in \sigma_c(T)$. Thus

$$\sigma(T) = \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T).$$

(9.5.11) Prove the relation (9.18).

Answer. If $T - \lambda I$ is not invertible and it is not bounded below, then $\lambda \in \sigma_{ap}(T)$. If it is bounded below, then it is injective and its range is closed. As such, it cannot be dense: if it were, $T - \lambda I$ would be bijective with a bounded inverse and so it would be invertible, a contradiction. So, we have shown that $\sigma(T) \setminus \sigma_{ap}(T) \subset \sigma_r(T)$, so $\sigma(T) = \sigma_{ap}(T) \cup \sigma_r(T)$.

(9.5.12) Construct an example where $\sigma_{ap}(T) \cap \sigma_r(T) \neq \varnothing$.

Answer. Take $T \in \mathcal{B}(\mathcal{X} \oplus \mathcal{X})$ to be $T = S \oplus \lambda I_{\mathcal{X}}$, where S is chosen so that there exists $\lambda \in \sigma_{ap}(S) \setminus \sigma_p(S)$—for instance, S could be the unilateral shift. Then $T - \lambda I = (S - \lambda I) \oplus 0$ is not bounded below, and its range is not dense. Thus $\lambda \in \sigma_{ap}(T) \cap \sigma_r(T)$.

(9.5.13) In the context of Example 9.5.9 calculate explicitly $S(T - \lambda I)$ and $(T - \lambda I)S$ for $\lambda \notin [0, 1]$.

Answer. Fix $f \in C[0,1]$. We have

$$\begin{aligned}
[S(T-\lambda I)f](x) &= \frac{xf(x)}{x-\lambda} + \frac{1}{x-\lambda}\int_0^x f(s)\,ds - \frac{1}{(x-\lambda)^2}\int_0^x sf(s)\,ds \\
&\quad - \frac{1}{(x-\lambda)^2}\int_0^x\int_0^t f(s)\,ds\,dt] \\
&\quad - \frac{\lambda f(x)}{x-\lambda} + \frac{\lambda}{(x-\lambda)^2}\int_0^x f(t)\,dt \\
&= \frac{xf(x)}{x-\lambda} + \frac{1}{x-\lambda}\int_0^x f(s)\,ds - \frac{1}{(x-\lambda)^2}\int_0^x sf(s)\,ds \\
&\quad - \frac{1}{(x-\lambda)^2}\int_0^x\int_s^x f(s)\,dt\,ds \\
&\quad - \frac{\lambda f(x)}{x-\lambda} + \frac{\lambda}{(x-\lambda)^2}\int_0^x f(t)\,dt \\
&= f(x) + \frac{1}{(x-\lambda)^2}\left[\int_0^x xf(s)\,ds - sf(s)\,ds - (x-s)f(s)\,ds\right] \\
&= f(x).
\end{aligned}$$

Hence $S(T-\lambda I) = I$. Similarly, denoting $C = [(T-\lambda I)Sf](x)$

$$\begin{aligned}
C &= (T-\lambda I)\left[\frac{f(x)}{x-\lambda} - \frac{1}{(x-\lambda)^2}\int_0^x f(s)\,ds\right] \\
&= f(x) - \frac{1}{x-\lambda}\int_0^x f(t)\,dt + \int_0^x \frac{f(t)}{t-\lambda}\,dt \\
&\quad - \int_0^x \frac{1}{(s-\lambda)^2}\int_0^s f(t)\,dt\,ds \\
&= f(x) - \frac{1}{x-\lambda}\int_0^x f(t)\,dt + \int_0^x \frac{f(t)}{t-\lambda}\,dt - \int_0^x f(t)\int_t^x \frac{1}{(s-\lambda)^2}\,ds\,dt \\
&= f(x) - \frac{1}{x-\lambda}\int_0^x f(t)\,dt + \int_0^x \frac{f(t)}{t-\lambda}\,dt - \int_0^x f(t)\left[\frac{1}{t-\lambda} - \frac{1}{x-\lambda}\right]dt \\
&= f(x).
\end{aligned}$$

Thus $(T-\lambda I)S = I$.

(9.5.14) With T the left unilateral shift as in Example 9.5.10, show that

$$\dim\ker(T-\lambda I) = 1$$

for all $\lambda \in \mathbb{D}$.

Answer. Suppose that $Tx = \lambda x$. In coordinates, this is

$$(x_2, x_3, x_4, \ldots) = (\lambda x_1, \lambda x_2, \lambda x_3, \ldots).$$

If $\lambda = 0$, this gives $x_2 = x_3 = \cdots = 0$, so $\ker T = \mathbb{C}e_1$. When $\lambda \neq 0$, we get $x_2 = \lambda x_1$, $x_3 = \lambda x_2 = \lambda^2 x_1$, and in general $x_{k+1} = \lambda^k x_1$. Hence $x \in \ker(T - \lambda I)$ if and only if

$$x = x_1(\lambda, \lambda^2, \lambda^3, \ldots).$$

The condition $|\lambda| < 1$ guarantees that $x \in \ell^p(\mathbb{N})$. And we have $\ker(T - \lambda I) = \mathbb{C}(\lambda, \lambda^2, \ldots)$.

(9.5.15) Let $\mathcal{X} = L^2(-\infty, \infty)$ and let $T \in \mathcal{B}(\mathcal{X})$ the translation operator

$$(Tf)(x) = f(x+1).$$

Find the norm and the parts of the spectrum of T.

Answer. From

$$\|Tf\|_2^2 = \int_{-\infty}^{\infty} |f(x+1)|^2\,dx = \int_{-\infty}^{\infty} |f(x)|^2\,dx = \|f\|_2^2$$

we get that $\|T\| = 1$. More than that, T is an isometry. As T is invertible (with $(T^{-1}f)(x) = f(x-1)$) we have that $\sigma(T) \subset \mathbb{T}$ by Proposition 9.5.15. Let us try to find eigenvalues. If $\lambda \in \mathbb{T}$ and $Tf = \lambda f$, then we have $f(x+1) = \lambda f(x)$ a.e. This forces $f(x+k) = \lambda^k f(x)$ for all $k \in \mathbb{Z}$, and if we write $\lambda = e^{i\theta}$,

$$\|f\|_2^2 = \lim_{m\to\infty} \int_{-m}^{m} |f|^2 = \sum_{k=-m}^{m-1} \int_k^{k+1} |f|^2 = \sum_{k=-m}^{m-1} \int_0^1 |f(x+k)|^2\,dx$$

$$= \sum_{k=-m}^{m-1} \lambda^{2k} \int_0^1 |f(x)|^2\,dx = \frac{2\cos 2m\theta}{e^{i\theta} - 1} \int_0^1 |f(x)|^2\,dx,$$

But

$$\lim_{m\to\infty} \frac{2\cos 2m\theta}{e^{i\theta} - 1}$$

does not exist (the computation is different if $\theta = 0$, that is $\lambda = 1$, but in that case the integral is simply unbounded). So $\sigma_p(T) = \varnothing$. Let us now adapt the trick from Example 9.5.12. Fix $\mu \in \mathbb{T}$. Let $V_\mu \in \mathcal{B}(\mathcal{X})$ be the operator

$$V_\mu f = \sum_{n\in\mathbb{Z}} \mu^n\, f\, 1_{[n,n+1)}.$$

As $|\mu| = 1$ we get that V is a bijective isometry, i.e. a unitary. And we have

$$\begin{aligned}(V^*TV)f(x) &= V^*T\sum_{n\in\mathbb{Z}}\mu^n\, f(x)\, 1_{[n,n+1)}(x)\\ &= V^*\sum_{n\in\mathbb{Z}}\mu^n\, f(x+1)\, 1_{[n,n+1)}(x+1)\\ &= V^*\sum_{n\in\mathbb{Z}}\mu^n\, f(x+1)\, 1_{[n-1,n)}(x)\\ &= V^*\sum_{n\in\mathbb{Z}}\mu^{n+1}\, f(x+1)\, 1_{[n,n+1)}(x)\\ &= \mu(Tf)(x).\end{aligned}$$

This gives us that $\sigma(T) = \sigma(V^*TV) = \mu\sigma(T)$. So $\sigma(T)$ is invariant for rotations, and hence $\sigma(T) = \mathbb{T}$. As every point is a boundary point, $\sigma_{ac}(T) = \mathbb{T}$. Since T is a unitary then $T^{-1} = T^*$. For any $\lambda \in \mathbb{T}$,

$$\overline{\operatorname{ran}(T-\lambda I)} = \ker(T-\lambda I)^\perp = \{0\}^\perp = \mathcal{X}.$$

So $T - \lambda I$ has dense range. In summary,

$$\begin{aligned}\sigma(T) &= \mathbb{T}\\ \sigma_p(T) &= \varnothing\\ \sigma_{ap}(T) &= \mathbb{T}\\ \sigma_r(T) &= \varnothing\\ \sigma_c(T) &= \mathbb{T}\end{aligned}$$

(9.5.16) Let $\mathcal{X} = L^2(0,\infty)$ and let $T \in \mathcal{B}(\mathcal{X})$ the translation operator

$$(Tf)(x) = f(x+1).$$

Find the norm and the parts of the spectrum of T.

Answer. We can use several ideas from the answer to (9.5.15). From

$$\|Tf\|_2^2 = \int_0^\infty |f(x+1)|^2\,dx = \int_1^\infty |f(x)|^2\,dx \le \|f\|_2^2$$

we get that $\|T\| \le 1$; if we take any $f \in L^2(0,\infty)$ and supported on $[1,\infty)$, then $\|Tf\|_2 = \|f\|_2$, and so $\|T\| = 1$. As for the spectrum, initially we know that $\sigma(T) \subset \overline{\mathbb{D}}$. If we look for eigenvalues, if $\lambda \in \mathbb{D}$ and

$$f = \sum_{k=0}^\infty \lambda^k\, 1_{[k,k+1)}$$

then for $x > 0$

$$Tf(x) = \sum_{k=0}^{\infty} \lambda^k 1_{[k,k+1)}(x+1) = \sum_{k=1}^{\infty} \lambda^k 1_{[k-1,k)}(x)$$

$$= \sum_{k=0}^{\infty} \lambda^{k+1} 1_{[k,k+1)}(x) = \lambda f(x).$$

So $\lambda \in \sigma_p(T)$. Then $\mathbb{D} \subset \sigma(T) \subset \overline{\mathbb{D}}$ and hence $\sigma(T) = \overline{\mathbb{D}}$. As every boundary point is an approximate eigenvalue (Proposition 9.5.5), we have $\sigma_{ap}(T) = \overline{\mathbb{D}}$.

If $\lambda \in \mathbb{T}$ and $Tf = \lambda f$, then we have $f(x+1) = \lambda f(x)$ a.e. This forces $f(x+k) = \lambda^k f(x)$ for all $k \in \mathbb{N}$, and if we write $\lambda = e^{i\theta}$,

$$\|f\|_2^2 = \lim_{m\to\infty} \int_0^m |f|^2 = \sum_{k=0}^{m-1} \int_k^{k+1} |f|^2 = \sum_{k=0}^{m-1} \int_0^1 |f(x+k)|^2 \, dx$$

$$= \sum_{k=0}^{m-1} \lambda^{2k} \int_0^1 |f(x)|^2 \, dx = \frac{2\cos 2m\theta}{e^{i\theta}-1} \int_0^1 |f(x)|^2 \, dx,$$

so

$$\|f\|_2^2 = \lim_{m\to\infty} \int_0^m |f|^2 = \lim_{m\to\infty} \frac{2\cos 2m\theta}{e^{i\theta}-1}$$

does not exist (the computation is different if $\theta = 0$, that is $\lambda = 1$, but in that case the integral is simply unbounded). So $\sigma_p(T) = \mathbb{D}$. For any $\lambda \in \mathbb{T}$,

$$\overline{\operatorname{ran}(T-\lambda I)} = \ker(T-\lambda I)^{\perp} = \{0\}^{\perp} = \mathcal{X}.$$

So $T - \lambda I$ has dense range. In summary,

$$\sigma(T) = \overline{\mathbb{D}}$$
$$\sigma_p(T) = \mathbb{D}$$
$$\sigma_{ap}(T) = \overline{\mathbb{D}}$$
$$\sigma_r(T) = \varnothing$$
$$\sigma_c(T) = \mathbb{T}$$

(9.5.17) Let $p \in (1,\infty)$ and $T \in \mathcal{B}(L^p[0,1])$ be the Hardy operator as in Example 9.5.22. Show that $\|T\| = q$ and find $\sigma(T)$, $\sigma_p(T)$, $\sigma_r(T)$, $\sigma_c(T)$, and $\sigma_{ap}(T)$. *(This exercise is mostly computational, but nailing the right ideas and performing all the computations will possibly not be a trivial task; so this exercise and Exercise 9.5.18 should be seen more as a minor project rather than a couple of exercises)*

Answer. It was proven in Example 9.5.22 that $\|T\| \leq q$. We will see below that $\operatorname{spr}(T) = q$, and thus $\|T\| = q$, since $q = \operatorname{spr}(T) \leq \|T\| \leq q$.

Let us look for eigenvalues first. If $Tf = 0$, then $f = 0$ a.e. by Exercise 2.5.6 (or Exercise 2.11.2); so T is injective. Since T is injective, any eigenvalue will have to be nonzero. So suppose that $\lambda \in \mathbb{C} \setminus \{0\}$ and $Tf = \lambda f$ for some nonzero f. We can write, when $0 < y < x$,

$$\begin{aligned}|f(x) - f(y)| &= \left| \frac{1}{\lambda x} \int_0^x f - \frac{1}{\lambda y} \int_0^y f \right| \\ &\leq \frac{1}{|\lambda x|} \int_y^x |f| + \frac{1}{|\lambda|} \left| \frac{1}{x} - \frac{1}{y} \right| \int_0^y |f| \\ &\leq \frac{1}{|\lambda|} \left((x-y)^{1/q} + \frac{x-y}{xy}. \right) \|f\|_p\end{aligned}$$

This shows that f is continuous for all $x > 0$. Going back to $\lambda f = Tf$, now the integral is differentiable by the Fundamental Theorem of Calculus, and hence f is differentiable for all $x > 0$. We can now differentiate

$$\lambda x f(x) = \int_0^x f,$$

to get

$$\lambda f(x) + \lambda x f'(x) = f(x),$$

which we can rewrite as

$$(1 - \lambda) f(x) - \lambda x f'(x) = 0.$$

This is a first-order linear differential equation, with solution $f(x) = x^{1/\lambda - 1}$ (and multiples of it, of course). For λ to be an eigenvalue of T we need this f to be in $L^p[0,1]$. And for this we need $\operatorname{Re} p(1/\lambda - 1) > -1$. Equivalently,

$$\operatorname{Re} \frac{1}{\lambda} > 1 - \frac{1}{p} = \frac{1}{q}. \tag{AB.9.1}$$

Writing $\lambda = a + ib$, then the inequality becomes

$$\frac{a}{a^2 + b^2} > \frac{1}{q}.$$

This in turn is $a^2 + b^2 < qa$, or $(a - \frac{q}{2})^2 + b^2 < \frac{q^2}{4}$. In terms of λ, this is $|\lambda - \frac{q}{2}| < \frac{q}{2}$, or $\lambda \in B_{q/2}(q/2)$. That is,

$$\sigma_p(T) = B_{q/2}(q/2).$$

This shows that T is not compact, for its set of eigenvalues is uncountable (see Theorem 9.6.13). Now let us try to find $(T - \lambda I)^{-1}$. Necessarily, $\lambda \notin \overline{B_{q/2}(q/2)}$, which is equivalent to $\operatorname{Re} \frac{1}{\lambda} < \frac{1}{q}$; this we write as $\operatorname{Re} -\frac{1}{\lambda} + 1 >$

$\frac{1}{p} > 0$. Suppose that $g = (T - \lambda I)f$; we want to express f in terms of g. Assume initially that f and g are differentiable. We have

$$g(x) = -\lambda f(x) + \frac{1}{x}\int_0^x f,$$

which we may write as

$$xg(x) = -\lambda x\, f(x) + \int_0^x f.$$

Differentiating,

$$(xg(x))' = (1-\lambda)f(x) - \lambda x\, f'(x).$$

Using the integrating factor $x^{-1/\lambda}$,

$$x^{-1/\lambda}(xg(x))' = (1-\lambda)x^{-1/\lambda}f(x) - \lambda x^{-1/\lambda+1}\, f'(x) = [-\lambda x^{-1/\lambda+1}f(x)]'.$$

Integrating (and using that $-\mathrm{Re}\, 1/\lambda + 1 > -1/q + 1 = 1/p > 0$ to evaluate the right-hand-side at $x = 0$)

$$\int_0^x t^{-1/\lambda}[tg(t)]'\, dt = -\lambda x^{-1/\lambda+1}\, f(x).$$

Solving for f and integrating by parts,

$$\begin{aligned} f(x) &= -\frac{1}{\lambda}\, x^{1/\lambda-1}\left[t^{-1/\lambda+1}g(t)\Big|_0^x + \frac{1}{\lambda}\int_0^x t^{-1/\lambda}g(t)\, dt\right] \\ &= -\frac{1}{\lambda}\, x^{1/\lambda-1}\left[x^{-1/\lambda+1}g(x) + \frac{1}{\lambda}\int_0^x t^{-1/\lambda}g(t)\, dt\right] \\ &= -\frac{1}{\lambda}\, g(x) - \frac{1}{\lambda^2}\, x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\, dt. \end{aligned}$$

This last expression does not require g to be differentiable, and $\mathrm{Re}\,(-1/\lambda) > -1/q$ together with $g \in L^p[0,1]$ guarantee—via Hölder—that the integral exists. It is also in $L^p[0,1]$, for the first term is, and the second terms satisfies the following (we use Minkowski's Integral inequality (2.49)):

$$\begin{aligned}\left(\int_0^1 \left|x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\,dt\right|^p dx\right)^{1/p} &= \left(\int_0^1 \left|\int_0^x x^{1/\lambda-1}t^{-1/\lambda}g(t)\,dt\right|^p dx\right)^{1/p}\\ &= \left(\int_0^1 \left|\int_0^1 v^{-1/\lambda}g(vx)\,dv\right|^p dx\right)^{1/p}\\ &\leq \int_0^1 \left(\int_0^1 v^{-\operatorname{Re} p/\lambda}|g(vx)|^p dx\right)^{1/p} dv\\ &= \int_0^1 v^{-\operatorname{Re} 1/\lambda}\left(\int_0^1 |g(vx)|^p dx\right)^{1/p} dv\\ &= \int_0^1 v^{-\operatorname{Re} 1/\lambda}\left(\int_0^v v^{-1}\,|g(t)|^p\,dt\right)^{1/p} dv\\ &\leq \int_0^1 v^{-1/p-\operatorname{Re} 1/\lambda}\left(\int_0^1 |g(t)|^p\,dt\right)^{1/p} dv\\ &= \|g\|_p \int_0^1 v^{-1/p-\operatorname{Re} 1/\lambda}\,dv\\ &= \frac{1}{\frac{1}{q}-\operatorname{Re}\frac{1}{\lambda}}\,\|g\|_p\end{aligned}$$

Note that, since $\operatorname{Re}\frac{1}{\lambda} < \frac{1}{q}$,

$$-\frac{1}{p}-\operatorname{Re}\frac{1}{\lambda} > -\frac{1}{p}-\frac{1}{q} = -1,$$

which justifies the evaluation of the integral.

We claim that, for $\lambda \notin \overline{B_{q/2}(q/2)}$,

$$(T-\lambda I)^{-1}g(x) = -\frac{1}{\lambda}\,g(x) - \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\,dt \qquad \text{(AB.9.2)}$$

We have just shown that this is a bounded operator on $L^p[0,1]$. Let us apply $T-\lambda I$ to it. If h denotes the expression in (AB.9.2), with Fubini's use to be justified below,

$$\begin{aligned}(-\lambda I+T)h &= g(x)+\frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\,dt-\frac{1}{\lambda x}\int_0^x g\\&\qquad-\frac{1}{\lambda^2 x}\int_0^x s^{1/\lambda-1}\int_0^s t^{-1/\lambda}g(t)\,dt\,ds\\&= g(x)+\frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\,dt-\frac{1}{\lambda x}\int_0^x g\\&\qquad-\frac{1}{\lambda^2 x}\int_0^x\int_t^x s^{1/\lambda-1}t^{-1/\lambda}g(t)\,ds\,dt\\&= g(x)+\frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\,dt-\frac{1}{\lambda x}\int_0^x g\\&\qquad-\frac{1}{\lambda x}\int_0^x (x^{1/\lambda}-t^{1/\lambda})t^{-1/\lambda}g(t)\,dt\\&= g(x).\end{aligned}$$

Also, if $g = (T - \lambda I)f$,

$$\begin{aligned}
[(T-\lambda I)^{-1}g](x) &= -\frac{1}{\lambda}\,[(T-\lambda I)f](x) - \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda}[(T-\lambda I)f](t)\,dt \\
&= -\frac{1}{\lambda}\,x^{-1}\int_0^x f(t)\,dt + f(x) \\
&\qquad - \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda-1}\int_0^t f(s)\,ds\,dt \\
&\qquad + \frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda} f(t)\,dt \\
&= -\frac{1}{\lambda}\,x^{-1}\int_0^x f(t)\,dt + f(x) \\
&\qquad - \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_0^x \int_s^x t^{-1/\lambda-1} f(s)\,dt\,ds \\
&\qquad + \frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda} f(t)\,dt \\
&= -\frac{1}{\lambda}\,x^{-1}\int_0^x f(t)\,dt + f(x) \\
&\qquad + \frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x (x^{-1/\lambda-1} - s^{-1/\lambda-1}) f(s)\,ds \\
&\qquad + \frac{1}{\lambda}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda} f(t)\,dt \\
&= f(x).
\end{aligned}$$

So (AB.9.2) is indeed an expression for $(T - \lambda I)^{-1}$.

We have thus shown that $B_{q/2}(q/2) = \sigma_p(T) \subset \sigma(T) \subset \overline{B_{q/2}(q/2)}$, and hence $\sigma(T) = \overline{B_{q/2}(q/2)}$.

The justification for two uses of Fubini above (as in Theorem 2.7.16) comes from

$$\begin{aligned}
\int_0^x \int_0^s |s^{1/\lambda-1}|\,|t^{-1/\lambda}|\,|g(t)|\,dt\,ds &= \int_0^x \int_0^s s^{\operatorname{Re} 1/\lambda-1}\, t^{-\operatorname{Re} 1/\lambda}\,|g(t)|\,dt\,ds \\
&\le \|g\|_p \int_0^x s^{\operatorname{Re} 1/\lambda-1} \left(\int_0^s t^{-\operatorname{Re} q/\lambda}\,dt \right)^{1/q} ds \\
&\le \frac{\|g\|_p}{(1-\operatorname{Re} q/\lambda)^{1/q}} \int_0^x s^{\operatorname{Re} 1/\lambda-1}\, s^{1/q-\operatorname{Re} 1/\lambda}\,ds \\
&= \frac{q\,\|g\|_p}{(1-\operatorname{Re} q/\lambda)^{1/q}}\, x^{1/q} < \infty,
\end{aligned}$$

and

$$\int_0^x \int_0^t |t^{-1/\lambda-1}|\,|f(s)|\,ds\,dt = \int_0^x t^{-\operatorname{Re} 1/\lambda-1} \int_0^t |f(s)|\,ds\,dt$$
$$\leq \int_0^x t^{-\operatorname{Re} 1/\lambda-1}\, t^{1/q}\Big(\int_0^t |f(s)|^p\,ds\Big)^{1/p} dt$$
$$\leq \|g\|_p \int_0^1 t^{-\operatorname{Re} 1/\lambda-1+1/q}\,dt = \frac{\|g\|_p}{\frac{1}{q}-\operatorname{Re}\frac{1}{\lambda}}.$$

When $\lambda \in \partial B_{q/2}(q/2)$, we know that $T-\lambda I$ is injective since $\lambda \notin \sigma_p(T)$. For any $n \in \{0\}\cup\mathbb{N}$ we have $T(x^n) = \frac{1}{n+1}\,x^n$. So, as long as $\lambda \neq \frac{1}{n+1}$—which cannot happen when $\lambda \in \partial B_{q/2}(q/2)$, as the only real values are 0 and $q>1$—the operator $T-\lambda I$ maps $\mathbb{C}[x]$ onto itself. In particular, it has dense range, since $\mathbb{C}[x]$ is uniformly dense in $C[0,1]$ by Stone–Weierstrass (Theorem 7.4.20) and $C[0,1]$ is dense in $L^p[0,1]$ by Proposition 2.8.18. Thus

$$\sigma(T) = \overline{B_{q/2}(q/2)}, \qquad \sigma_p(T) = B_{q/2}(q/2),$$
$$\sigma_c(T) = \partial B_{q/2}(q/2), \qquad \sigma_r(T) = \varnothing.$$

Also, $\sigma_{ap}(T) = \overline{B_{q/2}(q/2)}$ since every point is either in the point spectrum or in the boundary (Proposition 9.5.5).

(9.5.18) Let $p \in (1,\infty)$ and $S \in \mathcal{B}(L^p[0,\infty))$ be the Hardy operator as in Example 9.5.22. Show that $\|S\| = q$ and find $\sigma(S)$, $\sigma_p(S)$, and $\sigma_{ap}(S)$. *(See the disclaimer in Exercise 9.5.17)*

Answer. As usual, we write $q = \frac{p}{p-1}$. First thing is to check that $Sf \in L^p[0,\infty)$ for any $f \in L^p[0,\infty)$, and that S is bounded. We have, using substitution first and Minkowski's Integral inequality second,

$$\|Sf\|_p = \Big(\int_0^\infty \Big|\frac{1}{x}\int_0^x f(t)\,dt\Big|^p dx\Big)^{1/p} = \Big(\int_0^\infty \Big|\int_0^1 f(sx)\,ds\Big|^p dx\Big)^{1/p}$$
$$\leq \int_0^1 \Big(\int_0^\infty |f(sx)|^p\,dx\Big)^{1/p} ds = \int_0^1 \Big(\frac{1}{s}\int_0^\infty |f(t)|^p\,dt\Big)^{1/p} ds$$
$$= \|f\|_p \int_0^1 s^{-1/p}\,ds = \frac{p}{p-1}\,\|f\|_p = q\,\|f\|_p.$$

We will see below that $\operatorname{spr}(S) = q$, and thus $\|S\| = q$.

The computation for the eigenvalues is exactly the same as in the case of $L^p[0,1]$ (Exercise 9.5.17). So any eigenvector will be a multiple $f(x) = x^{1/\lambda-1}$. But no power of x can be integrable on $[0,\infty)$, and hence $\sigma_p(S) = \varnothing$.

Now let us try to find $(S-\lambda I)^{-1}$. Suppose that $g=(S-\lambda I)f$; we want to express f in terms of g. Assume initially that f and g are differentiable. We can repeat the argument from the case $L^p[0,1]$, and so

$$x^{-1/\lambda}(xg(x))' = [-\lambda x^{-1/\lambda+1}f(x)]'. \tag{AB.9.3}$$

Consider first the case where $\operatorname{Re}\frac{1}{\lambda} < \frac{1}{q}$ (that is, λ is outside of the disk $\overline{B_{q/2}(q/2)}$). In this case we have $t^{-1/\lambda}$ integrable at 0, since $-\operatorname{Re} 1/\lambda + 1 > 1/p > 0$, and hence

$$\int_0^x t^{-1/\lambda}[tg(t)]'\,dt = -\lambda x^{-1/\lambda+1}\,f(x).$$

Solving for f and integrating by parts,

$$\begin{aligned} f(x) &= -\frac{1}{\lambda}\,x^{1/\lambda-1}\left[t^{-1/\lambda+1}g(t)\Big|_0^x + \frac{1}{\lambda}\int_0^x t^{-1/\lambda}g(t)\,dt\right] \\ &= -\frac{1}{\lambda}\,x^{1/\lambda-1}\left[x^{-1/\lambda+1}g(x) + \frac{1}{\lambda}\int_0^x t^{-1/\lambda}g(t)\,dt\right] \\ &= -\frac{1}{\lambda}\,g(x) - \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda}g(t)\,dt. \end{aligned}$$

This last expression does not require g to be differentiable. It is also in $L^p[0,\infty)$, for the first term is by definition, and the second term satisfies (we use Minkowski's Integral inequality (2.49))

$$\begin{aligned}
\left(\int_0^\infty \left| x^{1/\lambda-1}\int_0^x t^{-1/\lambda} g(t)\,dt\right|^p dx\right)^{1/p} &= \left(\int_0^\infty \left|\int_0^x x^{1/\lambda-1} t^{-1/\lambda} g(t) dt\right|^p dx\right)^{1/p} \\
&= \left(\int_0^\infty \left|\int_0^1 v^{-1/\lambda} g(vx)\,dv\right|^p dx\right)^{1/p} \\
&\le \int_0^1 \left(\int_0^\infty v^{-\operatorname{Re} p/\lambda} |g(vx)|^p\,dx\right)^{1/p} dv \\
&= \int_0^1 v^{-\operatorname{Re} 1/\lambda}\left(\int_0^\infty |g(vx)|^p\,dx\right)^{1/p} dv \\
&= \int_0^1 v^{-\operatorname{Re} 1/\lambda}\left(\int_0^\infty v^{-1}\,|g(t)|^p\,dt\right)^{1/p} dv \\
&= \int_0^1 v^{-1/p-\operatorname{Re} 1/\lambda}\left(\int_0^\infty |g(t)|^p\,dt\right)^{1/p} dv \\
&= \|g\|_p \int_0^1 v^{-1/p-\operatorname{Re} 1/\lambda}\,dv \\
&= \frac{1}{\frac{1}{q} - \operatorname{Re}\frac{1}{\lambda}}\,\|g\|_p.
\end{aligned}$$

The condition $\operatorname{Re}\frac{1}{\lambda} < \frac{1}{q}$ guarantees that we can evaluate the integral at the end. So we claim that, for $\lambda \notin \overline{B_{q/2}(q/2)}$,

$$(S-\lambda I)^{-1}g(x) = -\frac{1}{\lambda}\,g(x) - \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_0^x t^{-1/\lambda} g(t)\,dt \qquad \text{(AB.9.4)}$$

We have just shown that this is a bounded operator on $L^p[0,\infty)$. The computation that $(S-\lambda I)(S-\lambda I)^{-1} = (S-\lambda I)^{-1}(S-\lambda I) = I$ is exactly the same as was done in Exercise 9.5.17, so we omit it.

So far, this shows that $\sigma(S) \subset \overline{B_{q/2}(q/2)}$. For $\lambda \in B_{q/2}(q/2)$, we now have $\operatorname{Re}\frac{1}{\lambda} > \frac{1}{q}$, which makes the antiderivative of the right-hand-side of (AB.9.3) vanish at ∞. This suggests integrating between x and ∞. So we get

$$(S-\lambda I)^{-1}g(x) = -\frac{1}{\lambda}\,g(x) + \frac{1}{\lambda^2}\,x^{1/\lambda-1}\int_x^\infty t^{-1/\lambda} g(t)\,dt \qquad \text{(AB.9.5)}$$

Let us apply $S - \lambda I$ to it. If h denotes the expression in (AB.9.5), with Fubini's use and that $h \in L^p[0,\infty)$ to be justified afterwards (note also that the region where we apply Fubini requires us to split it in two integrals),

$$
\begin{aligned}
(-\lambda I + S)h &= g(x) - \frac{1}{\lambda}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda} g(t)\, dt - \frac{1}{\lambda x} \int_0^x g \\
&\quad + \frac{1}{\lambda^2 x} \int_0^x s^{1/\lambda-1} \int_s^\infty t^{-1/\lambda} g(t)\, dt\, ds \\
&= g(x) - \frac{1}{\lambda}\, x^{1/\lambda-1} \int_0^x t^{-1/\lambda} g(t)\, dt - \frac{1}{\lambda x} \int_0^x g \\
&\quad + \frac{1}{\lambda^2 x} \int_0^x \int_0^t s^{1/\lambda-1} t^{-1/\lambda} g(t)\, ds\, dt \\
&\quad + \frac{1}{\lambda^2 x} \int_x^\infty \int_0^x s^{1/\lambda-1} t^{-1/\lambda} g(t)\, ds\, dt \\
&= g(x) - \frac{1}{\lambda}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda} g(t)\, dt - \frac{1}{\lambda x} \int_0^x g + \frac{1}{\lambda x} \int_0^x g(t)\, dt \\
&\quad + \frac{1}{\lambda x} \int_x^\infty x^{1/\lambda} t^{-1/\lambda} g(t)\, ds\, dt \\
&= g(x).
\end{aligned}
$$

Similarly, if $g = (S - \lambda I)f$,

$$
\begin{aligned}
(S - \lambda I)^{-1} g &= -\frac{1}{\lambda}\, [(T - \lambda I)f](x) + \frac{1}{\lambda^2}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda} [(T - \lambda I)f](t)\, dt \\
&= f(x) - \frac{1}{\lambda x} \int_0^x f(t)\, dt + \frac{1}{\lambda^2}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda-1} \int_0^t f(s)\, ds\, dt \\
&\quad - \frac{1}{\lambda}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda} f(t)\, dt \\
&= f(x) - \frac{1}{\lambda x} \int_0^x f(t)\, dt + \frac{1}{\lambda^2}\, x^{1/\lambda-1} \int_0^x \int_x^\infty t^{-1/\lambda-1} f(s)\, dt\, ds \\
&\quad + \frac{1}{\lambda^2}\, x^{1/\lambda-1} \int_x^\infty \int_s^\infty t^{-1/\lambda-1} f(s)\, dt\, ds \\
&\quad - \frac{1}{\lambda}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda} f(t)\, dt \\
&= f(x) - \frac{1}{\lambda x} \int_0^x f(t)\, dt + \frac{1}{\lambda}\, x^{1/\lambda-1} \int_0^x x^{-1/\lambda} f(s)\, ds \\
&\quad + \frac{1}{\lambda}\, x^{1/\lambda-1} \int_x^\infty s^{-1/\lambda} f(s)\, ds - \frac{1}{\lambda}\, x^{1/\lambda-1} \int_x^\infty t^{-1/\lambda} f(t)\, dt \\
&= f(x).
\end{aligned}
$$

This means that $\sigma(S) \subset \partial B_{q/2}(q/2)$.

We need to justify Fubini's last use, and that $h \in L^p[0,\infty)$. As we are only concerned with the second term in (AB.9.5) we have, using the same substitutions as before and Minkowski's Integral Inequality,

$$\begin{aligned}
\left(\int_0^\infty \left| \int_x^\infty x^{1/\lambda-1} t^{-1/\lambda}\, g(t)\, dt\right|^p dx\right)^{1/p} &= \left(\int_0^\infty \left| \int_1^\infty v^{-1/\lambda}\, g(vx)\, dv\right|^p dx\right)^{1/p} \\
&\leq \int_1^\infty \left(\int_0^\infty v^{-\operatorname{Re} p/\lambda}\, |g(vx)|^p\, dx\right)^{1/p} dv \\
&= \int_1^\infty v^{-\operatorname{Re} 1/\lambda} \left(\int_0^\infty |g(vx)|^p\, dx\right)^{1/p} dv \\
&= \int_1^\infty v^{-1/p-\operatorname{Re} 1/\lambda} \left(\int_0^\infty |g(t)|^p\, dt\right)^{1/p} dv \\
&= \|g\|_p \int_1^\infty v^{-1/p-\operatorname{Re} 1/\lambda}\, dv \\
&= \frac{\|g\|_p}{\operatorname{Re} \frac{1}{\lambda} - \frac{1}{q}} < \infty,
\end{aligned}$$

so $h \in L^p[0,\infty)$.

As for Fubini (used both times as in Theorem 2.7.16),

$$\begin{aligned}
&\int_0^x \int_s^\infty |s^{1/\lambda-1} t^{-1/\lambda} g(t)|\, dt\, ds \\
&\qquad = \int_0^x s^{\operatorname{Re} 1/\lambda-1} \int_s^\infty t^{-\operatorname{Re} 1/\lambda}\, |g(t)|\, dt\, ds \\
&\qquad \leq \int_0^x s^{\operatorname{Re} 1/\lambda-1} \left(\int_s^\infty t^{-\operatorname{Re} q/\lambda}\, dt\right)^{1/q} \left(\int_s^\infty |g(t)|^p\, dt\right)^{1/p} ds \\
&\qquad \leq \|g\|_p \int_0^x s^{\operatorname{Re} 1/\lambda-1} \left(\frac{s^{1-\operatorname{Re} q/\lambda}}{\operatorname{Re} \frac{q}{\lambda} - 1}\right)^{1/q} ds \\
&\qquad = \frac{\|g\|_p}{\left(\operatorname{Re} \frac{q}{\lambda} - 1\right)^{1/q}} \int_0^x s^{-1/p}\, ds = \frac{q\|g\|_p\, x^{1/q}}{\left(\operatorname{Re} \frac{q}{\lambda} - 1\right)^{1/q}} < \infty
\end{aligned}$$

and

$$\int_x^\infty \int_0^t |t^{-1/\lambda-1} f(s)|\,ds\,dt = \int_x^\infty t^{-\mathrm{Re}\,1/\lambda-1} \int_0^t |f(s)|\,ds\,dt$$
$$\le \int_x^\infty t^{-\mathrm{Re}\,1/\lambda-1}\, t^{1/q}\, \|f\|_p$$
$$= \|f\|_p \int_x^\infty t^{-\mathrm{Re}\,1/\lambda-1/p}\; dt$$
$$= \frac{\|f\|_p\, x^{1/q-\mathrm{Re}\,1/\lambda}}{\mathrm{Re}\,\frac{1}{\lambda}-\frac{1}{q}} < \infty.$$

Finally, let us show that $\sigma(S) = \partial B_{q/2}(q/2)$. Fix $\lambda \in B_{q/2}(q/2)$; that is, $\mathrm{Re}\,\frac{1}{\lambda} > \frac{1}{q}$. Let $g = t^{-1}\,1_{[1,\infty)} \in L^p[0,\infty)$. We have

$$\int_0^\infty \left| \int_x^\infty x^{1/\lambda-1} t^{-1/\lambda} g(t)\,dt \right|^p dx = \int_0^\infty \left| \int_{\max\{x,1\}}^\infty x^{1/\lambda-1} t^{-1-1/\lambda}\,dt \right|^p dx$$
$$= \int_0^\infty x^{\mathrm{Re}\,p/\lambda-p} \left| \int_{\max\{x,1\}}^\infty t^{-1-1/\lambda}\,dt \right|^p dx$$
$$= \int_0^\infty x^{\mathrm{Re}\,p/\lambda-p} \Big|\lambda \max\{x,1\}^{-1/\lambda}\Big|^p dx$$
$$\ge |\lambda|^p \int_0^1 x^{\mathrm{Re}\,p/\lambda-p}\,dx = \frac{|\lambda|^p}{p(\mathrm{Re}\,\frac{1}{\lambda}-1+\frac{1}{p})}$$
$$= \frac{|\lambda|^p}{p(\mathrm{Re}\,\frac{1}{\lambda}-\frac{1}{q})}.$$

Hence, if we express (AB.9.5) as $(S-\lambda I)g = -\lambda^{-1}g + \lambda^{-2}h$, where we just estimated the p-norm of h,

$$\|(S-\lambda I)^{-1}\| \ge \|g\|_p^{-1}\, \|(S-\lambda I)^{-1}g\|_p = \|g\|_p^{-1}\, \|-\lambda^{-1}g + \lambda^{-2}h\|_p$$
$$\ge \|g\|_p^{-1}\big(|\lambda|^{-2}\|h\|_p - |\lambda|^{-1}\,\|g\|_p\big)$$
$$\ge \|g\|_p^{-1}\left(\frac{|\lambda|^{-1}}{p^{1/p}\big(\mathrm{Re}\,\frac{1}{\lambda}-\frac{1}{q}\big)^{1/p}} - |\lambda^{-1}|\,\|g\|_p\right) \tag{AB.9.6}$$

Now consider $\lambda \in \partial B_{q/2}(q/2)$; that is, $\mathrm{Re}\,1/\lambda = 1/q$ or $\lambda = 0$. Choose a sequence $\{\lambda_n\} \subset B_{q/2}(q/2)$ with $\lambda_n \to \lambda$. If we assume that $(S-\lambda I)^{-1}$ is bounded we have, using Lemma 9.2.11,

$$\|(S-\lambda_n)^{-1} - (S-\lambda I)^{-1}\| = \|(S-(\lambda+(\lambda_n-\lambda))I - (S-\lambda I)^{-1}\|$$
$$\le |\lambda_n-\lambda|\, \frac{\|(S-\lambda I)^{-1}\|^2}{1-|\lambda_n-\lambda|\,\|(S-\lambda I)^{-1}\|} \xrightarrow[n\to\infty]{} 0.$$

If follows that $\{\|(S-\lambda_n I)^{-1}\|\}$ is uniformly bounded for n big enough, contradicting (AB.9.6). So $S-\lambda I$ is not invertible and $\lambda\in\sigma(S)$. Thus $\sigma(S)=\partial B_{q/2}(q/2)$. We have found that

$$\sigma(S)=\partial B_{q/2}(q/2),\qquad \sigma_p(S)=\varnothing,\qquad \sigma_{ap}(S)=\partial B_{q/2}(q/2),$$

the last equality due to Proposition 9.5.5.

(9.5.19) Let $\mathcal{X}$ be a Banach space and $T\in\mathcal{B}(\mathcal{X})$, surjective. Let $\mathcal{Y}=\ker T$. Show that $\tilde T:\mathcal{X}/\mathcal{Y}\to\mathcal{X}$ given by $\tilde T(x+\mathcal{Y})=Tx$ is linear, bijective, and bounded.

Answer. Linearity is automatic since T is linear: for

$$\tilde T(\alpha x+y+\mathcal{Y})=T(\alpha x+y)=\alpha Tx+Ty=\alpha\tilde T(x+\mathcal{Y})+\tilde T(y+\mathcal{Y}).$$

If $\tilde T(x+\mathcal{Y})=0$, this is $Tx=0$ and so $x\in\ker T=\mathcal{Y}$, so $x+\mathcal{Y}=0$ and $\tilde T$ is injective. Also, since T is surjective, for any $y\in\mathcal{X}$ there exists $x\in\mathcal{X}$ with $y=Tx$, and so $y=Tx=\tilde T(x+\mathcal{Y})$, and $\tilde T$ is surjective.

It remains to show that $\tilde T$ is bounded. Fix $x\in\mathcal{X}$ and $\varepsilon>0$. Then there exists $y\in\mathcal{Y}$ such that $\|x+y\|<\|x+\mathcal{Y}\|+\varepsilon$. Hence

$$\|Tx\|=\|T(x+y)\|\le\|T\|\,\|x+y\|\le\|T\|\,(\|x+\mathcal{Y}\|+\varepsilon).$$

As this can be done for all $\varepsilon>0$, we have shown that $\|\tilde T(x+\mathcal{Y})\|\le\|T\|\,\|x+\mathcal{Y}\|$. Hence $\tilde T$ is bounded and $\|\tilde T\|\le\|T\|$.

(9.5.20) Let X be a compact Hausdorff space and $\psi:X\to X$ continuous. Let $T:C(X)\to C(X)$ be given by $Tf=f\circ\psi$. Show that

(i) T is injective if and only if ψ is surjective;

(ii) T is surjective if and only if ψ is injective.

Answer. Suppose that ψ is not surjective. Because ψ is continuous and X is compact, $\psi(X)$ is compact. So $X\setminus\psi(X)$ is a nonempty open set. Choose a continuous function f with $f=0$ on $\psi(X)$ and $f\neq0$ on $X\setminus\psi(X)$; then $Tf=0$ and T is not injective. Conversely, if ψ is surjective and $Tf=0$, this is $f\circ\psi=0$ and so $f=0$, making T injective.

If ψ is not injective, there exist t_0,t_1 such that $\psi(t_0)=\psi(t_1)$. If $f\in C(X)$ is such that $f(t_0)\neq f(t_1)$ (which exists by Urysohn's Lemma), then $f\neq Tg$ for any g; thus T is not surjective. Conversely, if ψ is injective, then ψ is a homeomorphism $X\to\psi(X)$ (Exercise 1.8.38). By Tietze's Extension

Theorem (2.6.9) there exists $\eta : X \to X$ continuous with $\eta \circ \psi(x) = x$ for all $x \in X$. Given $f \in C(X)$, let $g = f \circ \eta \in C(X)$. Then $Tg = f$, and thus T is surjective.

(9.5.21) With the notation of Example 9.5.20, show that

$$f_\lambda(t) = \sum_{k=0}^{\infty} \lambda^k R^k g_0 = \sum_{k=0}^{\infty} \lambda^k \left(2^{k+1} t - \frac{1-2\lambda}{1-\lambda}\right) 1_{[2^{-k-1},2^{-k}]}(t).$$

Answer. We will need the following computation:

$$\begin{aligned}\sum_{j=1}^{\infty} \lambda^j\, 1_{(2^{-j},1]}(t) &= \sum_{j=1}^{\infty} \lambda^j \sum_{k=0}^{j-1} 1_{(2^{-k-1},2^{-k}]}(t) \\ &= \sum_{k=0}^{\infty} 1_{(2^{-k-1},2^{-k}]}(t) \sum_{j=k+1}^{\infty} \lambda^j \\ &= \sum_{k=0}^{\infty} 1_{(2^{-k-1},2^{-k}]}(t) \frac{\lambda^{k+1}}{1-\lambda}.\end{aligned}$$

Then, using that

$$(R^k g_0)(t) = (2^{k+1} t - 1)\, 1_{(2^{-k-1},2^{-k}]} + 1_{(2^{-k},1]}(t),$$

we have

$$\begin{aligned}f_\lambda(t) = \sum_{k=0}^{\infty} \lambda^k R^k g_0(t) &= (2t-1)\, 1_{[1/2,1]}(t) \\ &\quad + \sum_{k=1}^{\infty} \lambda^k \left[(2^{k+1} t - 1)\, 1_{(2^{-k-1},2^{-k}]}(t) + 1_{(2^{-k},1]}(t)\right] \\ &= \sum_{k=0}^{\infty} \lambda^k\, (2^{k+1} t - 1)\, 1_{(2^{-k-1},2^{-k}]} + \sum_{j=1}^{\infty} \lambda^j\, 1_{(2^{-j},1]}(t) \\ &= \sum_{k=0}^{\infty} \lambda^k\, (2^{k+1} t - 1)\, 1_{(2^{-k-1},2^{-k}]}(t) + \sum_{k=0}^{\infty} 1_{(2^{-k-1},2^{-k}]}(t) \frac{\lambda^{k+1}}{1-\lambda} \\ &= \sum_{k=0}^{\infty} \lambda^k \left(2^{k+1} t - 1 + \frac{\lambda}{1-\lambda}\right) 1_{(2^{-k-1},2^{-k}]}(t) \\ &= \sum_{k=0}^{\infty} \lambda^k \left(2^{k+1} t - \frac{1-2\lambda}{1-\lambda}\right) 1_{(2^{-k-1},2^{-k}]}(t).\end{aligned}$$

9.6. Compact Operators

(9.6.1) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces with at least one of them equal to $\ell^1(\mathbb{N})$, and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that T is completely continuous.

Answer. Suppose first that $\mathcal{X} = \ell^1(\mathbb{N})$. Let $\{x_n\} \subset \mathcal{X}$ with $x_n \xrightarrow{\text{weak}} 0$. By Proposition 7.1.22, $x_n \to 0$; as T is bounded, $Tx_n \to 0$. So T is completely continuous.

When $\mathcal{Y} = \ell^1(\mathbb{N})$, let $\{x_n\} \subset \mathcal{X}$ with $x_n \xrightarrow{\text{weak}} 0$. As T is bounded, $Tx_n \xrightarrow{\text{weak}} 0$ (because $\varphi \circ T \in \mathcal{X}^*$ for all $\varphi \in \mathcal{Y}^*$). By Proposition 7.1.22, $Tx_n \to 0$, and T is completely continuous.

(9.6.2) Let $(X, \mathcal{A}, \mu)$ be a measure space with finite measure, $1 < p < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. If $k : X \times X \to \mathbb{C}$ is an $\mathcal{A} \boxdot \mathcal{A}$-measurable function such that

$$\sup \left\{ \int_X |k(x,y)|^q \, d\mu(y) : x \in X \right\} < \infty,$$

show that

$$(Kf)(x) = \int_X k(x,y) f(y) \, d\mu(y)$$

defines a compact operator on $L^p(\mu)$. *(Hint: use complete continuity and reflexivity)*

Answer. Let

$$c = \sup \left\{ \int_X |k(x,y)|^q \, d\mu(y) : x \in X \right\}.$$

Suppose that $f_n \xrightarrow{\text{weak}} 0$. This means that for every $g \in L^q$,

$$\int_X f_n g \, d\mu \to 0.$$

In particular, for each x

$$\int_X k(x,y) \, f_n(y) \, d\mu(y) \to 0$$

Since a weakly convergent sequence is bounded (Proposition 7.1.11), there exists $b > 0$ with $\|f_n\|_p < b$ for all n. By Hölder's Inequality,

$$\left| \int_X k(x,y)\, f_n(y)\, d\mu(y) \right| \le \left(\int_X |k(x,y)|^q \, d\mu(y) \right)^{1/q} \|f_n\|_p < bc^{1/q}.$$

Because we are in a finite-measure space, a bounded measurable function is integrable. Then, using Dominated Convergence,

$$\|Kf_n\|_p^p = \int_X \left| \int_X k(x,y)\, f_n(y)\, d\mu(y) \right|^p d\mu(x) \to 0. \qquad (9.1)$$

so K is completely continuous, and as L^p is reflexive, K is compact by Proposition 9.6.5.

(9.6.3) Consider the multiplication operator M_b as in Example 9.5.6. Show that M_b is compact if and only if $b \in c_0$.

Answer. Suppose first that $b \in c_0$. For each $m \in \mathbb{N}$ let b_m be the truncation of b to its first m coordinates. Fix $\varepsilon > 0$ and choose m such that $|b(j)| < \varepsilon$ for all $j \ge m$. Then M_{b_m} is finite-rank and

$$[(M_b - M_{b_m})x](n) = \begin{cases} b(n)x(n), & n \ge m \\ 0, & \text{otherwise} \end{cases}$$

It follows that $\|(M_b - m_{b_m})x\| \le \varepsilon \|x\|$ for all x, and so $\|M_b - M_{b_m}\| \le \varepsilon$. This shows that M_b is a limit of finite-rank operators, and hence compact by Proposition 9.6.2.

Conversely, suppose that $b \notin c_0$. Then there exists $\delta > 0$ and a subsequence $\{b_{n_k}\}$ such that $|b_{n_k}| \ge \delta$ for all k. For each k, $e_{n_k} = M_b\left(\frac{1}{b_{n_k}} e_{n_k}\right)$. As $\left\|\frac{1}{b_{n_k}} e_{n_k}\right\| \le \delta^{-1}$, the sequence $\{e_{k_n}\}$ is in the image of the ball $M_b(B_{\delta^{-1}}(0))$. As $\|e_{n_k} - e_{n_j}\| = 1$ for all $k \ne j$, the sequence does not admit a convergent subsequence and hence M_b is not compact.

(9.6.4) Let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the linear operator induced by

$$Te_n = \tfrac{1}{n+1}\, e_{n+1}, \qquad n \in \mathbb{N}.$$

(i) Show that T is bounded on c_{00}, so that it extends to all of $\ell^2(\mathbb{N})$ and is bounded.

(ii) Show that T is compact.

(iii) Show that T is quasinilpotent.

Answer.

(i) If $x = \sum_{n=1}^m x_n e_n$, then

$$\|Tx\|^2 = \sum_{n=1}^{m} \frac{|x_n|^2}{(n+1)^2} \leq \|x\|^2.$$

So $\|T\| \leq 1$ and, since c_{00} is dense in $\ell^2(\mathbb{N})$, $T \in \mathcal{B}(\ell^2(\mathbb{N}))$ by Proposition 6.1.9.

(ii) For $m \in \mathbb{N}$, define

$$T_m x = \sum_{n=1}^{m} \frac{x_n}{n+1} e_n.$$

Then T_m is finite-rank for all m, and

$$\begin{aligned}\|(T - T_m)x\|^2 = \left\| \sum_{n=m+1}^{\infty} \frac{x_n}{n+1} e_{n+1} \right\|^2 &= \sum_{n=m+1}^{\infty} \frac{|x_n|^2}{(n+1)^2} \\ &\leq \|x\|^2 \sum_{n=m+1}^{\infty} \frac{1}{(n+1)^2},\end{aligned}$$

so

$$\|T - T_m\| \leq \sum_{n=m+1}^{\infty} \frac{1}{(n+1)^2}.$$

Therefore T is a limit of finite-rank operators, and thus compact.

(iii) Since T is compact, by Theorem 9.6.13 any nonzero element of its spectrum has to be an eigenvalue. If $Tx = \lambda x$ with $\lambda \neq 0$ and $x \neq 0$, this means that $x_1 = 0$ and

$$\frac{x_n}{n+1} = \lambda x_{n+1}.$$

Inductively, this forces $x_n = 0$ for all n. So λ is not an eigenvalue. Thus $\sigma(T) = \{0\}$. As T is injective, 0 is not an eigenvalue either.

(9.6.5) Show that $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ are subspaces of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

Answer. Since a scalar multiple of a set is compact if and only if the set is compact, it is clear that for nonzero α, αT is compact if and only if T is compact. Now suppose that $S, T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$. Since $\overline{(S+T)B_1(0)} \subset \overline{SB_1(0)} + \overline{TB_1(0)}$ and a closed subset of a compact is compact, all we need to do is show that a sum of compact sets is compact. So suppose that K_1, K_2 are compact. We can proceed in two ways here. One is to notice that $K_1 + K_2 = g(K_1 \times K_2)$, where g is the continuous function $g(x, y) = x + y$ and $K_1 \times K_2$ is compact (easily checked, or we can use Tychonoff for overkill); and then use the fact that a continuous image of compact is compact (Exercise 1.8.37). Another way is to use Proposition 1.8.19. Indeed, if $\{x_j\}$ is a bounded net in $B_1(0)^{\mathcal{X}}$ then there exists a subnet $\{x_{j'}\}$ such that $\{Sx_{j'}\}$ is convergent. Now we can use the compactness of T to obtain one further subnet $\{x_{j''}\}$ such that $\{Tx_{j''}\}$ is convergent. Then $(S+T)x_{j''}\}$ is convergent and $S+T$ is compact.

As for finite rank, a sum of finite-dimensional spaces is finite-dimensional, so a sum of finite-rank operators is finite-rank. And scalar multiples do not change the rank.

(9.6.6) Show that if $R \in \mathcal{B}(\mathcal{Z},\mathcal{X})$ and $T \in \mathcal{K}(\mathcal{X},\mathcal{Y})$, then $TR \in \mathcal{K}(\mathcal{Z},\mathcal{Y})$. And if $S \in \mathcal{B}(\mathcal{Y},\mathcal{Z})$ then $ST \in \mathcal{K}(\mathcal{X},\mathcal{Z})$. Show also that analog results hold with $T \in \mathcal{F}(\mathcal{X},\mathcal{Y})$.

Answer. Consider first the case $T \in \mathcal{F}(\mathcal{X},\mathcal{Y})$. As $\operatorname{ran} TR \subset \operatorname{ran} T$, we get that $TR \in \mathcal{F}(\mathcal{Z},\mathcal{Y})$. And since $\dim \operatorname{ran} T < \infty$ and linear dependence is preserved by a linear operator, $ST \in \mathcal{F}(\mathcal{X},\mathcal{Z})$.

If T is compact, using that R is bounded we have $RB_1(0) \subset B_{\|R\|}(0)$. Then

$$\overline{TR\,B_1(0)} \subset \overline{TB_{\|R\|}(0)} = \|R\|\,\overline{TB_1(0)}.$$

The set on the right is compact since T is compact, and then the set on the left is a closed subset of a compact set, so compact. Thus TR is compact. As for ST,

$$\overline{ST\big(B_1(0)\big)} \subset \overline{S\overline{TB_1(0)}}.$$

As S is continuous, it maps compact sets to compact sets (Exercise 1.8.37) so the set on the right is compact. The set on the left is thus a closed subset of a compact set, thus compact; so ST is compact.

(9.6.7) Let $\mathcal{X}$ be a normed space and $\mathcal{J}$ a (not necessarily closed) nonzero ideal. Show that $\mathcal{F}(\mathcal{X}) \subset \mathcal{J}$.

Answer. Let $T \in \mathcal{J}$ be nonzero. This means that there exists $x \in \mathcal{X}$ such that $Tx \neq 0$. Let $R \in \mathcal{F}(\mathcal{X})$ be rank-one. Necessarily R is of the form $Rz = \varphi(z)\,w$, for some $\varphi \in \mathcal{X}^*$ and nonzero $w \in \mathcal{X}$. Use Hahn–Banach (as in Corollary 5.7.6) to obtain $\psi \in \mathcal{X}^*$ with $\psi(Tx) = 1$. Let $S \in \mathcal{B}(\mathcal{X})$ be the operator $Sz = \varphi(z)\,x$, and V the operator $Vz = \psi(z)w$. Then

$$VTSz = \varphi(z)\,VTx = \varphi(z)\,\psi(Tx)w = \varphi(z)w = Rz$$

for all $z \in \mathcal{X}$. And so $R = VTS \in \mathcal{J}$, showing that $\mathcal{J}$ contains all rank-one operators. Since any finite-rank operator is a sum of rank-one operators (this can be easily proven directly, but at this stage we can also use the polar decomposition and the Spectral Theorem 9.6.13), $\mathcal{F}(\mathcal{X}) \subset \mathcal{J}$.

(9.6.8) Let $T \in \mathcal{B}(\mathcal{X})$. Show that if $\operatorname{ran} T$ is finite-dimensional, then T is of the form (9.32).

Answer. Let $x_1, \ldots, x_n$ be a basis of $\operatorname{ran} T$. Given $x \in \mathcal{X}$, we have $Tx = \sum_{j=1}^n c_j(x)\, x_j$, where the coefficients $c_j(x)$ are uniquely determined by the linear independence. Since

$$\sum_{j=1}^n c_j(\alpha x + y)\, x_j = T(\alpha x + y) = \alpha Tx + Ty = \sum_{j=1}^n (\alpha c_j(x) + c_j(y))\, x_j,$$

we get again from the linear independence that $c_j(\alpha x + y) = \alpha c_j(x) + c_j(y)$, so each c_j is linear. We also have that c_j is bounded, because $c_j = \pi_j \circ T$, where π_j is the map $\pi_j(\sum_{j=1}^n r_j x_j) = r_j$; this map is bounded because it is a linear map on a finite-dimensional space (Exercise 9.1.2).

(9.6.9) Let $\mathcal{X}, \mathcal{Y}$ be infinite-dimensional Banach spaces, $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$. Show that T is not bounded below on any infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$.

Answer. Suppose that $\|Tx\| \geq c\|x\|$ for all $x \in \mathcal{X}_0$ and some $c > 0$. By Theorem 5.2.9 the unit ball $\overline{B_1^{\mathcal{X}_0}(0)}$ is not compact, so there exists $\delta > 0$ and a sequence $\{x_n\} \subset \mathcal{X}$ with $\|x_n\| = 1$ and $\|x_n - x_m\| \geq \delta$ for all n, m. Then

$$\|Tx_n - Tx_m\| \geq c\|x_n - x_m\| \geq c\delta > 0$$

for all n, m, showing that $\{Tx_n\}$ is not Cauchy. Hence T is not compact.

(9.6.10) Let $V \in \mathcal{B}(L^2[0,1])$ be the Volterra operator defined in Example 9.6.15,

$$(Vf)(x) = \int_0^x f.$$

As mentioned, knowing that V is compact and has no eigenvalues, it follows immediately that $\sigma(V) = \{0\}$. Prove this fact explicitly by calculating $(T - \lambda I)^{-1}$ for any $\lambda \neq 0$.

Answer. We want to solve the equation $g = (T - \lambda I)f$ in terms of f. We will initially assume that g is differentiable. Differentiating the equality we get the differential equation

$$g' = f - \lambda f'.$$

We can rewrite this as

$$g(x)' = -\lambda\Big(-\frac{1}{\lambda}f(x) + f'(x)\Big) = -\lambda e^{x/\lambda}\left(e^{-x/\lambda}f(x)\right)'$$

Hence

$$f(x) = e^{x/\lambda}\left(c - \frac{1}{\lambda}\int_0^x e^{-t/\lambda}g'(t)\,dt\right)$$

for some scalar c. Evaluating at $x = 0$ we get $c = f(0)$. From $g = (T - \lambda I)f$ we have $g(0) = -\lambda f(0)$. Integrating by parts,

$$\begin{aligned} f(x) &= e^{x/\lambda}\left(-\frac{1}{\lambda}g(0) - \frac{1}{\lambda}\int_0^x e^{-t/\lambda}g'(t)\,dt\right) \\ &= e^{x/\lambda}\left(-\frac{1}{\lambda}g(0) - \frac{1}{\lambda}e^{-t/\lambda}g(t)\Big|_0^x - \frac{1}{\lambda^2}\int_0^x e^{-t/\lambda}g(t)\,dt\right) \\ &= e^{x/\lambda}\left(-\frac{1}{\lambda}g(0) - \frac{1}{\lambda}e^{-x/\lambda}g(x) + \frac{1}{\lambda}g(0) - \frac{1}{\lambda^2}\int_0^x e^{-t/\lambda}g(t)\,dt\right) \\ &= -\frac{1}{\lambda}g(x) - \frac{e^{x/\lambda}}{\lambda^2}\int_0^x e^{-t/\lambda}g(t)\,dt. \end{aligned}$$

That is,

$$[(T - \lambda I)^{-1}f](x) = -\frac{1}{\lambda}f(x) - \frac{e^{x/\lambda}}{\lambda^2}\int_0^x e^{-t/\lambda}f(t)\,dt.$$

The expression works for any g, differentiable or not. It is bounded, for if $\operatorname{Re} 1/\lambda \neq 0$

$$\begin{aligned} \|(T - \lambda I)^{-1}g\|_2^2 &\leq \frac{2}{|\lambda|^2}\int_0^1 |g(x)|^2\,dx + \frac{2}{|\lambda|^2}\int_0^1 \left|e^{x/\lambda}\int_0^x e^{-t/\lambda}g(t)\,dt\right|^2 dx \\ &\leq \frac{2}{\lambda|^2}\,\|g\|_2^2 + \frac{2}{|\lambda|^2}\,\|g\|_2^2\int_0^1 |e^{2x/\lambda}|\int_0^x |e^{-2t/\lambda}|\,dt\,dx \\ &= \frac{2}{|\lambda|^2}\,\|g\|_2^2 + \frac{2}{|\lambda|^2}\,\|g\|_2^2\int_0^1 e^{2x\operatorname{Re} 1/\lambda}\int_0^x e^{-2t\operatorname{Re} 1/\lambda}\,dt\,dx \\ &= \frac{2}{|\lambda|^2}\,\|g\|_2^2 - \frac{1}{\operatorname{Re}\frac{1}{\lambda}\,|\lambda|^2}\,\|g\|_2^2\int_0^1 e^{2x\operatorname{Re} 1/\lambda}\,(e^{-2x\operatorname{Re} 1/\lambda} - 1)\,dx \\ &= \frac{2}{|\lambda|^2}\,\|g\|_2^2 - \frac{1}{\operatorname{Re}\frac{1}{\lambda}\,|\lambda|^2}\,\|g\|_2^2\Big(1 - \frac{1}{2\operatorname{Re}\frac{1}{\lambda}}\big(e^{2\operatorname{Re} 1/\lambda} - 1\big)\Big) \\ &\leq \left(\frac{2}{|\lambda|^2} + \frac{1}{|\operatorname{Re}\frac{1}{\lambda}|\,|\lambda|^2} + \frac{e^{2|\operatorname{Re} 1/\lambda|} + 1}{4|(\operatorname{Re}\frac{1}{\lambda}|)^2\,|\lambda|^2}\right)\|g\|_2^2. \end{aligned}$$

And when $\operatorname{Re} 1/\lambda = 0$, the exponentials all have absolute value one, and we get

$$\|(T-\lambda I)^{-1}g\|_2^2 \le \frac{2}{|\lambda|^2}\,\|g\|_2^2.$$

So $(T-\lambda I)^{-1}$ is bounded. Finally, although it follows from the derivation, a straightforward computation shows that $(T-\lambda I)^{-1}(T-\lambda I)f = f$.

(9.6.11) Let $\mathcal{X} = \ell^2(\mathbb{N})$ and $\{\lambda_n\} \subset \mathbb{C}$ with $\lambda_n \to 0$. Show that there exists $T \in \mathcal{K}(\mathcal{X})$ with $\sigma(T) = \{\lambda_n\}$.

Answer. Let

$$T\left(\sum_{k=1}^{\infty} \alpha_k e_k\right) = \sum_{k=1}^{\infty} \lambda_k \alpha_k e_k.$$

Because the sequence $\{\lambda_k\}$ is convergent, it is bounded; there exists $c > 0$ with $|\lambda_k| \le c$ for all k. Then, for $x = \sum_{k=1}^{\infty} \alpha_k e_k$,

$$\|Tx\|^2 = \left\|\sum_{k=1}^{\infty} \lambda_k \alpha_k e_k\right\|^2 = \sum_{k=1}^{\infty} |\lambda_k|^2\,|\alpha_k|^2 \le c^2\,\|x\|^2.$$

So T is bounded. If T_n is the finite-rank operator

$$T_n\left(\sum_{k=1}^{\infty} \alpha_k e_k\right) = \sum_{k=1}^{n} \lambda_k \alpha_k e_k,$$

then

$$\|(T-T_n)x\|^2 = \sum_{k=n+1}^{\infty} |\lambda_k|^2\,|\alpha_k|^2 \le \sup\{|\lambda_k|^2 :\ k \ge n+1\}\,\|x\|^2.$$

Thus $\|T - T_n\| \le \sup\{|\lambda_k| :\ k \ge n+1\}$ which shows that $T = \lim_n T_n$. By Proposition 9.6.2, T is compact. We have

$$Te_n = \lambda_n e_n,$$

so $\{\lambda_n\} \subset \sigma(T)$. If $\lambda \ne 0$ and $\lambda \notin \{\lambda_n\}$, there exists $\delta > 0$ with $|\lambda - \lambda_n| \ge \delta$ for all n. Let S be the operator

$$S\left(\sum_{k=1}^{\infty} \alpha_k e_k\right) = \sum_{k=1}^{\infty} \frac{1}{\lambda_k - \lambda}\alpha_k e_k.$$

Then S is bounded with $\|S\| = \sup\{|\lambda_n - \lambda|^{-1} :\ n\} \le \delta^{-1}$. And

$$S(T-\lambda I)x = S(T-\lambda I)\left(\sum_{k=1}^{\infty} \alpha_k e_k\right) = \sum_{k=1}^{\infty} \frac{1}{\lambda_k - \lambda}(\lambda_k - \lambda)\alpha_k e_k = x.$$

Similarly, $(T-\lambda I)S = I$, and so $\lambda \notin \sigma(T)$. Hence $\sigma(T) = \{\lambda_n\}$.

(9.6.12) Let $\mathcal{X}$ be a normed space and $E \in \mathcal{B}(\mathcal{X})$ an idempotent. Show that E has finite-rank if and only if it is compact.

Answer. If E is finite-rank, then E is compact by Proposition 9.6.2. Conversely, if E is not finite-rank, use Riesz's Lemma (Lemma 5.2.8) as in the proof of Theorem 5.2.9 to produce a sequence $\{Ex_n\} \in E\mathcal{X}$ with $\|Ex_n\| = 1$ and $\|Ex_n - Ex_m\| \geq 1/2$ for all n, m. Let $z_n = Ex_n$, $n \in \mathbb{N}$. Then $\{z_n\} \subset \overline{B_1^{\mathcal{X}}(0)}$ and $\|Ez_n - Ez_m\| = \|Ex_n - Ex_m\| \geq 1/2$ so the sequence does not admit a convergence subsequence. Hence E is not compact.

(9.6.13) Let $\mathcal{X}$ be a normed space and $\mathcal{X}_0 \subset \mathcal{X}$ a nonzero finite-dimensional proper subspace. Show that there are uncountably many idempotents $E \in \mathcal{B}(\mathcal{X})$ with $X_0 = E\mathcal{X}$.

Answer. The answer does not depend on $\mathcal{X}$ being infinite-dimensional; it works for any proper subspace of any normed space. Fix a basis $\{e_1, \ldots, e_n\}$ of $\mathcal{X}_0$. Use the argument from Exercise 5.7.2 to construct $\alpha_1, \ldots, \alpha_n \in \mathcal{X}^*$ such that $\alpha_k(e_j) = \delta_{kj}\, e_k$. Since $\mathcal{X}_0$ is proper, there exists a unit vector $y \in \mathcal{X} \setminus X_0$. For each $t \in [0, 2\pi)$, define $\beta_t(cy) = e^{it}c$, $\beta_t(e_k) = 0$ for $k = 1, \ldots, n$, and extend to $\beta_t \in \mathcal{X}^*$ by Hahn–Banach. Now let E_t be the idempotent

$$E_t(x) = \sum_{k=1}^{n} (\alpha_k(x) + \beta_t(x))\, e_k.$$

Then $E_t \in \mathcal{B}(\mathcal{X})$, $E_t\mathcal{X} = X_0$, $E_t^2 = E_t$. And for $s - t \notin 2\pi\mathbb{Z}$ we have $E_t y = e^{it}y \neq e^{is}y = E_s y$, so $E_t \neq E_s$.

(9.6.14) Let $\mathcal{X}$ be a Banach space and $E \in \mathcal{B}(\mathcal{X})$ an idempotent. Show that $\operatorname{ran} E$ and $\operatorname{ran}(I - E)$ are closed.

Answer. Let $\{Ex_n\}$ be Cauchy. Then there exists $x \in \mathcal{X}$ such that $Ex_n \to x$. As E is bounded, $Ex_n = E(Ex_n) \to Ex$. Hence $\operatorname{ran} E$ is closed. And $I - E$ is also an idempotent, so its range is closed too.

(9.6.15) Let $E \in \mathcal{B}(\mathcal{X})$ be a finite-rank idempotent. Show that E^* is a finite-rank idempotent and that $\dim \operatorname{ran} E = \dim \operatorname{ran} E^*$.

Answer. For $g \in \mathcal{X}^*$ and $x \in \mathcal{X}$, we have

$$(E^*E^*g)x = (E^*g)(Ex) = g(E^2x) = g(Ex) = (E^*g)x.$$

As this holds for all $x \in \mathcal{X}$ and all $g \in \mathcal{X}^*$, we have the equality $(E^*)^2 = E^*$.

If $\dim E = n$, fixing a basis $\{e_1, \ldots, e_n\}$ of $\operatorname{ran} E$ we have

$$Ex = \sum_{k=1}^{n} \alpha_k(x) e_k$$

for certain coefficients $\alpha_1(x), \ldots, \alpha_n(x)$. The linear independence of the basis (or, equivalently, the uniqueness of the representation of an element of $\operatorname{ran} E$ in the basis) gives us that each α_k is linear. Let $\beta_k \in \mathcal{X}^*$ such that $\beta_k(e_j) = \delta_{kj}$ (as in the answer to Exercise 5.7.2, one defines these functionals on $\operatorname{ran} E$ and extends by Hahn–Banach). Then

$$|\alpha_k(x)| = |\beta_k(Ex)| \leq \|\beta_k\| \, \|E\| \, \|x\|,$$

so $\alpha_k \in \mathcal{X}^*$ for all k. Now

$$(E^*g)x = g(Ex) = \sum_{k=1}^{n} \alpha_k(x) g(e_k),$$

for all $g \in \mathcal{X}^*$ and all $x \in \mathcal{X}$. Thus

$$E^*g = \sum_{k=1}^{n} g(e_k)\, \alpha_k.$$

So $\operatorname{ran} E^* \subset \operatorname{span}\{\alpha_1, \ldots, \alpha_n\}$. We also have

$$E^*\beta_k = \alpha_k,$$

showing that $\operatorname{ran} E^* = \operatorname{span}\{\alpha_1, \ldots, \alpha_n\}$. So $\dim \operatorname{ran} E^* = n$.

Conversely, suppose $\dim \operatorname{ran} E^* < \infty$. By the above, $\dim \operatorname{ran} E^{**} = \dim \operatorname{ran} E^*$. By Proposition 9.4.6 and Exercise 7.3.5,

$$\operatorname{ran} E^{**} = (\ker E^*)^o = (\operatorname{ran} E)^{oo} = \overline{J_{\mathcal{X}} \operatorname{ran} E}^{w^*}$$

(note that $\operatorname{ran} E^{**}$ is closed, being finite-dimensional). Since the closure of $J_{\mathcal{X}} \operatorname{ran} E$ is finite-dimensional, then $J_{\mathcal{X}} \operatorname{ran} E$ is finite-dimensional. This means it is closed, and hence

$$\operatorname{ran} E^{**} = J_{\mathcal{X}} \operatorname{ran} E.$$

As $J_{\mathcal{X}}$ is injective it preserves dimension, so $\dim \operatorname{ran} E = \dim \operatorname{ran} E^{**} = \dim \operatorname{ran} E^*$.

(9.6.16) Let $\mathcal{X}$ be a vector space and $P, Q : \mathcal{X} \to \mathcal{X}$ be idempotents. Show that if $\ker Q \subset \ker P$ and $\operatorname{ran} Q \subset \operatorname{ran} P$, then $P = Q$.

Answer. From $\operatorname{ran} Q \subset \operatorname{ran} P$ we have $PQx = Qx$ for all $x \in \mathcal{X}$, so $Q = PQ$. Since $(I - Q)x \in \ker Q \subset \ker P$, we get that $P(I - Q)x = 0$. Then $P = PQ$. It follows that $Q = PQ = P$.

9.7. Fredholm Operators

(9.7.1) In the proof of Corollary 9.7.8, show that $T_{11} : \mathcal{X}_1 \to \mathcal{Y}_1$ is a bounded linear bijection.

Answer. Since T_{11} is a restriction, $\|T_{11}\| \leq \|T\|$. Suppose that $x \in \mathcal{X}_1$ and $T_{11}x = 0$; this is $Tx = 0$, so $x \in \ker T = \mathcal{X}_2$. As $\mathcal{X}_1 \cap \mathcal{X}_2 = \{0\}$ we get that $x = 0$ and T_{11} is injective. Given $y \in \mathcal{Y}_1 = \operatorname{ran} T$, there exists $x = x_1 + x_2$ with $Tx = y$ and $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. Then $y = Tx_1$, with $x_1 \in \mathcal{X}_1$. So T_{11} is linear, bijective, and bounded.

(9.7.2) In the proof of Corollary 9.7.8, show that $\operatorname{ran} Q = \ker T$.

Answer. We have, for $x \in \ker R$,

$$TQx = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} -T_{11}^{-1}T_{12}x \\ x \end{bmatrix} = \begin{bmatrix} -T_{12}x + T_{12}x \\ -T_{21}T_{11}^{-1}T_{12}x + T_{22}x \end{bmatrix} \begin{bmatrix} 0 \\ Rx \end{bmatrix} = 0.$$

So $Q(\ker R) \subset \ker T$. Given $x \in \ker T$, by (9.36) we have that $x_2 \in \ker R$. And by (9.35) we have that $x_1 = -T_{11}^{-1}T_{12}x_2$. Then

$$Qx_2 = \begin{bmatrix} -T_{11}^{-1}T_{12}x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x.$$

(9.7.3) Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$ Fredholm, $R \in \mathcal{B}(\mathcal{X})$ invertible. Show that RT and TR are Fredholm.

Answer. We have $\ker RT = \ker T$, $\ker TR = R^{-1} \ker T$, so in each case the dimensions agree. Similarly, $\operatorname{ran} RT = R(\operatorname{ran} T)$, $\operatorname{ran} TR = \operatorname{ran} T$ and again the dimensions agree. But we need equality of dimensions of the quotients. It is not true in general that taking the quotient over two isomorphic subspaces will lead to isomorphic quotients (for infinite-dimensional $\mathcal{X}$ we could take $\mathcal{X}_0 = \mathcal{X}$ and $\mathcal{X}_1 \subset \mathcal{X}$ a proper subspace with $\mathcal{X}_1 \simeq \mathcal{X}$). But here we have R, and automorphism of the whole subspace. This gives

$$\mathcal{X}/\operatorname{ran} RT = \mathcal{X}/(R\operatorname{ran} T) = R\mathcal{X}/(R\operatorname{ran} T) \simeq \mathcal{X}/\operatorname{ran} T.$$

Another way is to notice that by Proposition 9.7.7 there exists S such that $I-ST$ and $I-TS$ are compact. Then $I-(SR^{-1})RT = I-ST$ is compact, as is $I-RT(SR^{-1}) = R(I-ST)R^{-1}$. Similarly, $I-TR(R^{-1}S) = I-TS$ is compact, as is $I-R^{-1}STR = R^{-1}(I-ST)R$.

(9.7.4) Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and $Z \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be Fredholm. Show that ZT is Fredholm.

Answer. By Proposition 9.7.7 there exist $S \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ and $R \in \mathcal{B}(\mathcal{Z}, \mathcal{Y})$ and compact operators K_1, K_2, K_3, K_4 such that

$$ST = I_{\mathcal{X}} + K_1, \qquad TS = I_{\mathcal{Y}} + K_2, \qquad RZ = I_{\mathcal{Y}} + K_3, \qquad ZR = I_{\mathcal{Z}} + K_4.$$

Then

$$\begin{aligned}(SR)ZT &= S(RZ)T = S(I_{\mathcal{Y}} + K_3)T \\ &= ST + SK_3T = I_{\mathcal{X}} + (K_1 + SK_3T),\end{aligned}$$

so SR is a left-inverse for ZT modulo the compacts. Similarly,

$$\begin{aligned}ZT(SR) &= Z(TS)R = Z(I_{\mathcal{Y}} + K_2)R \\ &= ZR + ZK_2R = I_{\mathcal{Z}} + (K_4 + ZK_2R).\end{aligned}$$

Then Proposition 9.7.7 guarantees that ZT is Fredholm.

(9.7.5) Let $\mathcal{X}$ be a Banach space, and $\mathcal{X}_0, \mathcal{X}_1$ subspaces with $\mathcal{X}_0 \cap \mathcal{X}_1 = \varnothing$ and $\mathcal{X} = \mathcal{X}_0 + \mathcal{X}_1$ (that is, $\mathcal{X}$ is the direct sum of $\mathcal{X}_0$ and $\mathcal{X}_1$). Let $T \in \mathcal{B}(\mathcal{X})$ such that $T\mathcal{X}_0 \subset \mathcal{X}_0$, $T\mathcal{X}_1 \subset \mathcal{X}_1$. Let $T_0 = T|_{\mathcal{X}_0}$, $T_1 = T|_{\mathcal{X}_1}$. Show that T is Fredholm if and only if T_0 and T_1 are Fredholm, and that

$$\operatorname{Ind} T = \operatorname{Ind} T_0 + \operatorname{Ind} T_1.$$

Answer. We have $\ker T = \ker T_0 + \ker T_1$. Indeed, if $x \in \ker T$, we have $x = x_0 + x_1$ for unique $x_0 \in \mathcal{X}_0$, $x_1 \in \mathcal{X}_1$. Then $0 = Tx = T_0x_0 + T_1x_1$. As $T_0x_0 \in \mathcal{X}_0$, $T_1x_1 \in \mathcal{X}_1$, the uniqueness of the decomposition (coming from $\mathcal{X}_0 \cap \mathcal{X}_1 = \{0\}$) gives us that $Tx_0 = Tx_1 = 0$; that is $\ker T \subset \ker T_0 + \ker T_1$. Conversely, if $x_0 \ker T_0$, $x_1 \in \ker T_1$, then $T(x_0 + x_1) = T_0x_0 + T_1x_1 = 0 + 0 = 0$, and so $\ker T_0 + \ker T_1 \subset \ker T$.

We also have $\operatorname{ran} T = \operatorname{ran} T_0 + \operatorname{ran} T_1$, with $\operatorname{ran} T_0 \cap \operatorname{ran} T_1 = \varnothing$. Indeed, $Tx = T_0 x_0 + T_1 x_1$.

Also,

$$\mathcal{X}/(\operatorname{ran} T) \simeq \mathcal{X}/\operatorname{ran} T_0 \oplus \mathcal{X}/\operatorname{ran} T_1.$$

For this, consider the map $\gamma : \mathcal{X}/(\operatorname{ran} T) \to \mathcal{X}/\operatorname{ran} T_0 \oplus \mathcal{X}/\operatorname{ran} T_1$ given by

$$\gamma(x + \operatorname{ran} T) = \big(x_0 + \operatorname{ran} T_0, x_1 + \operatorname{ran} T_1\big).$$

This is well-defined: if $x - y \in \operatorname{ran} T$, then $x_0 - x_0 \in \operatorname{ran} T_0$, $x_1 - y_1 \in \operatorname{ran} T_1$. It is linear. It is injective: if $x_0 \in \operatorname{ran} T_0$, $x_1 \in \operatorname{ran} T_1$, then $x_0 + x_1 \in \operatorname{ran} T$. And it is surjective: $(x_0 + \operatorname{ran} T_0, x_1 + \operatorname{ran} T_1) = \gamma(x_0 + x_1 + \operatorname{ran} T)$.

In summary, we have proven that

$$\dim \ker T = \dim \ker T_0 + \dim \ker T_1,$$

and

$$\dim \mathcal{X}/\operatorname{ran} T = \dim \mathcal{X}_0/\operatorname{ran} T_0 + \dim \mathcal{X}_1/\operatorname{ran} T_1.$$

It follows that T is Fredholm if and only if both T_0 and T_1 are Fredholm, and $\operatorname{Ind} T = \operatorname{Ind} T_0 + \operatorname{Ind} T_1$.

(9.7.6) Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$ Fredholm. Show that the following statements are equivalent:

(i) there exist $S, K \in \mathcal{B}(\mathcal{X})$, with S invertible and K compact, such that $T = S + K$;

(ii) $\operatorname{Ind} T = 0$.

Answer.

Suppose first that $T = S + K$, with S invertible and K compact, by Corollary 9.7.14 we have $\operatorname{Ind} T = \operatorname{Ind}(S + K) = \operatorname{Ind} S = 0$, since $\ker S = \{0\}$ and $\operatorname{coker} S = \{0\}$.

Conversely, suppose that $\operatorname{Ind} T = 0$, then $\dim \ker T = \dim \operatorname{coker} T$. By Corollary 9.7.8 we have decompositions $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2 = \mathcal{Y}_1 \oplus \mathcal{Y}_2$, with $T|_{\mathcal{X}_1} : \mathcal{X}_1 \to \mathcal{Y}_1$ invertible and $\dim \mathcal{X}_2 = \dim \mathcal{Y}_2$. Let $V : \mathcal{X}_2 \to \mathcal{Y}_2$ be a linear bijection, and form $S : \mathcal{X} \to \mathcal{X}$ by

$$Sx = \begin{cases} Tx, & x \in \mathcal{X}_1 \\ Vx, & x \in \mathcal{X}_2 \end{cases}$$

Then S is a linear bijection, and it is bounded because T and V are. By the Inverse Function Theorem (6.3.6), S is invertible. Let $K = T - S$. Then

$$K = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} - \begin{bmatrix} T_{11} & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} 0 & T_{12} \\ T_{21} & T_{22} - V \end{bmatrix},$$

is compact, as it is finite-rank.

9.8. Schauder Bases and Basic Sequences

(9.8.1) Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis. Show that P_n is linear and bounded.

Answer. By the unique representation property of E, if

$$x = \sum_k c_k e_k, \qquad y = \sum_k d_k e_k$$

and $\lambda \in \mathbb{C}$, then

$$\lambda x + y = \sum_k (\lambda c_k + d_k) e_k.$$

Then

$$P_n(\lambda x + y) = \sum_{k=1}^n (\lambda c_k + d_k) e_k = \lambda \sum_{k=1}^n c_k e_k + \sum_{k=1}^n d_k e_k = \lambda P_n x + P_n y.$$

So P is linear. As P_n is finite-rank, it is bounded (Proposition 9.6.2).

(9.8.2) Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis. Show that $b_E < \infty$.

Answer. Fix $x \in \mathcal{X}$. Then $P_n x \to x$. This implies that $\|P_x\| \to \|x\|$, so the sequence $\{\|P_n x\|\}$ is bounded. By the Uniform Boundedness Principle, $\sup\{\|P_n\| : n\} < \infty$.

(9.8.3) In the proof of Proposition 9.8.2, show that $\|\cdot\|_b$ is a seminorm.

Answer. We have $\|x\|_b \geq 0$ by definition. Given $\lambda \in \mathbb{C}$,

$$\|\lambda x\|_b = \sup\{\|P_n \lambda x\| : n\} = |\lambda| \{\|P_n x\| : n\} = |\lambda| \, \|x\|_b.$$

If $x, y \in \mathcal{X}$ we have $\|P_n(x+y)\| \leq \|P_x\| + \|P_n y\|$ by the linearity of P_n and the triangle inequality. As the supremum is subadditive, we get $\|x + y\|_b \leq \|x\|_b + \|y\|_b$.

(9.8.4) Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis. Show that E is also a Schauder basis for the Banach space $(\mathcal{X}, \|\cdot\|_b\|)$.

Answer. Since $\|\cdot\|_b$ is equivalent to $\|\cdot\|$ by Proposition 9.8.2, we get that $(\mathcal{X}, \|\cdot\|_b)$ is a Banach space and that the series $\sum_k c_k e_k$ converges in $\|\cdot\|_b$ precisely when it converges in $\|\cdot\|$. So the same coefficients will represent x in terms of E both in $\|\cdot\|$ and $\|\cdot\|_b$, and the uniqueness is preserved. So E is a Schauder basis for $(\mathcal{X}, \|\cdot\|_b)$.

(9.8.5) Fix $t_1 = 0$, $t_2 = 1$, and $t_3 < \cdots < t_n \in (0, 1)$. Let $D_n \subset C[0, 1]$ be set of piecewise linear functions with nodes at each t_j. Show that D_n is a subspace and that $\dim D_n = n$.

Answer. It is enough to find a basis of n elements. Fix $f \in D_n$. Let $g_1 = 1$ and $b_1 = f(t_1)$. Then $f - b_1 g_1 \in D_n$ and $(f_1 - b_1 g_1)(0) = 0$. Let g_2 be the piecewise linear function with $g_2(t_1) = 0$, $g_2(t_2) = 1$. Then if $b_2 = f(t_2)$ we get $f - b_1 g_1 - b_2 g_2 \in D_n$ and $f - b_1 g_1 - b_2 g_2 \in D_n$ is 0 at t_1 and at t_2. Continuing like that, each time one more point is brought to zero. So after n steps.

$$f = \sum_{k=1}^{n} f(t_k) g_k.$$

(9.8.6) Let $\mathcal{X}$ be a Banach space and $\{e_n\}$ a sequence of nonzero elements satisfying (9.39). Show that $\{e_n\}$ is linearly independent.

Answer. Suppose that $\sum_{j=1}^{n} a_j e_j = 0$. By (9.39) we have

$$|a_1| \, \|e_1\| = \|a_1 e_1\| \le c \Big\| \sum_{j=1}^{n} a_j e_j \Big\| = 0.$$

So $a_1 = 0$. Repeating the argument we obtain successively that $a_j = 0$ for all j. So $\{e_n\}$ is linearly independent.

(9.8.7) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $\mathcal{X}_0 \subset \mathcal{X}$, $\mathcal{Y}_0 \subset \mathcal{Y}$ (not necessarily closed) subspaces, and $T : \mathcal{X}_0 \to \mathcal{Y}_0$ a bounded linear invertible operator. Show that there exists a unique extension $\tilde{T} : \overline{\mathcal{X}_0} \to \overline{\mathcal{Y}_0}$ that is still invertible. In other words, a bounded isomorphism of subspaces extends to an isomorphism of their closures.

Answer. The existence of the extension comes from Proposition 6.1.9. All we need to check is that this extension is invertible if T is. So fix $x \in \overline{\mathcal{X}_0}$. Then there exists a sequence $\{x_n\} \subset \mathcal{X}_0$ with $x_n \to x$. We have, since the extensions are bounded,

$$\widetilde{T^{-1}}\tilde{T}x = \lim_n \widetilde{T^{-1}}\tilde{T}x_n = \lim_n \widetilde{T^{-1}}Tx_n = \lim_n T^{-1}Tx_n = \lim_n x_n = x.$$

So $\widetilde{T^{-1}}\tilde{T} = \mathrm{id}_{\mathcal{X}_0}$, and similarly we get that $\tilde{T}\widetilde{T^{-1}} = \mathrm{id}_{\mathcal{Y}_0}$. Hence $\tilde{T}$ is invertible.

(9.8.8) Show that any subsequence of a basic sequence is a basic sequence.

Answer. It is enough to show that a subsequence of a Schauder basis is a basic sequence. Let $\{e_n\}$ be a Schauder basis for the Banach space $\mathcal{X}$, and let $\{e_{n_k}\}$ be a subsequence. Put $\mathcal{X}_0 = \overline{\operatorname{span}}^{\|\cdot\|}\{e_{n_k} : k\}$. If $\{P_n\}$ are the basis projections for $\{e_n\}$, let $Q_k = P_{n_k}|_{\mathcal{X}_0} \in \mathcal{B}(\mathcal{X}_0)$; this requires checking that Q_k acts on $\mathcal{X}_0$, which should be clear since

$$Q_k\left(\sum_{j=1}^{\ell} c_j e_{n_j}\right) = P_{n_k}\left(\sum_{j=1}^{\ell} c_j e_{n_j}\right) = \sum_{j=1}^{\ell} c_j P_{n_k} e_{n_j}$$
$$\in \operatorname{span}\{e_{n_1}, \ldots, e_{n_\ell}\} \in \mathcal{X}_0,$$

since $P_{n_k}(e_{n_j})$ is either 0 or e_{n_j}. Being a subsequence of a sequence of basis projections, $\{Q_k\}$ satisfies (ii) in Lemma 9.8.3. And since

$$Q_k\mathcal{X}_0 = \operatorname{span}\{e_{n_1}, \ldots, e_{n_k}\},$$

we also have $\dim Q_k\mathcal{X}_0 = k$. So by Lemma 9.8.3 $\{Q_k\}$ are basis projections for the Schauder basis $\{e_{n_k}\}_k$ of $\mathcal{X}_0$. This means that $\{e_{n_k}\}$ is a basic sequence.

(9.8.9) In the proof of (i) $\Longrightarrow$ (ii) in Proposition 9.8.9, show that S is bijective.

Answer. The equivalence of E and F guarantees that if $\sum_j a_j e_j$ converges, so does $\sum_j a_j f_j$. Together with the uniqueness of the coefficients, this guarantees that S is well-defined. The uniqueness of the coefficients also allows us to similarly define $T : \sum_j a_j f_j \longmapsto \sum_j a_j e_j$, which is an inverse for S. Hence S is bijective.

(9.8.10) Show that

(i) $c_0 \simeq c_0 \oplus_\infty c_0$, where $\|(x, y)\|_\infty = \max\{\|x\|_\infty, \|y\|_\infty\}$;

(ii) $\ell^p(\mathbb{N}) \simeq \ell^p(\mathbb{N}) \oplus_p \ell^p(\mathbb{N})$, $1 \leq p \leq \infty$, where

$$\|(x, y)\|_p = (\|x\|_p^p + \|y\|_p^p)^{1/p}.$$

In both cases the isomorphism can be made isometric.

Answer. This is a particular case of Exercise 5.3.7, but let us write and ad-hoc proof here.

Consider $\gamma : c_0 \to c_0 \oplus c_0$ given by

$$\gamma(x_1, x_2, \ldots) = (x_1, x_3, \ldots) \oplus (x_2, x_4, \ldots).$$

Linearity is easy to check. For the isometry, recall that the sup norm on c_0 is actually a maximum.

$$\begin{aligned}\|\gamma(x)\| &= \max\big\{\max\{|x_{2k-1}| :\ k\},\ \max\{|x_{2k}| :\ k\}\big\}\\ &= \max\{|x_k| :\ k\} = \|x\|.\end{aligned}$$

We can use the same γ in the $\ell^p(\mathbb{N})$ case. Now we can do

$$\begin{aligned}\|\gamma(x_1, x_2, \ldots)\| &= \Big(\sum_{k=1}^\infty |x_{2k-1}|^p + \sum_{k=1}^\infty |x_{2k}|^p\Big)^{1/p}\\ &= \Big(\sum_{k=1}^\infty |x_k|^p\Big)^{1/p} = \|x\|.\end{aligned}$$

(9.8.11) Show that

(i) $c_0 \simeq \Big(\bigoplus_{n\in\mathbb{N}} c_0\Big)_{c_0}$, where $\Big\|\bigoplus_n x_n\Big\|_\infty = \max\{\|x_n\|_\infty :\ n\}$;

(ii) $\ell^p(\mathbb{N}) \simeq \bigoplus_{n\in\mathbb{N}} \ell^p(\mathbb{N})$, $1 \le p \le \infty$, where

$$\Big\| \bigoplus_n x_n \Big\|_p = \Big(\sum_{n=1}^{\infty} \|x_n\|_p^p \Big)^{1/p}.$$

In both cases the isomorphism can be made isometric. *The only difficulty of this exercise compared to Exercise 9.8.10 is having to deal with sequences of sequences and not getting swamped by the notation. If needed, it might help to use $x(k)$ for the k^{th} entry of x.*

Answer.

(i) Let $\mathbb{N} = \bigcup_n L_n$ be a partition of $\mathbb{N}$ into countably many infinite sets, and for each n let $\alpha_n : \mathbb{N} \to L_n$ be a bijection. Let $\gamma : c_0 \to \bigoplus_{n\in\mathbb{N}} c_0$ be given by

$$\gamma(x) = \bigoplus_n x \circ \alpha_n.$$

Linearity of γ is automatic, since it occurs component wise. For the isometry,

$$\begin{aligned}\|\gamma(x)\| &= \sup\{\|x \circ \alpha_n\| :\ n \in \mathbb{N}\} = \sup_n \sup_k |x(\alpha_n(k))| \\ &= \sup\{|x(k)| :\ k\} = \|x\|.\end{aligned}$$

For the surjectivity, if $z : \mathbb{N} \to c_0$ satisfies $\|z(n)\|_\infty \to 0$, let x be given by

$$x(k) = z(n)(\alpha_n^{-1}(k)), \qquad \text{where } k \in L_n.$$

Then, noting that $\gamma(x)(n)(h) = x(\alpha_n(h))$,

$$\gamma(x)(n)(h) = x(\alpha_n(h)) = z(n)(\alpha_n^{-1}(\alpha_n(h))) = z(n)(h),$$

and so $\gamma(x) = z$ and γ is surjective. The fact that $\|z(n)\|_\infty \to 0$ is used to guarantee that $x \in c_0$.

(ii) We can use the same γ. Now

$$\|\gamma(x)\|_p = \Big(\sum_n \|x \circ \alpha_n\|_p^p \Big)^{1/p} = \Big(\sum_n \sum_k |x(\alpha_n(k)|^p \Big)^{1/p} = \|x\|_p.$$

(9.8.12) Corollary 9.8.13 is a result due to Mazur, but the proof offered is not the original one. Here we will outline the original argument.

(i) Given a normed space $\mathcal{X}$ and $V \subset \mathcal{X}$ a finite-dimensional subspace, and given $\varepsilon \in (0,1)$, show that there exists $x \in \mathcal{X}$ with $\|x\| = 1$ and such that

$$\|z\| \leq (1+\varepsilon)\|z + \lambda x\|, \qquad z \in V,\ \lambda \in \mathbb{C}. \tag{9.47}$$

(Hint: find an $\varepsilon/2$-net for $\partial B_1^V(0)$, bounded linear functionals that are 1 at each point of the net, and take x in the intersection of the kernels)

(ii) Given $\varepsilon > 0$, show that there exists $\delta > 0$ such that

$$\prod_{k=1}^{\infty}(1 + \delta\, 2^{-k}) \leq 1 + \varepsilon.$$

(iii) Use Proposition 9.8.6 to prove Corollary 9.8.13, by using (i) inductively.

Answer.

(i) Since $\partial B_1^V(0)$ is compact (by Corollary 5.2.4), there exist $v_1, \ldots, v_r$ such that $\partial B_1^V(0) \subset \bigcup_j B_{\varepsilon/2}(v_j)$. By Corollary 5.7.7 there exist $\varphi_1, \ldots, \varphi_r \in \mathcal{X}^*$ with $\varphi_j(v_j) = 1$, $j = 1, \ldots, r$. By Proposition 5.5.12, there exists $x \in \bigcap_j \ker \varphi_j$ with $\|x\| = 1$.

Since $0 < \varepsilon < 1$ we have $\varepsilon - \varepsilon^2 > 0$. Then $2 + 2\varepsilon - \varepsilon - \varepsilon^2 > 2$, which is $(2-\varepsilon)(1+\varepsilon) > 2$, and it can be written $1 - \varepsilon/2 > 1/(1+\varepsilon)$.

Given $z \in V$ with $\|z\| = 1$ there exists j with $\|z - v_j\| < \varepsilon/2$. Then

$$\begin{aligned}\|z + \lambda x\| &\geq \|v_j + \lambda x\| - \|z - v_j\| \geq |\varphi_j(v_j + \lambda x) - \frac{\varepsilon}{2} \\ &= 1 - \frac{\varepsilon}{2} > \frac{1}{1+\varepsilon},\end{aligned}$$

yielding (9.47) (since $\|z\| = 1$). For $z = 0$ there is nothing to prove, and for arbitrary $z \in V$ we have, applying (9.47) to $z/\|z\|$ and $\lambda/\|z\|$,

$$1 < (1+\varepsilon)\left\|\frac{z}{\|z\|} + \frac{\lambda}{\|z\|}\,x\right\|$$

and the general form of (9.47) follows.

(ii) Taking logarithm,

$$\log \prod_{k=1}^{n}(1 + \delta\, 2^{-k}) = \sum_{k=1}^{n} \log(1 + \delta\, 2^{-k}) \leq \sum_{k=1}^{n} \delta\, 2^{-k} \leq \delta.$$

So, for any n,

$$\prod_{k=1}^{n}(1 + \delta\, 2^{-k}) \leq e^{\delta}.$$

Now choose $\delta > 0$ and small enough so that $e^{\delta} \leq 1+\varepsilon$.

(iii) Choose $\delta > 0$ such that $\prod_{k=1}^{n}(1+\delta\, 2^{-k}) \leq 1+\varepsilon$. Fix $x_1 \in \mathcal{X}$ with $\|x_1\| = 1$. Apply (9.47) to $V = \operatorname{span}\{x_1\}$ to obtain $x_2 \in \mathcal{X}$ with $\|x_2\| = 1$ and $\|z\| \leq (1+\delta 2^{-1})\,\|z+\lambda x_2\|$ for all $z \in \operatorname{span}\{x_1\}$. Inductively, given $x_1, \ldots, x_n$ unit vectors with $\|z\| \leq (1+\delta\, 2^{-k})\|z+\lambda x_k\|$ for $z \in \operatorname{span}\{x_1, \ldots, x_{k-1}\}$, we get $x_{n+1} \in \mathcal{X}$ with $\|x_{n+1}\| = 1$ and $\|z\| \leq (1+\delta\, 2^{-n-1})\|z+\lambda x_{n+1}\|$ for all $z \in \operatorname{span}\{x_1, \ldots, x_n\}$. Now, for any $c \in \mathbb{C}^{\mathbb{N}}$ and $n \leq m$,

$$\begin{aligned}\Big\|\sum_{k=1}^{n} c_k x_k\Big\| &\leq (1+\delta\, 2^{-n})\Big\|\sum_{k=1}^{n+1} c_k x_k\Big\| \\ &\leq \cdots \leq \prod_{k=n}^{m}(1+\delta\, 2^{-k})\,\Big\|\sum_{k=1}^{m} c_k x_k\Big\| \\ &\leq (1+\varepsilon)\Big\|\sum_{k=1}^{m} c_k x_k\Big\|.\end{aligned}$$

By Proposition 9.8.6 we get that $E = \{x_n\}$ is a basic sequence and the estimate gives $b_E \leq 1+\varepsilon$.

(9.8.13) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $\mathcal{X}_0 \subset \mathcal{X}$ a subspace with $\dim \mathcal{X}_0 = \infty$, and $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$. Show that there exists a normalized basic sequence $X = \{x_n\} \subset \mathcal{X}_0$ with $b_X < 2$ and such that $\|Tx_n\| < 2^{-n}$ for all n.

Answer. In the answer to Exercise 9.8.12 the elements x_n of the basic sequence come out of $\bigcap_{j=1}^{n} \ker \varphi_j$. From Exercise 5.5.12 we know that said intersection is always infinite-dimensional. As T is strictly singular, it is not bounded below on $\bigcap_{j=1}^{n} \ker \varphi_j$ and so we can choose x_n with $\|x_n\| = 1$ and $\|Tx_n\| \leq 2^{-n}$.

(9.8.14) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$ if and only if for every infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$ there exists an infinite-dimensional subspace $\mathcal{X}_1 \subset \mathcal{X}_0$ with $T|_{\mathcal{X}_1}$ compact.

Answer. Suppose first that $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$. Fix an infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$. By Exercise 9.8.13 there exists a normalized basic sequence $\{x_n\} \subset \mathcal{X}_0$ with $\|Tx_n\| \leq 2^{-n}$ for all n. Let $\mathcal{X}_1 = \overline{\operatorname{span}}^{\|\cdot\|}\{x_n : n\}$; then

$\{x_n\}$ is a Schauder basis for it. If $\{P_n\}$ are the basis projections for $\{x_n\}$, let $T_n = TP_n$. Then T_n is finite-rank, so compact. Given $x = \sum_{j=1}^\infty c_j x_j$ with $\|x\| = 1$, by Proposition 9.8.2

$$|c_j| = |e_j^*(x)| \le 2b_X < 4,$$

and then

$$\|(T - T_n)x\| = \left\| \sum_{j=n+1}^\infty c_j T x_j \right\| \le \sum_{j=n+1}^\infty |c_j| \, \|Tx_j\| \le \sum_{j=n+1}^\infty 2^{-j+2} = 2^{-n+2}.$$

So $\|T - T_n\| \to 0$ and $T|_{\mathcal{X}_1}$ is compact.

Conversely, suppose that T is not strictly singular. Then there exists an infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$ with T bounded below on $\mathcal{X}_0$. This makes T bounded below on any infinite-dimensional subspace of $\mathcal{X}_0$, and hence it cannot be compact there.

(9.8.15) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. Show that

(i) $\mathcal{SS}(\mathcal{X}, \mathcal{Y})$ is a norm-closed subspace;

(ii) if $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$ and $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $R \in \mathcal{B}(\mathcal{Z}, \mathcal{X})$, then $ST \in \mathcal{SS}(\mathcal{X}, \mathcal{Z})$ and $TR \in \mathcal{SS}(\mathcal{Z}, \mathcal{X})$;

(iii) $\mathcal{K}(\mathcal{X}, \mathcal{Y}) \subset \mathcal{SS}(\mathcal{X}, \mathcal{Y})$.

Answer.

(i) Given nonzero $\lambda \in \mathbb{C}$ and $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$, it is clear that T is bounded below on a subspace if and only if λT is, so $\lambda T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$. And the operator 0 is strictly singular, so $\lambda = 0$ works too. If $S, T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{X}_0 \subset \mathcal{X}$ is an infinite-dimensional subspace, by Exercise 9.8.14 there exists an infinite-dimensional subspace $\mathcal{X}_1 \subset \mathcal{X}_0$ such that $T|_{\mathcal{X}_1}$ is compact. Applying again Exercise 9.8.14 but now to S and $\mathcal{X}_1$, there exists an infinite-dimensional subspace $\mathcal{X}_2 \subset \mathcal{X}_1$ such that $S|_{\mathcal{X}_2}$ is compact. Then $(T + S)|_{\mathcal{X}_2}$ is compact, and so by Exercise 9.8.14 we get that $T + S \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$.

It remains to show that $\mathcal{SS}(\mathcal{X}, \mathcal{Y})$ is closed. Suppose that $\{T_n\} \subset \mathcal{SS}(\mathcal{X}, \mathcal{Y})$ is Cauchy. By $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ being a Banach space there exists $T = \lim T_n$. We want to show that $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$. By passing to a subsequence we may assume that $\|T - T_n\| < 2^{-n-1}$. Fix $\mathcal{X}_0 \subset \mathcal{X}$, infinite-dimensional subspace. For each n, as T_n is strictly singular there exists $x_n \in \mathcal{X}_0$ with $\|x_n\| = 1$ and $\|T_n x_n\| < 2^{-n-1}$. Then

$$\|Tx_n\| \le \|(T - T_n)x_n\| + \|Tx_n\| \le 2^{-n-1} + 2^{-n-1} = 2^{-n}.$$

So T is not bounded below in $\mathcal{X}_0$. As this can be done for any $\mathcal{X}_0$, $T \in \mathcal{SS}(\mathcal{X},\mathcal{Y})$, and thus $\mathcal{SS}(\mathcal{X},\mathcal{Y})$ is closed.

(ii) If $T \in \mathcal{SS}(\mathcal{X},\mathcal{Y})$ and $S \in \mathcal{B}(\mathcal{Y},\mathcal{Z})$, fix an infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$. Then there exists $\{x_n\} \subset \mathcal{X}_0$ with $\|x_n\| = 1$ for all n and $Tx_n \to 0$, since T is not bounded below. Then $STx_n \to 0$, so $ST \in \mathcal{SS}(\mathcal{X},\mathcal{Y})$. Suppose that $TR \notin \mathcal{SS}(\mathcal{Z},\mathcal{X})$; then there exist an infinite-dimensional subspace $\mathcal{Z}_0 \subset \mathcal{Z}$ and $c > 0$ such that $\|TRz\| \geq c\|z\|$ for all $z \in \mathcal{Z}_0$. So R is injective on $\mathcal{Z}_0$ and $\dim RS\mathcal{Z}_0 = \infty$. We have

$$\|T(Rzx)\| \geq c\|z\| \geq \frac{c}{\|R\|}\,\|Rz\|,$$

so T is bounded below on $R\mathcal{Z}_0$, a contradiction. Hence $TR \in \mathcal{SS}(\mathcal{Z},\mathcal{X})$.

(iii) That $\mathcal{K}(\mathcal{X},\mathcal{Y}) \subset \mathcal{SS}(\mathcal{X},\mathcal{Y})$ is Exercise 9.6.9.

(9.8.16) Show that $\mathcal{B}(c_0)$ has a unique non-trivial (closed, double-sided) ideal.

Answer. Let $\mathcal{J} \subset \mathcal{B}(c_0)$ be an ideal, and let $T \in \mathcal{J} \setminus \mathcal{SS}(c_0)$. By definition, this means that there exists an infinite-dimensional subspace $\mathcal{X}_0 \subset c_0$ such that $T|_{\mathcal{X}_0}$ is bounded below. By Proposition 9.8.21 there exists an infinite-dimensional subspace $\mathcal{X}_1 \subset \mathcal{X}_0$ that is isomorphic to c_0. As $T|_{\mathcal{X}_1}$ is bounded below, $\operatorname{ran} T|_{\mathcal{X}_1}$ is closed (Exercise 9.4.1) and so $T|_{\mathcal{X}_1}$ is invertible onto its range by the Inverse Mapping Theorem (6.3.6). Then $\operatorname{ran} T|_{\mathcal{X}_1} \simeq \mathcal{X}_1 \simeq c_0$, and so $\operatorname{ran} T|_{\mathcal{X}_1}$ is complemented by Corollary 9.8.34. Let $P \in \mathcal{B}(c_0)$ be a bounded projection onto $\operatorname{ran} T|_{\mathcal{X}_1}$ and $S \in \mathcal{B}(c_0)$ be given by $Sx = (T|_{\mathcal{X}_1})^{-1}Px$. Let $R : c_0 \to \mathcal{X}_1$ be an isomorphism, so $R \in \mathcal{B}(c_0)$. And let $R' \in \mathcal{B}(c_0)$ be given by $R'x = R^{-1}Qx$, where $Q \in \mathcal{B}(c_0)$ is a projection onto $\mathcal{X}_1$. Then $R'STR \in \mathcal{J}$. But, as $Rx \in \mathcal{X}_1$,

$$R'STRx = R'(T|_{\mathcal{X}_1})^{-1}TRx = R'Rx = R^{-1}Rx = x = I_{c_0}x,$$

so $I_{c_0} \in \mathcal{J}$ and then $\mathcal{J} = \mathcal{B}(c_0)$.

We have thus shown that any closed proper ideal of $\mathcal{B}(c_0)$ satisfies $\mathcal{J} \subset \mathcal{SS}(c_0)$. As $\mathcal{J}$ necessarily contains the finite-rank operators, we get from Propositions 9.8.5 and 9.8.30 and Exercise 9.6.7 that

$$\mathcal{K}(c_0) = \overline{\mathcal{F}(c_0)} \subset \mathcal{J} \subset \mathcal{SS}(c_0) = \mathcal{K}(c_0).$$

Hence $\mathcal{J} = \mathcal{SS}(c_0) = \mathcal{K}(c_0) = \overline{\mathcal{F}(c_0)}$ is the unique non-trivial closed ideal in $\mathcal{B}(c_0)$.

(9.8.17) Show that if $\sum_k x_k$ and $\sum_k y_k$ converge unconditionally, then so does $\sum_k (ax_k + y_k)$ for $a \in \mathbb{C}$.

Answer. Let $\varepsilon > 0$, and let x, y be the limits of the series respectively By hypothesis there exist finite sets F_0, G_0, such that for any finite sets $F \supset F_0$, $G \supset G_0$,

$$\Big\|x - \sum_{k\in F} x_k\Big\| < \frac{\varepsilon}{2|a|} \qquad \Big\|y - \sum_{k\in G} x_k\Big\| < \frac{\varepsilon}{2}.$$

Then, if $H \supset F_0 \cup G_0$,

$$\Big\|ax + y - \sum_{k\in H}(ax_k + y_k)\Big\| \le |a|\Big\|x - \sum_{k\in H} x_k\Big\| + \Big\|y - \sum_{k\in H} y_k\Big\| < \frac{|a|\varepsilon}{2|a|} + \frac{\varepsilon}{2} = \varepsilon.$$

(9.8.18) Let $\mathcal{X}$ be a Banach space, and $\{x_k\} \subset \mathcal{X}$. Show that if $\sum_k x_k$ converges absolutely, then it converges unconditionally.

Answer. Fix $\varepsilon > 0$. Then there exists n_0 such that $\sum_{k>n_0} \|x_k\| < \varepsilon$. Let $F_0 = \{1, \ldots, n_0\}$. For any $F_1 \subset \mathbb{N} \setminus F_0$,

$$\Big\|\sum_{k\in F_1} x_k\Big\| \le \sum_{k\in F_1} \|x_k\| \le \sum_{k>n_0} \|x_k\| < \varepsilon.$$

So the series converges unconditionally.

(9.8.19) Let $\mathcal{X}$ be a finite-dimensional Banach space, and $\{x_k\} \subset \mathcal{X}$. Show that if $\sum_k x_k$ converges unconditionally, then it converges absolutely.

Answer. Consider first the case where $\mathcal{X} = \mathbb{R}$. Let $G = \{k : \ x_k \ge 0\}$. Then

$$\sum_k x_k^+ = \sum_{k\in G} x_k$$

converges by (v) in Proposition 9.8.27. Similarly, $\sum_k x_k^-$ converges. Then

$$\sum_k |x_k| = \sum_k x_k^+ + x_k^-$$

converges. When $\mathcal{X} = \mathbb{C}$, we have $|\operatorname{Re} x_k| \le |x_k|$; then $\sum_k \operatorname{Re} x_k$ converges unconditionally, and similarly for $\sum_k \operatorname{Im} x_k$. Then

$$\sum_k x_k = \sum_k \operatorname{Re} x_k + i \sum_k \operatorname{Im} x_k$$

converges unconditionally. Finally, consider the case where $\mathcal{X}$ is an arbitrary finite-dimensional Banach space. By Theorem 5.2.2 we may choose a norm that suits us. Let $e_1, \ldots, e_n$ be a basis for $\mathcal{X}$, and consider the norm

$$\Big\| \sum_{j=1}^n a_j e_j \Big\| = \sum_{j=1}^n |a_j|.$$

Let us write

$$x_k = \sum_{j=1}^n a_{k,j} e_j, \qquad x = \sum_{j=1}^n a_j e_j.$$

Using the definition of unconditional convergence, we will then have

$$\Big| a_h - \sum_{k \in F} a_{k,h} \Big| \le \sum_{j=1}^n \Big| a_j - \sum_{k \in F} a_{k,h} \Big| = \Big\| \sum_{j=1}^n a_j e_j - \sum_{k \in F} a_{k,j} e_j \Big\|$$
$$= \Big\| x - \sum_{k \in F} x_k \Big\| < \varepsilon.$$

Thus the series $\sum_k a_{k,j}$ is unconditionally convergent for each j. By the previous part, it is absolutely convergent. Thus $\displaystyle\sum_k |a_{k,j}| < \infty$ and then

$$\sum_k \|x_k\| = \sum_k \sum_{j=1}^n |a_{k,j}| = \sum_{j=1}^n \sum_k |a_{k,j}| < \infty.$$

(9.8.20) Let $p \in (1, \infty)$. Find a series in $\ell^p(\mathbb{N})$ that converges unconditionally but not absolutely.

Answer. Consider the sequence $\{z_n\}$, where $z_n = \displaystyle\sum_{k=1}^n \frac{1}{k} e_k$. Then the sequence converges to

$$z = \sum_k \frac{1}{k} e_k,$$

since

$$\|z - z_n\|_p^p = \sum_{k > n} \frac{1}{k^p} \xrightarrow[n \to \infty]{} 0.$$

But the series does not converge absolutely, for

$$\sum_k \left\|\frac{1}{k}\,e_k\right\|_p = \sum_k \frac{1}{k} = \infty.$$

9.9. A Brief Excursion into Injectivity

(9.9.1) Let $\mathcal{X}$ be a Banach space and $\mathcal{Y} \subset \mathcal{X}$ a closed subspace that is not complemented. Let $\mathcal{Z} = (\mathcal{X} \oplus_1 \mathcal{X})/\mathcal{K}$, where $\mathcal{K} = \{(y, -y) :\ y \in \mathcal{Y}\}$. Show that $\mathcal{X} \oplus 0$ and $0 \oplus \mathcal{X}$ are complemented in $\mathcal{Z}$ but their intersection is not.

Answer. To be considered as a subspace of $\mathcal{Z}$, by $\mathcal{X}\oplus 0$ we mean the subspace $\{(x + \mathcal{K}, 0) :\ x \in \mathcal{X}\} \subset \mathcal{Z}$; and similarly for $0 \oplus \mathcal{X}$.

Let $P : \mathcal{Z} \to \mathcal{X} \oplus 0$ be given by $P((a, b) + \mathcal{K}) = (a + b, 0) + \mathcal{K}$. To see that P is well-defined, for all $y \in \mathcal{Y}$ we have that $(a + y) + (b - y) = a + b$. The linearity is automatic. And

$$\begin{aligned}\|P((a, b) + \mathcal{K})\| &= \|(a + b, 0) + \mathcal{K}\| = \inf\{\|(a + b + y, -y)\| :\ y \in \mathcal{K}\}\\ &= \inf\{\|a + b + y\| + \|y\| :\ y \in \mathcal{Y}\} = \|a + b\|\\ &\leq \inf\{\|a + y\| + \|b - y\| :\ y \in \mathcal{Y}\} = \|(a, b) + \mathcal{K}\|.\end{aligned}$$

So P is a norm-one projection. Then $\mathcal{X} \oplus 0 = P\mathcal{Z}$ is complemented and a similar argument shows that $0 \oplus \mathcal{X}$ is also complemented.

Let us look at $(\mathcal{X}\oplus 0)\cap(0\oplus\mathcal{X})$ in $\mathcal{Z}$. These would be the classes $(a, b)+\mathcal{K}$ such that there exist $x_1, x_2 \in \mathcal{X}$ with $(a, b) + \mathcal{K} = (x_1, 0) + \mathcal{K} = (0, x_2) + \mathcal{K}$. That is, $b = x_1 - a \in \mathcal{Y}$, $a = x_2 - b \in \mathcal{Y}$. This can be rephrased as $a, b, x_1, x_2 \in \mathcal{Y}$. In other words,

$$(\mathcal{X} \oplus 0) \cap (0 \oplus \mathcal{X}) = \mathcal{Y} \oplus 0 = 0 \oplus \mathcal{Y}.$$

Suppose that $Q : \mathcal{Z} \to \mathcal{Y} \oplus 0$ is a bounded projection. Fix $x \in \mathcal{X}$ and let $(a, b) + \mathcal{K} = Q((x, 0) + \mathcal{K})$. As $(a, b) + \mathcal{K} \in \mathcal{Y} \oplus 0$, we have $a + b \in \mathcal{Y}$. If we consider two elements of the form $(a, 0) + \mathcal{K} = (a', 0) + \mathcal{K}$, we have $(a - a', 0) \in \mathcal{K}$, which implies that $a - a' = 0$. So there is a unique $Rx \in \mathcal{Y}$ with $Q((x, 0) + \mathcal{K}) = (Rx, 0) + \mathcal{K}$. The uniqueness gives us $R^2 = R$ and, together with the linearity of Q, it forces R to be linear. Also, by the triangle inequality

$$\|(Rx,0)+\mathcal{K}\| = \inf\{\|Rx+y\|+\|y\| : \ y\in\mathcal{Y}\} = \|Rx\|.$$

So

$$\begin{aligned}\|Rx\| &= \|(Rx,0)+\mathcal{K}\| = \|Q((x,0)+k)\| \\ &\leq \|Q\|\,\|(x,0)+\mathcal{K}\| = \|Q\|\,\inf\{\|x+y\|+\|y\| : \ y\in\mathcal{Y}\} \\ &= \|Q\|\,\|x\|\end{aligned}$$

and R is bounded. This would make $\mathcal{Y}$ complemented in $\mathcal{X}$, a contradiction. Thus $(\mathcal{X}\oplus 0)\cap(0\oplus\mathcal{X})$ is not complemented in $\mathcal{Z}$.

(9.9.2) Let $\mathcal{X}$ be a Banach space, $(j,\mathcal{J})$ an injective envelope for $\mathcal{X}$, and $\mathcal{K}$ a Banach space with $g:\mathcal{J}\to\mathcal{K}$ an isometric isomorphism. Show that $(g\circ j,\mathcal{K})$ is an injective envelope for $\mathcal{X}$.

Answer. Let $\psi:\mathcal{K}\to\mathcal{K}$ be a contractive linear map with $\psi\circ g\circ j = g\circ j$. Then $g^{-1}\circ\psi\circ g:\mathcal{J}\to\mathcal{J}$ is a contractive linear map with

$$(g^{-1}\circ\psi\circ g)\circ j = g^{-1}\circ(\psi\circ g\circ j) = g^{-1}\circ g\circ j = j.$$

As $(j,\mathcal{J})$ is an injective envelope for $\mathcal{X}$, we get that $g^{-1}\circ\psi\circ g = \mathrm{id}_{I(\mathcal{X})}$. Then $\psi = g\circ g^{-1} = \mathrm{id}_{\mathcal{K}}$ and hence $(g\circ j,\mathcal{K})$ is an injective envelope for $\mathcal{X}$.

(9.9.3) Show that i, as in the proof of Theorem 9.9.16, is a linear isometry.

Answer. We have $\mathcal{X}=\ell^\infty(\mathbb{N})$, $\mathcal{Y}=L^\infty[0,1]$, with $i:\mathcal{X}\to\mathcal{Y}$ given by

$$i(x) = \sum_{n=1}^\infty x_n\,1_{\left[\frac{1}{2n+1},\frac{1}{2n}\right]}.$$

Given $x,y\in\mathcal{X}$,

$$\begin{aligned} i(x+y)\,1_{\left[\frac{1}{2n+1},\frac{1}{2n}\right]} &= (x_n+y_n)\,1_{\left[\frac{1}{2n+1},\frac{1}{2n}\right]} \\ &= i(x)\,1_{\left[\frac{1}{2n+1},\frac{1}{2n}\right]} + i(y)\,1_{\left[\frac{1}{2n+1},\frac{1}{2n}\right]}.\end{aligned}$$

This works for all n, so $i(x+y)=i(x)+i(y)$. Multiplication by scalars works similarly. For $t\in\left[\frac{1}{2n+1},\frac{1}{2n}\right]$, we have $i(x)(t)=x_n$. Thus $\|i(x)\|_\infty\leq\|x\|_\infty$.

Given $\varepsilon > 0$, there exists n with $|x_n| \geq \|x\|_\infty - \varepsilon$. Then, with $t \in \left[\frac{1}{2n+1}, \frac{1}{2n}\right]$,

$$\|x\|_\infty - \varepsilon \leq |x_n| = i(x)(t) \leq \|i(x)\|_\infty.$$

As this can be done for all $\varepsilon > 0$, we obtain that $\|i(x)\|_\infty = \|x\|_\infty$ for all $x \in \mathcal{X}$.

CHAPTER

10

Bounded operators on a Hilbert space: Part I

10.1. Adjoints

(10.1.1) Prove the uniqueness of the adjoint of $T \in \mathcal{B}(\mathcal{H})$.

Answer. Suppose that R, S are adjoints for T. This means that

$$\langle \xi, R\eta \rangle = \langle T\xi, \eta \rangle = \langle \xi, S\eta \rangle, \qquad \xi, \eta \in \mathcal{H}.$$

So $\langle R\eta - S\eta, \xi \rangle = 0$ for all $\xi \in \mathcal{H}$. Then $R\eta = S\eta$ by Lemma 10.1.1. And this occurs for all $\eta \in \mathcal{H}$, so $R = S$.

(10.1.2) Let $\mathcal{H} = \mathbb{C}^2$. Let $\{\xi_1, \xi_2\}$ be the canonical basis, and T the operator induced $T\xi_1 = 0$, $T\xi_2 = \xi_1$.

(i) Find the matrix form of T and T^* with respect to the canonical basis. Confirm that T^* is the conjugate transpose.

(ii) Find the matrix form of T and T^* with respect to the basis $\eta_1 = \xi_1$, $\eta_2 = \xi_1 + \xi_2$. Is the matrix of T^* the conjugate transpose of the matrix of T? Why?

Answer.

(i) When we represent an operator as a matrix with respect to a basis, the columns are the coefficients of the images of each element of the basis. Let us first find T^*. We have $T\xi_1 = 0$, $T\xi_2 = \xi_1$, so

$$\begin{aligned}\langle T^*(\alpha\xi_1 + \beta\xi_2), \gamma\xi_1 + \delta\xi_2\rangle &= \langle \alpha\xi_1 + \beta\xi_2, T(\gamma\xi_1 + \delta\xi_2)\rangle \\ &= \langle \alpha\xi_1 + \beta\xi_2, \delta\xi_1\rangle \\ &= \alpha\bar{\delta} = \langle \alpha\xi_2, \gamma\xi_1 + \delta\xi_2\rangle,\end{aligned}$$

so $T^*\xi_1 = \xi_2$, $T^*\xi_2 = 0$. Hence

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad T^* = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

(ii) We now have $T\eta_1 = T\xi_1 = 0$, and $T\eta_2 = T(\xi_1+\xi_2) = \xi_2 = -\eta_1+\eta_2$. Also, $T^*\eta_1 = T^*\xi_1 = \xi_2 = -\eta_1 + \eta_2$, and $T^*\eta_2 = T^*(\xi_1 + \xi_2) = \xi_2 = -\eta_1 + \eta_2$. Therefore

$$T = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}, \qquad T^* = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}.$$

The matrix for T^* is not the conjugate transpose of the matrix of T, because we are dealing with a basis that is not orthonormal.

(10.1.3) Let $\mathcal{H}$ a Hilbert space and $\mathcal{H}_0 \subset \mathcal{H}$ a dense subspace. Let $B : \mathcal{H}_0 \times \mathcal{H}_0 \to \mathbb{C}$ be a bounded sesquilinear form. Show that B admits a unique extension to a bounded sesquilinear form $\tilde{B} : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, with $\|\tilde{B}\| = \|B\|$.

Answer. Let $\xi, \eta \in \mathcal{H}$ and $\{\xi_n\}, \{\eta_n\} \subset \mathcal{H}_0$ sequences with $\xi_n \to \xi$ and $\eta_n \to \eta$. Let $c > 0$ with $\|\xi_n\| \leq c$ and $\|\eta_n\| \leq c$ for all n. We have

$$\begin{aligned}|B(\xi_n, \eta_n) - B(\xi_m, \eta_m)| &\leq |B(\xi_n, \eta_n) - B(\xi_n, \eta_m)| + |B(\xi_n, \eta_m) - B(\xi_m, \eta_m)| \\ &= |(B(\xi_n, \eta_n - \eta_m)| + |B(\xi_n - \xi_m, \eta_m)| \\ &\leq \|B\|\,\|\xi_n\|\,\|\eta_n - \eta_m\| + \|B\|\,\|\eta_m\|\,\|\xi_n - \xi_m\| \\ &\leq c\,\|B\|\,\big(\|\eta_n - \eta_m\| + \|\xi_n - \xi_m\|\big).\end{aligned}$$

It follows that the sequence of numbers $\{B(\xi_n,\eta_n)\}$ is Cauchy, so its limit exists and we denote this limit by $\tilde{B}(\xi,\eta)$. This is well-defined, in the sense that if $\xi_n'\to\xi$ and $\eta_n'\to\eta$, then the same estimate as above gives us

$$|B(\xi_n,\eta_n)-B(\xi_n',\eta_n')|\leq c\,\|B\|\left(\|\eta_n-\eta_n'\|+\|\xi_n-\xi_n'\|\right)$$

and so the limit does not depend on the sequences and only on ξ and η. The sesquilinearity of $\tilde{B}$ is now automatic, for we have

$$\begin{aligned}\tilde{B}(\xi+\lambda\nu,\eta)&=\lim_n B(\xi_n+\lambda\nu_n,\eta_n)\\&=\lim_n B(\xi_n,\eta_n)+\lambda\lim_n B(\nu_n,\eta_n)\\&=B(\xi,\eta)+\lambda B(\nu,\eta)\end{aligned}$$

and a similar computation for the conjugate linearity on the second factor. The fact that $\tilde{B}$ extends B is a consequence of the independence of the limit, for given ξ and or η in $\mathcal{H}_0$ we can choose the respective constant sequences for them. Finally,

$$|\tilde{B}(\xi,\eta)|=\lim_n|B(\xi_n,\eta_n)|\leq\limsup_n\|B\|\,\|\xi_n\|\,\|\eta_n\|=\|B\|\,\|\xi\|\,\|\eta\|.$$

As $\tilde{B}$ extends B, this shows that $\|\tilde{B}\|=\|B\|$.

(10.1.4) Let $T\in\mathcal{B}(\mathcal{H})$, where $\dim\mathcal{H}<\infty$. Show that $T^*T=I$ if and only if $TT^*=I$.

Answer. Since $T^*T=I$, T is injective. Indeed, if $T\xi=0$, then $\xi=T^*T\xi=0$. So T maps an orthonormal basis to a basis, which means that $\dim\operatorname{ran}T=\dim\mathcal{H}$. Thus T is surjective. As T is invertible, from $T^*T=I$ we get $T^{-1}=T^*TT^{-1}=T^*$. So $T^*=T^{-1}$ and in particular $TT^*=TT^{-1}=I$.

The argument for $TT^*=I$ is the same, since this can be seen as $S^*S=I$, where $S=T^*$. And T^* is invertible if and only if T is.

(10.1.5) Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces of the same dimension, and $\{\xi_j\}_{j\in J}$ and $\{\eta_j\}_{j\in J}$ orthonormal bases for $\mathcal{H}$ and $\mathcal{K}$ respectively. Show that the assignment $U:\xi_j\longmapsto f_j$ induces a unique bounded linear operator $U\in\mathcal{B}(\mathcal{H},\mathcal{K})$ and that U is a unitary.

Answer. Since U has to be linear we have, for each finite $F\subset J$,

$$U\Big(\sum_{j\in F}c_j\,\xi_j\Big)=\sum_{j\in F}c_j\,\eta_j.$$

Then

$$\Big\|U\Big(\sum_{j\in F} c_j\,\xi_j\Big)\Big\| = \sum_{j\in F} |c_j|^2 = \Big\|\sum_{j\in F} c_j\,\eta_j\Big\|.$$

So U is isometric on the dense subspace $\text{span}\{e_j : j \in J\}$ and therefore it extends uniquely to $U \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ by Proposition 6.1.9. The extension will also be isometric by continuity. Finally, given $\sum_j c_j f_j \in \mathcal{K}$, we have

$$\sum_j c_j f_j = U\Big(\sum_j c_j e_j\Big),$$

so U is surjective. That is, U is a unitary.

(10.1.6) Let $K \subset \mathcal{H}$ be a closed subspace and $i : K \to \mathcal{H}$ the inclusion. Show that i^* is the orthogonal projection onto K.

Answer. Since $i : K \to \mathcal{H}$, we have $i^* : \mathcal{H} \to K$. Let $\xi \in K^\perp$. Then for any $\eta \in K$

$$\langle i^*\xi, \eta\rangle = \langle \xi, i(\eta)\rangle = \langle \xi, \eta\rangle = 0.$$

So $i^* = 0$ on $K^\perp$. If $\xi \in K$ and $\eta \in K$,

$$\langle i^*\xi, \eta\rangle = \langle \xi, \eta\rangle.$$

So $\langle \xi - i^*\xi, \eta\rangle = 0$ for all $\eta \in K$, which shows that $\xi - i^*\xi \in K^\perp$. Then $\xi - i^*\xi \in K \cap K^\perp = \{0\}$, showing that $i^*\xi = \xi$ for all $\xi \in K$.

(10.1.7) Let $\mathcal{H}$ be a Hilbert space and $\xi, \eta \in \mathcal{H}$ with $\|\xi\| = \|\eta\|$. Show that there exists a unitary $U \in \mathcal{B}(\mathcal{H})$ such that $U\xi = \eta$.

Answer. Let $\{\xi_j\}$ be an orthonormal basis such that $\xi_{j_0} = \xi$ for some j_0 (for instance, complete ξ to an orthonormal basis by choosing an orthonormal basis of $\{\xi\}^\perp$) and let $\{\eta_j\}$ another orthonormal basis with $\eta_{j_0} = \eta$. Then the unitary induced by $U\xi_j = \eta_j$ has $U\xi = \eta$.

(10.1.8) Let $\mathcal{H}$ be a Hilbert space, and consider the direct sum $\mathcal{H} \oplus \mathcal{H}$. But instead of considering the natural Hilbert space norm $\|(\xi, \eta)\| = (\|\xi\|^2 + \|\eta\|^2)^{1/2}$, we consider the norm $\|(\xi, \eta)\|_1 = \|\xi\| + \|\eta\|$. Since these two norms are equivalent, $\mathcal{H} \oplus_1 \mathcal{H}$ (with $\|\cdot\|_1$) is a Banach space. Show that there exist elements $(\xi, \eta), (\xi', \eta') \in \mathcal{H} \oplus_1 \mathcal{H}$ with $\|(\xi, \eta)\| = \|(\xi', \eta')\|$ but such that no linear isometry maps $(\xi,\eta)\longmapsto(\xi',\eta')$ (compare with Exercise 7.5.19).

Answer. We will first characterize the linear isometries on $\mathcal{H}\oplus_1\mathcal{H}$. Let V be such an isometry. Being a linear operator on $\mathcal{H}\oplus\mathcal{H}$, we can think of V as a 2×2 matrix of operators. That is,

$$V=\begin{bmatrix}A & B\\ C & D\end{bmatrix},$$

with $A,B,C,D\in\mathcal{B}(\mathcal{H})$. That is,

$$V(\xi,\eta)=(A\xi+B\eta,C\xi+D\eta),\qquad \xi,\eta\in\mathcal{H}.$$

Note that since $\|(\xi,\eta)\|_2\le\|(\xi,\eta)\|_1\le\sqrt{2}\,\|(\xi,\eta)\|_2$, the bounded operators on $\mathcal{H}\oplus\mathcal{H}$ are the same regardless of the norm.

Since V is an isometry, we have

$$\begin{aligned}\|\xi\|+\|\eta\|=\|V(\xi,\eta)\|&=\|(A\xi+B\eta,C\xi+D\eta)\|\\&=\|A\xi+B\eta\|+\|C\xi+D\eta\|.\end{aligned}$$

Taking $\eta=0$,

$$\|\xi\|=\|A\xi\|+\|C\xi\|,\qquad \xi\in\mathcal{H}. \tag{AB.10.1}$$

Taking $\xi=0$,

$$\|\eta\|=\|B\eta\|+\|D\eta\|,\qquad \eta\in\mathcal{H}. \tag{AB.10.2}$$

Taking $\eta=\lambda\xi$ with $\|\xi\|=1$, and $|\lambda|=1$,

$$2=\|(A+\lambda B)\xi\|+\|(C+\lambda D)\xi\|,\qquad \xi\in\mathcal{H}. \tag{AB.10.3}$$

Combining (AB.10.1), (AB.10.2), and (AB.10.3) for $\xi\in\mathcal{H}$ with $\|\xi\|=1$ we have

$$\|A\xi\|+\|B\xi\|+\|C\xi\|+\|D\xi\|\le 2=\|(A+\lambda B)\xi\|+\|(C+\lambda D)\xi\|.$$

This implies equality in both triangle inequalities, so for any $\xi\in\mathcal{H}$ and $|\lambda|=1$

$$\begin{aligned}\|A\xi\|+\|B\xi\|&=\|(A+\lambda B)\xi\|,\\ \|C\xi\|+\|D\xi\|&=\|(C+\lambda D)\xi\|.\end{aligned} \tag{AB.10.4}$$

We now work with A,B since the computations for C,D are entirely analogous. Squaring, expanding, and cancelling square norms in (AB.10.4),

$$\operatorname{Re}\lambda\langle B\xi,A\xi\rangle=\|B\xi\|\,\|A\xi\|,\qquad \xi\in\mathcal{H},\ |\lambda|=1.$$

By using $\lambda\in\mathbb{T}$ such that $\lambda\langle B\xi,A\xi\rangle=|\langle B\xi,A\xi\rangle|$ we get that

$$\|B\xi\|\,\|A\xi\|=0,\qquad \xi\in\mathcal{H}.$$

and analogously

$$\|C\xi\|\,\|D\xi\|=0,\qquad \xi\in\mathcal{H}. \tag{AB.10.5}$$

When $A\xi = 0$, we obtain from (AB.10.1) that $\|C\xi\| = \|\xi\|$; then (AB.10.5) implies that $D\xi = 0$. Similarly, when $B\xi = 0$ we get that $\|D\xi\| = \|\xi\|$ and then $C\xi = 0$. So either

$$\|A\xi\| = \|D\xi\| = 0, \qquad \|C\xi\| = \|B\xi\| = \|\xi\|, \tag{AB.10.6}$$

or

$$\|A\xi\| = \|D\xi\| = \|\xi\|, \qquad \|C\xi\| = \|B\xi\| = 0. \tag{AB.10.7}$$

Suppose that (AB.10.6) occurs for a certain ξ and (AB.10.7) for η, both nonzero. For $\xi + \eta$, we have

$$\|A(\xi+\eta)\| = \|A\eta\| = \|\eta\|, \qquad \|B(\xi+\eta)\| = \|B\xi\| = \|\xi\|,$$

so $\xi+\eta$ satisfies neither (AB.10.6) nor (AB.10.7). This proves that V satisfies either (AB.10.6) or (AB.10.7) for all ξ. In other words, the possibilities for V are

$$V = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}, \qquad \text{or} \qquad V = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}$$

with $A, B, C, D \in B(\mathcal{H})$ isometries.

Now it is easy to find the counterexample. Fix $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ and consider the elements $(\xi, 0)$ and $(\xi/2, \xi/2)$. With V of the first form we need to have

$$(\xi/2, \xi/2) = V(\xi, 0) = (A\xi, 0),$$

and this forces $\xi = 0$. With V of the second form the problem is the same:

$$(\xi/2, \xi/2) = V(\xi, 0) = (0, C\xi)$$

and we get $\xi = 0$. So no linear isometry (surjective or not) can map $(\xi, 0)$ to $(\xi/2, \xi/2)$.

10.2. Numerical Range and Numerical Radius

(10.2.1) Let $T \in \mathcal{B}(\mathcal{H})$. Show that $W(T^*) = \{\bar{\lambda} : \lambda \in W(T)\}$, and $\omega(T) = \omega(T^*)$.

Answer. We have $\langle T^*\xi, \xi\rangle = \overline{\langle T\xi, \xi\rangle}$, which gives us the first equality. As the two sets $W(T)$ and $W(T^*)$ contain the conjugates of the other set,

$$\omega(T^*) = \sup\{|\lambda| : \lambda \in W(T^*)\} = \sup\{|\lambda| : \lambda \in W(T)\} = \omega(T).$$

(10.2.2) Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Show that $\omega(T)$ defines a norm on $\mathcal{B}(\mathcal{H})$, equivalent to the operator norm; concretely,

$$\omega(T) \leq \|T\| \leq 2\omega(T).$$

The second inequality requires Proposition 10.3.3 and a basic knowledge of selfadjoint operators.

Answer. We have $\omega(T) \geq 0$ for all T by definition. If $\omega(T) = 0$, then $\langle T\xi, \xi\rangle = 0$ for all $\xi \in \mathcal{H}$, and then $T = 0$ by Lemma 10.1.1.

Also,

$$\omega(\lambda T) = \sup\{|\langle \lambda T\xi, \xi\rangle| : \ \xi \in \mathcal{H}, \ \|\xi\| = 1\} = |\lambda|\, \omega(T).$$

From $|\langle (T_1 + T_2)\xi, \xi\rangle| \leq |\langle T_1\xi, \xi\rangle| + |\langle T_2\xi, \xi\rangle|$,

$$\omega(T_1 + T_2) \leq \omega(T_1) + \omega(T_2).$$

As for the inequalities, from $|\langle T\xi, \xi\rangle| \leq \|T\|\,\|\xi\|^2$ we get $\omega(T) \leq \|T\|$. Since $\operatorname{Re} T$ is selfadjoint,

$$\|\operatorname{Re} T\| = \sup\{|\langle \operatorname{Re} T\,\xi, \xi\rangle| : \ \|\xi\| = 1\} = \omega(\operatorname{Re} T).$$

Similarly, $\|\operatorname{Im} T\| = \omega(\operatorname{Im} T)$. Also, from

$$|\langle \operatorname{Re} T\xi, \xi\rangle| = |\operatorname{Re}\langle T\xi, \xi\rangle| \leq |\langle T\xi, \xi\rangle|$$

we get $\omega(\operatorname{Re} T) \leq \omega(T)$ and similarly $\omega(\operatorname{Im} T) \leq \omega(T)$. Then

$$\|T\| = \|\operatorname{Re} T + i\operatorname{Im} T\| \leq \|\operatorname{Re} T\| + \|\operatorname{Im} T\|$$
$$= \omega(\operatorname{Re} T) + \omega(\operatorname{Im} T) \leq 2\omega(T).$$

(10.2.3) Let $B \in \mathcal{B}(\mathcal{H})^{\mathrm{sa}}$ such that $\|I_{\mathcal{H}} + iB\| \leq 1$. Show that $B = 0$.

Answer. Let $\xi \in \mathcal{H}$ with $\|\xi\| = 1$. Then

$$|1 + i\langle B\xi, \xi\rangle| = |\langle (I_{\mathcal{H}} + iB)\xi, \xi\rangle| \leq \|I_{\mathcal{H}} + iB\| \leq 1.$$

Hence $\langle B\xi, \xi\rangle = 0$. As ξ was arbitrary (after scaling), polarization gives us that $B = 0$.

10.3. Selfadjoint Operators

(10.3.1) Let $T \in \mathcal{B}(\mathcal{H})$ and $\alpha, \beta \in \mathbb{C}$. Show that $W(\alpha T + \beta I) = \alpha W(T) + \beta$.

Answer. If $\|\xi\| = 1$,

$$\langle (\alpha T + \beta I)\xi, \xi \rangle = \alpha \langle T\xi, \xi \rangle + \beta.$$

(10.3.2) Let $T \in \mathcal{B}(\mathcal{H})$ and $\lambda \in \mathbb{C}$. Show that T is normal if and only if $T - \lambda I$ is normal.

Answer. Suppose that T is normal. Then

$$\begin{aligned}(T-\lambda I)^*(T-\lambda I) &= (T^* - \bar{\lambda} I)(T - \lambda I) = T^*T + |\lambda|^2\, I - 2\mathrm{Re}\,\lambda T^* \\ &= TT^* + |\lambda|^2\, I - 2\mathrm{Re}\,\lambda T^* = (T-\lambda I)(T^* - \bar{\lambda} I) \\ &= (T-\lambda I)(T-\lambda I)^*.\end{aligned}$$

So $T - \lambda I$ is normal. Conversely, if we know that $T - \lambda I$ is normal, then by the above computation $T = (T - \lambda I) - (-\lambda)I$ is normal.

(10.3.3) Prove Proposition 10.1.8.

Answer. For any $\xi, \eta \in \mathcal{H}$,

$$\begin{aligned}\langle (T+\alpha S)^*\xi, \eta \rangle &= \langle \xi, (T+\alpha S)\eta \rangle = \langle \xi, T\eta \rangle + \bar{\alpha}\langle \xi, S\eta \rangle \\ &= \langle T^*\xi, \eta \rangle + \langle \bar{\alpha} S^*\xi, \eta \rangle = \langle (T^* + \bar{\alpha} S^*)\xi, \eta \rangle,\end{aligned}$$

so $(T + \alpha S)^* = T^* + \bar{\alpha} S^*$. Also,

$$\langle (T^*)^*\xi, \eta \rangle = \langle \xi, T^*\eta \rangle = \langle T\xi, \eta \rangle,$$

so $(T^*)^* = T$. For TS,

$$\langle (TS)^*\xi, \eta \rangle = \langle \xi, TS\eta \rangle = \langle S^*T^*\xi, \eta \rangle,$$

showing that $(TS)^* = S^*T^*$. If $ST^* = T^*S = I$, taking adjoints we get $TS^* = S^*T = I$, so T is invertible; and similarly T invertible implies T^* invertible. The same computation shows that $(T^*)^{-1} = S = (T^{-1})^*$. And from this we get that $T - \lambda I$ is invertible if and only if $T^* - \bar{\lambda} I$ is invertible.

(10.3.4) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\xi \in \mathcal{H}$ and $\mathcal{H}_1 \subset \mathcal{H}$ be the subspace

$$\mathcal{H}_1 = \overline{\{p(T,T^*)\xi : \ p \in \mathbb{C}[x,y]\}}.$$

Show that $\mathcal{H}_1^\perp$ is invariant for both T and T^*.

Answer. Fix $\eta \in \mathcal{H}_1^\perp$ and $p \in \mathbb{C}[x,y]$. Then, with $q \in \mathbb{C}[x,y]$ given by $q(x,y) = yp(x,y)$,

$$\langle T\eta, p(T,T^*)\xi\rangle = \langle \eta, T^*p(T,T^*)\xi\rangle = \langle \eta, q(T,T^*)\xi\rangle = 0.$$

So $T\eta \in \mathcal{H}_1^\perp$. The argument for T^* is entirely similar.

(10.3.5) Recall that an operator $T \in \mathcal{B}(\mathcal{H})$ is said to be **bounded below** if there exists $k > 0$ with $\|T\xi\| \geq k\|\xi\|$ for every $\xi \in \mathcal{H}$. Show that the following statements are equivalent:

(i) T is bounded below;

(ii) T admits a left inverse;

(iii) T is injective and has closed range.

(Hint: use the Inverse Mapping Theorem).

Answer. (i) $\implies$ (iii) If $T\xi = 0$, then $\|\xi\| \leq \frac{1}{k}\|T\xi\| = 0$, so T is injective. If $\{T\xi_n\}$ is Cauchy, then

$$\|\xi_n - \xi_m\| \leq \tfrac{1}{k}\|T\xi_n - T\xi_m\|,$$

so $\{\xi_n\}$ is Cauchy. By the completeness of $\mathcal{H}$ there exists $\xi = \lim_n \xi_n$. As T is bounded, $T\xi = \lim_n T\xi_n$. So the range of T is closed.

(iii) $\implies$ (ii) T is bounded and bijective onto its closed range. By the Inverse Mapping Theorem there exists $S : \operatorname{ran} T \to \mathcal{H}$ such that $ST = I$. We can consider $S \in \mathcal{B}(\mathcal{H})$ by taking $S = 0$ on $(\operatorname{ran} T)^\perp$.

(ii) $\implies$ (i) If $S \in \mathcal{B}(\mathcal{H})$ and $ST = I$, then $\|\xi\| = \|ST\xi\| \leq \|S\|\,\|T\xi\|$, so T is bounded below.

(10.3.6) Let T be a diagonal operator. Find T^* and show that T is normal.

Answer. We have $T\xi_j = \alpha_j\xi_j$ for a certain orthonormal basis $\{\xi_j\}$. Then

$$\langle T^*\xi_j, \xi_k\rangle = \langle \xi_j, T\xi_k\rangle = \langle \xi_j, \alpha_j\xi_k\rangle = \bar{\alpha}_j\langle \xi_j, \xi_k\rangle = \langle \bar{\alpha}_j\xi_j, \xi_k\rangle.$$

As this can be done for all j, k, T^* is the multiplication operator by $\{\bar{\alpha}_j\}$. And then

$$\langle T^*T\xi_j, \xi_k\rangle = |\alpha_j|^2\langle \xi_j, \xi_k\rangle = \langle TT^*\xi_j, \xi_k\rangle.$$

(10.3.7) Show that a diagonal operator T is selfadjoint if and only if all its diagonal entries are real.

Answer. If $T = T^*$, from Exercise 10.3.6 we conclude that $\bar{\alpha}_j = \alpha_j$ for all j, so $\alpha_j \in \mathbb{R}$ for all j. Conversely, if $\alpha_j = \bar{\alpha}_j$ for all j then $T^* = T$.

(10.3.8) Let T as in (10.3). Show that T is injective, selfadjoint, with $\sigma(T) = [0, 1]$ and $\sigma_p(T) = \varnothing$.

Answer. This was done in Example 9.5.8. With $g(t) = t$, we have that $\sigma(T) = g([0, 1]) = [0, 1]$ and $\sigma_p(T) = \varnothing$ since g is not constant on any set of positive measure.

(10.3.9) Let $\mathcal{H}$ be a Hilbert space and $\xi, \eta \in \mathcal{H}$.

(i) Show that it is possible to choose ξ, η in such a way that no selfadjoint T satisfies $T\xi = \eta$, even if both are nonzero.

(ii) Show that there exists T normal with $T\xi = \eta$.

Answer.

(i) We can take $\eta = i\xi$, and then the equality $T\xi = \eta = i\xi$ would force i to be an eigenvalue for T, making it impossible for T to be selfadjoint.

(ii) Let $\xi_1 = \xi/\|\xi\|$ and $\alpha = \langle \eta, \xi_1\rangle$, $\beta = \|\eta - \alpha\xi_1\|$. Then $\xi_2 = (\eta - \alpha\xi_1)/\beta$ is a unit vector orthogonal to ξ_1, and $\eta = \alpha\xi_1 + \beta\xi_2$.

We want $T\xi = \eta$, so we must have

$$T\xi_1 = \frac{1}{\|\xi\|}T\xi = \frac{1}{\|\xi\|}(\alpha\xi_1 + \beta\xi_2).$$

This means that, as a 2×2 matrix with respect to the orthonormal basis $\{\xi_1, \xi_2\}$ of $\operatorname{span}\{\xi, \eta\}$, the first column of T is $\alpha/\|\xi\|$, $\beta/\|\xi\|$. This suggests we define

$$T\xi_2 = \frac{\bar\beta}{\|\xi\|}\xi_1 + \frac{\alpha}{\|\xi\|}\xi_2.$$

This way we have $T = \frac{\alpha}{\|\xi\|}I_2 + S$, where S is the operator $S\xi_1 = \frac{\beta}{\|\xi\|}\xi_2$, $S\xi_2 = \frac{\bar\beta}{\|\xi\|}\xi_1$. So

$$S = \frac{1}{\|\xi\|}\begin{bmatrix} 0 & \beta \\ \bar\beta & 0 \end{bmatrix}$$

is selfadjoint. We have written T as a sum of a normal and a selfadjoint, and hence T is normal.

The choices we made above show that if $\langle \eta, \xi \rangle \in \mathbb{R}$ then T can be chosen to be selfadjoint, as $\alpha \in \mathbb{R}$. Said condition is necessary and sufficient for selfadjoint T to exist.

10.4. Positive operators

(10.4.1) Show that if $T : \mathcal{H} \to \mathcal{H}$ is linear and $\langle T\xi, \xi \rangle \geq 0$ for all $\xi \in \mathcal{H}$, then T is bounded.

Answer. By the usual identification between $\mathcal{H}$ and $\mathcal{H}^*$, our operator T satisfies the hypotheses in Proposition 6.3.15. So T is bounded.

Alternatively, we can repeat the argument from Proposition 6.3.15 in the current context. This is not a bad idea, due to the fact that the differences between Banach and Hilbert space adjoints can create the doubt of whether the argument does work or not for Hilbert spaces. Suppose that $\xi_n \to \xi$ and $T\xi_n \to \eta$. If we show that $\eta = T\xi$, then the Closed Graph Theorem guarantees that T is bounded.

Given $\zeta \in \mathcal{H}$,

$$\begin{aligned} 0 &\leq \langle T(\xi_n - \zeta), \xi_n - \zeta \rangle \\ &= \langle T\xi_n, \xi_n \rangle + \langle T\zeta, \zeta \rangle - \langle T\xi_n, \zeta \rangle - \langle T\zeta, \xi_n \rangle. \end{aligned}$$

Taking limit over n, and noting that convergent sequences are bounded (so we can take limits on both arguments of the inner product at once),

$$0 \leq \langle \eta, \xi \rangle + \langle T\zeta, \zeta \rangle - \langle \eta, \zeta \rangle - \langle T\zeta, \xi \rangle = \langle \eta, \xi - \zeta \rangle - \langle T\zeta, \xi - \zeta \rangle.$$

So $\langle T\zeta, \xi - \zeta \rangle \leq \langle \eta, \xi - \zeta \rangle$. Note that ζ was arbitrary; if we take $\nu = \xi - \zeta$, then $\langle T(\xi - \nu), \nu \rangle \leq \langle \eta, \nu \rangle$, which we re-write as

$$\langle T\xi - \eta, \nu \rangle \leq \langle T\nu, \nu \rangle, \qquad \nu \in \mathcal{H}.$$

In particular, $\langle T\xi - \eta, \nu \rangle \in \mathbb{R}$. Given any $\nu \in \mathcal{H}$, the inequality also works for $\pm\nu/n$, which gives us

$$\pm n \langle T\xi - \eta, \nu \rangle \leq \langle T\nu, \nu \rangle, \qquad \nu \in \mathcal{H},\ n \in \mathbb{N}.$$

This forces $\langle T\xi - \eta, \nu \rangle = 0$. As this occurs for all $\nu \in \mathcal{H}$, we have shown that $T\xi = \nu$.

(10.4.2) Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive with $S + T = 0$. Show that $S = T = 0$.

Answer. We have $0 \leq S \leq S + T = 0$, so $S = 0$. Properly, these inequalities show that $\langle S\xi, \xi \rangle = 0$ for all ξ, and then $\|S^{1/2}\xi\|^2 = 0$ for all ξ, which gives $S^{1/2} = 0$ and hence $S = 0$. Or we can use polarization. The same argument works for T.

(10.4.3) Let $A \in \mathcal{B}(\mathcal{H})$ be positive and invertible. Show that $A^{1/2}$ is invertible.

Answer. By Proposition 10.4.4 there exists a unique $B \in \mathcal{B}(\mathcal{H})$ with $B^2 = A$. We have $B(BA^{-1}) = AA^{-1} = I$, so B has a right-inverse. Similarly $(A^{-1}B)B = A^{-1}A = I$ and B has a left inverse. Then B is invertible by Exercise 1.1.7.

(10.4.4) Prove Proposition 10.4.3.

Answer.

(i) We have $\langle A\xi, \xi \rangle \leq \|A\| \, \|\xi\|^2 = \langle \|A\| \, \xi, \xi \rangle$.

(ii) For any $\xi \in \mathcal{H}$,

$$\langle TAT^*\xi, \xi \rangle = \langle AT^*\xi, T^*\xi \rangle \leq \langle BT^*\xi, T^*\xi \rangle = \langle TBT^*\xi, \xi \rangle.$$

(iii) Let $\xi \in \mathcal{H}$. We have, with $\eta = A^{-1}\xi$,
$$\langle A^{-1}\xi, \xi\rangle = \langle \eta, A\eta\rangle \geq 0.$$

(iv) If $A \leq I$, then $\langle A\xi, \xi\rangle \leq \langle \xi, \xi\rangle = \|\xi\|^2$. By Proposition 10.3.3, $\|A\| \leq 1$. The converse is (i).

(v) Suppose that $A \geq cI$ with $c > 0$. Then $\langle A\xi, \xi\rangle \geq \langle c\,\xi, \xi\rangle = c\,\|\xi\|^2$ for all $\xi \in \mathcal{H}$, and so $W(A) \subset [c, \infty)$. By Proposition 10.2.3, $\sigma(A) \subset [c, \infty)$ and so A is invertible. We also have, since $A^{1/2}$ is invertible and $A \geq cI$,
$$\begin{aligned} A^{-1} &= c^{-1/2}A^{-1/2}(cI)(c^{-1/2}A^{-1/2}) \\ &\leq c^{-1/2}A^{-1/2}\,A\,(c^{-1/2}A^{-1/2}) = c^{-1}\,I. \end{aligned}$$

(10.4.5) Let $T, S \in \mathcal{B}(\mathcal{H})$, with S invertible. Show that $T \geq 0$ if and only if $S^*TS \geq 0$.

Answer. If $T \geq 0$, then $S^*TS \geq 0$. Conversely, if $S^*TS \geq 0$ with S invertible, then for any $\xi \in \mathcal{H}$
$$\langle T\xi, \xi\rangle = \langle TS(S^{-1}\xi), SS^{-1}\xi\rangle = \langle S^*TS(S^{-1}\xi), S^{-1}\xi\rangle \geq 0.$$
So $T \geq 0$.

(10.4.6) Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. Show that T^* exists by using block matrices over $\mathcal{H} \oplus \mathcal{K}$ and Theorem 10.1.6. For this, consider $X \in \mathcal{B}(\mathcal{H} \oplus \mathcal{K})$ with $X_{21} = T$ and $X_{11} = 0$, $X_{12} = 0$, $X_{22} = 0$.

Answer. Let $X = \begin{bmatrix} 0 & 0 \\ T & 0 \end{bmatrix} \in \mathcal{B}(\mathcal{H} \oplus \mathcal{K})$. By Theorem 10.1.6, X^* exists. We cannot finish by using Proposition 10.4.12, for we don't know that T^* exists. Write
$$X^* = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$
Since $X = P_2XP_1$ we get $X^* = P_1X^*P_2$, which tells us that only $(X^*)_{12}$ is nonzero. That is, there exists $S \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ such that
$$X^* = \begin{bmatrix} 0 & S \\ 0 & 0 \end{bmatrix}.$$

Given $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$,

$$\langle S\eta, \xi\rangle = \left\langle \begin{bmatrix} 0 & S \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \begin{bmatrix} \xi \\ \eta \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ T & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} \right\rangle = \langle \eta, T\xi\rangle.$$

So $S = T^*$.

(10.4.7) Let $T \in \mathcal{B}(\mathcal{H})$ with $0 \leq T \leq I$. Show that $T^2 \leq T$.

Answer. From $I - T \geq 0$ we have $T^{1/2}(I-T)T^{1/2} \geq 0$. This is $T - T^2 \geq 0$.

(10.4.8) Show an example of positive $S, T \in \mathcal{B}(\mathcal{H})$ such that $S \leq T$ but $S^2 \not\leq T^2$. *(Hint: examples already exist on dimension 2)*

Answer. The matrices will have to fail to commute. We can take

$$S = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad T = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then $T - S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \geq 0$ (it is positive because it is selfadjoint and its eigenvalues are $0, 2$). But

$$T^2 - S^2 = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}.$$

From $\det(T^2 - S^2) = -1$ we know that one eigenvalue is negative, so $T^2 - S^2$ is not positive.

(10.4.9) Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive and such that $ST = TS$. Show that $S^{1/2}T^{1/2} = T^{1/2}S^{1/2}$.

Answer. From Proposition 10.4.4, $T^{1/2}$ is a limit of polynomials on $I - T$; it follows that $ST^{1/2} = T^{1/2}S$. And now, using that $S^{1/2}$ is a limit of polynomials on $(I - S)$, $S^{1/2}T^{1/2} = T^{1/2}S^{1/2}$.

(10.4.10) Show an example of $T \in \mathcal{B}(\mathcal{H})$ such that $\operatorname{ran} T$ is not closed.

Answer. Fix an orthonormal basis $\{e_n\}$ and let $Te_n = \frac{1}{n}\,e_n$. Then $\operatorname{ran} T$ is dense, since $e_n \in \operatorname{ran} T$ for all n, but $\operatorname{ran} T$ is not closed (for instance, $(1/n)_n \notin \operatorname{ran} T$, for it would have to come from $\xi = \sum_n e_n$).

(10.4.11) Let $T \in \mathcal{B}(\mathcal{H})$. Show that $\|T\xi\| = \|\,|T|\xi\|$ for all $\xi \in \mathcal{H}$, and that $|T|$ is the only positive operator in $\mathcal{B}(\mathcal{H})$ with that property.

Answer. We have

$$\|T\xi\|^2 = \langle T\xi, T\xi\rangle = \langle T^*T\xi, \xi\rangle = \langle |T|^2\xi, \xi\rangle = \|\,|T|\xi\|^2.$$

Now suppose that $\|T\xi\| = \|S\xi\|$ for all $\xi \in \mathcal{H}$ and that $S \geq 0$. With the same computations as above we obtain

$$\langle (T^*T - S^2)\xi, \xi\rangle = 0, \qquad \xi \in \mathcal{H}.$$

By polarization $T^*T - S^2 = 0$, so $S^2 = T^*T$. And $S \geq 0$, so $S = (T^*T)^{1/2} = |T|$.

(10.4.12) Let $V \in \mathcal{B}(\mathcal{H})$ be a partial isometry with $P = V^*V$ and $Q = VV^*$. Show that $V = QVP$.

Answer. This can be deduced from Proposition 10.4.10, but here is a simple direct argument. We know from Proposition 10.1.10 that $\ker V = \ker V^*V = \ker P = (I_{\mathcal{H}} - P)\mathcal{H}$. So $V(I_{\mathcal{H}} - P) = 0$. That is, $V = VP$. And then $V = VP = VV^*V = QV$.

(10.4.13) Show that when $\dim \mathcal{H} < \infty$ the partial isometry in Proposition 10.4.11 can be chosen to be a unitary, at the cost of losing the condition on the range.

Answer. Write $T = V|T|$ as in Proposition 10.4.11. Since $V : \operatorname{ran} T^* \to \operatorname{ran} T$ is a surjective isometry (due to the finite-dimension, dense range equals surjective), we also have that their complements have equal dimension. Let $W : (\operatorname{ran} T^*)^\perp \to (\operatorname{ran} T)^\perp$ be a unitary, and form $U = V + W$. As V and W are partial isometries with orthogonal initial and final spaces, U is a unitary. And $U|T| = V|T| = T$.

(10.4.14) Show that if $T = WZ = VS$ with W, V partial isometries with $V^*V = [\overline{\text{ran}}Z] = [\overline{\text{ran}}S] = W^*W$, $VV^* = WW^* = [\overline{\text{ran}}T]$, and $Z, S \geq 0$, then $W = V$ and $Z = S$.

Answer. We have $T^*T = ZW^*WZ = Z^2$, and similarly with S; so $S = Z = |T|$ by the uniqueness of the positive square root (Proposition 10.4.4). Then $W = V$ by Proposition 10.4.11.

(10.4.15) Let $T \in \mathcal{B}(\mathcal{H})$ be invertible. Show that if $T = V|T|$ is the Polar Decomposition, then $|T|$ is invertible and V is a unitary.

Answer. We know that T^* is invertible too. Since V is a partial isometry with initial space $\overline{\text{ran}\,T^*} = \mathcal{H}$ and final space $\overline{\text{ran}\,T} = \mathcal{H}$, we get that V is an isometry: $V^*V = I$. Same argument but with V^* shows that $VV^* = I$, so V is a unitary. Now $|T| = V^*T$, invertible.

(10.4.16) Give an example of $T \in \mathcal{B}(\mathcal{H})$ and decompositions $T = VR = WS$, with $R, S \geq 0$ and V, W partial isometries, such that $R \neq S$ and $V \neq W$. Explain why this does not contradict the uniqueness in Proposition 10.4.11.

Answer. Let $\mathcal{H} = \mathbb{C}^3$ and $T = E_{11}$. Then Proposition 10.4.11 gives $T = V|T|$ with $|T| = V = E_{11}$. Now let $S = E_{11} + E_{22}$ and $W = E_{11} + E_{33}$. There is no contradiction because W is not a partial isometry between the ranges of S^* (which is equal to S in this case) and S.

(10.4.17) Let $T \in \mathcal{B}(\mathcal{H})$. Show T can be expressed as $T = WA$ with A positive and W a unitary if and only if $\dim\ker T = \dim\ker T^*$.

Answer. Assume first $\dim\ker T = \dim\ker T^*$. By the Polar Decomposition, we have $T = UA$ with $A = |T| \geq 0$, and $U \in \mathcal{B}(\mathcal{H})$ is a partial isometry with $\ker T = \ker U = \ker A$, $\operatorname{ran} U = \overline{\text{ran}}T$.

The operator $U^*U \in \mathcal{B}(\mathcal{H})$ is a projection with

$$\ker U^*U = \ker T = \ker A.$$

So $\operatorname{ran} U^*U = (\ker U^*U)^\perp = (\ker A)^\perp = \overline{\text{ran}}A$. Thus

$$\ker U^* = (\operatorname{ran} U)^\perp = (\operatorname{ran} T)^\perp = \ker T^*. \tag{AB.10.8}$$

Since $\dim\ker T = \dim\ker T^*$, and by mapping an orthonormal basis to another, we can construct a partial isometry $V : \ker T \to \ker T^*$. Define

$$W = U + V.$$

By (AB.10.8) we have that $U^*V = 0$. Then

$$W^*W = U^*U + V^*V + 2\operatorname{Re} U^*V = U^*U + V^*V = I_{\mathcal{H}}$$

The equality with the identity is due to U^*U being the projection onto $\overline{\operatorname{ran}}A$, and V^*V being the projection onto $\ker T = \ker A = (\operatorname{ran} A)^\perp$. Similarly, since $\operatorname{ran} U^* = (\ker U)^\perp = (\ker T)^\perp$, we have $VU^* = 0$. Then

$$WW^* = UU^* + VV^* = I_{\mathcal{H}}.$$

The last equality now holds because UU^* is the projection onto

$$\operatorname{ran} UU^* = (\ker UU^*)^\perp = (\ker U^*)^\perp = (\ker T^*)^\perp,$$

and VV^* is the projection onto $\ker T^*$. So W is a unitary.

Finally, since $VU^* = 0$ we have $VA = VU^*UA = 0$, so

$$WA = (U + V)A = UV.$$

For the converse, if $T = WA$ with W a unitary, we have

$$A^2 = AW^*WA = T^*T$$

and so $A = |T|$. The equality $T = W|T|$ shows that $W|_{\overline{\operatorname{ran}}|T|}$ is an isomorphism between $\overline{\operatorname{ran}}|T|$ and $\operatorname{ran} T$. Then it also an isomorphism between $(\operatorname{ran}|T|)^\perp = \ker|T| = \ker T^*T = \ker T$ and $(\operatorname{ran} T)^\perp = \ker T^*$. Therefore $\dim\ker T^* = \dim\ker T$.

(10.4.18) Let $T \in \mathcal{B}(\mathcal{H})$, invertible. Show that $\|T\| = \|T^{-1}\| = 1$ if and only if T is a unitary.

Answer. If T is a unitary, then $\|T^{-1}\| = \|T^*\| = \|T\| = 1$.

For the converse, we have

$$\|T^*T\| = \|T\|^2 = \|T^{-1}\|^2 = \|T^{-1}(T^{-1})^*\|$$
$$= \|T^{-1}(T^*)^{-1}\| = \|(T^*T)^{-1}\|.$$

So $S = T^*T$ is a positive invertible operator with $\|S\| = \|S^{-1}\| = 1$. Since S is positive, we have $\|S\| = \max\sigma(S)$ (Proposition 10.3.3). Also,

$$1 = \|S^{-1}\| = \max\sigma(S^{-1}) = \max\{\lambda^{-1} : \lambda \in \sigma(S)\} = (\min\sigma(S))^{-1}.$$

So $\max\sigma(S) = 1 = \min\sigma(S)$. Then $\sigma(S) = \{1\}$. As $\|S\| = 1$, we have for any $\xi \in \mathcal{H}$ with $\|\xi\| = 1$,

$$\langle S\xi, \xi\rangle \le 1 = \langle \xi, \xi\rangle.$$

So $I - S \geq 0$. Also, $\sigma(I - S) = \{0\}$ (Exercise 9.5.1). Then $\|I - S\| = 0$ by Proposition 10.3.3. So $S = I$; that is $T^*T = I$. Now we can repeat the argument for T^*, to obtain $TT^* = I$; thus, T is a unitary.

A more direct argument is the following: given $\xi \in \mathcal{H}$ with $\|\xi\| = 1$,

$$1 = \langle \xi, \xi \rangle = \langle T^{-1}T\xi, T^{-1}T\xi \rangle = \|T^{-1}T\xi\|^2 \leq \|Tx\|^2 = \langle T\xi, T\xi \rangle.$$

Also,

$$\langle T\xi, T\xi \rangle = \|T\xi\|^2 \leq \|\xi\|^2 = \langle \xi, \xi \rangle = 1.$$

Thus $\langle T\xi, T\xi \rangle = \langle \xi, \xi \rangle$ for all ξ (as we can always scale to 1). This shows that $T^*T = I$. Repeating the argument for T^* we get that $TT^* = I$.

(10.4.19) Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive. Show that $\sigma(ST) \subset [0, \infty)$. Does this imply that ST is positive?

Answer. By Proposition 9.2.15 we have that

$$\sigma(ST) \subset \sigma(S^{1/2}TS^{1/2}) \cup \{0\},$$

and so $\sigma(ST) \subset [0, \infty)$ since $S^{1/2}TS^{1/2} \geq 0$ by Proposition 10.4.7. It is not necessary that ST is positive, though, even in finite dimension. For instance let

$$S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad T = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Then $S \geq 0$, $T \geq 0$, but

$$ST = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

is not even selfadjoint.

(10.4.20) Let $T \in \mathcal{B}(\mathcal{H})^+$ with $\|T\| \leq 1$, and $\xi_0 \in \mathcal{H}$ such that $\|T\xi_0\| = \|\xi_0\|$. Show that $T\xi_0 = \xi_0$, and that $T = P + T_0$, where P is a projection and $T_0 \in \mathcal{B}(\mathcal{H})$ satisfies $\|T_0\xi\| < \|\xi\|$ for all nonzero $\xi \in \mathcal{H}$.

Answer.

Since T is selfadjoint, $T^2 \geq 0$. We also have $T^2 \leq I$ by Proposition 10.4.3. Now

$$0 \leq \|(I - T^2)^{1/2}\xi\|^2 = \langle (I - T^2)\xi, \xi \rangle = \|\xi\|^2 - \|T\xi\|^2 = 0.$$

Thus $(I - T^2)^{1/2}\xi = 0$ and so $(I - T^2)\xi = 0$. This we can write as

$$(I + T)(I - T)\xi = 0.$$

By Exercise 9.5.1 we have that $\sigma(I+T) \subset [1,\infty)$, so $I+T$ is invertible. Hence $(I-T)\xi = 0$, which is $T\xi = \xi$.

Let

$$L = \{\eta : \ \|T\xi\| = \|\xi\|\} = \{\eta : \ T\eta = \eta\} = \ker(I-T) \subset \mathcal{H}.$$

By hypothesis this is a nonempty subspace, and it is closed by the continuity of T. Let P be the orthogonal projection onto L. Define $T_0 = T(I-P)$. Then $T = TP + T(I-P) = P + T_0$. For nonzero $\eta \in L^\perp = (I-P)\mathcal{H}$, we have $\|T_0\eta\| = \|T\eta\| < \|\eta\|$, for otherwise if $\|T\eta\| = \|\eta\|$ then $\eta \in L$.

(10.4.21) Let $S, T \in \mathcal{B}(\mathcal{H})$, both positive, and with $\|S\| \leq 1$ and $\|T\| \leq 1$. Show that $\|S-T\| \leq 1$.

Answer. We have $0 \leq S, T \leq I$ by Proposition 10.4.3. Then

$$-I \leq -T \leq S - T \leq S \leq I.$$

Given $\xi \in \mathcal{H}$, this means that

$$-\langle \xi, \xi \rangle \leq \langle (S-T)\xi, \xi \rangle \leq \langle \xi, \xi \rangle.$$

Then Proposition 10.2.3 implies $\sigma(S-T) \subset \overline{\text{conv}}\,\sigma(S-T) \subset \overline{W(S-T)} \subset [-1,1]$. Therefore $\|S-T\| = \text{spr}(S-T) \leq 1$ by Proposition 10.3.3.

(10.4.22) Prove Proposition 10.4.12.

Answer. We have $(T^*)_{kj} = P_k T^* P_j = (P_j T P_k)^* = (T_{jk})^*$.

10.5. Projections

(10.5.1) Let $P \in \mathcal{B}(\mathcal{H})$. Show that $P^*P = P$ if and only if $P = P^* = P^2$.

Answer. If $P^*P = P$, then $P^* = (P^*P)^* = P^*P = P$, and

$$P^2 = P^*P = P.$$

Conversely, if $P = P^* = P^2$ then $P^*P = P^2 = P$.

(10.5.2) Let $P \in \mathcal{B}(\mathcal{H})$ be a projection. Show that $\ker P = P^\perp \mathcal{H}$ and that $P^\perp$ is the orthogonal projection onto $(P\mathcal{H})^\perp$.

Answer. If $P\xi = 0$, then $P^\perp \xi = \xi - P\xi = \xi$, so $\xi \in P^\perp\mathcal{H}$. Conversely, if $\xi = P^\perp\xi$, then $P\xi = P(I-P)\xi = 0$, so $\xi \in \ker P$. Hence $\ker P = P^\perp\mathcal{H}$

We have $(P\mathcal{H})^\perp = (\operatorname{ran} P)^\perp = \ker P^* = \ker P = P^\perp\mathcal{H}$.

(10.5.3) When $\mathcal{H} = \mathbb{C}^2$, we can identify $\mathcal{B}(\mathcal{H})$ with $M_2(\mathbb{C})$. Find all orthogonal projections and all idempotents.

Answer. Let us start with the projections. We could play with equations as we will do with the idempotents, but let us try something else. The only rank-0 projection is 0, and the only rank-2 projection is the identity I_2. It remains to characterize the rank-1 projections. These are rank-one operators, so they are of the form $P = xy^*$, with x, y nonzero. From $P^* = P$ we obtain $yx^* = xy^*$. Evaluating at x we have $\|x\|^2 \, y = (y^*x)\, x$; as neither x nor y is zero, this tells us that x and y are colinear. Write $y = \lambda x$. Then $P = \lambda x x^*$. From $P^2 = P$, we have $\lambda xx^* = \lambda^2 \|x\|^2 xx^*$. So $\lambda = 1/\|x\|^2$, which is the same as assuming that $P = xx^*$ with $\|x\| = 1$.

The vectors $x \in \mathbb{C}^2$ with $\|x\| = 1$ are of the form

$$x = (\sqrt{t}\, e^{i\theta}, \sqrt{1-t}\, e^{i\gamma}), \qquad t \in [0,1].$$

Then

$$\begin{aligned} P = xx^* &= \begin{bmatrix} \sqrt{t}\, e^{i\theta} \\ \sqrt{1-t}\, e^{i\gamma} \end{bmatrix} \begin{bmatrix} \sqrt{t}\, e^{-i\theta} & \sqrt{1-t}\, e^{-i\gamma} \end{bmatrix} \\ &= \begin{bmatrix} t & \sqrt{t-t^2}\, e^{i(\theta-\gamma)} \\ \sqrt{t-t^2}\, e^{-i(\theta-\gamma)} & 1-t \end{bmatrix} \\ &= \begin{bmatrix} t & \lambda\sqrt{t-t^2} \\ \bar\lambda\sqrt{t-t^2} & 1-t \end{bmatrix} \qquad t \in [0,1],\ \lambda \in \mathbb{T}. \end{aligned}$$

For an idempotent $E = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the equation $E^2 = E$ translates to

$$a = a^2 + bc, \qquad (a+d)b = b, \qquad (a+d)c = c, \qquad d = d^2 + bc.$$

We consider two cases.

- If $b=c=0$, then $a,d=\pm 1$. So
$$E=\begin{bmatrix}\pm 1 & 0\\ 0 & \pm 1\end{bmatrix}.$$
- If b or c is not zero, we get $a+d=1$. Both a and d are solutions of the quadratic equation $t^2-t+bc=0$. Then
$$E=\begin{bmatrix}\frac12\pm\frac12\sqrt{1-4bc} & b\\ c & \frac12\mp\frac12\sqrt{1-4bc}\end{bmatrix},\qquad b,c\in\mathbb{C}.$$
This works even if $4bc>1$, by interpreting the $\pm$ and $\mp$ and giving the two distinct complex roots of $1-4bc$.

(10.5.4) Let $T\in\mathcal{B}(\mathcal{H})$ be normal, and $T\neq 0,\neq I$. Show that T is a projection if and only if $\sigma(T)=\{0,1\}$.

Answer. If $T^2=T$, then by Spectral Mapping (Proposition 9.2.9) we have that $\lambda^2=\lambda$ for all $\lambda\in\sigma(T)$. Thus $\sigma(T)\subset\{0,1\}$. If $\sigma(T)=\{0\}$, then $T=0$ by (ii) in Proposition 10.3.3, a contradiction. If $\sigma(T)=\{1\}$, then $T=I$ again by Proposition 10.3.3, this time applied to $I-T$. Hence $\sigma(T)=\{0,1\}$.

Conversely, suppose that $\sigma(T)=\{0,1\}$. By Spectral Mapping (Proposition 9.2.9), T^2-T is a normal operator with $\sigma(T^2-T)=\{0\}$. Thus $T^2-T=0$ by Proposition 10.3.3.

(10.5.5) Show that the projections are extreme points in the set $\mathcal{B}(\mathcal{H})_1^+$ of positive operators with norm at most 1 (it is actually true that the projections are the only extreme points, but that will have to wait until Exercise 12.4.4).

Answer. Let P be a projection and suppose that $P=tA+(1-t)B$ with $A,B\geq 0$ $\|A\|\leq 1$, $\|B\|\leq 1$, and $t\in[0,1]$. If $\xi=P\xi$ and $\|\xi\|=1$,
$$\begin{aligned}1=\langle P\xi,\xi\rangle&=t\langle A\xi,\xi\rangle+(1-t)\langle B\xi,\xi\rangle\\ &\leq t\|A\|+(1-t)\|B\|\leq 1.\end{aligned}$$
As all terms are nonnegative, the equality in the inequalities forces $\langle A\xi,\xi\rangle=1$. This is $\|A^{1/2}\xi\|^2=\|\xi\|^2$. By Exercise 10.4.20, $A^{1/2}\xi=\xi$ and then $A\xi=\xi$. Similarly, if $P\xi=0$ then
$$0\leq\langle A\xi,\xi\rangle+(1-t)\langle B\xi,\xi\rangle=\langle P\xi,\xi\rangle=0.$$

Arguing as above, we get $\langle A\xi,\xi\rangle = 0$ and thus $A\xi = 0$ when $P\xi = 0$. As $\mathcal{H} = P\mathcal{H} \oplus \ker P$, we get that $A = P$. This forces $B = P$ and so P is extreme.

(10.5.6) Show that the converse of Proposition 10.5.6 is false. That is, find projections $P, Q \in \mathcal{B}(\mathcal{H})$, unitarily equivalent, and such that $\|P - Q\| = 1$.

Answer. Let $P = E_{11}$, $Q = E_{22}$, $U = E_{12} + E_{21} + \sum_{n\geq 3} E_{nn}$. Then $UPU^* = Q$ and $\|P - Q\| = 1$.

(10.5.7) Prove Proposition 10.5.9.

Answer. Let $P = \bigwedge_j P_j$. By Exercise 10.5.2 we know that $P^\perp$ is the orthogonal projection onto $(P\mathcal{H})^\perp$. That is, $P^\perp$ is the orthogonal projection onto (using Exercises 10.5.2 and 4.3.12)

$$\left(\bigcap_j P_j\mathcal{H}\right)^\perp = \overline{\operatorname{span}} \bigcup_j P_j^\perp \mathcal{H}.$$

Hence $P^\perp = \displaystyle\bigvee_j P_j^\perp$. The second equality follows by taking $\perp$ on the first one.

(10.5.8) Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. Is it true that $P = P \wedge Q + P \wedge Q^\perp$? Provide either a proof of a counterexample.

Answer. No, it's not true in general. Let $\mathcal{H} = \mathbb{C}^2$,

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \qquad Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then $P \wedge Q = P \wedge Q^\perp = 0$, so the equality does not hold.

The equality $P = P \wedge Q + P \wedge Q^\perp$ occurs precisely when P and Q commute. Indeed, if $P = P \wedge Q + P \wedge Q^\perp$, multiplying by Q on the left and separately on the right, we get $QP = P \wedge Q = PQ$. Conversely, if $PQ = QP$, then $PQ = P \wedge Q$ (this can be seenby showing directly that PQ is the orthogonal projection onto $P\mathcal{H} \cap Q\mathcal{H}$), so

$$P \wedge Q + P \wedge Q^\perp = PQ + PQ^\perp = P(Q + Q^\perp) = P.$$

(10.5.9) Let $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ a family of projections. Show that

$$U\Big(\bigvee_j P_j\Big)U^* = \bigvee_j UP_jU^*.$$

Answer. Let $P = \bigvee_j P_j$. This is the projection onto $\overline{\text{span}}\bigcup_j P_j\mathcal{H}$. We have

$$\overline{\text{span}}\bigcup_j UP_jU^*\mathcal{H} = \overline{\text{span}}\bigcup UP_j\mathcal{H} = U\,\overline{\text{span}}\bigcup_j P_j\mathcal{H}.$$

So $P' = \bigvee_j UP_jU^*$ is the orthogonal projection onto $UP\mathcal{H}$. As UPU^* is also an orthogonal projection onto $UP\mathcal{H}$, the uniqueness gives us $P' = UPU^*$.

Another way to prove the equality is to use the fact that P is the supremum of the family of projections. Then one sees that UPU^* is the least upper bound for $\{UP_jU^*\}$.

(10.5.10) Let $\mathcal{H}$ be a Hilbert space and $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ an increasing net of projections such that $\bigvee_j P_j = I_\mathcal{H}$. Show that for all $\xi \in \mathcal{H}$ we have $\lim_j P_j\xi = \xi$.

Answer. Write $\mathcal{H}_j = P_j\mathcal{H}$. Then the hypothesis is that $\bigcup_j \mathcal{H}_j$ is dense in $\mathcal{H}$. Given $\varepsilon > 0$ there exists j_0 and $\xi_0 \in \mathcal{H}_{j_0}$ with $\|\xi - \xi_0\| < \varepsilon$. When $k \geq j$ we have $P_kP_j = P_j$ by Proposition 10.5.3. Hence, for $j \geq j_0$

$$P_j\xi_0 = P_jP_{j_0}\xi_0 = P_{j_0}\xi_0 = \xi_0;$$

so $\xi_0 \in \mathcal{H}_j$. Using Proposition 4.3.8,

$$\|\xi - P_j\xi\| \leq \|\xi - \xi_0\| < \varepsilon.$$

Thus $P_j\xi \to \xi$.

(10.5.11) Let $P, R \in \mathcal{B}(\mathcal{H})$ be projections. Show that

$$P \vee R + (I_\mathcal{H} - P) \wedge (I_\mathcal{H} - R) = I_\mathcal{H}.$$

Answer. We have $(P\mathcal{H} \cup R\mathcal{H})^\perp = (P\mathcal{H})^\perp \cap (R\mathcal{H})^\perp$. This says that

$$I_\mathcal{H} - P \vee R = (I_\mathcal{H} - P) \wedge (I_\mathcal{H} - R).$$

Hence $P \vee R + (I_\mathcal{H} - P) \wedge (I_\mathcal{H} - R) = I_\mathcal{H}$.

(10.5.12) Let $\mathcal{H}$ be a Hilbert space. Show that there exists an increasing net $\{P_j\}$ of finite-rank projections with $\bigvee_j P_j = I_{\mathcal{H}}$.

Answer. Let $\{\xi_k\}_{k\in K}$ be an orthonormal basis for $\mathcal{H}$, and put $J = \{F \subset K,\ \text{finite}\}$, ordered by inclusion. Define, for $j \in J$, $\mathcal{H}_j = \operatorname{span}\{\xi_k : \ k \in j\}$. Then $\{\xi_k\} \subset \bigcup_j \mathcal{H}_j$, so $\bigvee_j P_j = I_{\mathcal{H}}$.

(10.5.13) Let $\mathcal{H}$ be a Hilbert space and $\{P_j\}$ an increasing net of projections with $P_j\xi \to \xi$ for all $\xi \in \mathcal{H}$. Show that for all $T \in \mathcal{B}(\mathcal{H})$,

$$\|T\| = \sup\{\|P_jTP_j\| : \ j\}.$$

Answer. We always have $\|P_jTP_j\| \leq \|P_j\|^2\|T\| = \|T\|$. Fix $\varepsilon > 0$. Then there exists $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ and $\|T\xi\| > \|T\| - \varepsilon$. By hypothesis there exists j with $\|\xi - P_j\xi\| < \varepsilon$. Then if j is big enough so that $\|P_j\xi - \xi\| < \varepsilon$ and $\|P_jT\xi - T\xi\| < \varepsilon$,

$$\begin{aligned}
\|P_jTP_j\xi\| &= \big\|T\xi + \big(P_jTP_j\xi - P_jT\xi\big) + \big(P_jT\xi - T\xi\big)\big\| \\
&\geq \|T\xi\| - \|P_jTP_j\xi - P_jT\xi\| - \|P_jT\xi - T\xi\| \\
&\geq \|T\xi\| - \|T\|\,\|P_j\xi - \xi\| - \|P_jT\xi - T\xi\| \\
&\geq \|T\xi\| - (1 + \|T\|)\varepsilon \\
&> \|T\| - (2 + \|T\|)\varepsilon.
\end{aligned}$$

As this can be done for all $\varepsilon > 0$, we get that $\sup\{\|P_jTP_j\| : \ j\} \geq \|T\|$, and thus we have the equality.

(10.5.14) Let $T, P \in \mathcal{B}(\mathcal{H})$ with P a projection such that $(I_{\mathcal{H}} - P)T(I_{\mathcal{H}} - P) = 0$. Show that $|\langle T\xi, \xi\rangle| \leq 2\|T\|\,\|P\xi\|$ for all $\xi \in \mathcal{H}$ with $\|\xi\| \leq 1$.

Answer. We have

$$\begin{aligned}
|\langle T\xi, \xi\rangle| &\leq |\langle T\xi, P\xi\rangle| + |\langle T\xi, (I_{\mathcal{H}} - P)\xi\rangle| \\
&\leq |\langle T\xi, P\xi\rangle| + |\langle TP\xi, (I_{\mathcal{H}} - P)\xi\rangle| + |\langle T(I_{\mathcal{H}} - P)\xi, (I_{\mathcal{H}} - P)\xi\rangle| \\
&= |\langle T\xi, P\xi\rangle| + |\langle TP\xi, (I_{\mathcal{H}} - P)\xi\rangle \\
&\leq \|T\|\,\|P\xi\| + \|T\|\,\|P\xi\|\,\|(I_{\mathcal{H}} - P)\xi\| \leq 2\|T\|\,\|P\xi\|.
\end{aligned}$$

(10.5.15) Let $T \in \mathcal{B}(\mathcal{H})^+$, $\alpha, \beta, \gamma \in \mathbb{R}$ with $0 < \gamma < \alpha < \beta$, and $P, Q \in \mathcal{B}(\mathcal{H})$ two projections that commute with T and with each other, and such that

$$\alpha P \leq PT \leq \beta P, \qquad 0 \leq (I_{\mathcal{H}} - P)T \leq \gamma(I_{\mathcal{H}} - P),$$
$$\alpha Q \leq QT \leq \beta Q, \qquad 0 \leq (I_{\mathcal{H}} - Q)T \leq \gamma(I_{\mathcal{H}} - Q).$$

Show that $P = Q$.

Answer. We have

$$\alpha(I_{\mathcal{H}} - P)Q \leq (I_{\mathcal{H}} - P)TQ \leq \gamma(I_{\mathcal{H}} - P)Q.$$

As $\gamma < \alpha$, $(I_{\mathcal{H}} - P)Q = 0$. That is $Q = PQ$, which is to say that $Q \leq P$. As the roles of P and Q can be reversed, we also have $P \leq Q$ and hence $P = Q$.

(10.5.16) Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections.

(i) Let $T = QP + (I - Q)(I - P)$. Show that $TP = QT$.

(ii) Show that if $\|P - Q\| < 1$, then T is invertible.

(iii) Show that if $\|P - Q\| < 1$ then

$$\dim \operatorname{ran} P = \dim \operatorname{ran} Q, \ \dim \operatorname{ran}(I - P) = \dim \operatorname{ran}(I - Q).$$

(iv) Conclude that there exists a unitary U with $UPU^* = Q$.

This proves the result of Proposition 10.5.6, but without writing U as an expression depending on P and Q.

Answer.

(i) We have $TP = QP$ directly, since $(I - Q)(I - P)P = 0$.

(ii) This is the exact computation from Proposition 10.5.6. That is, $I - T = (2P - I)(P - Q)$, and then one uses the fact that $2P - I$ is a unitary to conclude that

$$\|I - T\| = \|(2P - I)(P - Q)\| = \|P - Q\| < 1,$$

so T is invertible by Proposition 6.2.3.

(iii) We have just shown that T is invertible. Hence $P = T^{-1}QT$, and

$$\begin{aligned}\dim \operatorname{ran} P &= \dim P\mathcal{H} = \dim T^{-1}QT\mathcal{H} \\ &= \dim QT\mathcal{H} = \dim Q\mathcal{H} = \dim \operatorname{ran} Q.\end{aligned}$$

We also have $\|(I-P)-(I-Q)\| = \|P-Q\| < 1$, so the above applies to show that $\dim\operatorname{ran}(I-P) = \dim\operatorname{ran}(I-Q)$.

(iv) Let $\{\xi_j\}_{j\in J}$, $\{\xi_k'\}_{k\in K}$, $\{\eta_j\}_{j\in J}$, $\{\eta_k'\}_{k\in K}$ be orthonormal bases for $P\mathcal{H}$, $(I-P)\mathcal{H}$, $Q\mathcal{H}$, $(I-Q)\mathcal{H}$ respectively. The equalities of the ranks allow us to use the same index sets J and K for P and Q. Let U be the linear map induced by $U\xi_j = \eta_j$ and $U\xi_k' = \eta_k'$ for all j, k. Then U is a unitary by Exercise 10.1.5. We have

$$UP\xi_j = U\xi_j = \eta_j = Q\eta_j = QU\xi_j.$$

Also $UP\xi_k' = 0$ and $QU\xi_k' = Q\eta_k' = 0$. Therefore $UP = QU$. As U is a unitary, $Q = UPU^*$.

(10.5.17) Regarding Remark 10.5.7, show that $U = (P+Q-I)|P+Q-I|^{-1}$ is a unitary and that $UPU^* = Q$.

Answer. Since $P+Q-I$ is selfadjoint, we have

$$\begin{aligned} U^*U &= |P+Q-I|^{-1}(P+Q-I)^2|P+Q-I|^{-1} \\ &= |P+Q-I|^{-1}|P+Q-I|^2|P+Q-I|^{-1} = I \end{aligned}$$

and, since a selfadjoint operator commutes with its absolute value,

$$\begin{aligned} UU^* &= (P+Q-I)|P+Q-I|^{-2}(P+Q-I) = |P+Q-I|^{-2}(P+Q-I)^2 \\ &= |P+Q-I|^{-2}|P+Q-I|^2 = I. \end{aligned}$$

When we see U as $\begin{bmatrix} C & S \\ S & -C \end{bmatrix}$, the equality $UPU^* = Q$ is a direct computation. To check the other expression,

$$\begin{aligned} (P+Q-I)^2 &= \begin{bmatrix} C^2 & CS \\ CS & S^2-I \end{bmatrix}^2 = \begin{bmatrix} C^2 & CS \\ CS & -C^2 \end{bmatrix}^2 \\ &= \begin{bmatrix} C^4 + C^2S^2 & 0 \\ 0 & C^2S^2 + C^4 \end{bmatrix} \\ &= \begin{bmatrix} C^2(C^2+S^2) & 0 \\ 0 & C^2(S^2+C^2) \end{bmatrix} = C^2\, I_2. \end{aligned}$$

Hence $|P+Q-I| = C$, since $C \geq 0$. Also $P+Q-I = -(I-(P-Q))$ is invertible since $\|P-Q\| < 1$, so C is invertible. Then

$$(P+Q-I)|P+Q-I|^{-1} = \begin{bmatrix} C^2 & CS \\ CS & -C^2 \end{bmatrix} C^{-1} = \begin{bmatrix} C & S \\ S & -C \end{bmatrix}.$$

10.6. Compact operators

(10.6.1) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\xi \in \mathcal{H}$. Show that $\|T^*\xi\| = \|T\xi\|$. Use an example to show that the equality is not necessarily true when T is not normal.

Answer. We have

$$\|T^*\xi\|^2 = \langle T^*\xi, T^*\xi\rangle = \langle TT^*\xi, \xi\rangle = \langle T^*T\xi, \xi\rangle = \langle T\xi, T\xi\rangle = \|T\xi\|^2.$$

The typical counterexample was given in the text:

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \eta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then $T\eta = \eta$, but $T^*\eta = 0$.

(10.6.2) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Prove the inequality $\|T\xi\|^2 \leq \|T^2\xi\|\,\|\xi\|$, and use it to show that if T^2 is compact, then T is compact. Show, with an example, that normality is crucial as a hypothesis for both assertions.

Answer. We have, using the Cauchy–Schwarz in equality and Exercise 10.6.1,

$$\|T\xi\|^2 = \langle T^*T\xi, \xi\rangle \leq \|T^*T\xi\|\,\|\xi\| = \|T^2\xi\|\,\|\xi\|.$$

Let $\{T\xi_n\}$ be a sequence with $\|\xi_n\| \leq 1$ for all n. Since T^2 is compact, the sequence $\{T^2\xi_n\}$ admits a convergent subsequence $\{T^2\xi_{n_k}\}$ with the limit in the range of T^2, say $T^2\xi_{n_k} \to T^2\eta$. Now

$$\begin{aligned} \|T\xi_{n_k} - T\eta\|^2 &= \|T(\xi_{n_k} - \eta)\|^2 \\ &\leq \|T^2(\xi_{n_k} - \eta)\|\,\|\xi_{n_k} - \eta\| \\ &\leq 2\|T^2\xi_{n_k} - T^2\eta\|, \end{aligned}$$

which shows that $T\xi_{n_k} \to T\eta$. Thus T is compact.

When normality is dropped, consider any non-compact T with $T^2 = 0$. For instance in any infinite-dimensional Hilbert space fix an orthonormal basis and consider its associated matrix units. Let $T = \sum_n E_{2n+1,2n}$. Then T is isometric, so non-compact, and $T^2 = 0$, compact.

(10.6.3) Prove the identity (10.10).

Answer. Given $\nu \in \mathcal{H}$,

$$(\xi_1 \otimes \eta_1)(\xi_2 \otimes \eta_2)\nu = \langle \nu, \eta_2\rangle\,(\xi_1 \otimes \eta_1)\xi_2$$
$$= \langle \nu, \eta_2\rangle\,\langle \xi_2, \eta_1\rangle\,\xi_1 = \langle \xi_2, \eta_1\rangle\,(\xi_1 \otimes \eta_2)\,\nu.$$

(10.6.4) Given a rank-one operator $\xi\eta^*$, show that

$$(\xi\eta^*)^* = \eta\xi^*. \tag{10.18}$$

Answer. Given $\nu \in \mathcal{H}$,

$$\langle(\xi\eta^*)^*\nu, \nu\rangle = \langle \nu, (\xi\eta^*)\nu\rangle = \overline{\langle \nu, \eta\rangle}\,\langle \nu, \xi\rangle = \langle(\langle \nu, \xi\rangle)\eta, \nu\rangle = \langle(\eta\xi^*)\nu, \nu\rangle.$$

Then polarization gives (10.18).

(10.6.5) Write a complete proof of Proposition 10.6.1.

Answer. In all of (ii),(iii),(iv),(v) it is clear that the image of T is finite-dimensional, so they all imply (i).

(iv) $\Longrightarrow$ (v) Using Gram–Schmidt on $\xi_1', \ldots, \xi_n'$, we obtain an orthonormal basis $\xi_1, \ldots, \xi_n$. The coefficients come from $\xi_j' = \sum_k c_{kj}\xi_k$.

(i) $\Longrightarrow$ (iii) Let $\xi_1, \ldots, \xi_n$ be an orthonormal basis of the image of T. Then

$$T\xi = \sum_{k=1}^n \langle T\xi, \xi_k\rangle\,\xi_k, \qquad \xi \in \mathcal{H}.$$

As T is bounded, the map $\xi \longmapsto \langle T\xi, \xi_k\rangle$ is bounded. Then, by the Riesz Representation Theorem (Theorem 4.5.4) there exist $\eta_1', \ldots, \eta_n'$ with $\langle T\xi, \xi_k\rangle = \langle \xi, \eta_k'\rangle$.

(iii) $\Longrightarrow$ (ii) Trivial.

(iii) $\Longrightarrow$ (iv) By hypothesis we can write

$$T = \sum_{k=1}^n \xi_k\eta_k^*$$

with $\{\xi_1, \ldots, \xi_n\}$ orthonormal. Then, by Exercise 10.6.4,

$$T^* = \sum_{k=1}^n \eta_k\xi_k^*.$$

This has the desired form.

(10.6.6) Show that the operator T from (ii) in Examples 10.6.5 is well-defined, and that it is a limit of finite-rank operators.

Answer. On a linear combination $\xi = \sum_{j=1}^{n} c_j \xi_j$ we have

$$\|T\xi\| = \left\| T\left(\sum_{j=1}^{n} c_j \xi_j \right) \right\|^2 = \left\| \sum_{j=1}^{n} \frac{1}{j}\, c_j \xi_j \right\|^2 = \sum_{j=1}^{n} \frac{|c_j|^2}{j^2} \le \sum_{j=1}^{n} |c_j|^2 = \|\xi\|^2.$$

It follows that $\|T\| \le 1$ on the dense subspace $\operatorname{span}\{\xi_j : j\}$ and so it admits a unique bounded extension to $\mathcal{H}$ by Proposition 6.1.9. If we let T_n be given by

$$T_n \xi = \sum_{j=1}^{n} \frac{1}{j} \langle \xi, \xi_j \rangle\, \xi_j$$

then $T_n \in \mathcal{F}(\mathcal{H})$ and

$$\|(T - T_n)\xi\|^2 = \left\| T\left(\sum_{j=n+1}^{\infty} c_j \xi_j \right) \right\|^2 = \sum_{j=n+1}^{\infty} \frac{|c_j|^2}{j^2} \le \frac{1}{n^2} \|\xi\|.$$

Hence $\|T - T_n\| \le \frac{1}{n}$ and $T = \lim_n T_n$.

(10.6.7) Fix an orthonormal basis $\{\xi_n\}$ for $\mathcal{H}$ and define

$$T\xi = \sum_{k=1}^{\infty} \frac{1}{k} \langle \xi, \xi_k \rangle\, \xi_{k+1}.$$

(i) Show that $T \in \mathcal{K}(\mathcal{H})$;

(ii) find T^*;

(iii) show that $\ker T^*T = \{0\}$ and that $\ker TT^* = \mathbb{C}\,\xi_1$.

Answer.

(i) We need to show that T is a limit of finite-rank operators; this will be true if the series converges in norm. For this,

$$\left\| \sum_{k=m}^{n} \frac{1}{k} \langle \xi, \xi_k \rangle\, \xi_{k+1} \right\|^2 = \sum_{k=m}^{n} \frac{1}{k^2} |\langle \xi, \xi_k \rangle|^2 \le \|\xi\| \sum_{k=m}^{n} \frac{1}{k^2}.$$

As the series for $\frac{1}{k^2}$ converges its tails are Cauchy and hence the series for T converges in norm.

(ii) Using Exercise 10.6.4 and the fact that taking adjoints is conjugate-linear and continuous,

$$T^* = \Big(\sum_{k=1}^\infty \frac{1}{k}\,\xi_{k+1}\xi_k^*\Big)^* = \sum_{k=1}^\infty \frac{1}{k}\,\xi_k\xi_{k+1}^*.$$

That is,

$$T^*\xi = \sum_{k=1}^\infty \frac{1}{k}\,\langle \xi, \xi_{k+1}\rangle\,\xi_k.$$

(iii) We have

$$(\xi_1\eta_1^*)^*(\xi_2\eta_2^*)\nu = (\eta_1\xi_1^*)\langle \nu, \eta_2\rangle\,\xi_2 = \langle \nu, \eta_2\rangle\,\langle \xi_1, \xi_2\rangle\,\eta_1.$$

We can now calculate directly,

$$T^*T\xi = \sum_{k,j} \frac{1}{kj}\,(\xi_k\xi_{k+1}^*)(\xi_{j+1}\xi_j^*)\xi = \sum_k \frac{1}{k^2}\,\langle \xi, \xi_k\rangle\,\xi_k.$$

So $T^*T\xi = 0$ if and only if $\langle \xi, \xi_k\rangle = 0$ for all k, which is equivalent to $\xi = 0$. That is, $\ker T^*T = \{0\}$. On the other hand,

$$TT^*\xi = \sum_{k=1}^\infty \frac{1}{k^2}\,\langle \xi, \xi_{k+1}\rangle\,\xi_{k+1} = \sum_{k=2}^\infty \frac{1}{k^2}\,\langle \xi, \xi_k\rangle\,\xi_k.$$

It follows that $\xi_1 \in \ker TT^*$. And if $TT^*\xi = 0$, we get that $\langle \xi, \xi_k\rangle = 0$ for all $k \geq 2$, so $\xi \in \mathbb{C}\,\xi_1$. Hence $\ker TT^* = \mathbb{C}\xi_1$.

(10.6.8) If $T \in \mathcal{F}(\mathcal{H})$, show that $T^* \in \mathcal{F}(\mathcal{H})$.

Answer. By Proposition 10.6.1, we can write $T = \sum_{k=1}^n \xi_k\eta_k^*$. Using Exercise 10.6.4 and the additivity of the adjoint operation we have that $T^* = \sum_{k=1}^n \eta_k\xi_k^*$. Then $T^* \in \mathcal{F}(\mathcal{H})$ by Proposition 10.6.1.

(10.6.9) Let $\mathcal{H}$ be a Hilbert space and $\{\eta_j\}_{j\in J}$ an orthonormal basis. For each finite $F \subset J$, let P_F be the orthogonal projection onto $\operatorname{span}\{\xi_j : j \in F\}$. Show that for each $\xi \in \mathcal{H}$, $\|(T - P_FT)\xi\| \to 0$, $\|(T - TP_F)\xi\| \to 0$, and $\|(T - P_FTP_F)\xi\| \to 0$.

Answer. Write $\xi = \sum_j c_j\eta_j$ with $c \in \ell^2(J)$. We have, since $\{|\langle T\xi, \eta_j\rangle|\} \in \ell^2(J)$,

$$\|(T - P_F T)\xi\|^2 = \|(I_{\mathcal{H}} - P_F)T\xi\|^2 = \sum_{j\notin F} |\langle T\xi, \eta_j\rangle|^2 \xrightarrow[F]{} 0.$$

And

$$\begin{aligned}\|(T - TP_F)\xi\|^2 &= \|T(I_h - P_F)\xi\|^2 \\ &\leq \|T\|^2\,\|(I_{\mathcal{H}} - P_F)\xi\|^2 = \|T\|^2 \sum_{j\notin F} |\langle \xi, \eta_j\rangle|^2 \xrightarrow[F]{} 0.\end{aligned}$$

Now we get

$$\begin{aligned}\|(T - P_F TP_F)\xi\| &\leq \|(T - P_F T)\xi\| + \|(P_F T - P_F TP_F)\xi\| \\ &\leq \|(T - P_F T)\xi\| + \|(T - TP_F)\xi\| \to 0.\end{aligned}$$

(10.6.10) Show that for $T \in \mathcal{K}(\mathcal{H})$ and $\{\xi_j\}$ an orthonormal basis,

$$\lim_j T\xi_j = 0.$$

Answer. Assume first that $T = \xi\eta^*$. Then

$$\|(\xi\eta^*)\,\xi_j\| = \|\langle \xi_j, \eta\rangle\,\xi\| \leq \|\xi\|\,|\langle \eta, \xi_j\rangle| \xrightarrow[j\to\infty]{} 0$$

by Parseval's Equality (4.13). It follows automatically, via Proposition 10.6.1 that $T\xi_j \to 0$ for any $T \in \mathcal{F}(\mathcal{H})$. For arbitrary $T \in \mathcal{K}(\mathcal{H})$, given $\varepsilon > 0$ there exists $S \in \mathcal{F}(\mathcal{H})$ with $\|T - S\| < \varepsilon$. Then

$$\|T\xi_j\| \leq \|(T - S)\xi_j\| + \|S\xi_j\| \leq \varepsilon + \|S\xi_j\|.$$

Then $\limsup_{j\to\infty} \|T\xi_j\| \leq \varepsilon$. By the Limsup Routine the limit exists and equals zero.

(10.6.11) Prove that whenever $\mathcal{H}$ is infinite-dimensional, any finite-rank operator $T \in \mathcal{F}(\mathcal{H})$ has non-trivial kernel.

Answer. If T has trivial kernel, then it is injective. An injective linear map takes linearly independent sets to linearly independent sets. So given an infinite orthonormal set $\{\xi_k\} \subset \mathcal{H}$, the image $\{T\xi_k\}$ is a linearly independent subset of the image of T. So $T \notin \mathcal{F}(\mathcal{H})$.

Another way to prove this is by using Proposition 10.6.1. If

$$T = \sum_{k=1}^{n} \langle \cdot, \eta_k \rangle \, \xi_k,$$

we get that $\{\eta_1, \ldots, \eta_n\}^\perp \subset \ker T$.

(10.6.12) Let $\mathcal{H} = \ell^2(\mathbb{N})$. Define operators

$$T(a_1, a_2, \ldots) = \left(a_1, \tfrac{a_2}{2}, \tfrac{a_3}{3}, \ldots\right),$$

$$S(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots).$$

Prove that T is an injective compact operator with

$$\sigma(T) = \{0, 1, 1/2, 1/3, \ldots\},$$

and that $R = ST$ is an injective compact operator with $\sigma(R) = \{0\}$. Conclude that $Z = R^*$ is a non-injective compact operator with dense range and $\sigma(Z) = \{0\}$.

Answer. If $Ta = 0$, then $a_k/k = 0$ for all k, so $a_k = 0$ for all k and $a = 0$; so T is injective. Since

$$T = \sum_{k=1}^{\infty} \frac{1}{k} \, E_{kk},$$

we get from Lemma 10.6.11 that $\sigma(T) = \{0, 1, 1/2, 1/3, \ldots\}$. We have

$$Ra = STa = \left(0, a_1, \tfrac{a_2}{2}, \tfrac{a_3}{3}, \ldots\right).$$

Then R is injective (direct proof, or we notice that R here is the T from Exercise 10.6.6 and $\{0\} = \ker R^*R = \ker R$). Being compact, the nonzero spectrum of R has to consist of eigenvalues, but if $Ra = \lambda a$, this gives

$$0 = \lambda a_1, \;\; a_1 = \lambda a_2, \;\; a2 = 2\lambda a_3, \;\; \cdots$$

and so $a = 0$; this shows that $\sigma(R) = \{0\}$.

The operator R^* is not injective, for

$$\langle R^* e_1, a \rangle = \langle e_1, Ra \rangle = 0,$$

so $e_1 \in \ker R^*$. Hence R^* is compact (adjoint of compact), not injective, and $\sigma(R^*) = \overline{\sigma(R)} = \{0\}$. As for the range, $\overline{\operatorname{ran}} R^* = (\ker R)^\perp = \{0\}^\perp = \mathcal{H}$, so R^* has dense range.

(10.6.13) Let $\mathcal{H}$ be a Hilbert space with $\dim \mathcal{H} > k$ and

$$A = \{0, \lambda_1, \ldots, \lambda_k\} \subset \mathbb{C}.$$

Show that there exists $T \in \mathcal{K}(\mathcal{H})$ with $\sigma(T) = A$.

Answer. Let $\xi_1, \ldots, \xi_k$ be pairwise orthogonal unit vectors in $\mathcal{H}$ and for each j let P_j be the orthogonal projection onto $\mathbb{C}\,\xi_j$. Then

$$T = \sum_{j=1}^{k} \lambda_j\, P_j$$

is a finite-rank operator with $\sigma(T) = A$. The reason $0 \in \sigma(T)$ is that T cannot be invertible, for it has rank at most k and $\dim \mathcal{H} > k$, so T is not surjective; it is not injective either, for $\{\xi_1, \ldots, \xi_k\}^\perp \subset \ker T$.

(10.6.14) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space and $\{\lambda_k\} \subset \mathbb{C}$ be a sequence with $\lim \lambda_k = 0$. Let $A = \{0\} \cup \{\lambda_k\}$. Show that there exists $T \in \mathcal{K}(\mathcal{H})$ with $\sigma(T) = A$.

Answer. Let $\{\xi_k\}$ be an orthonormal basis and form $T = \sum_{j=1}^{\infty} \lambda_j\, P_j$, with P_j the orthogonal projection onto $\mathbb{C}\,\xi_j$. Then $T \in \mathcal{K}(\mathcal{H})$ and $\sigma(T) = A$ by Lemma 10.6.11.

(10.6.15) What is the relation between the operator in Exercise 10.6.6 and the operator in Exercise 10.6.12?

Answer. The operator T from Exercise 10.6.6 and the operator R from Exercise 10.6.12 are the same.

(10.6.16) Write the explicit form of (10.11) in the case where T is a diagonal matrix.

Answer. If T is diagonal, then $T = \sum_{j=1}^{n} T_{jj}\, E_{jj}$. That is precisely the form the Spectral Theorem gives.

(10.6.17) Show that if T is normal then $\sigma_k(T) = |\lambda_k(T)|$.

Answer. By the Spectral Theorem we can write

$$T = \sum_k \lambda_k(T)\, P_k.$$

Then

$$|T| = (T^2)^{1/2} = \sum_k |\lambda_k(T)|\, P_k.$$

By Corollary 10.6.28, $\sigma_k(T) = |\lambda_k(T)|$ for all k.

(10.6.18) Prove Proposition 10.6.19.

Answer. Write T as in (10.11) and $f(T)$ as in (10.12). The continuity of f at 0 is enough to guarantee that $f(\sigma(T))$ is compact. Indeed, the case where $\sigma(T)$ is finite is trivial, so let us assume that $\sigma(T)$ is infinite. Let $\{V_j\}$ be an open cover of $f(\sigma(T))$. As $0 = f(0)$, there exists j_0 with $0 \in V_{j_0}$. As f is continuous at 0, there exists $\delta > 0$ such that $|t| < \delta$ implies $f(t) \in V_{j_0}$. Because $\{\lambda_k\}$ is a sequence that converges to 0 (Theorem 9.6.13) there exists k_0 such that $|\lambda_k| < \delta$ for all $k > k_0$. For each of $1, 2, \ldots, k_0$ choose j_r such that $f(\lambda_r) \in V_{j_r}$. Then $V_{j_0}, V_{j_1}, \ldots, V_{j_{k_0}}$ is a finite subcover, showing that $f(\sigma(T))$ is compact.

Let $\mu \notin f(\sigma(T))$. As $f(\sigma(T))$ is compact, there exists $\delta > 0$ such that $\mu - f(\lambda_k) > \delta$ for all k. Then one can readily check that

$$[f(T) - \mu I]^{-1} = \sum_{k=0}^{\infty} \frac{1}{f(\lambda_k) - \mu}\, P_k$$

is bounded, since $\frac{1}{|f(\lambda_k)-\mu|} < \frac{1}{\delta}$, and it is the inverse of $f(T) - \mu I$. Thus $\mu \notin \sigma(f(T))$, showing that $\sigma(f(T)) \subset f(\sigma(T))$.

Conversely, if $\mu \notin \sigma(f(T))$, then $f(T) - \mu I$ is invertible; this implies $\mu \neq f(\lambda_k)$ for all k. It also implies that $\mu \neq f(0) = \lim_k f(\lambda_k)$ (the limit exists and it is the only accumulation point of the sequence by the continuity of f), because otherwise $f(T) - \mu I = f(T)$ would be compact and thus not invertible; therefore $\mu \notin f(\sigma(T))$.

(10.6.19) Let $T \in \mathcal{B}(\mathcal{H})$ be a positive compact operator. Show that

$$\xi \in \operatorname{ran} T \iff \text{there exists } C > 0 \text{ such that}$$

$$|\langle \xi, \eta\rangle| \leq C\|T\eta\|, \qquad \eta \in \mathcal{H}.$$

Show by example that the condition can fail for $\xi \in \overline{\operatorname{ran} T}$.

Answer.

If $\xi \in \operatorname{ran} T$, then there exists ξ' with $\xi = T\xi'$. Thus

$$|\langle \xi, \eta \rangle| = |\langle T\xi', \eta \rangle| = |\langle \xi', T\eta \rangle| \leq \|\xi'\| \, \|T\eta\|.$$

Conversely, suppose that there exists $C > 0$ with

$$|\langle \xi, \eta \rangle| \leq C \|T\eta\| \tag{AB.10.9}$$

for all $\eta \in \mathcal{H}$. By Theorem 10.6.12 there exists an orthonormal set $\{\eta_k\}$ such that for all ξ we have

$$T\xi = \sum_k \lambda_k \langle \xi, \eta_k \rangle \, \eta_k$$

with $\lambda_k \neq 0$ for all k. If $\eta' \perp \eta_k$ for all k, then $T\eta' = 0$, which by (AB.10.9) implies $\langle \xi, \eta' \rangle = 0$. So we can write $\xi = \sum_k x_k \eta_k$ for appropriate coefficients x_k. We have

$$\begin{aligned}
\sum_{k=1}^n \frac{|x_k|^2}{\lambda_k^2} &= \Big\langle \sum_{k=1}^n \frac{x_k}{\lambda_k} \, \eta_k, \sum_{k=1}^n \frac{x_k}{\lambda_k} \, \eta_k \Big\rangle \\
&= \Big\langle \sum_{k=1}^n x_k \, \eta_k, \sum_{k=1}^n \frac{x_k}{\lambda_k^2} \, \eta_k \Big\rangle = \Big\langle \xi, \sum_{k=1}^n \frac{x_k}{\lambda_k^2} \, \eta_k \Big\rangle \\
&\leq C \left\| T\Big(\sum_{k=1}^n \frac{x_k}{\lambda_k^2} \, \eta_k \Big) \right\| = C \left\| \sum_{k=1}^n \frac{x_k}{\lambda_k} \, \eta_k \right\| = C \left(\sum_{k=1}^n \frac{|x_k|^2}{\lambda_k^2} \right)^{1/2}
\end{aligned}$$

This implies that

$$\sum_{k=1}^n \frac{|x_k|^2}{\lambda_k^2} \leq C^2$$

for all n, and hence the full series is convergent. This allows us to define $\xi' = \displaystyle\sum_k \frac{x_k}{\lambda_k} \, \eta_k \in \mathcal{H}$, and it is immediate that $T\xi' = \xi$.

As for the example where the condition can fail, fix an orthonormal basis $\{\eta_k\}$ and let $T = \sum_k \frac{1}{k} \langle \cdot, \eta_k \rangle \, \eta_k$. Then T is compact, positive, with dense range since it is injective. The element $\xi = \sum_k \frac{1}{k} \, \eta_k$ is not in $\operatorname{ran} T$, as $\xi = T\xi'$ would require the coefficients of ξ' with respect to the basis $\{\eta_k\}$ to be all 1, which is impossible. If (AB.10.9) held for all η, we would have

$$\sum_{k=n}^\infty \frac{1}{k^2} = \Big\langle \xi, \sum_{k=n}^\infty \frac{1}{k} \, \eta_k \Big\rangle \leq C \left\| \sum_{k=n}^\infty \frac{1}{k} \, T\eta_k \right\|$$

$$= C \left\| \sum_{k=n}^\infty \frac{1}{k^2} \, \eta_k \right\| = C \left(\sum_{k=n}^\infty \frac{1}{k^4} \right)^{1/2}.$$

But, comparing via integrals,

$$\frac{\left(\sum_{k=n}^{\infty}\frac{1}{k^2}\right)^2}{\sum_{k=n}^{\infty}\frac{1}{k^4}} \geq \frac{\left(\int_{n+1}^{\infty}\frac{1}{x^2}\,dx\right)^2}{\int_{n}^{\infty}\frac{1}{x^4}\,dx} = \frac{\frac{1}{(n+1)^2}}{\frac{1}{3n^3}} = \frac{3n^3}{(n+1)^2}$$

which is unbounded. So C cannot exist.

(10.6.20) Show that if V is the Volterra operator on $L^2[0,1]$, its adjoint V^* is given by

$$V^*f(s) = \int_s^1 f(t)\,dt, \qquad f \in L^2[0,1].$$

Answer. We have

$$\begin{aligned}\langle V^*f, g\rangle = \langle f, Vg\rangle &= \int_0^1 f(s)\int_0^s g(t)\,dt\,ds \\ &= \int_0^1\int_t^1 f(s)\,ds\,g(t)\,dt = \left\langle \int_t^1 f\,g\right\rangle.\end{aligned}$$

So

$$V^*f(s) = \int_s^1 f(t)\,dt, \qquad f \in L^2[0,1].$$

(10.6.21) Let $T \in \mathcal{K}(\mathcal{H})$ and let $T = V|T|$ be its Polar Decomposition. Show that $|T| \in \mathcal{K}(\mathcal{H})$.

Answer. Because V^*V is the projection onto the closure of the range of T^* and $|T| = (T^*T)^{1/2}$, we have $V^*V|T| = |T|$. Then $|T| = V^*V|T| = V^*T \in \mathcal{K}(\mathcal{H})$ since $\mathcal{K}(\mathcal{H})$ is an ideal.

(10.6.22) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Use the Polar Decomposition to show that the unit ball of $\mathcal{K}(\mathcal{H})$ has no extreme points.

Answer. Let $T \in \mathcal{K}(\mathcal{H})$ with $\|T\| \leq 1$. By the Polar Decomposition and Exercise 10.6.21 we can write $T = V|T|$, with $|T|$ positive and compact. By the Spectral Theorem (10.6.12) we can write

$$|T| = \sum_{k=1}^{\infty} \lambda_k P_k,$$

where $\{P_k\}$ are nonzero pairwise orthogonal projections with $\sum_k P_k = I_{\mathcal{H}}$, and the sequence $\{\lambda_k\}$ converges to zero. So there exists j such that $0 \leq \lambda_j < 1$. Let $\delta = (1 - \lambda_j)/2$. Then $|T| = \frac{1}{2}\,T_1 + \frac{1}{2}\,T_2$, where

$$T_1 = (\lambda_j - \delta)P_j + \sum_{k \neq j} \lambda_k P_k, \qquad T_2 = (\lambda_j + \delta)P_j + \sum_{k \neq j} \lambda_k P_k.$$

Both T_1 and T_2 are selfadjoint compact operators with eigenvalues in $[-1, 1]$, so $\|T_1\| \leq 1$ and $\|T_2\| \leq 1$. And then we can write

$$T = V|T| = \frac{1}{2}\,VT_1 + \frac{1}{2}\,VT_2.$$

Since $V^*V|T| = |T|$ (as seen in the answer to Exercise 10.6.21), if we had $VT_1 = VT_2$ we would have $V^*VT_1 = V^*VT_2$. After cancelling the parts that are equal, this leads us to $(\lambda_j - \delta)V^*VP_j = (\lambda_j + \delta)V^*VP_j$, which reduces to $V^*VP_j = 0$. This gives $V^*V|T| \neq |T|$, a contradiction. So $T_1 \neq T_2$ and T is not extreme.

(10.6.23) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Show that $\mathcal{K}(\mathcal{H})$ is not a dual by using Exercise 10.6.22 and Krein–Milman (Theorem 7.5.11).

Answer. Suppose that $\mathcal{K}(\mathcal{H}) = \mathcal{X}^*$ for some Banach space $\mathcal{X}$. Then we have a weak*-topology on $\mathcal{K}(\mathcal{H})$ and in particular the unit ball is compact in this topology (by Banach–Alaoglu, Theorem 7.2.13). Then Krein–Milman (Theorem 7.5.11) shows that the unit ball of $\mathcal{K}(\mathcal{H})$ is the closed convex hull of its extreme points; as the unit ball in $\mathcal{K}(\mathcal{H})$ has no extreme points (Exercise 10.6.22), we conclude that $\mathcal{K}(\mathcal{H})$ is not a dual.

(10.6.24) Let $T \in \mathcal{K}(\mathcal{H})$ be normal. Show that $\overline{\operatorname{conv}}\,\sigma(T) = \overline{W(T)}$.

Answer. We know from Proposition 10.2.3 that $\overline{\operatorname{conv}}\,\sigma(T) \subset \overline{W(T)}$. Now fix $\lambda \in W(T)$. So $\lambda = \langle T\xi, \xi \rangle$ for some unit vector ξ. By the Spectral Theorem (Theorem 10.6.12) there exists an orthonormal basis $\{\xi_k\}$ such that

$$T = \sum_k \lambda_k\, P_k,$$

where P_k is the rank-one projection onto $\mathbb{C}\xi_k$. Then

$$\lambda = \langle T\xi, \xi\rangle = \sum_{k,j} \lambda_k \langle \xi, \xi_k\rangle \overline{\langle \xi, \xi_j\rangle} \langle \xi_k, \xi_j\rangle = \sum_k |\langle \xi, \xi_k\rangle|^2 \lambda_k.$$

From $\|\xi\| = 1$ we get that $\sum_k |\langle \xi, \xi_k\rangle|^2 = 1$. Fix $\varepsilon > 0$ and choose m such that $\sum_{k>m} |\langle \xi, \xi_k\rangle|^2 < \varepsilon$. Let $\alpha_k = |\langle \xi, \xi_k\rangle|^2$ for $k = 1, \dots, m$ and $\alpha_{m+1} = \sum_{k>m} |\langle \xi, \xi_k\rangle|^2$. Then $\alpha_1, \dots, \alpha_{m+1}$ are convex coefficients and

$$\begin{aligned}\Big|\lambda - \sum_{k=1}^{m+1} \alpha_k\lambda_k\Big| &= \Big|\sum_{k>m} |\langle \xi, \xi_k\rangle|^2 (1 - \lambda_{m+1})\Big| \\ &\le |1 - \lambda_{m+1}|\, \varepsilon \le (1 + \|T\|)\, \varepsilon.\end{aligned}$$

As $\varepsilon > 0$ can be chosen arbitrarily small, this shows that $\lambda \in \overline{\operatorname{conv}}\sigma(T)$. So $W(T) \subset \overline{\operatorname{conv}}\sigma(T)$ and therefore $\overline{\operatorname{conv}}\sigma(T) = \overline{W(T)}$.

(10.6.25) Let $\mathcal{H}$ be a Hilbert space with an orthonormal basis $\{\eta_j\}_{j\in J}$. For each $k, j \in J$, let E_{kj} be the rank-one operator $E_{kj}\xi = \langle \xi, \eta_j\rangle \eta_k$. These operators are called **matrix units** and satisfy the relations

$$E_{kj}E_{ab} = \delta_{j,a}\, E_{kb}. \qquad E_{kj}^* = E_{jk}. \tag{10.19}$$

Using notation we have also discussed, $E_{kj} = \eta_k\eta_j^*$.

(i) Prove the matrix unit relations (10.19).

(ii) Show every E_{kj} is a partial isometry and $\{E_{kk}\}$ are pairwise orthogonal mutually equivalent projections.

(iii) Let $T \in \mathcal{B}(\mathcal{H})$. Show that there exist unique numbers $\{t_{kj}\}_{k,j\in J}$ such that

$$T = \sum_{k,j} t_{kj}\, E_{kj}, \tag{10.20}$$

where the series converges pointwise (if coming from a later chapter, the series converges sot).

Answer.

(i) This was done in (10.10) and Exercise 10.6.4. Here is the argument in the given notation. Given $\xi \in \mathcal{H}$,

$$\begin{aligned}E_{kj}E_{ab}\xi &= \langle E_{ab}\xi, \eta_j\rangle \eta_k = \langle \xi, \eta_b\rangle \langle \eta_a, \eta_j\rangle \eta_k \\ &= \delta_{j,a}\langle \xi, \eta_b\rangle \eta_k = \delta_{j,a}\, E_{kb}\xi.\end{aligned}$$

And

$$\langle E_{kj}^*\xi, \eta\rangle = \langle \xi, E_{kj}\eta\rangle = \overline{\langle \eta, \eta_j\rangle}\,\langle \xi, \eta_k\rangle$$
$$= \langle \langle \xi, \eta_k\rangle\,\eta_j, \eta\rangle = \langle E_{jk}\xi, \eta\rangle.$$

As $\xi, \eta \in \mathcal{H}$ are arbitrary, $E_{kj}^* = E_{jk}$.

(ii) From (10.19),

$$E_{kj}^* E_{kj} = E_{jk}E_{kj} = E_{jj}, \qquad E_{kj}E_{kj}^* = E_{kj}E_{jk} = E_{kk}.$$

By Proposition 10.4.10, E_{kj} is a partial isometry and E_{kk} is a projection (the projection part can obtained directly from the matrix unit relations).

(iii) For each $k, j \in J$, let $t_{kj} = \langle T\eta_j, \eta_k\rangle$. Fix $\xi \in \mathcal{H}$, so $\xi = \sum_j \langle \xi, \eta_j\rangle\,\eta_j$. The idea of what we need to do is in Proposition 10.6.2. Fix $\varepsilon > 0$ and choose $F \in J$, finite. Let P_F be the orthogonal projection onto $\operatorname{span}\{\eta_j :\ j \in F\}$. Then

$$\begin{aligned}
\Big\|T\xi - \sum_{k,j\in F} t_{kj}\, E_{kj}\xi\Big\| &= \Big\|\sum_j \langle \xi, \eta_j\rangle T\eta_j - \sum_{k,j\in F}\langle T\eta_j, \eta_k\rangle\,\langle \xi, \eta_j\rangle\,\eta_k\Big\| \\
&= \Big\|\sum_{j,k}\langle \xi, \eta_j\rangle\,\langle T\eta_j, \eta_k\rangle\,\eta_k - \sum_{k,j\in F}\langle T\eta_j, \eta_k\rangle\,\langle \xi, \eta_j\rangle\,\eta_k\Big\| \\
&= \Big\|\sum_{j,k}\langle \xi, \eta_j\rangle\,\langle T\eta_j, \eta_k\rangle\,\eta_k - \sum_{k,j}\langle TP_F\eta_j, P_F\eta_k\rangle\,\langle \xi, \eta_j\rangle\,\eta_k\Big\| \\
&= \Big\|\sum_{j,k}\langle \xi, \eta_j\rangle\,\langle (T - P_FTP_F)\eta_j, \eta_k\rangle\,\eta_k\Big\| \\
&= \Big\|\sum_j \langle \xi, \eta_j\rangle\,(T - P_FTP_F)\eta_j\Big\| \\
&= \|(T - P_FTP_F)\xi\| \to 0
\end{aligned}$$

by Exercise 10.6.9. The double series above can be manipulated freely because we do not exchange the indices and everything converges.

For the uniqueness, if $T = \sum_{k,j} t_{kj}E_{kj}$ then

$$\langle T\eta_s, \eta_r\rangle = \Big\langle \sum_{k,j} t_{kj}E_{kj}\eta_s, \eta_r\Big\rangle = \Big\langle \sum_k t_{ks}\eta_k, \eta_r\Big\rangle = t_{rs}.$$

10.7. Trace-Class Operators

(10.7.1) Proving the inequality $|\operatorname{Tr}(T)| \leq \operatorname{Tr}(|T|)$ in (iii) of Proposition 10.7.5, it is tempting to get a direct proof by using

$$|\langle T\xi_n, \xi_n\rangle| \leq \langle |T|\xi_n, \xi_n\rangle,$$

from where the inequality would follow directly. Show that this inequality does not hold in general.

Answer. Let $\mathcal{H}$ be any Hilbert space with $\dim \mathcal{H} \geq 2$, and let $\{\xi_n\}$ be an orthonormal basis for $\mathcal{H}$, with the usual associated matrix units $\{E_{kj}\}$. Put

$$T = E_{12}, \qquad \xi = \frac{\sqrt{3}}{2}\,\xi_1 + \frac{1}{2}\,\xi_2.$$

Then $|T| = E_{22}$ and

$$\langle T\xi, \xi\rangle = \frac{\sqrt{3}}{4} > \frac{1}{4} = \langle |T|\xi, \xi\rangle.$$

(10.7.2) Let $\varphi : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ be linear and *positive*, that is $\varphi(T) \geq 0$ if $T \geq 0$. Prove the Cauchy–Schwarz inequality

$$|\varphi(S^*T)| \leq \varphi(S^*S)^{1/2}\varphi(T^*T)^{1/2}.$$

Answer. The form $[S,T] = \varphi(S^*T)$ is sesquilinear, and so the proof of Cauchy–Schwarz (Theorem 4.2.2) applies.

(10.7.3) Show that if $\dim \mathcal{H} = \infty$ then the inclusions $\mathcal{F}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H}) \subset \mathcal{K}(\mathcal{H})$ are proper.

Answer. Fix an orthonormal sequence $\{\xi_k\}$ for $\mathcal{H}$. Define the operators $S, T \in \mathcal{B}(\mathcal{H})$ by

$$S\xi = \sum_k \frac{1}{k^2}\,\langle \xi, \xi_k\rangle\,\xi_k, \qquad T\xi = \sum_k \frac{1}{k}\,\langle \xi, \xi_k\rangle\,\xi_k.$$

Then S is trace-class but not finite-rank, and T is compact but not trace-class. This can be verified directly because both S, T are positive, and then $\operatorname{Tr}(S) = \sum_k \frac{1}{k^2} < \infty$ and $\operatorname{Tr}(T) = \sum_k \frac{1}{k} = \infty$. Finally, S is not finite-rank because $\xi_k = S(k\xi_k) \in \operatorname{ran} S$ for all k.

(10.7.4) Show that $\mathcal{T}(\mathcal{H})^* = \mathcal{B}(\mathcal{H})$; that is, show that there is an isometric isomorphism $\Gamma : \mathcal{B}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})^*$ such that $\Gamma(T)(S) = \operatorname{Tr}(TS)$.

Answer. We define $\Gamma : \mathcal{B}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})^*$ by $\Gamma(T)(S) = \operatorname{Tr}(TS)$. This map is clearly linear. We have $|\operatorname{Tr}(TS)| \leq \|T\|\,\|S\|_1$, which shows that $\|\Gamma(T)\| \leq \|T\|$. If $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ and $\|T\xi\| \geq (1-\varepsilon)\|T\|$ for some $\varepsilon > 0$, let $S = \frac{1}{\|T\xi\|}\,\xi(T\xi)^*$. Then S is rank-one, so trace-class, with $\|S\| = 1$. And calculating over an orthonormal basis whose first element is ξ,

$$\operatorname{Tr}(ST) = \langle \frac{1}{\|T\xi\|}\,\xi(T\xi)^*T\xi, \xi\rangle = \|T\xi\| \geq (1-\varepsilon)\|T\|.$$

It follows that $\|\Gamma(T)\| = \|T\|$. So $\Gamma(T)$ is isometric. It remains to show that Γ is surjective. Fix $\psi \in \mathcal{T}(\mathcal{H})^*$. Consider the sesquilinear form $[\xi, \eta] = \psi(\xi\eta^*)$. Since $|[\xi,\eta]| \leq \|\psi\|\,\|\xi\|\,\|\eta\|$, by Proposition 10.1.5 there exists $T \in \mathcal{B}(\mathcal{H})$ with $\psi(\xi\eta^*) = \langle T\xi, \eta\rangle$. If $S \in \mathcal{T}(\mathcal{H})$ is positive, by the Spectral Theorem we can write

$$S = \sum_{j=1}^\infty \lambda_j\,\xi_j\xi_j^*$$

for an orthonormal basis of eigenvectors for S. Then

$$\begin{aligned}\psi(S) &= \sum_{j=1}^\infty \lambda_j\,\psi(\xi_j\xi_j^*) = \sum_{j=1}^\infty \lambda_j\,\langle T\xi_j, \xi_j\rangle \\ &= \sum_{j=1}^\infty \langle T\xi_j, \lambda_j\xi_j\rangle = \sum_{j=1}^\infty \langle T\xi_j, S\xi_j\rangle = \operatorname{Tr}(ST).\end{aligned}$$

As any $S \in \mathcal{T}(\mathcal{H})$ can be written as a linear combination of four positive trace-class operators, the equality $\operatorname{Tr}(TS) = \psi(S)$ holds for all $S \in \mathcal{T}(\mathcal{H})$.

(10.7.5) Let $T \in \mathcal{B}(\mathcal{H})$. Show that if $\lambda \neq 0$ then $\ker T \cap \ker(T-\lambda I)^n = \{0\}$ for all n.

Answer. If $\eta \in \ker T \cap \ker(T-\lambda I)^n$, then $(T-\lambda I)\eta = -\lambda\eta$, and iterating we get

$$(-1)^n\lambda^n\eta = (T-\lambda I)^n\eta = 0,$$

and so $\eta = 0$.

(10.7.6) Show that $\dim\ker(T - \lambda I)^n < \infty$ for any T compact, $\lambda \in \mathbb{C} \setminus \{0\}$, and $n \in \mathbb{N}$.

Answer. We have

$$(T - \lambda I)^n = (-1)^n \lambda^n I + \sum_{j=0}^{n-1} \binom{n}{j} (-1)^j \lambda^j T^{n-j} = (-1)^n \lambda^n I + S,.$$

with S compact. If $\mu = (-1)^{n+1}\lambda^n$, then $\ker(T - \lambda I)^n = \ker(S - \mu I)$ is finite-dimensional by Corollary 9.6.14.

(10.7.7) Show that if $\ker(T - \lambda I)^n = \ker(T - \lambda I)^{n+1}$ then $\ker(T - \lambda I)^{n+k} = \ker(T - \lambda I)^n$ for all $k \in \mathbb{N}$.

Answer. This was done in Exercise 9.1.6, but we include and ad-hoc argument here.

Suppose that $\ker(T-\lambda I)^n = \ker(T-\lambda I)^{n+1}$, and let $v \in \ker(T-\lambda I)^{n+k}$. Then, since $n = (n+1)+(k-1)$ we have that $(T-\lambda I)^{k-1}v \in \ker(T-\lambda I)^{n+1} = \ker(T - \lambda I)^n$, which implies that $v \in \ker(T - \lambda I)^{n+k-1}$. Iterating this we get that $v \in \ker(T - \lambda I)^n$. Hence $\ker(T - \lambda I)^{n+k} = \ker(T - \lambda I)^n$ for all $k \in \mathbb{N}$.

(10.7.8) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $T : \mathcal{H} \to \mathcal{K}$, $S : \mathcal{K} \to \mathcal{H}$ linear and such that $TS \in \mathcal{T}(\mathcal{H})$, $ST \in \mathcal{T}(\mathcal{K})$. Show that ST and TS have the same nonzero eigenvalues, with $\alpha_\lambda(ST) = \alpha_\lambda(TS)$ for all $\lambda \in \sigma(ST) \setminus \{0\}$. Conclude that $\operatorname{Tr}(TS) = \operatorname{Tr}(ST)$.

Answer. We know that ST and TS, being compact, have all their nonzero elements of the spectrum as eigenvalues (Theorem 10.6.8), and that nonzero elements of the spectrum are the same for both (Proposition 9.2.15; the proof there is phrased in the context of an algebra, but all that matters is that we can multiply).

Let $\xi \in \ker(ST - \lambda I)^k$ be nonzero. We can write this as

$$\sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j} (ST)^j \xi = 0.$$

Then $T\xi \neq 0$, for otherwise the equality above becomes $\lambda^k \xi = 0$, a contradiction. It follows that T is injective when restricted to $\bigcup_k \ker(ST - \lambda I)^k$.

We also have that $T\xi \in \bigcup_k \ker(TS - \lambda I)^k$. Indeed, from $(TS - \lambda I)T = T(ST - \lambda I)$ we get by induction $(TS - \lambda I)^k T = T(ST - \lambda I)^k$ for all k, and then $(TS - \lambda I)^k T\xi = T(ST - \lambda I)^k \xi = 0$.

Thus T is an injective map from $\bigcup_k \ker(ST - \lambda I)^k$ into $\bigcup_k \ker(TS - \lambda I)^k$, and hence $\alpha_\lambda(ST) \leq \alpha_\lambda(TS)$. As the roles of S, T can be exchanged, $\alpha_\lambda(ST) = \alpha_\lambda(TS)$.

The equality $\operatorname{Tr}(TS) = \operatorname{Tr}(ST)$ now follows from (10.27).

(10.7.9) Let $S \in \mathcal{B}(\mathcal{H})$. Show that $S^*S \in \mathcal{T}(\mathcal{H})$ if and only if $SS^* \in \mathcal{T}(\mathcal{H})$.

Answer. Let $S = U|S|$ be the polar decomposition of S. Suppose that $S^*S \in \mathcal{T}(\mathcal{H})$. Then

$$SS^* = U|S|^2U^* = US^*SU^* \in \mathcal{T}(\mathcal{H})$$

since $\mathcal{T}(\mathcal{H})$ is an ideal. The converse is shown by exchanging the roles of S and S^*.

(10.7.10) Let $S \in \mathcal{T}(\mathcal{H})$. Show that

$$\inf\{\|S + R\|_1 : \operatorname{Tr}(R) = 0\} = |\operatorname{Tr}(S)|. \tag{10.31}$$

Answer.

For any R with $\operatorname{Tr}(R) = 0$ we have

$$|\operatorname{Tr}(S)| = |\operatorname{Tr}(S + R)| \leq \|S + R\|_1.$$

Hence $|\operatorname{Tr}(S)| \leq \inf\{\|S + R\|_1 : \operatorname{Tr}(R) = 0\}$.

Fix $\varepsilon > 0$. By Proposition 10.7.9 there exists $F \in \mathcal{F}(\mathcal{H})$ with $\|S - F\|_1 < \varepsilon$. Writing F as in Proposition 10.6.1, we see that F acts on the subspace $\mathcal{H}_0 = \operatorname{span}\{\xi_k, \eta_k : k\}$. This allows us to apply Schur's Triangularization (Proposition 1.7.14) to get (after expansion to the whole space) an orthonormal basis in which F is triangular. Say the orthonormal basis is $\{\xi_k\}$, so

$$F\xi_k = \lambda_k \xi_k + \sum_{j=1}^{k-1} \alpha_{kj} \xi_j$$

and $\lambda_1, \ldots, \lambda_n$ are the nonzero eigenvalues of F. Now form the finite-rank operator R where

$$R\xi_k = -\lambda_k \xi_k - \sum_{j=1}^{k-1} \alpha_{kj}\xi_j, \qquad k = 2, \ldots, n$$

and

$$R\xi_1 = \Big(\sum_{j=2}^{n} \lambda_j\Big)\xi_1.$$

Then $\operatorname{Tr}(R) = 0$ and

$$\|F + R\|_1 = \big\| \operatorname{Tr}(F)E_{11}\big\| = |\operatorname{Tr}(F)|.$$

Then

$$\begin{aligned} |\operatorname{Tr}(S)| &\geq |\operatorname{Tr}(F)| - \varepsilon = \|F + R\|_1 - \varepsilon \geq \|S + R\|_1 - 2\varepsilon \\ &\geq -2\varepsilon + \inf\{\|S + R\|_1 : \operatorname{Tr}(R) = 0\}. \end{aligned}$$

As this can be done for any $\varepsilon > 0$, we get that $|\operatorname{Tr}(S)| \geq \inf\{\|S + R\|_1 : \operatorname{Tr}(R) = 0\}$.

(10.7.11) Let $\{\xi_n\}, \{\eta_n\} \subset \mathcal{H}$ such that

$$\sum_n \|\xi_n\|^2 < \infty, \qquad \sum_n \|\eta_n\|^2 < \infty.$$

Use Proposition 10.7.12 to show that the operator $S = \sum_n \xi_n\eta_n^*$ is trace-class.

Answer. First we check quickly that $S \in \mathcal{B}(\mathcal{H})$. This is because

$$\begin{aligned} \|S\xi\| &= \Big\|\sum_n \xi_n \langle \xi, \eta_n\rangle\Big\| \leq \sum_n \|\xi_n\|\,\|\eta_n\|\,\|\xi\| \\ &\leq \Big(\sum_n \|\xi_n\|^2\Big)^{1/2}\Big(\sum_n \|\eta_n\|^2\Big)^{1/2}\|\xi\|. \end{aligned}$$

Fix an orthonormal basis $\{\nu_n\}$. Then

$$\begin{aligned}\sum_k |\langle S\nu_k,\nu_k\rangle| = \sum_k \left|\sum_n \langle \xi_n,\nu_k\rangle\,\langle \nu_k,\eta_n\rangle\right| &\le \sum_k\sum_n |\langle \xi_n,\nu_k\rangle|\,|\langle \nu_k,\eta_n\rangle| \\ &= \sum_n\sum_k |\langle \xi_n,\nu_k\rangle|\,|\langle \nu_k,\eta_n\rangle| \\ &\le \sum_n \Big(\sum_k |\langle \xi_n,\nu_k\rangle|^2\Big)^{1/2}\Big(\sum_k |\langle \nu_k,\eta_n\rangle|^2\Big)^{1/2} \\ &= \sum_n \|\xi_n\|\,\|\eta_n\| \le \Big(\sum_n \|\xi_n\|^2\Big)^{1/2}\Big(\sum_n \|\eta_n\|^2\Big)^{1/2} < \infty.\end{aligned}$$

As the computation would be the same under any reordering of $\{\nu_n\}$, we have shown that $\sum_n \langle S\nu_n,\nu_n\rangle$ converges for any orthonormal basis $\{\nu_n\}$. Hence $S \in \mathcal{T}(\mathcal{H})$ by Proposition 10.7.12.

(10.7.12) Let TS be as in (10.28) and let $z_1,\ldots,z_n \in \mathbb{C}$ with $|z_j| \le \frac{1}{2}$ for all j. Show that there exists an orthonormal basis $\{\xi_k\}$ such that the diagonal of TS in said basis begins with $z_1,\ldots,z_n$.

Answer. We have $TS = \sum_k E_{2k-1,2k}$, where the matrix units come from the orthonormal basis $\{\eta_j\}$. Write $z_j = r_j e^{i\theta_j}$ the polar form. Let

$$\xi_j = r_j^{1/2} e^{-i\theta_j}\,\eta_{4j-3} + r_j^{1/2}\,\eta_{4j-2} + (1-2r_j)^{1/2}\,\eta_{4j-1}, \qquad j = 1,\ldots,n.$$

Then $\xi_1,\ldots,\xi_n$ are orthonormal and

$$\begin{aligned}\langle (TS)\xi_j,\xi_j\rangle &= \sum_k \langle E_{2k-1,2k}\xi_j,\xi_j\rangle \\ &= \langle r_j^{1/2}\eta_{4j-3}, r_j^{1/2} e^{-i\theta_j}\,\eta_{4j-3} + r_j^{1/2}\,\eta_{4j-2} + (1-2r_j)^{1/2}\,\eta_{4j-1}\rangle \\ &= r_j^{1/2}\overline{r_j^{1/2} e^{-i\theta_j}} = z_j.\end{aligned}$$

We finish by extending $\{\xi_1,\ldots,\xi_n\}$ to an orthonormal basis.

(10.7.13) Let $\xi,\eta \in \mathcal{H}$. Show that

$$\|\xi\eta^*\|_1 = \|\xi\|\,\|\eta\|. \qquad (10.32)$$

Answer. Let Q be the projection onto $\mathbb{C}\eta$; that is, $Q = \frac{1}{\|\eta\|^2}\,\eta\eta^*$. We have

$$|\xi\eta^*| = (\eta\xi^*\xi\eta^*)^{1/2} = \|\xi\|\,(\eta\eta^*)^{1/2} = \|\xi\|\,\|\eta\|\,Q^{1/2} = \|\xi\|\,\|\eta\|\,Q.$$

Then

$$\|\xi\eta^*\|_1 = \operatorname{Tr}(|\xi\eta^*|) = \|\xi\|\,\|\eta\|\,\operatorname{Tr}(Q) = \|\xi\|\,\|\eta\|.$$

(10.7.14) Let $\xi, \eta \in \mathcal{H}$ be unit vectors and $P = \xi\xi^*$, $Q = \eta\eta^*$ the corresponding rank-one projections. Show that

$$\|P - Q\|_1 = 2\sqrt{1 - |\langle \xi, \eta\rangle|^2}.$$

Answer. We have

$$(P - Q)^2 = P + Q - 2\operatorname{Re} PQ,$$

and

$$\operatorname{Tr}(PQ) = \operatorname{Tr}(\xi\xi^*\eta\eta^*) = \langle \xi\xi^*\eta\eta^*\,\eta, \eta\rangle = |\langle \xi, \eta\rangle|^2.$$

Since P, Q are rank-one, $P - Q$ has rank at most 2; so it has at most two nonzero eigenvalues, say α, β. The equality $\operatorname{Tr}(P-Q) = 0$ forces $\beta = -\alpha$. As $P-Q$ is selfadjoint, the eigenvalues λ_1, λ_2 of $|P-Q|$ are the square roots of the eigenvalues of $(P - Q)^2$; that is, $\lambda_1 = |\alpha^2|^{1/2} = |\alpha|$, $\lambda_2 = |\beta^2|^{1/2} = |\alpha| = \lambda_1$. Then

$$\begin{aligned}\|P - Q\|_1 &= \lambda_1 + \lambda_2 = 2\lambda_1 = \sqrt{2}\,(2\lambda_1^2)^{1/2}\\ &= \sqrt{2}\left(\lambda_1^2 + \lambda_2^2\right)^{1/2} = \sqrt{2}\,\operatorname{Tr}\left((P - Q)^2\right)^{1/2}\\ &= \sqrt{2}\,\operatorname{Tr}(P + Q - 2\operatorname{Re} PQ)^{1/2} = \sqrt{2}\left(2 - 2|\langle \xi, \eta\rangle|^2\right)^{1/2}\\ &= 2\sqrt{1 - |\langle \xi, \eta\rangle|^2}.\end{aligned}$$

(10.7.15) Let $\mathcal{H}$ be a Hilbert space. Show that the Banach space $\mathcal{T}(\mathcal{H})$ is separable if and only if $\mathcal{H}$ is separable.

Answer. Suppose that $\mathcal{H}$ is separable. By Proposition 10.7.9 it is enough to show that $\mathcal{F}(\mathcal{H})$ is separable. Let $\{\nu_n\}$ be a countable dense subset in $\mathcal{H}$. Fix $\varepsilon > 0$ and $T \in \mathcal{F}(\mathcal{H})$. By Proposition 10.6.1 we can write $T = \sum_{k=1}^{n} \xi_k\eta_k^*$ for certain $\xi_k, \eta_k \in \mathcal{H}$, $k = 1, \ldots, n$. Let $c = \max\{\|\xi_k\| + \|\eta_k\| : \ k\} + 1$. For each k choose n_k, m_k such that $\|\nu_{n_k} - \xi_k\| < \varepsilon/(2cn)$ and $\|\nu_{m_k} - \eta_k\| < \varepsilon/(2cn)$.

Then, using (10.32),

$$\begin{aligned}\Big\|T-\sum_{k=1}^n \nu_{n_k}\nu_{m_k}^*\Big\|_1 &\leq \sum_{k=1}^n \|\xi_k\eta_k^*-\nu_{n_k}\nu_{m_k}^*\|_1\\ &\leq \sum_{k=1}^n \|\xi_k\eta_k^*-\xi_k\nu_{m_k}^*\|_1+\|\xi_k\nu_{m_k}^*-\nu_{n_k}\nu_{m_k}^*\|_1\\ &= \sum_{k=1}^n \|\xi_k\|\,\|\eta_k-\nu_{m_k}\|+\|\nu_{m_k}\|\,\|\xi_k-\nu_{n_k}\|\\ &\leq \sum_{k=1}^n \frac{2c\varepsilon}{2cn}=\varepsilon.\end{aligned}$$

This shows that the countable set

$$C=\Big\{\sum_{k=1}^n \nu_{n_k}\nu_{m_k}^*\ :\ n,n_k,m_k\in\mathbb{N},\ k=1,\ldots,n\Big\}$$

is dense in $\mathcal{T}(\mathcal{H})$.

Conversely, if $\mathcal{H}$ is not separable then there is an uncountable orthonormal basis $\{\xi_j\}\subset\mathcal{H}$. Then the set of rank-one operators $\{\xi_j\xi_j^*\}$ is uncountable, and by Exercise 10.7.14

$$\|\xi_j\xi_j^*-\xi_k\xi_k^*\|_1=2,$$

so $\mathcal{T}(\mathcal{H})$ cannot be separable.

(10.7.16) Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces and $U:\mathcal{H}\to\mathcal{K}$ a unitary. Show that $A\in\mathcal{T}(\mathcal{H})$ if and only if $UAU^*\in\mathcal{T}(\mathcal{K})$.

Answer. Fix $A\in\mathcal{T}(\mathcal{H})$. This means that $\mathrm{Tr}(|A|)<\infty$. Because U is a unitary, $|UAU^*|=U|A|U^*$ (simply check that $(U|A|U^*)^2=(UAU^*)^*UAU^*$, and recall that the positive square root is unique). Let $\{\eta_j\}$ be an orthonormal basis of $\mathcal{K}$. Then $\{U^*\eta_j\}$ is an orthonormal basis for $\mathcal{H}$. Hence

$$\begin{aligned}\mathrm{Tr}(|UAU^*|)=\mathrm{Tr}(U|A|U^*)&=\sum_j\langle U|A|U^*\eta_j,\eta_j\rangle\\ &=\sum_j\langle |A|U^*\eta_j,U^*\eta_j\rangle=\mathrm{Tr}(|A|)<\infty.\end{aligned}$$

So $UAU^*\in\mathcal{T}(\mathcal{K})$. The converse follows immediately by using that U^* is a unitary.

(10.7.17) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $U : \mathcal{H} \to \mathcal{K}$ a unitary. Show that $\text{Tr}(UAU^*) = \text{Tr}(A)$ for all $A \in \mathcal{T}(\mathcal{H})$ *(this is slightly less trivial than it looks, since we are using the trace in two different spaces).*

Answer. This follows from Exercise 10.7.8 but we offer here a short ad-hoc proof. Fix $A \in \mathcal{T}(\mathcal{H})$. By Exercise 10.7.16, $UAU^* \in \mathcal{T}(\mathcal{K})$. Now we can do the same computation as in Exercise 10.7.16. Let $\{\eta_j\}$ be an orthonormal basis of $\mathcal{K}$. Then $\{U^*\eta_j\}$ is an orthonormal basis for $\mathcal{H}$. Hence

$$\text{Tr}(UAU^*) = \sum_j \langle UAU^*\eta_j, \eta_j\rangle = \sum_j \langle AU^*\eta_j, U^*\eta_j\rangle = \text{Tr}(A).$$

CHAPTER 11

C*-Algebras

11.1. C*-Algebra Basics

(11.1.1) Prove the statements made in Remark 11.1.3.

Answer.

- $M_n(\mathbb{C})$ is separable since it is finite dimensional.
- c_0 is separable, since we can form the countable dense subset
$$\text{span}\{(q+ip)e_n : \ q,p \in \mathbb{Q}, \ n \in \mathbb{N}\}.$$
- $\ell^\infty(\mathbb{N})$ is not separable because it contains the uncountable subset $\{0,1\}^{\mathbb{N}}$, and any two elements in it are at distance 1, so there cannot be a dense countable subset.
- We can embed $\ell^\infty(\mathbb{N})$ in $\mathcal{B}(\mathcal{H})$ as multiplication operators over a fixed orthonormal basis, so $\mathcal{B}(\mathcal{H})$ cannot be separable when $\dim \mathcal{H} = \infty$.

(11.1.2) Show that the norm defined in (11.1) is submultiplicative.

Answer. We have

$$\|(a,\lambda)(b,\mu)\| = \sup\{\|abc + \lambda bc + \mu ac + \lambda\mu c\| : \ \|c\| = 1\}.$$

Since $\|abc + \lambda bc + \mu ac + \lambda\mu c\| = \|a(bc+\mu c) + \lambda(bc+\mu c)\|$, we get that

$$\|(a,\lambda)(b,\mu)\| \leq \sup_{\|c\|\leq 1} \|(a,\lambda)\| \, \|bc+\mu c\| = \|(a,\lambda)\| \, \|(b,\mu)\|.$$

(11.1.3) Show that if $x \in \tilde{\mathcal{A}}$ satisfies $bx = 0$ for all $b \in \mathcal{A}$, then $x = 0$.

Answer. We have $x = (a,\lambda)$ for some $a \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. The condition $bx = 0$ looks like $ba + \lambda b = 0$. Then

$$\begin{aligned} \|(a^*,\bar\lambda)\| &= \sup\{\|a^*b + \bar\lambda b\| : \ b \in \mathcal{A}, \|b\| \leq 1\}] \\ &= \sup\{\|ba + \lambda b\| : \ b \in \mathcal{A}, \ \|b\| \leq 1\} = 0. \end{aligned}$$

Thus $(a,\lambda)^* = 0$, and then $(a,\lambda) = 0$.

(11.1.4) Show that if $\mathcal{A}$ is a C*-algebra, unital or not, then

$$\|(a,\lambda)\| = \max\big\{|\lambda|, \sup\{\|ab + \lambda b\| : \ \|b\| \leq 1\}\big\}.$$

defines a norm that makes $\tilde{\mathcal{A}}$ a C*-algebra, and that when $\mathcal{A}$ is not unital we recover $\tilde{\mathcal{A}}$ as in Proposition 11.1.4. *(Hint: at some point you will possibly need the fact that the norm on a C*-algebra is unique)*

Answer. If $\|(a,\lambda)\| = 0$, then $|\lambda| = 0$, and

$$0 = \sup\{\|ab + \lambda b\| : \ \|b\| \leq 1\} = \sup\{\|ab\| : \ \|b\| \leq 1\} \geq \|a\|,$$

so $a = 0$. The homogeneity follows easily since

$$\begin{aligned} \|\mu(a,\lambda)\| = \|(\mu a, \mu\lambda)\| &= \max\big\{|\mu\lambda|, \sup\{\|\mu ab + \mu\lambda b\| : \ \|b\| \leq 1\}\big\} \\ &= |\mu| \, \|(a,\lambda)\|. \end{aligned}$$

For the subadditivity,

$$\begin{aligned}\|(a,\lambda)+(a',\lambda')\| &= \|(a+a',\lambda+\lambda')\|\\ &= \max\big\{|\lambda+\lambda'|,\sup\{\|(a+a')b+(\lambda+\lambda')b\|:\ \|b\|\le 1\}\big\}\\ &\le \|(a,\lambda)\|+\|(a',\lambda')\|,\end{aligned}$$

where we are using that the absolute value, the norm of $\mathcal{A}$, and the supremum are subadditive. Now the submultiplicativity. We have

$$\begin{aligned}\|(a,\lambda)(a',\lambda')\| &= \|(aa'+\lambda a'+\lambda' a,\lambda\lambda')\|\\ &= \max\big\{|\lambda\lambda'|,\sup\{\|aa'b+\lambda a'b+\lambda' ab+\lambda\lambda' b\|:\ \|b\|\le 1\}\big\}\\ &= \max\big\{|\lambda|\,|\lambda'|,\sup\{\|a(a'b+\lambda' b)+\lambda(a'b+\lambda' b)\|:\ \|b\|\le 1\}\big\}\\ &\le \max\big\{|\lambda|\,|\lambda'|,\|(a,\lambda)\|\,\sup\{\|a'b+\lambda' b\|:\ \|b\|\le 1\}\big\}\\ &= \max\big\{|\lambda|\,|\lambda'|,\|(a,\lambda)\|\,\|(a',\lambda')\|\big\}\\ &\le \|(a,\lambda)\|\,\|(a',\lambda')\|.\end{aligned}$$

And now we have, with the same idea as in the proof of Proposition 11.1.4,

$$\begin{aligned}\|(a,\lambda)\|^2 &= \max\big\{|\lambda|^2,\sup\{\|ab+\lambda b\|^2:\ \|b\|\le 1\}\big\}\\ &= \max\big\{|\lambda|^2,\sup\{\|(ab+\lambda b)^*(ab+\lambda b)\|:\ \|b\|\le 1\}\big\}\\ &= \max\big\{|\lambda|^2,\sup\{\|b^*a^*ab+\bar\lambda b^*ab+\lambda b^*a^*b+|\lambda|^2b^*b\|:\ \|b\|\le 1\}\big\}\\ &= \max\big\{|\lambda|^2,\sup\{\|b^*(a^*ab+\bar\lambda ab+\lambda a^*b+|\lambda|^2b)\|:\ \|b\|\le 1\}\big\}\\ &\le \max\big\{|\lambda|^2,\sup\{\|a^*ab+\bar\lambda ab+\lambda a^*b+|\lambda|^2b\|:\ \|b\|\le 1\}\big\}\\ &= \|(a^*a+\bar\lambda a+\lambda a^*,|\lambda|^2)\| = \|(a,\lambda)^*(a,\lambda)\|\\ &\le \|(a,\lambda)^*\|\,\|(a,\lambda)\|.\end{aligned}$$

Then when (a,λ) is nonzero we can cancel and we get $\|(a,\lambda)\|\le\|(a,\lambda)^*\|$. As the roles can be exchanged, this becomes an equality and so

$$\|(a,\lambda)\|^2\le\|(a,\lambda)^*(a,\lambda)\|\le\|(a,\lambda)^*\|\,\|(a,\lambda)\|=\|(a,\lambda)\|^2.$$

Thus $\|(a,\lambda)\|^2=\|(a,\lambda)^*(a,\lambda)\|$.

When $\mathcal{A}$ is non-unital we know from Proposition 11.1.4 that $\sup\{\|ab+\lambda b\|:\ \|b\|\le 1\}$ defines a norm on $\mathcal{A}$. Being a subalgebra of $\tilde{\mathcal{A}}$, and because the norm on a C*-algebra is unique, we get

$$\max\big\{|\lambda|,\sup\{\|ab+\lambda b\|:\ \|b\|\le 1\}\big\}=\sup\{\|ab+\lambda b\|^2:\ \|b\|\le 1\}.$$

(11.1.5) Show that the inequality

$$|\lambda| \leq \sup\{\|ab + \lambda b\| : \ \|b\| \leq 1\}, \qquad a \in \mathcal{A},$$

holds when $\mathcal{A}$ is non-unital, and it fails when $\mathcal{A}$ is unital.

Answer. When $\mathcal{A}$ is non-unital, we have by Exercise 11.1.4 that $\sup\{\|ab + \lambda b\|^2 : \ \|b\| \leq 1\}$ defines a norm on $\mathcal{A}$ and

$$\max\left\{|\lambda|, \sup\{\|ab + \lambda b\| : \ \|b\| \leq 1\}\right\} = \sup\{\|ab + \lambda b\|^2 : \ \|b\| \leq 1\}.$$

In particular,

$$|\lambda| \leq \sup\{\|ab + \lambda b\|^2 : \ \|b\| \leq 1\}.$$

When $\mathcal{A}$ is unital we can take $a = I_{\mathcal{A}}$, $\lambda = -1$, and the inequality above becomes $1 = |\lambda| \leq 0$.

(11.1.6) Show that, for any $a \in \mathcal{A}$, $\sigma(a^*) = \overline{\sigma(a)}$.

Answer. If $a - \lambda I_{\mathcal{A}}$ is invertible, then $a^* - \bar{\lambda} I_{\mathcal{A}} = (a - \lambda I_{\mathcal{A}})^*$ is invertible by (11.3).

(11.1.7) Let $\{\mathcal{A}_j\}_{j \in J}$ be a family of C*-subalgebras of a C^*-algebra $\mathcal{A}$. Show that $\bigcap_j \mathcal{A}_j$ is a C*-subalgebra of $\mathcal{A}$.

Answer. Let $\mathcal{B} = \bigcap_j \mathcal{A}_j$. Being an intersection of algebras, $\mathcal{B}$ is an algebra. If $b \in \mathcal{B}$, then $b \in \mathcal{A}_j$ for all j, and so $b^* \in \mathcal{A}_j$ for all j, which means that $b^* \in \mathcal{B}$. So $\mathcal{B}$ is a *-algebra. As C*-algebras are closed, $\mathcal{B}$ is closed being an intersection of closed sets. So $\mathcal{B}$ is a closed $*$-subalgebra of $\mathcal{A}$; that is to say, a C*-subalgebra.

(11.1.8) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$. Show that $C^*(a)$ is equal to

$$\overline{\{p(a, a^*) : \ p \ \text{ non-commutative polynomial with } p(0,0) = 0\}}.$$

Answer. Let

$$\mathcal{B} = \overline{\{p(a, a^*) : \ p \ \text{ non-commutative polynomial with } p(0,0) = 0\}}.$$

Being the closure of a $*$-subalgebra of $\mathcal{A}$, it is a C*-subalgebra of $\mathcal{A}$ that contains a. Hence $C^*(a) \subset \mathcal{B}$. On the other hand $p(a, a^*) \in C^*(a)$ for any non-commutative polynomial with $p(0,0)=0$. So $\mathcal{B}\subset C^*(a)$, and therefore $C^*(a)=\mathcal{B}$.

(11.1.9) Let $\mathcal{A}$ be a C*-algebra such that every normal element has spectrum consisting of a single point. Prove that $\mathcal{A}=\mathbb{C}$.

Answer. Let $a\in\mathcal{A}$ be normal with $\sigma(a)=\{\lambda\}$. Working in the unitization if needed (which is how we define $\sigma(a)$), we have that $b=a-\lambda\,I_{\mathcal{A}}$ is normal and $\sigma(b)=\{0\}$. By Lemma 11.1.10, $\|b\|=\operatorname{spr}(b)=0$, so $b=0$ and hence $a=\lambda I_{\mathcal{A}}$. As the normal elements span $\mathcal{A}$ by Proposition 11.1.13, we get that $\mathcal{A}=\mathbb{C}$.

(11.1.10) Show that if $\mathcal{A}$ is not unital then $(\tilde{\mathcal{A}})^*\simeq\mathcal{A}^*\oplus_1\mathbb{C}$.

Answer. Let $\Gamma:\mathcal{A}^*\oplus_1\mathbb{C}\to(\tilde{\mathcal{A}})^*$ be given by

$$[\Gamma(\varphi,\lambda)](a,\mu)=\varphi(a)+\lambda\mu.$$

The map Γ is clearly linear, and it is injective, for if $\Gamma(\varphi,\lambda)=0$ this means that $\varphi(a)+\lambda\mu=0$ for all $a\in\mathcal{A}$ and all $\mu\in\mathbb{C}$. Taking $a=0$, $\mu=1$ we get that $\lambda=0$, and then $\varphi(a)=0$ for all a, so $\varphi=0$. Γ is also surjective: if $\psi\in(\tilde{\mathcal{A}})^*$, then we can define $\psi_{\mathcal{A}}(a)=\psi(a,0)$ and $\lambda=\psi(0,1)$ and we get $\Gamma(\psi_{\mathcal{A}},\lambda)=\psi$. And Γ is continuous. Indeed, the identity map $\operatorname{id}:\mathcal{A}\oplus_1\mathbb{C}\to\tilde{\mathcal{A}}$ satisfies

$$\|\operatorname{id}(a,\lambda)\|=\|(a,\lambda)\|_{\tilde{\mathcal{A}}}\leq\|a\|+|\lambda|=\|(a,\lambda)\|_1,$$

so it is a bounded bijection. By the Inverse Mapping Theorem (6.3.6) the inverse is bounded, which implies that there exists $c>0$ such that $\|a\|+|\lambda|\leq c\|(a,\lambda)\|$ for all $a\in\mathcal{A}$, $\lambda\in\mathbb{C}$. We also have, from the proof of Proposition 11.1.4 and Exercise 11.1.5,

$$\|a\|=\|(a,0)\|=\|(a,\lambda)+(0,-\lambda)\|\leq\|(a,\lambda)\|+|\lambda|\leq 2\|(a,\lambda)\|.$$

Then

$$\begin{aligned}|[\Gamma(\varphi,\lambda)](a,\mu)|&=|\varphi(a)+\lambda\mu|\\&\leq(\|\varphi\|+|\lambda|)\max\{\|a\|,|\mu|\}\\&\leq 2c\,\|(\varphi,\lambda)\|_1\,\|(a,\mu)\|,\end{aligned}$$

showing that $\|\Gamma(\varphi,\lambda)\|\leq c\,\|(\varphi,\lambda)\|_1$ and Γ is bicontinuous again by the Inverse Mapping Theorem.

11.2. Continuous Functional Calculus

(11.2.1) Let $a \in \mathcal{A}$ be selfadjoint and let $v = e^{ia}$. Show that $v^* = e^{-ia}$.

Answer. By definition, $v = \sum_{k=0}^{\infty} \frac{i^k a^k}{k!}$. The operation of taking the adjoint is an isometry; in particular, it is continuous. Hence

$$v^* = \left(\sum_{k=0}^{\infty} \frac{i^k a^k}{k!} \right)^* = \sum_{k=0}^{\infty} \frac{(i^k a^k)^*}{k!} \sum_{k=0}^{\infty} \frac{(-i^k) a^k}{k!} = e^{-ia}.$$

(11.2.2) Show that if $\mathcal{A}$ is an abelian C*-algebra then $\Gamma(A) = \{\Gamma(a) : a \in A\} \subset C_0(\Sigma(\mathcal{A}))$ is a closed, selfadjoint subalgebra that separates points and vanishes nowhere.

Answer. From Γ being an isometric $*$-homomorphism we know that $\Gamma(\mathcal{A})$ is a closed $*$-subalgebra of $C_0(\Sigma(\mathcal{A}))$. So it is closed and selfadjoint. It vanishes nowhere, because given $\tau \in \Sigma(\mathcal{A})$ it is nonzero and so there exists $a \in \mathcal{A}$ with $\tau(a) \neq 0$, which is $\hat{a}(\tau) \neq 0$. Finally, if $\tau_1 \neq \tau_2$, this means that there exists $a \in \mathcal{A}$ with $\tau_1(a) \neq \tau_2(a)$; this is $\hat{a}(\tau_1) \neq \hat{a}(\tau_2)$, so $\Gamma(\mathcal{A})$ separates points.

(11.2.3) Let $\mathcal{A}$ be a C*-algebra, $a \in \mathcal{A}$ and $f \in C(\sigma(a^*a) \cup \{0\})$. Show that $af(a^*a) = f(aa^*)a$.

Answer. We have $a(a^*a) = (aa^*)a$. Inductively, if $a(a^*a)^n = (aa^*)^n a$, then

$$a(a^*a)^{n+1} = a(a^*a)^n a^*a = (aa^*)^n aa^*a = (aa^*)^{n+1}a.$$

It follows that $ap(a^*a) = p(aa^*)a$ for all $p \in \mathbb{C}[x]$. Now if $f \in C(\sigma(a^*a) \cup \{0\}) = C(\sigma(aa^*) \cup \{0\})$, by Stone–Weierstrass (Corollary 7.4.23) there exists $\{p_n\} \subset \mathbb{C}[x]$ such that $p_n \to f$ uniformly. Then, by the continuity of the functional calculus,

$$af(a^*a) = \lim_n ap_n(a^*a) = \lim_n p_n(aa^*)a = f(aa^*)a.$$

(11.2.4) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$ be normal. Let $\lambda \in \mathbb{C} \setminus \sigma(a)$. Show that $\operatorname{dist}(\lambda, \sigma(a)) = \|(a - \lambda I_{\mathcal{A}})^{-1}\|^{-1}$.

Answer. As a is normal, $\|a\| = \max\{|\alpha| : \alpha \in \sigma(a)\}$.

Let $g(t) = 1/t$. Since $0 \notin \sigma(a - \lambda I_{\mathcal{A}})$, we have that $g \in C\big(\sigma(a - \lambda I_{\mathcal{A}})\big)$. Then

$$\begin{aligned}\|(a - \lambda I_{\mathcal{A}})^{-1}\|^{-1} &= \frac{1}{\|g(a - \lambda I_{\mathcal{A}})\|} = \frac{1}{\|g\|_\infty} = \frac{1}{\max\{|g(t)| : t \in \sigma(a - \lambda I_{\mathcal{A}})\}} \\ &= \min\left\{\frac{1}{|g(t)|} : t \in \sigma(a - \lambda I_{\mathcal{A}})\right\} \\ &= \min\{|t| : t \in \sigma(a - \lambda I_{\mathcal{A}})\} \\ &= \min\{|t - \lambda| : t \in \sigma(a)\} = \operatorname{dist}(\lambda, \sigma(a)).\end{aligned}$$

(11.2.5) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\rho : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Let $\hat{\mathcal{A}}$ denote $\mathcal{A}$ is $\mathcal{A}$ is unital, and $\tilde{\mathcal{A}}$ if $\mathcal{A}$ is not unital, and do similarly with $\mathcal{B}$. Show that there exists a unique $*$-homomorphism $\tilde{\rho} : \hat{\mathcal{A}} \to \hat{\mathcal{B}}$, unital onto its image, that extends ρ.

Answer. If $\mathcal{A}$ is unital, there is nothing to be done, as we take $\hat{\rho} = \rho$ and $\rho : \mathcal{A} \to \rho(\mathcal{A})$ is unital; the uniqueness is trivially true from $\hat{\mathcal{A}} = \mathcal{A}$. So we assume that $\mathcal{A}$ it not unital. We define

$$\tilde{\rho}(a, \lambda) = \rho(a) + \lambda\, I_{\hat{\mathcal{B}}}.$$

This is additive:

$$\begin{aligned}\tilde{\rho}(a + b, \lambda + \mu) &= \rho(a + b) + (\lambda + \mu) I_{\hat{\mathcal{B}}} \\ &= \rho(a) + \lambda\, I_{\hat{\mathcal{B}}} + \rho(b) + \mu\, I_{\hat{\mathcal{B}}} = \tilde{\rho}(a, \lambda) + \tilde{\rho}(b, \mu);\end{aligned}$$

and multiplicative, for

$$\begin{aligned}\tilde{\rho}\big((a, \lambda)(b, \mu)\big) &= \tilde{\rho}(ab + \lambda b + \mu a, \lambda\mu) = \rho(ab + \lambda b + \mu a) + \lambda\mu\, I_{\hat{\mathcal{B}}} \\ &= \big(\rho(a) + \lambda\, I_{\mathcal{B}}\big)\big(\rho(b) + \mu\, I_{\hat{\mathcal{B}}}\big) = \tilde{\rho}(a, \lambda)\tilde{\rho}(b, \mu).\end{aligned}$$

Also,

$$\tilde{\rho}\big((a, \lambda)^*\big) = \tilde{\rho}(a^*, \bar{\lambda}) = \rho(a^*) + \bar{\lambda}\, I_{\hat{\mathcal{B}}} = \big(\rho(a) + \lambda\, I_{\hat{\mathcal{B}}}\big)^* = \big(\tilde{\rho}(a, \lambda)\big)^*.$$

So $\tilde{\rho}$ is a $*$-homomorphism, and $\tilde{\rho}(a, 0) = \rho(a)$ by construction.

If $\nu : \tilde{\mathcal{A}} \to \hat{\mathcal{B}}$ is another unital $*$-homomorphism that satisfies $\nu(a,0) = \rho(a)$, then

$$\nu(a,\lambda) = \nu(a,0) + \lambda\nu(0,1) = \rho(a) + \lambda\, I_{\mathcal{B}} = \rho(a,\lambda)$$

so $\nu = \tilde{\rho}$.

(11.2.6) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\rho : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Prove that for all $a \in \mathcal{A}$, $\sigma(\rho(a)) \subset \sigma(a) \cup \{0\}$. The zero can be omitted when $\rho(\mathcal{A})$ is not unital, and also when $\mathcal{A}, \mathcal{B}$ are unital and ρ is unital. The zero cannot be omitted in general.

Answer. To talk about the spectrum we need to work on the unitization. By working on $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ as in Exercise 11.2.5, we get to assume that $\mathcal{A}, \mathcal{B}$ are unital and that ρ is unital onto its image. If $a - \lambda I_{\mathcal{A}}$ is invertible, then so is $\rho(a) - \lambda I_{\rho(\mathcal{A})} = \rho(a - \lambda I_{\mathcal{A}})$ in $\rho(\mathcal{A})$. Thus $\sigma_{\rho(\mathcal{A})}(\rho(a)) \subset \sigma(a)$.

If $\rho(I_{\mathcal{A}}) = I_{\mathcal{B}}$, from Proposition 11.1.12 we have the equality $\sigma_{\mathcal{B}}(\rho(a)) = \sigma_{\rho(\mathcal{A})}(\rho(a))$, and hence $\sigma_{\mathcal{B}}(\rho(a)) \subset \sigma(a)$.

And when $\rho(I_{\mathcal{A}}) \neq I_{\mathcal{B}}$, Proposition 11.1.12 gives us

$$\sigma_{\mathcal{B}}(\rho(a)) = \sigma_{\rho(\mathcal{A})}(a) \cup \{0\}.$$

For an example that the zero cannot be omitted in general, let $\mathcal{A} = \mathbb{C}$, $\mathcal{B} = \mathbb{C}^2$, and $\rho(a) = (a, 0)$. Then for any nonzero $a \in \mathbb{C}$ we have $\sigma(a) = \{a\}$, while $\sigma(\rho(a)) = \{0, a\}$.

(11.2.7) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\rho : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Given $a \in \mathcal{A}$ normal and $f \in C(\sigma(a) \cup \{0\})$, show that $f(\rho(a))$ makes sense and that $f(\rho(a)) = \rho(f(a))$.

Answer. By Exercise 11.2.6 we know that $\sigma(\rho(a)) \subset \sigma(a) \cup \{0\}$. So f is continuous on $\sigma(\rho(a))$ and thus $f(\rho(a))$ makes sense. From Stone–Weierstrass (Theorem 7.4.20), the two-variable complex polynomials on z and $\bar{z}$ are dense in $C(\sigma(a) \cup \{0\})$. So there is a sequence $\{p_n\}$ of such polynomials with $p_n \to f$ uniformly. As mentioned after Definition 11.2.7, we have

$$f(a) = \lim_n p_n(a, a^*),$$

and similarly with $\rho(a)$. By Proposition 11.2.10 ρ is bounded, so

$$\begin{aligned} f(\rho(a)) &= \lim_n p_n(\rho(a), \rho(a)^*) = \lim_n \rho(p_n(a, a^*)) \\ &= \rho(\lim_n p_n(a, a^*)) = \rho(f(a)). \end{aligned}$$

(11.2.8) The Banach algebra $\mathbb{A} = \{f \in C(\bar{\mathbb{D}}) : \text{ analytic on } \mathbb{D}\}$ is not a C*-subalgebra of $C(\bar{\mathbb{D}})$ because it does not contain adjoints for non-scalar functions. Show that $\mathbb{A}$ fails to be a C*-algebra in a deeper sense: it is not isomorphic to $C(X)$ for any compact Hausdorff X, even if the isomorphism is not required to be isometric.

Answer. The idea one can use is that $\mathbb{A}$ does not contain divisors of zero, while any C*-algebra of dimension at least 2 (that is, any C*-algebra that is not $\mathbb{C}$) contains divisors of zero. Indeed, if $\dim C(X) \geq 2$ then X has at least two points. Using Urysohn's Lemma (Theorem 2.6.5) we can construct $g_1, g_2 \in C(X)$ with disjoint supports. So $g_1 g_2 = 0$ everywhere and therefore the isomorphism would have to map one of them to zero.

(11.2.9) Let $v \in \mathcal{A}$ be a selfadjoint partial isometry (that is, $v^* = v$, and v^2 is a projection). Show that there exist projections $p, q \in \mathcal{A}$ with $pq = 0$ and $v = p - q$.

Answer. Since $r = v^2$ is a projection, from $r^2 = r$ we get that $\sigma(v^2) \subset \{0, 1\}$. Therefore $\sigma(v) \subset \{-1, 0, 1\}$. The function $f(t) = 1_{[0,\infty)}$ is continuous on $\sigma(v)$, so we can use functional calculus to define $p = f(v)$. By the Spectral Mapping Theorem (Corollary 11.2.8) we have $\sigma(p) = \{0, 1\}$, so p is a projection ($p^2 = p$ and p is selfadjoint by construction). We can similarly define $q = g(v)$, with $g(t) = 1_{(-\infty,0)}$; then q is a projection. As $f(t) - g(t) = t$ on $\sigma(v)$, we have $v = p - q$. The condition $pq = 0$ follows from $f(t)g(t) = 0$.

11.3. Positivity

(11.3.1) Let $a \in \mathcal{A}$. Show that $\|\mathrm{Re}\, a\| \leq \|a\|$ and $\|\mathrm{Im}\, a\| \leq \|a\|$.

Answer. We have $2\|\mathrm{Re}\, a\| = \|a + a^*\| \leq \|a\| + \|a^*\| = 2\|a\|$. Similarly, $2\|\mathrm{Im}\, a\| = \|a - a^*\| \leq \|a\| + \|a^*\| = 2\|a\|$.

(11.3.2) Let $\mathcal{A}$ be a C*-algebra. Show that $I_\mathcal{A}$ is an extreme point in the closed unit ball of $\mathcal{A}$.

Answer. Suppose that $I_\mathcal{A} = \frac{1}{2}(a+b)$ with $\|a\| \leq 1$ and $\|b\| \leq 1$. Considering the real part we have $I_\mathcal{A} = \frac{1}{2}(\operatorname{Re} a + \operatorname{Re} b)$. We get $0 = (I_\mathcal{A} - \operatorname{Re} a) + (I_\mathcal{A} - \operatorname{Re} b)$, and both terms are positive by Corollary 11.3.8 since $\|\operatorname{Re} a\| \leq \|a\| \leq 1$ and the same for b. It follows that $\operatorname{Re} a = I_\mathcal{A}$ and $\operatorname{Re} b = I_\mathcal{A}$. Then

$$1 = \|a\|^2 = \|a^*a\| = \|(I_\mathcal{A} + i\operatorname{Im} a)^*(I_\mathcal{A} + i\operatorname{Im} a)\| = \|I_\mathcal{A} + (\operatorname{Im} a)^2\|.$$

This implies that $I_\mathcal{A} + (\operatorname{Im} a)^2 \leq I_\mathcal{A}$, and so $\operatorname{Im} a = 0$. Similarly $\operatorname{Im} b = 0$, and so $a = b = I_\mathcal{A}$ and $I_\mathcal{A}$ is extreme.

(11.3.3) Let $a \in \mathcal{A}^+$. Show that its positive square root is unique.

Answer. Let $b = f(a)$, where $f(t) = t^{1/2}$.

If $c \in \mathcal{A}^+$ and $c^2 = a$, then $c = f(c^2) = f(a) = b$. The first equality follows from functional calculus and equality $t = (t^2)^{1/2}$ on $\sigma(c) \subset [0, \infty)$.

(11.3.4) Let $\mathcal{A}$ be a C*-algebra and $a, b \in \mathcal{A}^+$. Show that if $b \leq a$, then $b^{1/2} \leq a^{1/2}$.

Answer. We may assume without loss of generality that $\mathcal{A}$ is unital. Suppose first that b is invertible. Then $0 \notin \sigma(b)$ and, as the spectrum is compact, there exists $\delta > 0$ with $\sigma(b) \subset [\delta, \infty)$. This implies that $b \geq \delta I_\mathcal{A}$ since $b - \delta I_\mathcal{A}$ is selfadjoint and $\sigma(b - \delta I_\mathcal{A}) \subset [0, \infty)$. Then $a \geq b \geq \delta I_\mathcal{A}$, showing that $\sigma(a) \subset [\delta, \infty)$ and so a is invertible. Knowing that $a^{-1/2}$ exists, from $b \leq a$ we get

$$a^{-1/2} b a^{-1/2} \leq I_\mathcal{A}.$$

Hence $\|b^{1/2} a^{-1/2}\| \leq 1$ by the C*-identity and Corollary 11.3.8. By Proposition 9.2.15,

$$\sigma(a^{-1/4} b^{1/2} a^{-1/4}) = \sigma(b^{1/2} a^{-1/2}).$$

Thus

$$\begin{aligned} \|a^{-1/4} b^{1/2} a^{-1/4}\| &= \operatorname{spr}(a^{-1/4} b^{1/2} a^{-1/4}) = \operatorname{spr}(b^{1/2} a^{-1/2}) \\ &\leq \|b^{1/2} a^{-1/2}\| \leq 1. \end{aligned}$$

This implies $a^{-1/4} b^{1/2} a^{-1/4} \leq I_\mathcal{A}$, and so $b^{1/2} \leq a^{1/2}$.

When b is not invertible, we have shown above that $\big(b+\frac1n\big)^{1/2} \leq \big(a+\frac1n\big)^{1/2}$. So it is enough to show that $(b+\frac1n)^{1/2} \to b^{1/2}$, and that a limit of positive operators is positive. We have, since everything commutes,

$$\Big\|\Big(b+\frac{1}{n}\Big)^{1/2} - b^{1/2}\Big\| = \frac{1}{n}\,\Big\|\Big(\Big(b+\frac{1}{n}\Big)^{1/2} + b^{1/2}\Big)^{-1}\Big\| \leq \frac{1}{n}\sqrt{n} \to 0,$$

using that $\big(\frac1n\big)^{1/2} \leq \big(b+\frac1n\big)^{1/2} + b^{1/2}$ and so $\big((b+\frac1n)^{1/2}+b^{1/2}]big)^{-1} \leq \sqrt{n}$.

As for a limit of positive operators, if $b_n \geq 0$ for all n and $b = \lim b_n$, then for any character $\tau(b) = \lim_n \tau(b_n) \geq 0$, so $\sigma(b) \subset [0,\infty)$. Thus $b \geq 0$ by Lemma 11.3.2.

(11.3.5) Show by example that it is not true in general that if $0 \leq b \leq a$, then $b^2 \leq a^2$.

Answer. Let $\mathcal{A} = M_2(\mathbb{C})$, and

$$b = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad a = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

We have $b = b^*b$ and $a = b + e$, where e is the matrix with all entries equal to 1; so $e = \frac12 e^2 \geq 0$, and hence $a \geq 0$. We have

$$a^2 = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}.$$

Then

$$a^2 - b^2 = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix},$$

which is not positive (the determinant is negative, which implies that one eigenvalue is negative).

(11.3.6) Let $a \in \mathcal{A}$ be selfadjoint. Show that the following statements are equivalent:

(i) $a \geq 0$;

(ii) $\|\gamma I_{\mathcal{A}} - a\| \leq \gamma$ for all $\gamma \geq \|a\|$;

(iii) there exists $\gamma \in \mathbb{R}$ with $\gamma \geq \|a\|$ and $\|\gamma I_{\mathcal{A}} - a\| \leq \gamma$.

Answer. (i) $\implies$ (ii) Fix $\gamma \in \mathbb{R}$ with $\gamma \geq \|a\|$. By Theorem 9.2.12 we have $\sigma(a) \subset [0, \|a\|] \subset [0, \gamma]$. Then $a \leq \gamma I_{\mathcal{A}}$ by Corollary 11.3.6, so $\gamma I_{\mathcal{A}} - a \geq 0$. We also have $\gamma I_{\mathcal{A}} - a \leq \gamma I_{\mathcal{A}}$, for $\gamma I_{\mathcal{A}} - (\gamma I_{\mathcal{A}} - a) = a \geq 0$. Hence $0 \leq$

$\gamma I_{\mathcal{A}} - a \le \gamma I_{\mathcal{A}}$. By Corollary 11.3.6, $\sigma(\gamma I_{\mathcal{A}} - a) \subset [0, \gamma]$, and then

$$\|\gamma I_{\mathcal{A}} - a\| = \operatorname{spr}(\gamma I_{\mathcal{A}} - a) \le \gamma.$$

(ii) $\implies$ (iii) Trivial.

(iii) $\implies$ (i) We have $\operatorname{spr}(\gamma I_{\mathcal{A}} - a) = \|\gamma I_{\mathcal{A}} - a\| \le \gamma$. So $\sigma(\gamma I_{\mathcal{A}} - a) \subset [-\gamma, \gamma]$, and therefore $\sigma(a) \subset [0, 2\gamma]$. Then $a \ge 0$ by Lemma 11.3.2.

(11.3.7) Let $\mathcal{A}$ be a C*-algebra. A function $f : [0, \infty) \to [0, \infty)$ is **operator monotone** if $f(a) \le f(b)$ whenever $0 \le a \le b$.

(i) Use Corollary 11.3.7 to prove that $f_s(t) = st(1 + st)^{-1}$ is operator monotone for each $s > 0$.

(ii) Prove that $g(t) = t^\beta$ is operator monotone for all $\beta \in (0, 1)$ by considering $g(t) = \int_0^\infty f_s(t) s^{-\beta-1}\, ds$.

Answer.

(i) We have

$$f_s(t) = \frac{st}{1 + st} = 1 - \frac{1}{1 + st}.$$

If $0 \le a \le b$, then $I_{\mathcal{A}} + sa \le I_{\mathcal{A}} + sb$, so $(I_{\mathcal{A}} + sb)^{-1} \le (I_{\mathcal{A}} + sa)^{-1}$ by Corollary 11.3.7. Then

$$f_s(a) = I_{\mathcal{A}} - (I_{\mathcal{A}} + sa)^{-1} \le I_{\mathcal{A}} - (I_{\mathcal{A}} + sb)^{-1} = f_s(b).$$

(ii) We have, with the substitution $v = ts$,

$$\int_0^\infty f_s(t)\, s^{-\beta-1}\, ds = \int_0^\infty \frac{st}{(1 + st)s^{\beta+1}}\, ds = t^\beta \int_0^\infty \frac{1}{1 + v}\, \frac{1}{v^\beta}\, dv.$$

The improper integral converges both at 0 and ∞ because the integrand is of the form $v^{-\beta}$ and $v^{-\beta-1}$ respectively. So if

$$r = \int_0^\infty \frac{1}{1 + v}\, \frac{1}{v^\beta}\, dv,$$

we have

$$t^\beta = g(t) = \frac{1}{r} \int_0^\infty f_s(t)\, s^{-\beta-1}\, ds.$$

The inequality $f_s(a)s^{-\beta-1} \le f_s(b)s^{-\beta-1}$ for all s will survive through Riemann sums (valued on the C*-algebra) as all terms are positive. We will show that the Riemann sums of an integral of a continuous function over a closed interval converge uniformly. Indeed, if f is continuous on $[1/R, R]$ and $\varepsilon > 0$ is given, since f is uniformly continuous there exists $\delta > 0$ with $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$. Let $\{s_j\}$ be a partition of $[1/R, R]$ with $|s_j - s_{j-1}| < \delta$. Then

$$\left|\int_{1/R}^{R} f(s)\,ds - \sum_{j=1}^{n} f(s_j)\,\Delta_j\right| = \left|\sum_{j=1}^{n} \int_{s_{j-1}}^{s_j} \big(f(s) - f(s_j)\big)\,ds\right|$$
$$\leq \sum_{j=1}^{n} \int_{s_{j-1}}^{s_j} \big|f(s) - f(s_j)\big|\,ds \leq \varepsilon\,R.$$

If follows that if $h_R(t) = \frac{1}{r}\int_{1/R}^{R} f_s(t)\,s^{-\beta-1}\,ds$, then

$$h_R(a) = \lim_n \sum_{j=1}^{n} f_{s_j^{(n)}}(a)\,\Delta_j^{(n)}$$

uniformly, and hence $h_R(a) \leq h_R(b)$. We also have that

$$\left|\int_{0}^{\infty} f_s(t)\,s^{-\beta-1}\,ds - \int_{1/R}^{R} f_s(t)\,s^{-\beta-1}\,ds\right| \leq \left|\int_{0}^{1/R} \frac{st}{1+st}\,\frac{1}{s^{\beta+1}}\,ds\right|$$
$$+ \left|\int_{R}^{\infty} \frac{st}{1+st}\,\frac{1}{s^{\beta+1}}\,ds\right|$$
$$\leq t\int_{0}^{1/R} s^{-\beta}\,ds + \int_{R}^{\infty} s^{-\beta-1}\,ds$$
$$= \frac{t}{(1-\beta)R^{1-\beta}} + \frac{1}{\beta R^{\beta}},$$

and, as $t \leq \|b\|$, the limit goes to 0 uniformly on R; that is, $h_R \to g$ uniformly. Then

$$a^{\beta} = g(a) = \lim_{R\to\infty} h_R(a) \leq \lim_{R\to\infty} h_R(b) = g(t) = b^{\beta}.$$

(11.3.8) Let $\mathcal{A}$ be a C*-algebra and $a, p \in \mathcal{A}$ with $a \geq 0$ and p a projection ($p = p^2 = p^*$). Show that $a \leq p$ if and only if $\|a\| \leq 1$ and $a = pa = ap$. *(When a and p are projections in $\mathcal{B}(\mathcal{H})$ this was done in Proposition 10.5.3; now we need rather similar arguments, but they have to be entirely algebraic)*

Answer. We can work on the unitization if needed. If $a \leq p$, then

$$0 \leq (I_{\mathcal{A}} - p)a(I_{\mathcal{A}} - p) \leq (I_{\mathcal{A}} - p)p(I_{\mathcal{A}} - p) = 0.$$

Hence

$$0 = (I_{\mathcal{A}} - p)a(I_{\mathcal{A}} - p) = \big[a^{1/2}(I_{\mathcal{A}} - p)\big]^*\big[a^{1/2}(I_{\mathcal{A}} - p)\big],$$

so $a^{1/2}(I_{\mathcal{A}} - p) = 0$ and then $a(I_{\mathcal{A}} - p) = 0$ which is $a = ap$; taking adjoints, $a = ap = pa$. Also, from $0 \leq a \leq p$ we get $\|a\| \leq \|p\| = 1$.

Conversely, if $a = pa = ap$ and $\|a\| \leq 1$, then

$$a = pa = pap \leq \|a\| \, p^2 = p.$$

(11.3.9) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$ nonzero. Show that there exist $r \in \{1, -1\}$ and $s \in \{1, i\}$ such that $(\operatorname{Re} rsa)^+ \neq 0$.

Answer. Via Proposition 11.3.10 we can write

$$a = \operatorname{Re} a + i \operatorname{Im} a = a_1 - a_2 + i(a_3 - a_4),$$

with $a_1, a_2, a_3, a_4 \in \mathcal{A}^+$, $a_1 a_2 = a_3 a_4 = 0$. Since $a \neq 0$, there exists j with $a_j \neq 0$. If $j \in \{1, 2\}$ we take $s = 1$ for a_j is already in the real part; otherwise we take $s = i$. Now sa has real part with a_j either the positive part—in which case we take $r = 1$—or as the negative part—in which case we take $r = -1$. Thus $rsa = a_j - a_k + i(a_m - a_n)$ for $k, m, n \in \{1, 2, 3, 4\} \setminus \{j\}$ distinct, and thus $(\operatorname{Re} rsa)^+ = a_j$.

11.4. Ideals and $*$-Homomorphisms

(11.4.1) Let $\mathcal{A}$ be a C*-algebra. Show that $\mathcal{Z}(\mathcal{A})$ is a C*-subalgebra of $\mathcal{A}$.

Answer. If $a_1, a_2 \in \mathcal{Z}(\mathcal{A})$ and $b \in \mathcal{A}$, $\lambda \in \mathbb{C}$,

$$(a_1 + \lambda a_2)b = a_1 b + \lambda a_2 b = b(a_1 + \lambda a_2), \qquad \text{and} \qquad a_1 a_2 b = a_1 b a_2 = b a_1 a_2.$$

And from $a_1 b^* = b^* a_1$, taking adjoints we get $a_1^* b = b a_1^*$. Thus $\mathcal{Z}(\mathcal{A})$ is a $*$-subalgebra of $\mathcal{A}$. If $\{a_n\} \subset \mathcal{Z}(\mathcal{A})$ and $a_n \to a$, then

$$ab = \lim a_n b = \lim b a_n = b \lim a_n = ba.$$

So $\mathcal{Z}(\mathcal{A})$ is closed, and it is thus a C*-subalgebra.

(11.4.2) Show the equivalence (11.8).

Answer. If $\lim_j ae_j = a$ for all a, then for a fixed a we have $\lim_j a^*e_j = a^*$; as the adjoint operation is continuous, $\lim_j e_ja = a$. The converse is proven similarly.

(11.4.3) If $\mathcal{A}$ is separable, show that it admits a countable approximate unit *(Hint: consider an increasing sequence of finite sets with dense union)*

Answer. Since $\mathcal{A}$ is separable, there exist finite subsets $F_1 \subset F_2 \subset \cdots \subset \mathcal{A}$ with $\bigcup_n F_n$ dense. From $\mathcal{Z}$ as in the proof of Theorem 11.4.4, since it is an approximate unit choose $e_1 \in \mathcal{Z}$ with $\|a - ae_1\| < 1$ for all $a \in F_1$; inductively, given e_k, choose $e_{k+1} \in \mathcal{Z}$ with $e_{k+1} \geq e_k$ and $\|a - ae_{k+1}\| < 1/(k+1)$ for all $a \in F_{k+1}$. So, for any $a \in \bigcup_n F_n$, we have that $\lim_n \|a - ae_n\| = 0$. As $\bigcup_n F_n$ is dense, $\lim_n \|a - ae_n\| = 0$ for any $a \in \mathcal{A}$ via a typical $\varepsilon/3$-argument.

(11.4.4) Let $\mathcal{A}$ be a C*-algebra and $\{e_j\} \subset \mathcal{A}$ an approximate unit. Show that $\lim_j \|e_j\| = \lim_j \|e_j^2\| = 1$.

Answer. Since the approximate unit is monotone, if $j \leq k$ then $0 \leq e_j \leq e_k$, and so $\|e_j\| \leq \|e_k\|$ by Corollary 11.3.8. Then $c = \lim_j \|e_j\|$ exists, and $c \leq 1$ as $\|e_j\| \leq 1$ for all j. Let $a \in \mathcal{A}$ with $\|a\| = 1$. Then $\|ae_j\| \leq \|e_j\| \leq c$, and so

$$1 - c \leq \|a\| - \|ae_j\| \leq \|a - ae_j\| \to 0.$$

Therefore $c = 1$. For the squares, we simply note that $\|e_j^2\| = \|e_j\|^2$ since they are positive.

(11.4.5) Let $\mathcal{B} \subset \mathcal{A}$ be an inclusion of C*-algebras. Suppose that $\mathcal{B}$ has an approximate unit $\{e_j\}$ that is also an approximate unit for $\mathcal{A}$. Show that any other approximate unit of $\mathcal{B}$ is an approximate unit for $\mathcal{A}$.

Answer. Fix $a \in \mathcal{A}$ and let $\{f_k\}$ be another approximate unit for $\mathcal{B}$. Then

$$\begin{aligned}\|a - af_k\| &\leq \|a - ae_j\| + \|ae_j - ae_jf_k\| + \|ae_jf_k - af_k\| \\ &= \|a - ae_j\| + \|a(e_j - e_jf_k)\| + \|(ae_j - a)f_k\| \\ &\leq \|a - ae_j\| + \|a(e_j - e_jf_k)\| + \|ae_j - a\|\end{aligned}$$

It follows that

$$\limsup_k \|a - af_k\| \leq 2\|a - ae_j\|.$$

As we are free to choose j, we get by the Limsup Routine that $\lim_k \|a - af_k\| = 0$. An entirely similar argument (or, taking adjoints) shows that $\|a - f_ka\| \to 0$ for all $a \in \mathcal{A}$.

(11.4.6) Let $\mathcal{A}$ be a C*-algebra and $\mathcal{A}_0 \subset \mathcal{A}$ a dense $*$-subalgebra. Let $\mathcal{B}$ be a C*-algebra and $\pi : \mathcal{A}_0 \to \mathcal{B}$ an injective $*$-homomorphism. Show by example that π is not necessarily bounded.

Answer. Let $\mathcal{A} = C[0,1]$, $\mathcal{B} = \mathbb{C} \oplus C[0,1]$, and $\mathcal{A}_0 = \mathbb{C}[x]$, considered as a subalgebra of $\mathcal{A}$ (so $\|p\| = \sup\{|p(x)| : x \in [0,1]\}$). Define $\pi : \mathcal{A}_0 \to \mathcal{B}$ by

$$\pi(p) = p(2) \oplus p.$$

This is clearly an injective $*$-homomorphism. Moreover, $\pi(\mathcal{A}_0)$ is dense in $\mathcal{B}$: given $(\lambda, f) \in \mathcal{B}$, let $g \in C[0,2]$ with $g = f$ on $[0,1]$ and $g(2) = \lambda$. By Stone–Weierstrass (Theorem 7.4.20) there exists a sequence $\{p_n\}$ of polynomials with $p_n \to g$ uniformly on $[0,2]$. In particular, $\pi(p_n) \to (\lambda, f)$.

But π is not bounded: consider $q_n(x) = x^n$. Then $\|q_n\| = 1$ for all n, but $\|\pi(q_n)\| = 2^n$.

(11.4.7) Let $\mathcal{A}$ be approximately finite dimensional (AFD); that is, $\mathcal{A} = \overline{\bigcup_n \mathcal{A}_n}$, where $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ and $\dim \mathcal{A}_n < \infty$ for all n. Show that $\mathcal{A}$ has a countable approximate unit made out of projections. *(Hint: use that finite-dimensional C*-algebras are unital, Lemma 11.8.2)*

Answer. Let p_n be the identity of $\mathcal{A}_n$, which exists by Lemma 11.8.2. Since $p_n \in \mathcal{A}_{n+1}$, we have $p_{n+1}p_n = p_n$, so they commute (by taking adjoints) and—working momentarily on $\tilde{\mathcal{A}}$—$p_n = p_{n+1}p_np_{n+1} \leq p_{n+1}I_{\tilde{\mathcal{A}}}p_{n+1} = p_{n+1}$. Fix $\varepsilon > 0$. Given $a \in \mathcal{A}$, by hypothesis there exists n_0 and $b \in \mathcal{A}_{n_0}$ with

$\|a - b\| < \varepsilon$. Then, for $n \geq n_0$,

$$\|a - ap_n\| \leq \|a - b\| + \|b - bp_n\| + \|bp_n - ap_n\|$$
$$= \|a - b\| + \|(b - a)p_n\| \leq 2\|a - b\| < 2\varepsilon.$$

Thus $\lim_n ap_n = a$

(11.4.8) Let $\mathcal{A}$ be a C*-algebra and $\{e_s\}$ an approximate unit for $\mathcal{A}$. Consider linearly independent elements $b_1, \ldots, b_m \in \mathcal{A}$. Show that there exists s_0 such that for all $s \geq s_0$, $e_s b_1, \ldots, e_s b_m$ are linearly independent.

Answer. Suppose that there exists a subnet $\{s_t\}$ and, for every t, coefficients

$$\alpha_{t,1}, \ldots, \alpha_{t,m} \qquad \text{with} \qquad \sum_{k=1}^{m} \alpha_{t,k} e_{s_t} b_k = 0.$$

Since at least one coefficient is nonzero, we may assume that for each t the largest coefficient is 1. Then there exists at least one k with infinitely many $\alpha_{t,k} = 1$. By passing to a yet another subnet, we may assume that there exists k with $\alpha_{t,k} = 1$ for all t. Since each net $\{\alpha_{t,k}\}$ of coefficients is bounded, it admits a convergent subnet. After taking m subnets, we get $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$ with $\alpha_{t,k} \to \alpha_j$. Then

$$\sum_{k=1}^{m} \alpha_k b_k = \lim_t \alpha_{t,k} e_{s_t} b_k = 0.$$

So, if $b_1, \ldots, b_m$ are linearly independent, there has to exist s_0 such that

$$e_s b_1, \ldots, e_s b_m$$

is linearly independent for all $s \geq s_0$.

11.5. States

(11.5.1) Show that $\phi \in \mathcal{A}^*$ is Hermitian if and only if $\phi(a^*) = \overline{\phi(a)}$ for all $a \in \mathcal{A}$.

Answer. If $\phi(a^*) = \overline{\phi(a)}$, then for a selfadjoint we have $\overline{\phi(a)} = \phi(a^*) = \phi(a)$, so $\phi(a) \in \mathbb{R}$.

Conversely, suppose that $\phi(a) \in \mathbb{R}$ for all selfadjoint a. For arbitrary a, we have

$$\phi(a) + \phi(a^*) = \phi(a + a^*) \in \mathbb{R}, \qquad i(\phi(a) - \phi(a^*)) = \phi(i(a - a^*)) \in \mathbb{R}.$$

From the second equality we get that $\operatorname{Im}\phi(a^*) = -\operatorname{Im}\phi(a)$; and from the first one we get that $\operatorname{Re}\phi(a^*) = \operatorname{Re}\phi(a)$. Thus $\phi(a^*) = \overline{\phi(a)}$.

(11.5.2) Let $\varphi \in \mathcal{A}^*$. Put

$$\varphi^*(a) = \overline{\varphi(a^*)}.$$

Show that $\varphi^* \in \mathcal{A}^*$ and that $\varphi + \varphi^*$ is Hermitian.

Answer. We have

$$\begin{aligned}\varphi^*(\beta a + b) &= \overline{\varphi((\beta a + b)^*)} = \overline{\varphi(\overline{\beta} a^* + b^*)} \\ &= \beta\,\overline{\varphi(a^*)} + \overline{\varphi(b^*)} = \beta\varphi^*(a) + \varphi^*(b),\end{aligned}$$

so φ^* is linear. Also $|\varphi^*(a)| \leq \|\varphi\|\,\|a^*\| = \|\varphi\|\,\|a\|$, so φ^* is bounded and hence $\varphi^* \in \mathcal{A}^*$. If $a = a^*$, then

$$\varphi(a) + \varphi^*(a) = \varphi(a) + \overline{\varphi(a)} = 2\operatorname{Re}\varphi(a) \in \mathbb{R}.$$

So $\varphi + \varphi^*$ is Hermitian.

(11.5.3) Show that a positive linear functional is Hermitian.

Answer. Given $a \in \mathcal{A}$ selfadjoint, we can write $a = a^+ - a^-$ with $a^+, a^- \geq 0$ (Proposition 11.3.10), and then $\varphi(a) = \varphi(a^+) - \varphi(a^-) \in \mathbb{R}$.

(11.5.4) Let $\varphi \in \mathcal{A}^*$ be positive, and $a, b \in \mathcal{A}$ with $b \geq 0$. Show that $|\varphi(ab)| \leq \|a\|\varphi(b)$.

Answer. We use Cauchy–Schwarz in the following way:

$$\begin{aligned}|\varphi(ab)| &= |\varphi(ab^{1/2}\,b^{1/2})| \leq \varphi(b^{1/2}a^*ab^{1/2})^{1/2}\varphi(b)^{1/2} \\ &\leq \|a\|\,\varphi(b)^{1/2}\varphi(b)^{1/2} = \|a\|\,\varphi(b).\end{aligned}$$

(11.5.5) Let $\mathcal{A}$ be a unital C*-algebra and $\mathcal{A}_0 \subset \mathcal{A}$ a $*$-subalgebra with $I_{\mathcal{A}} \in \mathcal{A}_0$. Let $\varphi : \mathcal{A}_0 \to \mathbb{C}$ be linear and positive. Show that φ is bounded, $\|\varphi\| = \varphi(I_{\mathcal{A}})$, and φ extends to a positive $\tilde{\varphi} \in \mathcal{A}^*$ with $\|\tilde{\varphi}\| = \|\varphi\|$.

Answer. Compared to Proposition 11.5.4, the presence of the unit makes all the difference. Given $a \in \mathcal{A}_0$ with $\|a\| \leq 1$, we have that $a^*a \leq I_{\mathcal{A}}$. Then, using Cauchy–Schwarz,

$$|\varphi(a)|^2 \leq \varphi(I_{\mathcal{A}})\,\varphi(a^*a) \leq \varphi(I_{\mathcal{A}})^2.$$

Hence $|\varphi(a)| \leq \varphi(I_{\mathcal{A}})\,\|a\|$ for all $a \in \mathcal{A}$ and so $\|\varphi\| \leq \varphi(I_{\mathcal{A}})$. As $\varphi(I_{\mathcal{A}}) \leq \|\varphi\|$, we get $\|\varphi\| = \varphi(I_{\mathcal{A}})$. Now we extend by Hahn–Banach (Corollary 5.7.6) to get $\tilde{\varphi} : \mathcal{A} \to \mathbb{C}$ with

$$\|\tilde{\varphi}\| = \|\varphi\| = \varphi(I_{\mathcal{A}}) = \tilde{\varphi}(I_{\mathcal{A}}).$$

Then $\tilde{\varphi} \geq 0$ by Proposition 11.5.4.

(11.5.6) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$. Show that

$$\|a\| = \max\{\varphi(a^*a)^{1/2} : \ \varphi \in S(\mathcal{A})\}.$$

Answer. Suppose first that $a \geq 0$. By Corollary 11.5.8 there exists $\varphi \in S(\mathcal{A})$ with $\varphi(a) = \|a\|$. We also have, for any $\psi \in S(\mathcal{A})$, that $\psi(I_{\mathcal{A}}) = 1$ (Exercise 11.5.5 or Proposition 11.5.4) and then

$$|\psi(a)| = \psi(a) \leq \psi(\|a\|\, I_{\mathcal{A}}) = \|a\|,$$

so $\|a\| = \max\{\varphi(a) : \ \varphi \in S(\mathcal{A})\}$.

For arbitrary $a \in \mathcal{A}$,

$$\begin{aligned}\|a\| = \|a^*a\|^{1/2} &= \max\{\varphi(a^*a) : \ \varphi \in S(\mathcal{A})\}^{1/2} \\ &= \max\{\varphi(a^*a)^{1/2} : \ \varphi \in S(\mathcal{A})\}.\end{aligned}$$

(11.5.7) We outline here a proof of Proposition 11.5.16 and Corollary 11.5.17 that goes another way. So let $\mathcal{A}$ be a C*-algebra and $\varphi \in \mathcal{A}^*$ a Hermitian functional.

(i) Let $K = \{\psi \in \mathcal{A}^* : \ \psi \geq 0 \ \text{ and } \ \|\psi\| \leq 1\}$. Show that K is weak*-compact and it separates points.

(ii) Use the idea in Proposition 7.2.24 to show that $\mathcal{A}$ embeds isometrically in $C(K)$ in a way that preserves positivity.

(iii) With $\rho : \mathcal{A} \to C(K)$ the embedding, show that $\varphi \circ \rho^{-1} : \rho(\mathcal{A}) \to \mathbb{C}$ extends to a bounded Hermitian functional $\tilde{\varphi}$ on $C(K)$.

(iv) Use Theorem 5.6.15 and the Jordan decomposition of a measure to conclude that there exist $\varphi^+, \varphi^- \in S(\mathcal{A})$ and non-negative scalars α, β such that $\varphi = \alpha\varphi^+ - \beta\varphi^-$.

Answer.

(i) That K separates points is Corollary 11.5.8. For the weak*-compactness one can either repeat the proof of Proposition 11.5.15 now for $\{\psi_j\}$ with $\|\psi_j\| \leq 1$ instead of equal to 1; or we notice that $K = \gamma(\overline{\mathbb{D}} \times S(\mathcal{A}))$, a continuous image of a compact set, with $\gamma(\lambda, \psi) = \lambda\psi$.

(ii) Let $\rho : \mathcal{A} \to C(K)$ be given by $(\rho(a))(\psi) = \psi(a)$. This map is linear, and it is isometric by Corollary 11.5.8. If $a \geq 0$, the $\psi(a) \geq 0$ for all $\psi \in K$, so $\rho(a)$ is a positive function.

(iii) By Hahn–Banach (Corollary 5.7.6) there exists $\psi : C(K) \to \mathbb{C}$, linear and bounded, with $\|\psi\| = \|\varphi\|$ (note that $\|\varphi \circ \rho^{-1}\| = \|\varphi\|$ since ρ is isometric) and $\psi|_{\rho(\mathcal{A})} = \varphi \circ \rho^{-1}$. Now define

$$\tilde{\varphi}(g) = \frac{1}{2}\,(\psi(g) + \overline{\psi(g^*)}), \qquad g \in C(K).$$

Then $\tilde{\varphi}$ is Hermitian, $\|\tilde{\varphi}\| = \|\varphi\|$, and since φ is Hermitian we also have $\tilde{\varphi}|_{\rho(\mathcal{A})} = \varphi \circ \rho^{-1}$.

(iv) By Theorem 5.6.15 there exists a complex regular Borel measure μ such that

$$\varphi(a) = \tilde{\varphi}(\rho(a)) = \int_K \rho(a)\, d\mu.$$

Using the Jordan Decomposition on $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$ we find finite measures μ_r, with $r = 1, 2, 3, 4$ such that $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$. Because φ is Hermitian, if we evaluate on selfadjoints we see that we can take $\mu_3 - \mu_4 = 0$, so we may assume without loss of generality that $\mu = \mu_1 - \mu_2$. Now for $j = 1, 2$ we can define $\varphi'_j(a) = \int_K \rho(a)\, d\mu_j$, a positive bounded linear functional on $\mathcal{A}$, and normalizing we get $\varphi_j = \frac{1}{\alpha_j}\varphi'_j \in S(\mathcal{A})$ for all j, where $\alpha_j = \|\varphi'_j\|$. It follows that $\varphi = \alpha_1\varphi_1 - \alpha_2\varphi_2$ for the two states $\varphi_1, \varphi_2 \in S(\mathcal{A})$.

(11.5.8) Show that when $\mathcal{A}$ is non-unital Lemma 11.5.14 still holds if we replace condition (ii) by the existence, for every $\varepsilon > 0$, of positive $x, y \in \mathcal{A}$ with $x + y \leq I_{\mathcal{A}}$ and $\varphi(x) \geq \|\varphi\| - \varepsilon$, $\psi(y) \geq \|\psi\| - \varepsilon$.

Answer. (i) $\Longrightarrow$ (ii) Let w with $\|w\| = 1$ such that $\varphi(w) - \psi(w) > \|\varphi - \psi\| - \varepsilon$. Because $\varphi - \psi$ is Hermitian we may assume without loss of generality that $w = w^*$. Fix an approximate unit $\{e_j\}$ and fix j such that $\varphi(e_j) > \|\varphi\| - \frac{\varepsilon}{2}$ and $\psi(e_j) > \|\psi\| - \frac{\varepsilon}{2}$; also, because $w = \lim_j e_j w e_j$ and φ, ψ are continuous, by choosing j large enough we can also guarantee that $\varphi(e_j w e_j) - \psi(e_j w e_j) > \|\varphi - \psi\| - \varepsilon$. Put

$$x = \frac{1}{2}(e_j + e_j w e_j), \qquad y = \frac{1}{2}(e_j - e_j w e_j).$$

Then $x \geq 0$, $y \geq 0$, and $x + y = e_j \leq I_{\mathcal{A}}$. We have

$$\begin{aligned} 2\|\varphi\| - 2\varphi(x) + 2\|\psi\| - 2\psi(y) &= 2\|\varphi\| - \varphi(e_j + e_j w e_j) + 2\|\psi\| \\ &\qquad - \psi(e_j - e_j w e_j) \\ &= \|\varphi\| - \varphi(e_j) + \|\psi\| - \psi(e_j) + \|\varphi - \psi\| \\ &\qquad - (\varphi - \psi)(e_j w e_j) \\ &\leq 2\varepsilon. \end{aligned}$$

As the left-hand-side can be seen as the sum of the two non-negative terms $2\|\varphi\| - 2\varphi(x)$ and $2\|\psi\| - 2\psi(y)$, we conclude that

$$\|\varphi\| - \varphi(x) < \varepsilon, \qquad \|\psi\| - \psi(y) < \varepsilon$$

as desired.

(ii) $\Longrightarrow$ (i) Let $b = x + y$, $w = x - y$. We have $\|b\| \leq 1$ and $\|w\| \leq 1$ (the latter, from $-(x + y) \leq x - y \leq x + y$). And

$$\begin{aligned} \|\varphi - \psi\| &\leq \|\varphi\| + \|\psi\| \\ &\leq \|\varphi\| + \|\psi\| + 2\varepsilon + 2\varphi(x) - 2\|\varphi\| + 2\varepsilon + \psi(y) - 2\|\psi\| \\ &= 4\varepsilon + \varphi(b + w) + \psi(b - w) - \|\varphi\| - \|\psi\| \\ &= 4\varepsilon + (\varphi - \psi)(w) + \varphi(b) - \|\varphi\| + \psi(b) - \|\psi\| \\ &\leq 4\varepsilon + (\varphi - \psi)(w) \\ &\leq 4\varepsilon + \|\varphi - \psi\|. \end{aligned}$$

As $\varepsilon > 0$ was arbitrary, $\|\varphi - \psi\| = \|\varphi\| + \|\psi\|$.

(11.5.9) Let $\mathcal{A}$ be a C*-algebra and $p \in \mathcal{A}$ a projection. Show that $p\mathcal{A}p$ is a hereditary C*-subalgebra of $\mathcal{A}$.

Answer. If $x, y, z, w \in \mathcal{A}$ and $\lambda \in \mathbb{C}$,

$$pxp + \lambda pyp + (pzp)(pwp) = p[pxp + \lambda pyp + (pzp)(pwp)]p \in p\mathcal{A}p.$$

Together with $(pxp)^* = px^*p$, this shows that $p\mathcal{A}p$ is a $*$-algebra. It remains to show that it is closed and hereditary. For closed, if $\{px_np\}$ is Cauchy, then by the completeness of $\mathcal{A}$ there exists $x \in \mathcal{A}$ with $px_np \to x$. But then $pxp = \lim px_np = x$, so $x \in \mathcal{A}$. Finally, if $0 \leq y \leq pxp$, working on the unitization we get $(I_\mathcal{A} - p)y(I_\mathcal{A} - p) = 0$. This is $[y^{1/2}(I_\mathcal{A} - p)]^* y^{1/2}(I_\mathcal{A} - p) = 0$, so $y^{1/2}(I_\mathcal{A} - p) = 0$, from where $y(I_\mathcal{A} - p) = 0$; that is $y = yp$. Taking adjoints we also get $y = py$, and so $y = pyp \in p\mathcal{A}p$.

(11.5.10) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}^+$. Show that $\overline{a\mathcal{A}a}$ is a hereditary C*-subalgebra of $\mathcal{A}$. Show by example that the closure is needed in general.

Answer. If $x, y, z, w \in \mathcal{A}$ and $\lambda \in \mathbb{C}$,

$$axa + \lambda aya + (aza)(awa) = a[x + \lambda y + za^2w]a \in a\mathcal{A}a.$$

And $(axa)^* = ax^*a \in a\mathcal{A}a$, so $a\mathcal{A}a$ is a $*$-algebra; its closure is then a C*-subalgebra of $\mathcal{A}$. As for hereditary, if $0 \leq b \leq c$ and $c \in \overline{a\mathcal{A}a}$, let $\{e_j\}$ be an approximate unit for $\overline{a\mathcal{A}a}$. Working on the unitization, we have

$$0 \leq (I_\mathcal{A} - e_j)b(I_\mathcal{A} - e_j) \leq (I_\mathcal{A} - e_j)c(I_\mathcal{A} - e_j).$$

This gives us

$$\begin{aligned}\|b^{1/2} - b^{1/2}e_j\|^2 &= \|(I_\mathcal{A} - e_j)b(I_\mathcal{A} - e_j)\| \\ &\leq \|(I_\mathcal{A} - e_j)c(I_\mathcal{A} - e_j)\| = \|c^{1/2} - c^{1/2}e_j\|^2.\end{aligned}$$

Since $c^{1/2} = \lim_j c^{1/2}e_j$, the above inequality implies that $b^{1/2} = \lim_j b^{1/2}e_j$, and so $b = \lim_j e_jbe_j \in \overline{a\mathcal{A}a}$.

As for the example, let $\mathcal{A} = C[0,1]$ and $a(t) = t$. We claim that $a\mathcal{A}a$ is not closed. To see this, we will show that $f(t) = t^{1/2}$ is in $\overline{a\mathcal{A}a}$ but not in $a\mathcal{A}a$. Indeed, consider the algebra $\mathcal{B} = C_0(0,1]$. By Corollary 7.4.23, $a\mathcal{B}a = a^2\mathcal{B}$ is dense in $\mathcal{B}$. Hence $f \in \overline{a^2\mathcal{A}}$. But we cannot have $t^{1/2} = t^2g(t)$ for some $g \in C[0,1]$; for we would have, for $t > 0$, that $1 = t^{3/2}\,g(t)$, contradicting the fact that the right-hand-side goes to 0 as $t \to 0$.

Another example can be $\mathcal{A} = \mathcal{K}(\mathcal{H})$, and $a = \sum_k \frac{1}{k}\,E_{kk}$ for a fixed set of matrix units $\{E_{kj}\}$. Then $\overline{a\mathcal{A}a} = \mathcal{A}$, since a is strictly positive (alternatively, one can show that $E_{kj} \in a\mathcal{A}a$ for all k, j, and that the span of the matrix units is dense in $\mathcal{K}(\mathcal{H})$). But every element in $a\mathcal{A}a$ is trace-class, while $a \in \overline{a\mathcal{A}a}$ is not trace-class.

(11.5.11) Show that the ideal generated by the identity function in $C[0,1]$ is an example of a non-closed, non-hereditary ideal of a C*-algebra. Use the example to give an explanation for why the argument after Definition 11.5.19 fails in that case.

Answer. We have $\mathcal{J} = \{t \longmapsto tg(t) : \ g \in C[0,1]\}$. This is an ideal, since it is of the form $hC[0,1]$, where $h(t) = t$ is the identity function. We need to show that this is not closed. The norm closure of $\mathcal{J}$ is $\mathcal{J}_0 = \{f \in C[0,1] : f(0) = 0\}$. Indeed, given $f \in \mathcal{J}_0$ and $\varepsilon > 0$, choose $\delta > 0$ such that $|f(t)| < \varepsilon$ for all $t \in [0, \delta]$. Let $g_\varepsilon \in C[0,1]$ the continuous function with $g_\varepsilon(t) = f(t)/t$ for all $t \geq \delta$ and that goes linearly to 0 on $[0, \delta)$. Then $|f(t) - t\, g_\varepsilon(t)| = 0$ for all $t \geq \delta$. And for $t < \delta$,

$$|f(t) - t\, g_\varepsilon(t)| \leq |f(t)| + \delta|f(\delta)/\delta| \leq 2\varepsilon.$$

Hence $\|f - h\, g_\varepsilon\|_\infty < 2\varepsilon$, showing that $f \in \overline{\mathcal{J}}$. As $\mathcal{J} \subset \mathcal{J}_0$, this shows that $\overline{\mathcal{J}} = \mathcal{J}_0$.

Consider the function $g(t) = t \sin^2 1/t$. This is in $\mathcal{J}_0$. But it is not in $\mathcal{J}_0$, for if $g(t) = tf(t)$, then $f(t) = \sin^2 1/t$ for all $t > 0$ and it cannot be continuous at 0. And $g(t) \leq t$ for all t while $g \notin \mathcal{J}$, showing that $\mathcal{J}$ is not hereditary.

The reason the argument after Definition 11.5.19 does not work for our $\mathcal{J}$ is that even though the map $\rho : \mathcal{A} \to \mathcal{A}/\mathcal{J}$ is a $*$-homomorphism, it is not guaranteed that it maps positive elements to positive elements. The problem is that the argument that $\rho(a) \geq 0$ if $a \geq 0$ relies on writing $a = b^*b$, and this only works on a C*-algebra. The objects in the quotient are not functions, and we cannot rely on pointwise evaluation to assess positivity. The spectrum does not behave well on non-closed algebras either, so it cannot be used to define positivity.

(11.5.12) Let $\mathcal{A}$ be a non-unital C*-algebra and $\tilde{\mathcal{A}}$ its unitization. Show that $\psi((a, \lambda)) = \lambda$ defines a state on $\tilde{\mathcal{A}}$.

Answer. We have that ψ is unital by definition, and its linearity is straightforward. Also,

$$\psi\big((a, \lambda)^*(a, \lambda)\big) = \psi(a^*a + 2\mathrm{Re}\, \lambda a^*, |\lambda|^2) = |\lambda|^2 \geq 0,$$

so ψ is positive. Then $\|\psi\| = 1$ by (11.2).

(11.5.13) Let $\mathcal{B} \subset \mathcal{A}$ be C*-algebras with $\mathcal{B}$ hereditary and $\mathcal{A}$ non-unital. Show that $\mathcal{B}$ is hereditary in $\tilde{\mathcal{A}}$.

Answer. A positive element in $\tilde{\mathcal{A}}$ is of the form $a = (c + \lambda I_{\tilde{\mathcal{A}}})^*(c + I_{\tilde{\mathcal{A}}}) = c^*c + 2\mathrm{Re}\,\lambda c^* + |\lambda|^2 I_{\tilde{\mathcal{A}}}$, with $c \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. Suppose that $a \leq b$ for some $b \in \mathcal{B}$. If $\rho \in S(\tilde{\mathcal{A}})$ is the state given by $\rho(\alpha I_{\tilde{\mathcal{A}}}) = \alpha$ and $\rho|_{\mathcal{A}} = 0$ from Exercise 11.5.12, then $0 \leq \rho(b - a) = -|\lambda|^2$, so $\lambda = 0$. This gives us $a = c^*c$ and, as $\mathcal{B}$ is a hereditary subalgebra of $\mathcal{A}$, from $c^*c = a \leq b$ we get that $a = c^*c \in \mathcal{B}$. Hence $\mathcal{B}$ is hereditary in $\tilde{\mathcal{A}}$.

(11.5.14) Let $a \in \mathcal{A}$ positive. Show that a is strictly positive if and only if $\left\{a\left(\frac{1}{n} I_{\mathcal{A}} + a\right)^{-1}\right\}_n$ is an approximate unit for $\mathcal{A}$.

Answer. Write

$$e_n = a\left(\frac{1}{n} I_{\mathcal{A}} + a\right)^{-1}.$$

Suppose first that a is not strictly positive. Then there exists $\varphi \in S(\mathcal{A})$ with $\varphi(a) = 0$. We have, since $e_n \geq 0$ and it commutes with a,

$$\varphi(e_n) = \varphi\left(a^{1/2}\left(\frac{1}{n} I_{\mathcal{A}} + a\right)^{-1} a^{1/2}\right) \leq n\,\varphi(a) = 0,$$

so $\varphi(e_n) = 0$. Let $b \in \mathcal{A}$ with $\varphi(b) = 1$. Then

$$|\varphi(be_n)| = |\varphi(be_n^{1/2}e_n^{1/2})| \leq \varphi(e_n)^{1/2}\,\varphi(|be_n^{1/2}|^2)^{1/2} = 0.$$

Thus

$$\|b - be_n\| \geq |\varphi(b - be_n)| = \varphi(b) = 1.$$

So $\{be_n\}$ does not converge to b and $\{e_n\}$ is not an approximate unit.

Conversely, suppose that a is strictly positive. Then $\{e_n\}$ is an approximate unit by the proof of Proposition 11.5.24. Here is another argument, though. By Proposition 11.5.23, $\mathcal{A} = \overline{a\mathcal{A}a}$. Let $b \in A$; fix $\varepsilon > 0$ and let $c \in A$ such that $\|b - aca\| < \varepsilon$. We have

$$a\left(\frac{1}{n} + a\right)^{-1} b - b = a\left(\frac{1}{n} + a\right)^{-1} b - \left(\frac{1}{n} + a\right)\left(\frac{1}{n} + a\right)^{-1} b$$
$$= -\frac{1}{n}\left(\frac{1}{n} I_{\mathcal{A}} + a\right)^{-1} b.$$

Then, using that $\left\|\frac{1}{n}(\frac{1}{n}+a)^{-1}\right\| \leq 1$,

$$\begin{aligned}\left\|a\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}b-b\right\| &= \left\|\frac{1}{n}\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}b\right\| \\ &\leq \left\|\frac{1}{n}\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}aca\right\| + \left\|\frac{1}{n}\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}(aca-b)\right\| \\ &\leq \left\|\frac{1}{n}\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}aca\right\| + \|aca-b\| \\ &\leq \|ac\|\left\|\frac{1}{n}\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}a\right\| + \varepsilon \\ &\leq \frac{\|ac\|}{1+n}+\varepsilon,\end{aligned}$$

where the last estimate comes from functional calculus on a with the function $\frac{t}{n}(\frac{1}{n}+t)^{-1} = \frac{t}{1+nt} \leq \frac{1}{1+n}$ for $t \in [0,1]$.

Thus

$$\limsup_n \left\|\frac{1}{n}\Big(\frac{1}{n}\,I_{\mathcal A}+a\Big)^{-1}b-b\right\| \leq \varepsilon$$

and, as ε was arbitrary, we get from the Limsup Routine that

$$\lim_n \left\|\frac{1}{n}\Big(\frac{1}{n}+a\Big)^{-1}b-b\right\| = 0.$$

(11.5.15) Let $a \in \mathcal{A}$ positive with $\|a\| \leq 1$. Show that a is strictly positive if and only if $\{a^{1/n}\}_n$ is an approximate unit for $\mathcal{A}$.

Answer. Suppose first that a is not strictly positive. Then there exists $\varphi \in S(\mathcal{A})$ with $\varphi(a) = 0$. Using Cauchy–Schwarz we have, working on $\tilde{\mathcal{A}}$ with the unique extension of φ,

$$\varphi(a^{1/n}) \leq \varphi(a^{2/n})^{1/2} \leq \cdots \varphi(a^{2^m/n})^{2^{-m}}. \tag{AB.11.1}$$

If $r > 1$, then

$$0 \leq \varphi(a^r) = \varphi(a^{1/2}a^{r-1}a^{1/2}) \leq \varphi(a) = 0,$$

so $\varphi(a^r) = 0$. With m big enough so that $2^m/n > 1$, we have $\varphi(a^{2^m/n}) = 0$, and hence $\varphi(a^{1/n}) = 0$ by (AB.11.1). Given $b \in \mathcal{A}$ with $\varphi(b) = 1$,

$$|\varphi(ba^{1/n})| = |\varphi(ba^{1/2n}a^{1/2n})| \leq \varphi(a^{1/n})^{1/2}\varphi(|ba^{1/2n}|^2)^{1/2} = 0.$$

Then

$$\|b - ba^{1/n}\| \geq |\varphi(b - ba^{1/n})| = \varphi(b) = 1,$$

showing that $ba^{1/n}$ does not converge to b.

Conversely, suppose that a is strictly positive. By Proposition 11.5.23, $\mathcal{A} = \overline{a\mathcal{A}a}$. Let $b \in \mathcal{A}$; fix $\varepsilon > 0$ and let $c \in \mathcal{A}$ such that $\|b - aca\| < \varepsilon$. We have

$$\begin{aligned}\|a^{1/n}b - b\| &\leq \|a^{1/n}b - a^{1/n}aca\| + \|a^{1/n}aca - aca\| + \|aca - b\| \\ &\leq \|a^{1/n}\| \, \|b - aca\| + \|ca\| \, \|a^{1+1/n} - a\| + \|b - aca\| \\ &\leq 2\varepsilon + \|c\| \, \|a^{1+1/n} - a\|.\end{aligned}$$

We will be done if we show that $\|a^{1+1/n} - a\| \to 0$. For this, we use functional calculus. Let $f_n(t) = t^{1+1/n} - t$. Fix $\delta \in (0,1)$. When $t < \delta$,

$$|f_n(t)| = |t^{1+1/n} - t| \leq \delta^{1+1/n} + \delta \leq \delta^2 + \delta \leq 2\delta.$$

And when $t \geq \delta$, using the coarse inequality $|t^{1/n} - 1| \leq \frac{k}{n} e^{k/n}$ for $k = \max\{|\log \delta|, \|a\|\}$ and $t \in [\delta, \|a\|]$,

$$|f_n(t)| = \left|t^{1+\frac{1}{n}} - t\right| = t\left|t^{1/n} - 1\right|. \leq \frac{2ke^{k/n}}{n}$$

The two estimates together show that $f_n \to 0$ uniformly on $[0, \|a\|]$. Hence $f_n(a) \to 0$ in $\mathcal{A}$. That is, $\|a^{1+1/n} - a\| \to 0$. Then

$$\limsup_n \|a^{1/n}b - b\| \leq 2\varepsilon$$

and, as this can be done for all $\varepsilon > 0$, the Limsup Routine guarantees that $\lim_n \|a^{1/n}b - b\| = 0$.

11.6. Representations

(11.6.1) Let $\{\mathcal{H}_j\}$ be a family of Hilbert spaces and let $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ the ℓ^2 direct sum as in Definition 5.3.8. Show that $\mathcal{H}$ is a Hilbert space.

Answer. We are given the norm and not the inner product, but knowing that if $g \in \mathcal{H}$ the norm of g is $\|g\| = \sum_j \|g(j)\|^2$, it is not hard to guess that this norm comes from the inner product

$$\langle g, h \rangle = \sum_j \langle g(j), h(j) \rangle.$$

A double use of Cauchy–Schwarz shows that $|\langle g,h\rangle| \le \|g\|\,\|h\|$. That the product is sesquilinear follows from the sesquilinearity of the inner product in each $\mathcal{H}_j$ and linearity of limits. The completeness of $\mathcal{H}$ was addressed in Exercise 5.3.6.

(11.6.2) Show that in the proof of Theorem 11.6.2 the form $\langle a+L, b+L\rangle = [a,b]$ is well-defined and an inner product.

Answer. If $a'-a=l_1\in L$ and $b'-b=l_2\in L$, by (11.18)

$$[a',b'] = [a+l_1,b+l_2] = [a,b]+[l_1,b]+[a,l_2]+[l_1,l_2] = [a,b],$$

so the definition of the form does not depend on the representatives. The sesquilinearity follows from the sesquilinearity of the form:

$$\begin{aligned}[a_1+\alpha a_2, b_1+\beta b_2] &= \varphi((b_1+\beta b_2)^*(a_1+\alpha a_2))\\ &= \varphi(b_1^*a_1)+\overline{\beta}\varphi(b_2^*a_1)+\alpha\varphi(b_1^*a_2)+\overline{\beta}\alpha\varphi(b_2^*a_2)\\ &= [a_1,b_1]+\overline{\beta}[a_1,b_2]+\alpha[a_2,b_1]+\overline{\beta}\alpha[a_2,b_2].\end{aligned}$$

Finally, if $[a,a]=0$ this is $\varphi(a^*a)=0$, so $a\in L$ and $a+L=L$, so the form is an inner product.

(11.6.3) If $\varphi,\psi\in S(\mathcal{A})$ with $\varphi\sim\psi$, construct V and show the equality (11.19).

Answer. We have $\psi=\varphi(u\cdot u^*)$. Note that

$$L_\varphi u = \{au:\ \varphi(a^*a)=0\} = \{au:\ \psi(u^*a^*au)=0\} = L_\psi.$$

Define $V(a+L_\varphi)=au+L_\psi$; this is well-defined since $L_\psi=L_\varphi u$. We have

$$\|V(a+L_\varphi)\|^2 = \|au+L_\psi\|^2 = \psi(u^*a^*au) = \varphi(a^*a) = \|a+L_\varphi\|^2,$$

so V is an isometry. As it has dense range and it is isometric, V is a unitary. Also,

$$\begin{aligned}\pi_\psi(a)V(b+L_\varphi) &= \pi_\psi(a)(bu+L_\psi) = abu+L_\psi\\ &= V(ab+L_\varphi) = V\pi_\varphi(a)(b+L_\varphi).\end{aligned}$$

As $\mathcal{A}+L_\varphi$ is dense in $\mathcal{H}_\varphi$ and the operators involved are bounded, we get that $\pi_\psi(a)V=V\pi_\varphi(a)$ for all a. As V is a unitary this is $\pi_\psi(a)=V\pi_\varphi(a)V^*$.

(11.6.4) Find an example of $\mathcal{A}$, $a \in \mathcal{A}$, $\varphi \in S(\mathcal{A})$, such that $\varphi(a) = 0$ but $\pi_\varphi(a) \neq 0$.

Answer. The assertion $\pi_\varphi(a) = 0$ is equivalent to $\varphi(b^*a^*ab) = 0$ for all $b \in \mathcal{A}$. So for instance take $\mathcal{A} = M_2(\mathbb{C})$, and $\varphi(X) = X_{11}$, and $A = E_{12}$. Then $\varphi(A) = 0$, but for instance

$$\langle \pi_\varphi(A)(E_{21} + L), E_{11} + L\rangle = \varphi(E_{11}) = 1,$$

so $\pi_\varphi(A) \neq 0$.

(11.6.5) Let $\mathcal{A}$ be a simple C*-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a nonzero representation. Show that π is faithful.

Answer. Since π is bounded by Proposition 11.4.9, $\ker\pi$ is a closed bilateral ideal. It cannot be all of $\mathcal{A}$ because π is nonzero; so $\ker\pi = \{0\}$ and hence π is injective/faithful.

(11.6.6) Let $\varphi \in S(\mathcal{A})$, $a \in \mathcal{A}$. Show that $\pi_\varphi(a) = 0$ if and only if $\varphi(b^*a^*ab) = 0$ for all $b \in \mathcal{A}$. Conclude that φ faithful implies π_φ faithful. Show that the converse is not necessarily true.

Answer. If $\pi_\varphi(a) = 0$ then $\pi_\varphi(a^*a) = 0$ and this means that, for all $b \in \mathcal{A}$,

$$0 = \langle \pi_\varphi(a^*a)b\xi, b\xi\rangle = \varphi(b^*a^*ab). \tag{AB.11.2}$$

And the equality in (AB.11.2) also implies the converse. So, if φ is faithful and $\pi_\varphi(a) = 0$, in particular $0 = \varphi(a^*a) = 0$ and so $a = 0$, showing that π_φ is faithful.

Finally, we need to construct a non-faithful state with faithful GNS representation. This is actually easy if we take $\mathcal{A}$ to be simple: in that case any non-zero representation is faithful. We did this in Exercise 11.6.4 (and also in text, on page 805 of the Book) with $\mathcal{A} = M_2(\mathbb{C})$ and $\varphi(a) = a_{11}$. Then

$$L = \{a \in M_2(\mathbb{C}) : \ (a^*a)_{11} = 0\} = \left\{ \begin{bmatrix} 0 & w \\ 0 & z \end{bmatrix} : \ w, z \in \mathbb{C} \right\}.$$

So $\mathcal{H}_\varphi = \mathcal{A}/L$ is 2-dimensional and is spanned by the classes of E_{11} and E_{21}; also,

$$\langle E_{11}, E_{11}\rangle = \varphi(E_{11}) = 1,$$

$$\langle E_{21}, E_{21}\rangle = \varphi(E_{12}E_{21}) = 1,$$

$$\langle E_{11}, E_{21}\rangle = \varphi(E_{12}E_{11}) = 0.$$

So $\{E_{11}, E_{21}\}$ is an orthonormal basis of $\mathcal{H}_\varphi$. And, as a matrix, $\pi_\varphi(a)$ satisfies

$$\langle \pi_\varphi(a)E_{11}, E_{11}\rangle = \varphi(E_{11}aE_{11}) = a_{11},$$

$$\langle \pi_\varphi(a)E_{11}, E_{21}\rangle = \varphi(E_{12}aE_{11}) = a_{21},$$

$$\langle \pi_\varphi(a)E_{21}, E_{11}\rangle = \varphi(E_{11}aE_{21}) = a_{12},$$

$$\langle \pi_\varphi(a)E_{21}, E_{21}\rangle = \varphi(E_{12}aE_{21}) = a_{22}.$$

So using the orthonormal basis $\{E_{11}, E_{21}\}$ on $\mathcal{H}_\varphi$, we get that π_φ is the identity representation, which of course is faithful.

(11.6.7) Let $\varphi \in S(\mathcal{A})$ be faithful, and $a \in \mathcal{A}$. Show that if $\pi_\varphi(a)\xi_\varphi = 0$, then $a = 0$. *We will give a name to this property of ξ_φ in Section 12.5: **separating**.*

Answer. We have

$$\varphi(a^*a) = \langle \pi_\varphi(a^*a)\xi_\varphi, \xi_\varphi\rangle = \langle \pi_\varphi(a)\xi_\varphi, \pi_\varphi(a)\xi_\varphi\rangle = 0.$$

As φ is faithful, $a = 0$.

(11.6.8) Let $\mathcal{A}$ be a C*-algebra, $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation, and $\xi \in \mathcal{H}$. Let $P \in \mathcal{B}(\mathcal{H})$ be the orthogonal projection onto $\overline{\pi(\mathcal{A})\xi}$. Show that $P\pi(a) = \pi(a)P$ for all $a \in \mathcal{A}$.

Answer. We have

$$(P\pi(a)P)\pi(b)\xi = P\pi(ab)\xi = \pi(ab)\xi = \pi(a)\pi(b)\xi = \pi(a)P\pi(b)\xi.$$

As $\{\pi(b)\xi :\ b \in \mathcal{A}\}$ is dense in $P\mathcal{H}$, we get that $P\pi(a)P = \pi(a)P$ for all $a \in \mathcal{A}$. When a is selfadjoint, taking adjoints we get $P\pi(a) = P\pi(a)P = \pi(a)P$. And as the selfadjoint elements span $\mathcal{A}$, we get $P\pi(a) = \pi(a)P$ for all $a \in \mathcal{A}$.

(11.6.9) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$. Show that there exists a unique $b \in \mathcal{A}$ such that $a = bb^*b$.

Answer. We can think of $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$. Write $a = u|a|$ the polar decomposition. Recall that $|a| \in \mathcal{A}$ but that in general $u \notin \mathcal{A}$. Let $b = u|a|^{1/3}$. Since $u^*u|a| = |a|$ (because u^*u is the orthogonal projection onto the closure of the range of a^*), we obtain $u^*u|a|^m = |a|^m$ for all $m \in \mathbb{N}$, and hence $u^*up(a) = p(a)$ for every $p \in \mathbb{C}[x]$. Taking limits we get $u^*u|a|^{1/3} = |a|^{1/3}$. Then

$$bb^*b = u|a|^{1/3}|a|^{1/3}u^*u|a|^{1/3} = u|a| = a.$$

We haven't shown yet that $b \in \mathcal{A}$. Note, though, that

$$|b|^2 = b^*b = |a|^{1/3}u^*u|a|^{1/3} = |a|^{2/3}.$$

So $|b| = |a|^{1/3} \in \mathcal{A}$. We have that $u|a| = a \in \mathcal{A}$. Then $u|a|^m = a|a|^{m-1} \in \mathcal{A}$ for all $m \in \mathbb{N}$. Thus $up(|a|) \in \mathcal{A}$ for all $p \in \mathbb{C}[x]$. By Functional Calculus, if $\{p_n\}$ is a sequence of polynomials with $p_n(t) \to t^{1/3}$ uniformly, $b = u|a|^{1/3} = \lim_n up_n(|a|) \in \mathcal{A}$.

As for the uniqueness, suppose that $bb^*b = cc^*c$. Then

$$|b|^6 = (b^*bb^*)bb^*b = (c^*cc^*)cc^*c = |c|^6.$$

Functional calculus then gives us $|b| = |c|$. So $b|b| = c|b|$. Now since $\ker b = \ker b^*b = \ker |b|^2 = \ker |b|$, we have that $(\ker b)^\perp = \overline{\operatorname{ran}}|b|$. And from $|c| = |b|$ we get that $(\ker c)^\perp = (\ker b)^\perp$. Given $\xi \in \mathcal{H}$ we can write $\xi = \xi_0 + \xi_1$ with $\xi_0 \in \ker b$ and $\xi_1 \in \overline{\operatorname{ran}}|b|$. Write $\xi_1 = \lim_n |b|\eta_n$. Then

$$b\xi = b\xi_1 = \lim_n b|b|\eta_n = \lim_n c|b|\eta_n = c\xi_1 = c\xi.$$

Thus $c = b$.

The existence can also be shown without using representations and using instead Proposition 11.3.11 to see that $a = u(a^*a)^{1/3}$ for some $u \in \mathcal{A}$. The construction of u in Proposition 11.3.11 gives $u^*u = (a^*a)^{1/3}$, and then $uu^*u = a$.

Here is a third argument using block matrices. Let

$$r = \begin{bmatrix} 0 & a^* \\ a & 0 \end{bmatrix}, \qquad u = \begin{bmatrix} I_{\tilde{\mathcal{A}}} & 0 \\ 0 & -I_{\tilde{\mathcal{A}}} \end{bmatrix}.$$

Then u is a unitary and $u^*ru = -r$. As $f(t) = t^{1/3}$ is continuous everywhere, we have $-r^{1/3} = -f(r) = f(-r) = f(u^*ru) = u^*f(r)u = u^*r^{1/3}u$. So $r^{1/3}u = -ur^{1/3}$, which forces

$$r^{1/3} = \begin{bmatrix} 0 & b^* \\ b & 0 \end{bmatrix}$$

for some $b \in \mathcal{A}$. Since

$$\begin{bmatrix} 0 & a^* \\ a & 0 \end{bmatrix} = (r^{1/3})^3 = \begin{bmatrix} 0 & b^* \\ b & 0 \end{bmatrix}^3 = \begin{bmatrix} 0 & b^*bb^* \\ bb^*b & 0 \end{bmatrix},$$

it follows that $a = bb^*b$. For the uniqueness, if $a = cc^*c$, we form

$$z = \begin{bmatrix} 0 & c^* \\ c & 0 \end{bmatrix},$$

and then z is selfadjoint with $z^3 = r$. Then $z = f(z^3) = f(r) = r^{1/3}$, and hence $c = b$.

(11.6.10) Let $\mathcal{A}$ be a non-unital C*-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. Prove that there exists a (unique, if π is non-degenerate) representation $\tilde{\pi} : \tilde{\mathcal{A}} \to \mathcal{B}(\mathcal{H})$ that extends π. If π is faithful, so is $\tilde{\pi}$.

Answer. Let $\tilde{\pi}_{\mathcal{A}} : \tilde{\mathcal{A}} \to \mathcal{B}(\mathcal{H})$ be given by $\tilde{\pi}_{\mathcal{A}}(a, \lambda) = \pi_{\mathcal{A}}(a) + \lambda\, I_{\mathcal{H}}$. We have

$$\begin{aligned} \tilde{\pi}_{\mathcal{A}}(a_1 + a_2, \lambda_1 + \lambda_2) &= \pi_{\mathcal{A}}(a_1 + a_2) + (\lambda_1 + \lambda_2)\, I_{\mathcal{H}} \\ &= \tilde{\pi}_{\mathcal{A}}(a, \lambda) + \tilde{\pi}_{\mathcal{A}}(a, \lambda_2), \end{aligned}$$

so $\tilde{\pi}$ is linear. Similarly,

$$\begin{aligned} \tilde{\pi}\big((a_1, \lambda_1)(a_2, \lambda_2)\big) &= \tilde{\pi}(a_1a_2 + \lambda_2 a_1 + \lambda_1 a_2, \lambda_1\lambda_2) \\ &= \pi(a_1a_2) + \lambda_2\pi(a_1) + \lambda_1\pi(a_2) + \lambda_1\lambda_2\, I_{\mathcal{H}} \\ &= \big(\pi(a_1) + \lambda_1\, I_{\mathcal{H}}\big)\big(\pi(a_2) + \lambda_2\, I_{\mathcal{H}}\big) \\ &= \tilde{\pi}(a_1, \lambda_1)\tilde{\pi}(a_2, \lambda_2), \end{aligned}$$

so $\tilde{\pi}$ is multiplicative. And

$$\tilde{\pi}\big((a, \lambda)^*\big) = \tilde{\pi}(a^*, \bar{\lambda}) = \pi(a^*) + \bar{\lambda}\, I_{\mathcal{H}} = \tilde{\pi}(a, \lambda)^*.$$

Thus $\tilde{\pi}$ is a representation, and it extends π by construction. When π is faithful and $\tilde{\pi}(a, \lambda) = 0$, we have $\pi(a) = -\lambda I_{\mathcal{H}}$. If $\lambda \neq 0$, then for any $b \in \mathcal{A}$ we have

$$\pi((\lambda^{-1}a)b) = \lambda^{-1}\pi(a)\pi(b) = I_{\mathcal{H}}\pi(b) = \pi(b).$$

As π is faithful, $(\lambda^{-1}a)b = b$, and a similarly computation can be made on the other side, showing that $(\lambda^{-1}a) = I_{\mathcal{A}}$; but $\mathcal{A}$ was assumed not unital. It follows that $\lambda = 0$, and then $\pi(a) = 0$ which in turn implies $a = 0$ by the faithfulness of π. So $\tilde{\pi}$ is faithful.

As for the uniqueness, suppose that π is non-degenerate and $\rho : \tilde{\mathcal{A}} \to \mathcal{B}(\mathcal{H})$ is a representation with $\rho|_{\mathcal{A}} = \pi$. We have $\rho(a, \lambda) = \rho(a, 0) + \lambda\rho(0, 1) = \pi(a) + \lambda\rho(0, 1)$. In $\tilde{\mathcal{A}}$ we have, for every $a \in \mathcal{A}^+$ with $\|a\| \leq 1$, that $(a, 0) \leq (0, 1)$ (this is Corollary 11.3.8). Then $\pi(a) = \tilde{\pi}(a, 0) = \rho(a, 0) \leq \rho(0, 1)$. As

$\rho(0,1)$ is a projection, it follows by Exercise 11.3.8 that $\rho(0,1)\pi(a) = \pi(a)$ for all $a \in \mathcal{A}^+$, and a fortiori for all $a \in \mathcal{A}$. Then $\rho(0,1)\pi(a)\xi = \pi(a)\xi$ for all $a \in \mathcal{A}$ and $\xi \in \mathcal{H}$; and because π is non-degenerate, $\rho(0,1) = I_{\mathcal{H}}$.

Then, for any $a \in \mathcal{A}$ and $\lambda \in \mathbb{C}$,

$$\rho(a,\lambda) = \rho(a,0) + \lambda\rho(0,1) = \pi(a) + \lambda I_{\mathcal{H}} = \tilde{\pi}(a,\lambda).$$

When π is degenerate the extension is not unique because $\rho(0,1)$ can be assigned to be any projection with range containing $\pi(\mathcal{A})\mathcal{H}$.

11.7. Matrices over a C*-algebra

(11.7.1) Prove Proposition 11.7.1 (block matrices were already considered in Section 10.4).

Answer.

(i) Let j be such that $\|a_j\| = \max\{\|a_k\| : \ k\}$. Given $\tilde{\xi} \in \mathcal{H}^n$,

$$\Big\|\Big(\sum_{k=1}^n E_{kk} \otimes a_k\Big)\tilde{\xi}\Big\|^2 = \sum_{k=1}^n \|a_k\xi_k\|^2 \le \sum_{k=1}^n \|a_k\|^2 \|\xi_k\|^2$$
$$\le \|a_j\|^2 \sum_{k=1}^n \|\xi_k\|^2 = \|a_j\|^2 \|\tilde{\xi}\|^2.$$

So the norm is at most $\|a_j\|$. Now fix $\varepsilon > 0$ and choose $\xi \in \mathcal{H}$ such that $\|\xi\| = 1$ and $\|a_j\xi\| \ge \|a_j\| - \varepsilon$. Let $\tilde{\xi} \in \mathcal{H}^n$ be the vector with ξ in the j^{th} entry and zeros elsewhere. Then

$$\Big\|\Big(\sum_{k=1}^n E_{kk} \otimes a_k\Big)\tilde{\xi}\Big\|^2 = \|a_j\xi\|^2 \ge (\|a_j\| - \varepsilon)^2.$$

This shows that $\Big\|\sum_{k=1}^n E_{kk} \otimes a_k\Big\| \ge \|a_j\| - \varepsilon$. As ε was arbitrary, the reverse inequality is proven, and so $\Big\|\sum_{k=1}^n E_{kk} \otimes a_k\Big\| = \|a_j\|$.

(ii) We have, since the matrix is selfadjoint and using what we have just proven,

$$\left\| \begin{bmatrix} 0 & a \\ a^* & 0 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} 0 & a \\ a^* & 0 \end{bmatrix}^2 \right\| \left\| \begin{bmatrix} aa^* & 0 \\ 0 & a^*a \end{bmatrix} \right\|$$
$$= \max\{\|a^*a\|, \|aa^*\|\} = \|a\|^2.$$

(iii) If $\begin{bmatrix} I_{\mathcal{A}} & a \\ a^* & b \end{bmatrix} \geq 0$, then for any $\xi \in \mathcal{H}$

$$\langle (b - a^*a)\xi, \xi \rangle = \left\langle \begin{bmatrix} I_{\mathcal{A}} & a \\ a^* & b \end{bmatrix} \begin{bmatrix} -a\xi \\ \xi \end{bmatrix}, \begin{bmatrix} -a\xi \\ \xi \end{bmatrix} \right\rangle \geq 0,$$

so $a^*a \leq b$. Conversely, if $a^*a \leq b$, then for any $\xi, \eta \in \mathcal{H}$,

$$\left\langle \begin{bmatrix} I_{\mathcal{A}} & a \\ a^* & b \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \begin{bmatrix} \xi \\ \eta \end{bmatrix} \right\rangle = \|\xi\|^2 + \langle b\eta, \eta \rangle + 2\mathrm{Re}\, \langle a\eta, \xi \rangle$$
$$\geq \|\xi\|^2 + \langle b\eta, \eta \rangle - 2\langle a^*a\eta, \eta \rangle^{1/2} \|\xi\|$$
$$\geq \|\xi\|^2 + \langle b\eta, \eta \rangle - 2\langle b\eta, \eta \rangle^{1/2} \|\xi\|$$
$$= (\|\xi\| - \langle b\eta, \eta \rangle^{1/2})^2 \geq 0.$$

(iv) Let $\xi, \eta \in \mathcal{H}$. Let $\tilde{\xi} = (0, \ldots, 0, \xi, 0, \ldots, 0)$, $\tilde{\eta} = (0, \ldots, 0, \eta, 0, \ldots, 0)$, where ξ is in the j position and η in the k position. Then

$$|\langle a_{kj}\xi, \eta \rangle| = |\langle a\tilde{\xi}, \tilde{\eta} \rangle| \leq \|a\| \, \|\tilde{\xi}\| \, \|\tilde{\eta}\| = \|a\| \, \|\xi\| \, \|\eta\|.$$

As $\|a_{kj}\| = \max\{|\langle a_{kj}\xi, \eta \rangle| : \ \|\xi\| = \|\eta\| = 1\}$, we get that $\|a_{kj}\| \leq \|a\|$.

(11.7.2) Given a compact Hausdorff space T, show that the C*-algebras $\mathcal{A} = M_n(C(T))$ and $\mathcal{B} = C(T, M_n(\mathbb{C}))$ are canonically isomorphic, where the norm in $\mathcal{B}$ is given by

$$\|y\|_{\mathcal{B}} = \sup\{\|y(t)\| : \ t \in T\}.$$

Answer. We define $\Gamma : \mathcal{A} \to \mathcal{B}$ by

$$\Gamma(\tilde{a})(t) = \sum_{k,j} a_{kj}(t)\, E_{kj}.$$

We have

$$\Gamma(\tilde a+\tilde b)(t)=\sum_{k,j}\big(a_{kj}(t)+b_{kj}(t)\big)\,E_{kj}$$
$$=\sum_{k,j}a_{kj}(t)\,E_{kj}+\sum_{k,j}b_{kj}(t)\,E_{kj}$$
$$=\Gamma(\tilde a)(t)+\Gamma(\tilde b)(t).$$

Also,

$$\Gamma(\tilde a\tilde b)(t)=\sum_{k,j}a_{kj}(t)\,E_{kj}\sum_{r,s}b_{rs}(t)\,E_{rs}=\Gamma(\tilde a)(t)\,\Gamma(\tilde b)(t).$$

And

$$\Gamma(\tilde a^*)(t)=\sum_{k,j}\overline{a_{jk}}(t)\,E_{kj}=\Gamma(\tilde a)(t)^*.$$

So Γ is a $*$-homomorphism. If $\Gamma(\tilde a)=0$, this means that $a_{kj}(t)=0$ for all t and all k,j, so $a_{kj}=0$ for all k,j and hence $a=0$; thus Γ is injective. Given $y\in C(T,M_n(\mathbb{C}))$, let $\tilde a=\sum_{k,j}y_{kj}(t)\,E_{kj}$. Then $\Gamma(\tilde a)=y$ and Γ is surjective. Thus Γ is a C*-isomorphism.

(11.7.3) Let $\mathcal{A}$ be a C*-algebra and $\tilde{\mathcal{J}}\subset M_n(\mathcal{A})$. Show that $\tilde{\mathcal{J}}$ is an ideal if and only if $\tilde{\mathcal{J}}=M_n(\mathcal{J})$ for an ideal $\mathcal{J}\subset\mathcal{A}$.

Answer. Let

$$\mathcal{J}=\{(\tilde a)_{11}:\ a\in\tilde{\mathcal{J}}\}.$$

Since addition of matrices and multiplication by scalars are entrywise, $\mathcal{J}$ is a subspace. Given $b\in\mathcal{J}$ and $a\in\mathcal{A}$, let $\tilde b\in\tilde{\mathcal{J}}$ such that b is the $1,1$ entry of $\tilde b$. Then ab is the $1,1$ entry of $(E_{11}\otimes a)\tilde b$; that is, $ab\in\mathcal{J}$. Similarly, $ba\in\mathcal{J}$ and so $\mathcal{J}$ is an ideal. Assume initially that $\mathcal{A}$ is unital. Given any $\tilde b=\sum_{k,j}b_{kj}\otimes E_{kj}\in\tilde{\mathcal{J}}$, fix indices r and s. Then

$$\tilde{\mathcal{J}}\ni(E_{1r}\otimes I_{\mathcal{A}})\tilde b(E_{s1}\otimes I_{\mathcal{A}})=\sum_{k,j}E_{1r}E_{kj}E_{s1}\otimes b_{kj}=E_{11}\otimes b_{rs},\tag{AB.11.3}$$

showing that $b_{rs}\in\mathcal{J}$. As r,s were arbitrary, $\tilde b\in M_n(\mathcal{J})$. Given an arbitrary element $\tilde b\in M_n(\mathcal{J})$ and r,s, by (AB.11.3) there exists $\tilde c\in\tilde{\mathcal{J}}$ such that b_{rs} is the $1,1$ entry of $\tilde c$. Then

$$E_{11}\otimes b_{rs}=(E_{1r}\otimes I_{\mathcal{A}})\tilde c(E_{s1}\otimes I_{\mathcal{A}})\in\tilde{\mathcal{J}}.$$

Therefore
$$\tilde{b} = \sum_{k,j} E_{kj} \otimes b_{kj} = \sum_{k,j} (E_{k1} \otimes I_{\mathcal{A}})(E_{11} \otimes b_{kj})(E_{1j} \otimes I_{\mathcal{A}}) \in \tilde{\mathcal{J}},$$
and so $M_n(\mathcal{J}) = \tilde{J}$.

When $\mathcal{A}$ is not unital in the computations above, we replace $I_{\mathcal{A}}$ above with elements from an approximate identity; for instance we get a net of elements in $\tilde{\mathcal{J}}$ that converge to $E_{11} \otimes b_{rs}$; and as $\tilde{\mathcal{J}}$ is closed, we are done.

Conversely, if $\tilde{\mathcal{J}} = M_n(\mathcal{J})$ for an ideal $\mathcal{J} \subset \mathcal{A}$, when we perform the matrix product between and element $\tilde{a} \in \mathcal{A}$ and $\tilde{b} \in \tilde{\mathcal{J}}$, all the terms will have a factor from $\mathcal{J}$, and so the result stays in $M_n(\mathcal{J})$, which is thus an ideal. Note that we already know that $M_n(\mathcal{J})$ is a C*-algebra.

11.8. Finite-Dimensional C*-algebras

(11.8.1) Let $\mathcal{A}$ be a C*-algebra and $p \in \mathcal{A}$ a projection. Show that p is positive and that $0 \leq p \leq I_{\mathcal{A}}$.

Answer. By definition of projection we have $p = p^*p$, so p is positive. We also have $\|p\|^2 = \|p^*p\| = \|p\|$, so if p is nonzero we get that $\|p\| = 1$. Then $0 \leq p \leq I_{\mathcal{A}}$ by Corollary 11.3.8.

(11.8.2) Let $\mathcal{A}$ be a C*-algebra and $p_1, \ldots, p_m \in \mathcal{A}$ projections. Show that the following statements are equivalent:

(i) $p_1, \ldots, p_m$ are pairwise orthogonal;

(ii) $\displaystyle\sum_{k=1}^{m} p_k \leq I_{\mathcal{A}}$.

Answer. If $p_1, \ldots, p_m$ are pairwise orthogonal, let $p = \sum_k p_k$. Then $p^2 = p$ and $p^* = p$, so p is a projection. By Exercise 11.8.1 we have that $p \leq I_{\mathcal{A}}$.

The converse is proven in Proposition 10.5.5. The argument is entirely algebraic, so it applies in any C*-algebra.

(11.8.3) Let $\mathcal{A}$ be a C*-algebra, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ nonzero, and $p_1, \ldots, p_n \in \mathcal{A}$ pairwise orthogonal projections. Let

$$a = \sum_{j=1}^{n} \lambda_j p_j.$$

Show that a is normal, $\sigma(a) = \{\lambda_1, \ldots, \lambda_n\}$ when $\sum_j p_j = I_{\mathcal{A}}$ and $\sigma(a) = \{0\} \cup \{\lambda_1, \ldots, \lambda_n\}$ when $\sum_j p_j \neq I_{\mathcal{A}}$; and for any function $f : \mathbb{C} \to \mathbb{C}$,

$$f(a) = \sum_{j=1}^{n} f(\lambda_j) p_j.$$

Answer. The fact that $p_1, \ldots, p_n$ are pairwise orthogonal gives $a^*a = aa^*$ by a direct computation. From $(a - \lambda_j I_{\mathcal{A}})p_j = 0$ we get that $a - \lambda_j I_{\mathcal{A}}$ cannot be invertible, and so $\{\lambda_1, \ldots, \lambda_n\} \subset \sigma(a)$. Conversely, if $\lambda \notin \{\lambda_1, \ldots, \lambda_n\}$ is nonzero, let $q = I_{\mathcal{A}} - \sum_j p_j$, and put

$$b = -\frac{1}{\lambda}\, q + \sum_{j=1}^{n} \frac{1}{\lambda_j - \lambda}\, p_j \in \mathcal{A}.$$

Then $b(a - \lambda I_{\mathcal{A}}) = (a - \lambda I_{\mathcal{A}})b = I_{\mathcal{A}}$, so $\lambda \notin \sigma(a)$. When $\sum_j p_j = I_{\mathcal{A}}$ we get $q = 0$ in the computation above, so b can be defined even if $\lambda = 0$.

When f is any function, it is continuous on $\sigma(a)$ since it is a finite set. In fact f agrees with a polynomial on $\sigma(a)$. The pairwise orthogonality of the projections gives

$$a^k = \sum_{j=1}^{n} \lambda_j^k p_j,$$

for any k and by taking linear combinations we get that $g(a) = \sum_{j=1}^{n} g(\lambda_j) p_j$ for any polynomial g, and thus for f too.

(11.8.4) Let $\mathcal{A}$ be a C*-algebra, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ nonzero, and $p_1, \ldots, p_n \in \mathcal{A}$ pairwise orthogonal projections. Let

$$a = \sum_{j=1}^{n} \lambda_j p_j.$$

Show that if a is a projection, then $\lambda_1 = \cdots = \lambda_n = 1$.

Answer. Since a is a projection, $\sigma(a) \subset \{0,1\}$ (this follows from $a^2 = a$ and Spectral Mapping, Proposition 9.2.9). From Exercise 11.8.3 we know that $\lambda_1, \ldots, \lambda_n \in \sigma(a)$; since they are nonzero, necessarily $\lambda_j = 1$ for all j.

(11.8.5) Let $\mathcal{A}$ be a C*-algebra, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ distinct and nonzero, and also $\mu_1, \ldots, \mu_m \in \mathbb{C}$ distinct and nonzero. Consider projections $p_1, \ldots, p_n$, and $q_1, \ldots, q_m$ in $\mathcal{A}$ with $p_k p_j = q_k q_j = 0$ when $k \neq j$. Show that if

$$\sum_{k=1}^{n} \lambda_k p_k = \sum_{j=1}^{m} \mu_j q_j$$

then $n = m$ and $\lambda_j = \mu_j$, $p_j = q_j$ for all j.

Answer. We use Exercise 11.8.3. Let g be a polynomial with $g(\lambda_1) = 1$ and $g(\lambda_k) = 0$ for all $k \geq 2$. Then

$$p_1 = g(a) = \sum_{j=1}^{m} g(\mu_j) q_j.$$

By Exercise 11.8.4, $g(\mu_j) \in \{0,1\}$ for all j. If we have $\mu_j \neq \lambda_1$ for all j, we can choose g with $g(\mu_j) = 0$ for all j and $g(\lambda_1) = 1$, giving us the contradiction $p_1 = 0$. Hence there exists j with $\mu_j = \lambda_1$. By reordering if needed we may assume that $\mu_1 = \lambda_1$. Using a polynomial g with $g(\mu_1) = 1$ and $g(\mu_j) = 0$ for all $j \geq 2$ (possible since $\mu_1, \ldots, \mu_m$ are distinct), we get $p_1 = q_1$. We may now remove the first term, and repeat the argument with both sides starting from 2. Thus we inductively get $\mu_j = \lambda_j$ and $q_j = p_j$, after possibly reordering on each step. If one side runs out before the other we would get 0 on one side of the equality and a nonzero linear combination of projections on the other, leading to a contradiction; so $m = n$.

CHAPTER

12

Bounded Operators on a Hilbert Space: Part II

12.1. Locally Convex Topologies in $\mathcal{B}(\mathcal{H})$

(12.1.1) Let $\mathcal{H}$ be a Hilbert space. Show that the adjoint map $T \longmapsto T^*$ is wot-continuous on $\mathcal{B}(\mathcal{H})$.

Answer. Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ be a net such that $T_j \xrightarrow{\text{wot}} T \in \mathcal{B}(\mathcal{H})$. Fix $\xi, \eta \in \mathcal{H}$. Then

$$\langle T_j^*\xi, \eta\rangle = \overline{\langle T_j\eta, \xi\rangle} \to \overline{\langle T\eta, \xi\rangle} = \langle T^*\xi, \eta\rangle.$$

As this can be done for all choices of ξ, η, we get that $T_j^* \xrightarrow{\text{wot}} T$.

(12.1.2) Show that any bounded sot or wot Cauchy net in $\mathcal{B}(\mathcal{H})$ is convergent.

Answer. Since wot is weaker than sot, it is enough to show that a wot-Cauchy net is convergent. Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ be a bounded wot-Cauchy net with $\|T_j\| \leq c$ for all j. Given $\xi, \eta \in \mathcal{H}$ we can consider the wot neighbourhoods of 0 given by $N_\varepsilon = \{S \in \mathcal{B}(\mathcal{H}) : \ |\langle S\xi, \eta\rangle| < \varepsilon\}$, $\varepsilon > 0$. So for each $\varepsilon > 0$ there exists j_0 such that $T_j - T_k \in N_\varepsilon$ for all $j, k \geq j_0$. This implies that the net of numbers $\{\langle T_j\xi, \eta\rangle\}$ is Cauchy. Define $[\xi, \eta] = \lim_j \langle T_j\xi, \eta\rangle$. As limits are linear, this is a sesquilinear form. And it is bounded, for $|[\xi, \eta]| \leq c\,\|\xi\|\,\|\eta\|$. By Proposition 10.1.5 there exists $T \in \mathcal{B}(\mathcal{H})$ such that $\langle T\xi, \eta\rangle = \lim_j \langle T_j\xi, \eta\rangle$ for all $\xi, \eta \in \mathcal{H}$, so $T_j \xrightarrow{\text{wot}} T$.

(12.1.3) Let $\mathcal{H}$ be a separable Hilbert space. Show that the sot and wot are metrizable on the closed unit ball.

Answer. Since $\mathcal{H}$ is separable and subsets of separable metric spaces are separable (Proposition 1.8.5), $\overline{B_1^{\mathcal{H}}(0)}$ is separable. Let $\{\xi_n\} \subset \overline{B_1^{\mathcal{H}}(0)}$ be a countable dense subset and put

$$d_s(S, T) = \sum_{n=1}^{\infty} 2^{-n}\,\|(S - T)\xi_n\|.$$

This is a metric in $\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}$. Indeed, $d_s \geq 0$ and $d_s(T, T) = 0$; and the triangle inequality follows from the triangle inequality for the norm: $\|(S - T)\xi_n\| \leq \|(S - R)\xi_n\| + \|(R - T)\xi_n\|$. Since both the sot topology and the metric d_s are translation invariant, we only need to deal with convergence at 0. Suppose that $T_j \xrightarrow{\text{sot}} 0$. Fix $\varepsilon > 0$ and choose n_0 such that $2^{n_0} > \varepsilon^{-1}$. Then, since $\|T\xi_n\| \leq 1$ for all n,

$$\begin{aligned} d_s(T_j, 0) &= \sum_{n=1}^{\infty} 2^{-n}\,\|T_j\xi_n\| \leq \sum_{n=1}^{n_0} 2^{-n}\,\|T_j\xi_n\| + \sum_{n=n_0+1}^{\infty} 2^{-n}\,\|T_j\xi_n\| \\ &\leq \sum_{n=1}^{n_0} 2^{-n}\,\|T_j\xi_n\| + \sum_{n=n_0+1}^{\infty} 2^{-n} = \sum_{n=1}^{n_0} 2^{-n}\,\|T_j\xi_n\| + 2^{-n_0} \\ &< \sum_{n=1}^{n_0} 2^{-n}\,\|T_j\xi_n\| + \varepsilon. \end{aligned}$$

Thus $\limsup_j d_s(T_j, 0) \leq \varepsilon$ and, by the Limsup Routine, $\lim_j d_s(T_j, 0) = 0$. Conversely, suppose that $d_s(T_j, 0) \to 0$. Fix $\varepsilon > 0$ and $\xi \in \mathcal{H}$ with $\|\xi\| \leq 1$. As $\{\xi_n\}$ is dense, there exists n_0 such that $\|\xi_{n_0} - \xi\| < \varepsilon$. Choose j_0 such that $d_s(T_j, 0) < \varepsilon/2^{n_0}$. Then, for $j \geq j_0$,

$$\begin{aligned} \|T_j\xi\| &\leq \|T_j\xi_{n_0}\| + \|T_j(\xi - \xi_{n_0})\| \\ &\leq \|T_j\xi_{n_0}\| + \varepsilon \leq 2^{n_0} d_s(T_j, 0) + \varepsilon < 2\varepsilon. \end{aligned}$$

Then $\lim_j \|T_j\xi\| = 0$; as this works for all ξ, $T_j \xrightarrow{\text{sot}} 0$.

Now we need to consider the wot. With the same notation as above, we define

$$d_w(S,T) = \sum_{n,m=1}^{\infty} 2^{-n-m} |\langle (S-T)\xi_n, \xi_m\rangle|.$$

The fact that $\|S\|, \|T\|, \|\xi_n\| \in [0,1]$ for all n guarantees the convergence of the series. With the same argument as for d_s above we get that d_w is a metric. If $T_j \xrightarrow{\text{wot}} 0$ we argue as above to get

$$d_w(T_j, 0) \le \sum_{n,m=1}^{n_0} |\langle T_j\xi_n, \xi_m\rangle| + 3\varepsilon,$$

and we conclude that $\lim_j d_w(T_j, 0) = 0$ by the Limsup Routine. Conversely, if we have $d_w(T_j, 0) \to 0$ fix $\varepsilon > 0$ and $\xi, \eta \in \mathcal{H}$. We may assume without loss of generality that $\|\xi\| \le 1$ and $\|\eta\| \le 1$. We can get n_0 and m_0 with $\|\xi - \xi_{n_0}\| < \varepsilon$ and $\|\eta - \xi_{m_0}\| < \varepsilon$. Choose j_0 such that $d_w(T_j, 0) < \varepsilon/2^{n_0+m_0}$ for all $j \ge j_0$. Then

$$|\langle T_j\xi, \eta\rangle| \le 2\varepsilon + |\langle T_j\xi_{n_0}, \xi_{m_0}| \le 2^{n_0+m_0} d_w(T_j, 0) + 2\varepsilon.$$

It follows that $\lim_j |\langle T_j\xi, \eta\rangle| = 0$ by the Limsup Routine. As ξ, η were arbitrary, $T_j \xrightarrow{\text{wot}} 0$.

(12.1.4) Prove the equalities (12.1):

$$\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)} = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{sot}} = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{wot}}.$$

Answer. Because each topology is successively weaker than the previous one, we have

$$\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)} \subset \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{sot}} \subset \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{wot}}.$$

Suppose that $\{T_j\} \subset \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{wot}}$ is wot-Cauchy. Exercise 12.1.2 guarantees that there exists $T \in \mathcal{B}(\mathcal{H})$ with $T_j \xrightarrow{\text{wot}} T$. Then

$$|\langle T\xi, \eta\rangle| = \lim_j |\langle T_j\xi, \eta\rangle| \le \limsup_j \|T_j\|\,\|\xi\|\,\|\eta\| \le \|\xi\|\,\|\eta\|.$$

It follows that

$$\|T\| = \sup\{|\langle T\xi, \eta\rangle| : \ \|\xi\| = \|\eta\| = 1\} \le 1.$$

Thus $\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{wot}} \subset \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}$, completing the chain of equalities.

(12.1.5) Let $\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ such that $T_n \xrightarrow{\text{sot}} T \in \mathcal{B}(\mathcal{H})$. Show that $\{T_n\}$ is bounded.

Answer. The idea we need was already used in the proof of Corollary 6.3.17.

Since $\{T_n\xi\}$ is a convergent sequence in a normed space, it is bounded. Then

$$\sup\{\|T_n\xi\| : \ n \in \mathbb{N}\} < \infty$$

for each ξ. By the Uniform Boundedness Principle (Theorem 6.3.16), there exists $k > 0$ with $\|T_n\| \leq k$ for all n.

(12.1.6) Let $T \in \mathcal{B}(\mathcal{H})$, $\{\xi_j\} \subset \mathcal{H}$ with dense span, and $\{T_\alpha\} \subset \mathcal{B}(\mathcal{H})$ a net such that there exists $c > 0$ with $\|T_\alpha\| \leq c$ for all α. Show that if $T_\alpha\xi_j \to T\xi_j$ for al j, then $T_\alpha \xrightarrow{\text{sot}} T$.

Answer. Let $\xi \in \mathcal{H}$. Fix $\varepsilon > 0$. By hypothesis there exists $\eta \in \operatorname{span}\{\xi_j\}$ with $\|\xi - \eta\| < \varepsilon$. By linearity of the limit, $T_\alpha\eta \to T\eta$. And

$$\begin{aligned}\|(T - T_\alpha)\xi\| &\leq \|T\xi - T\eta\| + \|T\eta - T_\alpha\eta\| + \|T_\alpha\eta - T_\alpha\xi\| \\ &\leq 2c\|\xi - \eta\| + \|T\eta - T_\alpha\eta\| \leq 2c\varepsilon + \|T\eta - T_\alpha\eta\|.\end{aligned}$$

So $\limsup_\alpha \|(T - T_\alpha)\xi\| \leq 2c\varepsilon$. As ε was arbitrary, the Limsup Routine implies that $T_\alpha\xi \to T\xi$. This works for any ξ, and hence $T_\alpha \xrightarrow{\text{sot}} T$.

(12.1.7) Explain why the argument you used in Exercise 12.1.5 does not apply to nets (as guaranteed by Remark 12.1.11).

Answer. The argument in Exercise 12.1.5 needs the fact that a convergent sequence of numbers is bounded. The same is not true for nets of numbers, so the argument does not apply. For a simple example, consider the net $\{e^{-n}\}_{n\in\mathbb{Z}}$ with the usual order in the integers. For an example that "feels more like a net" let $\mathcal{F}$ the collection of all finite subsets of $\mathbb{N}$, and let

$$\alpha_F = \begin{cases} |F|, & 1 \notin F \\ 0, & 1 \in F \end{cases}$$

Then $\alpha_F \to 0$, as eventually $1 \in F$, but $|\alpha_F|$ can be arbitrary large.

(12.1.8) Let $K \in \mathcal{K}(\mathcal{H})$ and $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ be a bounded net with $T_j \xrightarrow{\text{sot}} 0$. Show that $\|T_jK\| \to 0$. Show by example that the assertion can fail if the net is unbounded.

Answer. Suppose first that K is rank-one. Then $K\xi = \langle \xi, \eta\rangle\,\nu$ for some $\eta, \nu \in \mathcal{H}$. We have

$$\|T_jK\xi\| = |\langle \xi, \eta\rangle|\,\|T_j\nu\| \leq \|T_j\nu\|\,\|\eta\|\,\|\xi\|.$$

Given $\varepsilon > 0$ choose j_0 such that $\|T_j\nu\| \leq \varepsilon/\|\eta\|$ for all $j \geq j_0$. Then for $j \geq j_0$ we have

$$\|T_jK\xi\| \leq \varepsilon\,\|\xi\|,$$

which means that $\|T_jK\| \to 0$. When K is finite-rank it is a sum of rank-one operators, and we also get $\|T_jK\| \to 0$ via the triangle inequality and linearity of limits. For K arbitrary, by Proposition 10.6.4 there exists a sequence $\{K_n\}$ of finite-rank operators with $\|K - K_n\| \to 0$. Fix c with $\|T_j\| \leq c$ for all j. Then

$$\|T_jK\| \leq \|T_j(K - K_n)\| + \|T_jK_n\| \leq c\|K - K_n\| + \|T_jK_n\|,$$

which gives

$$\limsup_j \|T_jK\| \leq c\|K - K_n\|.$$

As this can be done for all n, the Limsup Routine gives us that $\|T_jK\| \to 0$.

For an example, let $T_j\xi = \sqrt{n_j}\,\langle \xi, \xi_{n_j}\rangle\,\xi_{n_j}$ as in Remark 12.1.12. Let

$$K\xi = \sum_n n^{-1/4}\,\langle \xi, \xi_n\rangle\,\xi_n$$

for the same orthonormal basis as in the remark. Then

$$T_jK\xi = n_j^{1/4}\langle \xi, \xi_{n_j}\rangle\,\xi_{n_j}.$$

So $\|T_jK\| = n_j^{1/4}$ is unbounded and cannot converge to 0, even though $T_j \xrightarrow{\text{sot}} 0$.

(12.1.9) Show that $T_j \xrightarrow{\text{sot}} T$ if and only if $\operatorname{Tr}(S(T_j - T)^*(T_j - T)) \to 0$ for all $S \in \mathcal{F}(\mathcal{H})$.

Answer. Suppose first that $T_j \xrightarrow{\text{sot}} T$. Then $\|(T_j - T)\xi\| \to 0$ for all $\xi \in \mathcal{H}$. Since S is finite-rank, we may write $S = \sum_{k=1}^m \xi_k\eta_k^*$ with $\eta_1, \ldots, \eta_m$ orthonormal (Proposition 10.6.1). Then, using for the trace an orthonormal basis that begins with $\eta_1, \ldots, \eta_m$,

$$\begin{aligned}\operatorname{Tr}(S(T_j - T)^*(T_j - T)) &= \sum_{k=1}^m \operatorname{Tr}(\xi_k \eta_k^* (T_j - T)^*(T_j - T)) \\ &= \sum_{k=1}^m \operatorname{Tr}((T_j - T)^*(T_j - T)\xi_k \eta_k^*) \\ &= \sum_{k=1}^m \langle (T_j - T)\xi_k, (T_j - T)\eta_k \rangle \\ &\leq \sum_{k=1}^m \|(T_j - T)\xi_k\| \, \|(T_j - T)\eta_k\| \to 0.\end{aligned}$$

Conversely, if $\operatorname{Tr}(S(T_j - T)^*(T_j - T)) \to 0$ for all $S \in \mathcal{F}(\mathcal{H})$ and $\xi \in \mathcal{H}$, with $S = \xi\xi^*$ we have

$$\begin{aligned}\|(T_j - T)\xi\|^2 &= \langle (T_j - T)^*(T_j - T)\xi, \xi \rangle \\ &= \operatorname{Tr}(S(T_j - T)^*(T_j - T)) \to 0.\end{aligned}$$

(12.1.10) Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ with $\|T_j\| \leq c$ for all j and families $\{P_j\}$ and $\{Q_j\}$ of pairwise orthogonal projections such that $T_j = Q_j T_j P_j$ for all j. Show that $\sum_j T_j$ converges sot.

Answer. As projections are positive, the series $P = \sum_j P_j$ converges sot by Proposition 12.1.10. Then, for any $\xi \in \mathcal{H}$

$$\xi = \sum_j P_j \xi \qquad \text{and } \|\xi\|^2 = \sum_j \|P_j \xi\|^2.$$

Fix $\xi \in \mathcal{H}$ and $\varepsilon > 0$. Then there exists j_0 such that $\sum_{j \geq j_0} \|P_j \xi\|^2 < \varepsilon$. For any finite set $F \subset \{j : \ j \geq j_0\}$, as the Q_j are pairwise orthogonal,

$$\Big\| \sum_{j \in F} T_j \xi \Big\|^2 = \sum_{j \in F} \|T_j \xi\|^2 = \sum_{j \in F} \|T_j P_j \xi\|^2 \leq c \sum_{j \in F} \|P_j \xi\|^2 < c\varepsilon.$$

This shows that the tails of the series are arbitrarily small, and so $\sum_j T_j \xi$ converges for all ξ, and therefore $\sum_j T_j$ converges sot.

(12.1.11) Show that in Exercise 12.1.10 it is not possible to relax the hypothesis on the projections to just $T_j = Q_j T_j$ (that is, it is not enough for the ranges to be pairwise orthogonal, the domains should be too).

Answer. Let $\mathcal{H}$ be an infinite-dimensional Hilbert space, $\{\xi_j\}$ an orthonormal basis, and $\{E_{j,k}\}$ the corresponding matrix units. Fix k_0 and let $T_j = E_{j,k_0}$. Then the ranges of the T_j are pairwise orthogonal but the series $\sum_j T_j = \sum_j E_{j,k_0}$ cannot converge sot, as $E_{j,k_0}\xi_{k_0} = \xi_j$ for all j, and the series $\sum_j \xi_j$ cannot converge as all its tails are larger than 1.

(12.1.12) Prove Proposition 12.1.5. Use ideas from the proof of Proposition 12.1.4 and Exercise 10.4.13. The idea in Exercise 11.2.3 will be needed, too.

Answer. Fix a wot-neighbourhood $\mathcal{W}$ of 0: choose $\xi_1,\ldots,\xi_n,\eta_1,\ldots,\eta_n \in \mathcal{H}$ and put

$$\mathcal{W} = \{S \in \mathcal{B}(\mathcal{H}) : |\langle S\xi_k, \eta_k\rangle| < 1\}.$$

Let

$$L = \operatorname{span}\{\xi_1,\ldots,\xi_n,\eta_1,\ldots,\eta_n\}$$

Because $\dim L < \infty$ and $\dim \mathcal{H} = \infty$, there exists a subspace $L' \subset L^\perp$ with $\dim L' = \dim L$. We then have a natural identification of $L + L'$ with L^2, and we can consider operators on L^2 as 2×2 block matrices. Let Q be the orthogonal projection onto L, and R the orthogonal projection onto L'. Let $V : L \to L'$ be a partial isometry such that $V^*V = Q$, $VV^* = R$ (it exists because $\dim L' = \dim L$ and we can define V to map an orthonormal basis of L to an orthonormal basis of L'). Let $S = QTQ$. We define

$$U_0 = S + (Q - |S^*|^2)^{1/2}V^* - V(Q - |S|^2)^{1/2} + VS^*V^*,$$

$$U = U_0 + (I_{\mathcal{H}} - Q - R).$$

Then (to be checked at the end) $U_0^*U_0 = Q + R$, and since $Q + R = I_{L+L'}$ on the finite-dimensional space $L + L'$ we also get that $U_0U_0^* = I_{L+L'}$. Therefore $U^*U = UU^* = I_{\mathcal{H}}$. And, since $QUQ = QTQ$,

$$\langle (U-T)\xi_k, \eta_k\rangle = \langle (U-T)Q\xi_k, Q\eta_k\rangle = \langle Q(U-T)Q\xi_k, \eta_k\rangle = 0,$$

showing that $U - T \in \mathcal{W}$, which is $U \in T + \mathcal{W}$. We have shown that for every wot-neighbourhood $\mathcal{W}$ of T, there exists a unitary $U_{\mathcal{W}} \in T + \mathcal{W}$. That is, $T = \lim_{\mathcal{W}} U_{\mathcal{W}}$ is a wot-limit of unitaries.

As for the computation for U_0, using Exercise 11.2.3 and noting that

$$S^*V = SV = QV = VR = 0,$$

we have

$$
\begin{aligned}
U_0^*U_0 &= |S|^2 + S^*(Q-|S^*|^2)^{1/2}V^* + V(Q-|S^*|^2)^{1/2}S + V(Q-|S^*|^2)V^* \\
&\qquad + Q - |S|^2 - (Q-|S|^2)^{1/2}S^*V^* - VS(Q-|S|^2)^{1/2} + V|S^*|^2V^* \\
&= |S|^2 + S^*(Q-|S^*|^2)^{1/2}V^* + V(Q-|S^*|^2)^{1/2}S + V(Q-|S^*|^2)V^* \\
&\qquad + Q - |S|^2 - S^*(Q-|S^*|^2)^{1/2}V^* - V(Q-|S^*|^2)^{1/2}S + V|S^*|^2V^* \\
&= VQV^* + Q = R + Q.
\end{aligned}
$$

(12.1.13) Show that $T_j \xrightarrow{\text{wot}} T$ if and only if $\operatorname{Tr}(ST_j) \to \operatorname{Tr}(ST)$ for all $S \in \mathcal{F}(\mathcal{H})$.

Answer. By Proposition 10.6.1 any $S \in \mathcal{F}(\mathcal{H})$ is of the form $S = \sum_{k=1}^m \xi_k\eta_k^*$ with $\eta_1,\ldots,\eta_m$ orthonormal. If $T_j \to T$ wot, then for an orthonormal basis $\{\nu_n\}$ such that its first r elements are $\eta_1,\ldots,\eta_m$,

$$
\begin{aligned}
\operatorname{Tr}(ST_j) = \operatorname{Tr}(T_jS) &= \sum_n \langle T_jS\nu_n, \nu_n\rangle \\
&= \sum_{k=1}^m \langle T_jS\eta_k, \eta_k\rangle = \sum_{k=1}^m \langle T_j\xi_k, \eta_k\rangle \\
&\xrightarrow[j]{} \sum_{k=1}^m \langle T\xi_k, \eta_n\rangle \\
&= \operatorname{Tr}(TS) = \operatorname{Tr}(ST).
\end{aligned}
$$

Conversely, if $\operatorname{Tr}(ST_j) \to \operatorname{Tr}(ST)$ for all $S \in \mathcal{F}(\mathcal{H})$, given $\xi,\eta \in \mathcal{H}$ put $S = \xi\eta^*$ and then

$$
\langle T_j\xi, \eta\rangle = \operatorname{Tr}(ST_j) \to \operatorname{Tr}(ST) = \langle T\xi, \eta\rangle.
$$

So $T_j \xrightarrow{\text{wot}} T$.

(12.1.14) Prove Proposition 12.1.6.

Answer. We have

$$\begin{aligned}\|(P_j - P)\xi\|^2 &= \langle (P_j - P)^2\xi, \xi\rangle = \langle (P_j + P - PP_j - P_jP)\xi, \xi\rangle \\ &= \langle P_j\xi, \xi\rangle + \langle P\xi, \xi\rangle - \langle P_j\xi, P\xi\rangle - \langle P_jP\xi, \xi\rangle \\ &\to 2\langle P\xi, \xi\rangle - 2\langle P\xi, \xi\rangle = 0,\end{aligned}$$

so $P_j \xrightarrow{\text{sot}} P$. For the case of unitaries,

$$\begin{aligned}\|(U_j - U)\xi\|^2 &= \langle (U_j - U)^*(U_j - U)\xi, \xi\rangle \\ &= \langle (U_j^*U_j - U_j^*U - U^*U_j + U^*U)\xi, \xi\rangle \\ &= \langle \xi, \xi\rangle - \langle U\xi, U_j\xi\rangle - \langle U_j\xi, U\xi\rangle + \langle \xi, \xi\rangle \\ &\to 2\langle \xi, \xi\rangle - 2\langle U\xi, U\xi\rangle = 0,\end{aligned}$$

and so $U_j \xrightarrow{\text{sot}} U$.

(12.1.15) Find an example of a wot-convergent sequence of projections such that its limit is not a projection.

Answer. We cannot use Proposition 12.1.4 because it will not provide us with a sequence, but rather a net.

Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space with orthonormal basis $\{\xi_n\}$ and corresponding matrix units $\{E_{kj}\}$. Let

$$P_k = \frac{1}{2}(E_{11} + E_{1k} + E_{k1} + E_{kk}).$$

Then $P_k^*P_k = P_k$ for all k, and so P_k is an orthogonal projection. If $\xi = \sum_n c_n\xi_n$,

$$\begin{aligned}\Big\langle \Big(P_k - \frac{1}{2}\,E_{11}\Big)\xi, \xi\Big\rangle &= \frac{1}{2}\sum_{n,m=1}^{\infty} c_n\overline{c_m}\,\big\langle \big(E_{1k} + E_{k1} + E_{kk}\big)\xi_n, \xi_m\big\rangle \\ &= \frac{1}{2}\sum_{m=1}^{\infty} \overline{c_m}\,\big\langle c_k\xi_1 + c_1\xi_k + c_k\xi_k, \xi_m\big\rangle \\ &= \frac{1}{2}\left(c_k\overline{c_1} + c_1\overline{c_k} + |c_k|^2\right) \xrightarrow[k\to\infty]{} 0.\end{aligned}$$

So $P_k \xrightarrow{\text{wot}} \frac{1}{2}\,E_{11}$. In light of Proposition 12.1.6, the sequence $\{P_k\}$ does not converge sot.

(12.1.16) Find an example of a wot-convergent sequence of unitaries such that its limit is not a unitary.

Answer. We can use the exact same idea as in Exercise 12.1.15.

Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space with orthonormal basis $\{\xi_n\}$ and corresponding matrix units $\{E_{kj}\}$. Let

$$U_k = \frac{1}{\sqrt{2}}(E_{11} + E_{1k} + E_{k1} - E_{kk}).$$

Then $U_k^* U_k = U_k U_k^* = I_{\mathcal{H}}$ for all k, and so U_k is a unitary. If $\xi = \sum_n c_n \xi_n$, then

$$\begin{aligned}\Big\langle \Big(U_k - \frac{1}{\sqrt{2}}E_{11}\Big)\xi, \xi\Big\rangle &= \frac{1}{\sqrt{2}} \sum_{n,m=1}^{\infty} c_n \overline{c_m} \left\langle \big(E_{1k} + E_{k1} - E_{kk}\big)\xi_n, \xi_m\right\rangle \\ &= \frac{1}{\sqrt{2}} \sum_{m=1}^{\infty} \overline{c_m} \left\langle c_k\xi_1 + c_1\xi_k - c_k\xi_k, \xi_m\right\rangle \\ &= \frac{1}{\sqrt{2}} \left(c_k\overline{c_1} + c_1\overline{c_k} - |c_k|^2\right) \xrightarrow[k\to\infty]{} 0.\end{aligned}$$

So $U_k \xrightarrow{\text{wot}} \frac{1}{\sqrt{2}}E_{11}$, which is not a unitary. In light of Proposition 12.1.6, the sequence $\{U_k\}$ does not converge sot.

(12.1.17) Show that the sot-closure of $\mathcal{U}(\mathcal{H})$ is the set of isometries.

Answer. Suppose that $\{U_j\}$ is a net of unitaries and $U_j \xrightarrow{\text{sot}} V$. For any $\xi \in \mathcal{H}$,

$$\|V\xi\| = \lim_j \|U_j\xi\| = \|\xi\|,$$

so V is an isometry.

Conversely, let $V \in \mathcal{B}(\mathcal{H})$ be an isometry. If $\dim\mathcal{H} < \infty$ then V is a unitary and there is nothing to prove, so we assume $\dim\mathcal{H} = \infty$. Let $\mathcal{W} = \{T : \ \|(T-V)\xi_j\| < \varepsilon,\ j = 1,\ldots,n\}$ be a sot neighbourhood of V. Let $\mathcal{H}_1 = \operatorname{span}\{\xi_1,\ldots,\xi_n\}$ and $\mathcal{K}_1 = \operatorname{span}\{V\xi_1,\ldots,V\xi_n\}$. As $\mathcal{H}$ is infinite-dimensional, $\dim\mathcal{H}_1^\perp = \dim\mathcal{K}_1^\perp = \infty$. By mapping an orthonormal basis of $\mathcal{H}_1^\perp$ to an orthonormal basis of $\mathcal{K}_1^\perp$ we induce a unitary $W : \mathcal{H}_1^\perp \to \mathcal{K}_1^\perp$. Then $U = V|_{\mathcal{H}_1} \oplus W$ is a unitary; indeed, both $V|_{\mathcal{H}_1}$ and W are unitaries with orthogonal ranges, so

$$U^*U = (V|_{\mathcal{H}_1}^* + W^*)(V|_{\mathcal{H}_1} + W) = V|_{\mathcal{H}_1}^* V|_{\mathcal{H}_1} + W^*W = I_{\mathcal{H}_1} + I_{\mathcal{H}_1^\perp} = I_{\mathcal{H}},$$

and similarly $UU^* = I_{\mathcal{H}}$. We have

$$(U - V)\xi_j = V|_{\mathcal{H}_1}\xi_j - V\xi_j = 0.$$

So $U \in \mathcal{W}$. This shows that given the family $\{\mathcal{W}\}$ of sot-neighbourhoods of V, for each $\mathcal{W}$ we can construct a unitary U with $U \in \mathcal{W}$. Thus there is a net of unitaries that converges sot to V.

(12.1.18) Show that both sot and σ-weak are weaker than σ-sot, which is weaker than the norm topology.

Answer. It $\|T_j - T\| \to 0$, then for any $S \in \mathcal{T}(\mathcal{H})$ we have

$$|\operatorname{Tr}(S(T_j - T)^*(T_j - T))| \leq \|T_j - T\|^2 \operatorname{Tr}(|S|) \to 0.$$

So the σ-sot is weaker than the norm topology. If $T_j - T \xrightarrow{\sigma\text{-sot}} 0$ and $S \in \mathcal{T}(\mathcal{H})$ is positive, then

$$\begin{aligned}|\operatorname{Tr}(S(T - T_j))| &= |\operatorname{Tr}(S^{1/2}\,S^{1/2}(T - T_j))| \\ &\leq \operatorname{Tr}(S)^{1/2} \operatorname{Tr}\left((T - T_j)^* S(T - T_j)\right)^{1/2} \to 0.\end{aligned}$$

As the positive trace-class operators span $\mathcal{T}(\mathcal{H})$ (Lemma 10.7.4 and Proposition 10.7.5), we get that $T_j \xrightarrow{\sigma\text{-weak}} T$. Also, given $\xi \in \mathcal{H}$, with $S = \xi\xi^*$ we have

$$\|(T_j - T)\xi\|^2 = \operatorname{Tr}(S(T_j - T)^*(T_j - T)) \to 0.$$

So $T_j \xrightarrow{\text{sot}} T$.

(12.1.19) Show that the sot and σ-sot agree on bounded sets.

Answer. The sot is weaker than the σ-sot, so we need to show that if $T_j \xrightarrow{\text{sot}} T$ and $\|T_j\| \leq c$ for all j, then $T_j \xrightarrow{\sigma\text{-sot}} T$. Fix $\varepsilon > 0$ and let $S \in \mathcal{T}(\mathcal{H})$. By Proposition 10.7.9 there exists $S_0 \in \mathcal{F}(\mathcal{H})$ such that $\|S - S_0\|_1 < \varepsilon$. Then

$$\begin{aligned}|\operatorname{Tr}(S(T_j - T)^*(T_j - T))| &\leq \left|\operatorname{Tr}\left((S - S_0)(T_j - T)^*(T_j - T)\right)\right| \\ &\qquad + |\operatorname{Tr}(S_0(T - T_j)^*(T_j - T))| \\ &\leq \|S - S_0\|_1 \, \|T_j - T\|^2 \\ &\quad + |\operatorname{Tr}(S_0(T - T_j)^*(T_j - T))| \\ &\leq 4c^2\,\varepsilon + |\operatorname{Tr}(S_0(T - T_j)^*(T_j - T))|.\end{aligned}$$

Then $\limsup_j |\operatorname{Tr}(S(T_j - T)^*(T_j - T))| \leq 4c^2\varepsilon$ for all $\varepsilon > 0$. Therefore $\lim_j |\operatorname{Tr}(S(T_j - T)^*(T_j - T))| = 0$ by the Limsup Routine. That is, $T_j \xrightarrow{\sigma\text{-sot}} T$.

(12.1.20) Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. Show that

$$I_{\mathcal{H}} - P \wedge Q = (I_{\mathcal{H}} - P) \vee (I_{\mathcal{H}} - Q).$$

Answer. This is a particular case of Proposition 10.5.9 (proven in Exercise 10.5.7). We write an ad-hoc argument regardless.

Given a subset $\mathcal{K} \subset \mathcal{H}$, we use the notation $[\mathcal{K}]$ to mean the orthogonal projection onto $\operatorname{span} \mathcal{K}$. We have

$$\begin{aligned} I_{\mathcal{H}} - P \wedge Q &= [(P\mathcal{H} \cap Q\mathcal{H})^{\perp}] = [(P\mathcal{H})^{\perp} \cup (Q\mathcal{H})^{\perp}] \\ &= [(I_{\mathcal{H}} - P)\mathcal{H} \cup (I_{\mathcal{H}} - Q)\mathcal{H}] = (I_{\mathcal{H}} - P) \vee (I_{\mathcal{H}} - Q). \end{aligned}$$

(12.1.21) For each $k, j = 1, \ldots, n$ consider a net $\{T_{k,j,\alpha}\}_\alpha \subset \mathcal{B}(\mathcal{H})$. Form the $n \times n$ matrices $\tilde{T}_\alpha = [T_{k,j,\alpha}]$.

(i) Show that $\tilde{T}_\alpha \xrightarrow{\text{wot}} \tilde{T}$ in $\mathcal{B}(\mathcal{H}^n)$ if and only if $T_{k,j,\alpha} \xrightarrow{\text{wot}} T_{k,j}$ for each k, j.

(ii) Show that $\tilde{T}_\alpha \xrightarrow{\text{sot}} \tilde{T}$ in $\mathcal{B}(\mathcal{H}^n)$ if and only if $T_{k,j,\alpha} \xrightarrow{\text{sot}} T_{k,j}$ for each k, j.

Answer. By the linearity of the topologies we may assume without loss of generality that $T = 0$.

(i) Suppose first that $\tilde{T}_\alpha \xrightarrow{\text{wot}} 0$. Fix $\xi, \eta \in \mathcal{H}$, and let $\tilde{\xi} \in \mathcal{H}^n$ be the vector with ξ in the j^{th} entry and zeros elsewhere, and let $\tilde{\eta}$ with η in the k^{th} entry and zeros elsewhere. Then

$$\langle T_{k,j,\alpha}\xi, \eta\rangle = \langle \tilde{T}_\alpha \tilde{\xi}, \tilde{\eta}\rangle \to 0.$$

This can be done for any $\xi, \eta \in \mathcal{H}$, so $T_{k,j,\alpha} \xrightarrow{\text{wot}} 0$.

Conversely, if $T_{k,j,\alpha} \xrightarrow{\text{wot}} 0$ for each k, j, fix $\tilde{\xi}, \tilde{\eta} \in \mathcal{H}^n$. Then

$$\langle \tilde{T}_\alpha \tilde{\xi}, \tilde{\eta}\rangle = \sum_{k,j=1}^{n} \langle T_{k,j,\alpha}\xi_j, \eta_k\rangle \to 0.$$

Therefore $\tilde{T}_\alpha \xrightarrow{\text{wot}} 0$.

(ii) Suppose first that $\tilde{T}_\alpha \xrightarrow{\text{sot}} 0$. Fix $\xi \in \mathcal{H}$, and let $\tilde{\xi} \in \mathcal{H}^n$ be the vector with ξ in the j^{th} entry and zeros elsewhere. Then

$$\|T_{k,j,\alpha}\xi\|^2 \leq \sum_{h=1}^n \left\|T_{h,j,\alpha}\xi_j\right\|^2 = \|\tilde{T}_\alpha\tilde{\xi}\|^2 \to 0.$$

This can be done for any $\xi \in \mathcal{H}$, so $T_{k,j,\alpha} \xrightarrow{\text{sot}} 0$.

Conversely, if $T_{k,j,\alpha} \xrightarrow{\text{sot}} 0$ for each k, j, fix $\tilde{\xi} \in \mathcal{H}^n$. Then, using Minkowsky's Integral Inequality (Proposition 2.8.20)

$$\|\tilde{T}_\alpha\tilde{\xi}\| = \Big(\sum_{k=1}^n \Big\|\sum_{j=1}^n T_{k,j,\alpha}\xi_j\Big\|\Big)^{1/2} \leq \Big(\sum_{k=1}^n \Big[\sum_{j=1}^n \|T_{k,j,\alpha}\xi_j\|\Big]^2\Big)^{1/2}$$

$$\leq \sum_{j=1}^n \Big(\sum_{k=1}^n \|T_{k,j,\alpha}\xi_j\|^2\Big)^{1/2} \to 0.$$

Therefore $\tilde{T}_\alpha \xrightarrow{\text{sot}} 0$.

(12.1.22) Let $\{P_j\}_{j\in J} \subset \mathcal{B}(\mathcal{H})$ be a net of pairwise orthogonal projections. Show that the series $P = \sum_j P_j$ converges sot, and $P = \bigvee_j P_j$.

Answer. The convergence is a particular case of Exercise 12.1.10 (or Proposition 12.1.10), and the fact that a sot-limit of projections is a projection (we did this explicitly right before Proposition 12.1.4; it also follows from Proposition 12.1.13). The limit is positive (and hence selfadjoint) because already a wot limit of positives is positive.

If $\xi \in \bigcup_j P_j\mathcal{H}$, then there exists j with $P_j\xi = \xi$. This gives us $P\xi = PP_j\xi = P_j\xi = \xi$. Thus $\overline{\text{span}}\bigcup_j P_j\mathcal{H} \subset P\mathcal{H}$. Conversely, if $\xi \in \Big(\bigcup_j P_j\mathcal{H}\Big)^\perp$, then $P_j\xi = 0$ for all j. This gives us $\sum_{j\in F} P_j\xi = 0$ for all finite F, and taking limit $P\xi = 0$. That is, $\Big(\overline{\text{span}}\bigcup_j P_j\mathcal{H}\Big)^\perp \subset (P\mathcal{H})^\perp$, which is the inclusion $P\mathcal{H} \subset \overline{\text{span}}\bigcup_j P_j\mathcal{H}$.

Alternatively, here is a direct argument to show the convergence. Let $\mathcal{V} = \{T : \|T\xi_k\| < 1,\ k = 1, \ldots, m\}$ be a sot-neighbourhood of 0. Form an orthonormal basis $\{\eta_{j,\ell}\}_{j\in J,\ \ell\in L_j} \cup \{\nu_r\}_r$ where each $\{\eta_{j,\ell}\}_{\ell\in L_j}$ is an orthonormal basis for $P_j\mathcal{H}$. For each $k = 1, \ldots, m$, by Parseval there exists a finite set $F_k \subset \bigcup_{j\in J,\ \ell\in L_j}(j,\ell)$ such that

$$\sum_{(j,\ell)\notin F_k} |\langle \xi_k, \eta_{j,\ell}\rangle|^2 < 1.$$

Put $F = \{j :\ \exists \ell, k,\ (j,\ell) \in F_k\}$. Then for any $F' \supset F$

$$\Big\| \sum_{j\notin F'} P_j\xi_k \Big\|^2 \leq \sum_{j\notin F} \|P_j\xi_k\Big\|^2 \leq \sum_{(j,\ell)\notin F_k} |\langle \xi_k, \eta_{j,\ell}\rangle|^2 < 1.$$

for all k. This means that $\sum_{j\in F'} P_j \in \mathcal{V}$. As this can be done for any sot-neighbourhood of 0, we have that the series converges.

(12.1.23) Let $\mathcal{A}$ be a C*-algebra, $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a non-degenerate representation, and $\{e_j\}$ an approximate unit for $\mathcal{A}$. Show that $\pi(e_j) \xrightarrow{\text{sot}} I_{\mathcal{H}}$.

Answer. The net $\{\pi(e_j)\}$ is a bounded monotone increasing net of positive operators. By Proposition 12.1.10 $Q = \lim_{\text{sot}} \pi(e_j)$ exists in $\mathcal{B}(\mathcal{H})$. For any $a \in \mathcal{A}$ and $\xi \in \mathcal{H}$ we have

$$\|(I_{\mathcal{H}} - Q)\pi(a)\xi\| = \lim_j \|\pi(a - e_j a)\xi\| \leq \lim_j \|a - e_j a\|\,\|\xi\| = 0.$$

Thus Q is the identity on the dense subspace $\pi(\mathcal{A})\mathcal{H}$. Hence $Q = I_{\mathcal{H}}$.

(12.1.24) Let $\{Q_j\} \subset \mathcal{B}(\mathcal{H})$ be a non-increasing net of projections. Show that $\bigwedge_j Q_j = \lim_{\text{sot}} Q_j$.

Answer. This follows directly from Exercise 12.1.22 and Proposition 10.5.9. We provide an ad-hoc argument below.

Applying Proposition 12.1.10 to the non-decreasing net $\{I_{\mathcal{H}} - Q_j\}$ we get that $Q' = \lim_{\text{sot}} I_{\mathcal{H}} - Q_j$ exists. Then, with $Q = I_{\mathcal{H}} - Q'$, we have that $Q_j \xrightarrow{\text{sot}} Q$. This Q is a projection; indeed, $Q^2 = Q$ by an application of Proposition 12.1.13, and $Q^* = Q$ because taking adjoints is wot-continuous. For any index j we have that $Q_k \leq Q_j$ for all $k \geq j$. It follows that $Q_j - Q = \lim_k Q_j - Q_k \geq 0$. So $Q \leq Q_j$ for all j, and therefore $Q \leq \bigwedge_j Q_j$. Conversely, if $\xi \in \bigcap_j Q_j\mathcal{H}$ then $Q_j\xi = \xi$ for all j. Taking the limit, $Q\xi = \xi$. As Q is bounded, it follows that $\overline{\bigcap_j Q_j\mathcal{H}}$ is invariant for Q, and this is $Q\bigwedge_j Q_j = \bigwedge_j Q_j$. Hence $\bigwedge_j Q_j \leq Q$, which proves the equality.

12.2. Multiplication Operators

(12.2.1) Prove Proposition 12.2.2.

Answer.

(i) We have

$$\|M_f h\|_2^2 = \int_X |f|^2 \, |h|^2 \, d\mu \leq \|f\|_\infty^2 \, \|h\|_2^2.$$

So M_f is bounded and $\|M_f\| \leq \|f\|_\infty$. Now fix $\varepsilon > 0$ and choose E with $0 < \mu(E) < \infty$ and $|f| > \|f\|_\infty - \varepsilon$ on E. Then

$$\|M_f 1_E\|_2^2 = \int_E |f|^2 \, d\mu \geq (\|f\|_\infty - \varepsilon)^2 \, \mu(E) = (\|f\|_\infty - \varepsilon)^2 \, \|1_E\|^2.$$

Thus $\|M_f\| \geq (\|f\|_\infty - \varepsilon)^2$. As ε was arbitrary, $\|M_f\| \geq \|f\|_\infty$ and therefore the equality holds.

(ii) We have

$$(M_f + M_g)h = M_f h + M_g h = fg + fh = (f+g)h = M_{f+g} h$$

and

$$M_f M_g h = f \, M_g h = fgh = M_{fg} h$$

for all h, so $M_f + M_g = M_{f+g}$ and $M_f M_g = M_{fg}$.

(iii) We have

$$\langle M_f g, h\rangle = \int_X f g \, \bar{h} \, d\mu = \int_X g \, \overline{\bar{f} \, h} \, d\mu = \langle g, M_{\bar{f}} h\rangle.$$

So $M_f^* = M_{\bar{f}}$. If $M_f = M_f^*$, then $0 = M_f - M_{\bar{f}} = M_{f-\bar{f}}$, so $\|f - \bar{f}\|_\infty = 0$ and thus $f = \bar{f}$ a.e.; hence f is real a.e.

(iv) Suppose that $\lambda \in \operatorname{ess\,ran} f$. Then for any given $\varepsilon > 0$ there exists E with $\mu(E) > 0$ and $|f - \lambda| < \varepsilon$ on E. In particular, for each $n \in \mathbb{N}$ we have

$$\mu(f^{-1}(B_{1/n}(\lambda))) > 0.$$

The sets $f^{-1}(B_{1/n}(\lambda))$ decrease as n increases. Fix $X_0 \subset X$ with $0 < \mu(X_0) < \infty$ and $\mu(X_0 \cap f^{-1}(B_1(\lambda))) > 0$ (X_0 exists because μ is semifinite). Let

$$F_n = X_0 \cap f^{-1}(B_{1/n}(\lambda)), \qquad g_n = \frac{1}{\mu(F_n)^{1/2}} \, 1_{F_n}.$$

Then $\|g_n\|_2 = 1$, and

$$\|(M_f - \lambda I)g_n\|_2^2 = \frac{1}{\mu(F_n)} \int_{F_n} |f - \lambda|^2 \leq \frac{1}{n^2} \to 0.$$

Hence $M_f - \lambda I$ is not bounded below, which implies it is not invertible. So $\lambda \in \sigma(M_f)$.

Conversely, if $\lambda \notin \operatorname{ess\,ran} f$ then there exists $\varepsilon > 0$ such that

$$\mu(f^{-1}(B_\varepsilon(\lambda))) = 0.$$

Let $G = f^{-1}(B_\varepsilon(\lambda))$ and $g = \frac{1}{f-\lambda} 1_{X\setminus G}$. On $X\setminus G$ we have that $|f-\lambda| \geq \varepsilon$, so $\|g\|_\infty \leq \frac{1}{\varepsilon}$. And

$$(M_f - \lambda I)M_g h = |f - \lambda| \frac{1}{|f - \lambda|} 1_{X\setminus G}\, h = 1_{X\setminus G}\, h,$$

which is equal to h in $L^2(X)$ since $\mu(G) = 0$. As multiplier operators commute, M_g is the inverse of $M_f - \lambda I$ and so $\lambda \notin \sigma(M_f)$.

(12.2.2) Show that the semifinite hypothesis is crucial for (iv) to hold in Proposition 12.2.2.

Answer. Let $X = \{1, \infty\}$ with $\mu(\{1\}) = 1$ and $\mu(\{\infty\}) = \infty$. Let $f = \alpha\, 1_{\{1\}} + \beta\, 1_{\{\infty\}} \in L^\infty(X)$ with $\alpha \neq \beta$. Given $g \in L^2(X)$, since $\|g\|_2 < \infty$ we have $g(\infty) = 0$. Then $fg = \alpha g$, so $\sigma(M_f) = \{\alpha\} \subsetneq \{\alpha, \beta\} = \operatorname{ess\,ran} f$.

(12.2.3) Show that $\lambda \in \sigma(M_f)$ is an eigenvalue with multiplicity m if and only if $\{f = \lambda\}$ consists of exactly m atoms.

Answer. We work first with the case where $m < \infty$.

Suppose $\dim\ker(M_f - \lambda I) = m$. This subspace consists precisely of those functions $h \in L^2(X)$ such that $(f - \lambda)h = 0$. If $\{f = \lambda\} = \bigcup_{j=1}^{m+1} E_j$ with $\mu(E_j) > 0$ for all j and the E_j pairwise disjoint, then $\{1_{E_j}\}_{j=1}^{m+1}$ would be $m+1$ linearly independent functions in $\ker(M_f - \lambda I)$, contradicting that $\dim\ker(M_f - \lambda I) = m$. We cannot have any E_j infinitely divisible into sets of positive measure, because this would give us $\dim\ker(M_f - \lambda I) = \infty$. So there is a maximal partition $\ker(M_f - \lambda I) = \bigcup_{j=1}^{r} E_j$ with each E_j an atom for μ. If $r < m$ we also get a contradiction, because we cannot distinguish m linearly independent functions by using $r < m$ points (there would be two functions that agree at every point). We have shown that $\{f = \lambda\}$ consists of precisely m atoms.

Conversely, if $\{f=\lambda\}$ consists of precisely m atoms $\{E_j\}$, then $M_f 1_{E_j} = \lambda 1_{E_j}$, so $\dim\ker(M_f-\lambda I)\geq m$. As before the dimension cannot be more than m, because we would have $m+1$ linearly independent functions to be separated by m points. Thus $\dim\ker(M_f-\lambda I)=m$.

Now consider the case $m=\infty$. We use that

$$\ker(M_f-\lambda)=\{h\in L^2(X):\ \operatorname{ess\,sup} h\subset\{f=\lambda\}\}.$$

If $\dim\ker(M_f-\lambda I)=\infty$, then there are infinitely many linearly independent functions in $\ker(M_f-\lambda I)$, making it impossible for $\{f=\lambda\}$ to have finitely many atoms. Conversely, if $\{f=\lambda\}$ is arbitrarily divisible into partitions of sets with positive measure, we get that $\dim\ker(M_f-\lambda I)=\infty$.

12.3. Commutants and Double Commutants

(12.3.1) Show that $\mathcal{B}(\mathcal{H})'=\mathcal{K}(\mathcal{H})'=\mathbb{C}\,I_{\mathcal{H}}$, and $(\mathbb{C}\,I_{\mathcal{H}})'=\mathcal{B}(\mathcal{H})$.

Answer. We know that $\mathbb{C}\,I_{\mathcal{H}}\subset\mathcal{B}(\mathcal{H})'$. Suppose that $T\in\mathcal{B}(\mathcal{H})'$. In particular T commutes with all rank-one operators. So for any $\xi,\eta\in\mathcal{H}$ we have $T\xi\eta^*=\xi\eta^*T$. Applied to η with $\|\eta\|=1$, this gives us

$$T\xi=\langle\eta,\eta\rangle\,T\xi=T\xi\eta^*\eta=\xi\eta^*T\eta=\langle T\eta,\eta\rangle\xi.$$

If $T=0$ then $T=\lambda\,I$ with $\lambda=0$; and if $T\neq0$, then there exists ξ_0 with $T\xi_0\neq0$, which implies that $\lambda=\langle T\eta,\eta\rangle\neq0$, and thus $T\xi=\lambda\xi$ for all $\xi\in\mathcal{H}$. That is $T\in\mathbb{C}\,I_{\mathcal{H}}$.

The argument above only used finite-rank operators (in fact, rank-one), so it also proves that $\mathcal{K}(\mathcal{H})'=\mathbb{C}\,I_{\mathcal{H}}$.

The equality $(\mathbb{C}\,I_{\mathcal{H}})'=\mathcal{B}(\mathcal{H})$ is just the fact that $\lambda I_{\mathcal{H}}$ commutes with all $T\in\mathcal{B}(\mathcal{H})$, so $(\mathbb{C}I_{\mathcal{H}})'$ is a big as it can be.

(12.3.2) Let $\mathcal{H}=\ell^2(\mathbb{N})$. Let $\mathcal{A}=\{M_a:\ a\in\ell^\infty(\mathbb{N})\}$ be the algebra of multipliers. Show that $\mathcal{A}'=\mathcal{A}$.

Answer. Since $M_aM_b=M_{ab}=M_{ba}=M_bM_a$ for all a,b, $\mathcal{A}$ is abelian and so $\mathcal{A}\subset\mathcal{A}'$. Now let $T\in\mathcal{A}'$. Write $\{e_n\}\subset\mathcal{H}$ for the canonical basis, both as elements of $\mathcal{H}$ and of $\ell^\infty(\mathbb{N})$. Then we can consider the multiplication operators $\{M_{e_n}\}$. The operator M_{e_n} is precisely the projection onto $\mathbb{C}\,e_n$, since $M_{e_n}a = a_n e_n$. Then, since $T \in \mathcal{A}'$,

$$\begin{aligned}\langle Te_j, e_k\rangle &= \langle TM_{e_j}e_j, M_{e_k}e_k\rangle = \langle M_{e_k}TM_{e_j}e_j, e_k\rangle \\ &= \langle M_{e_k}M_{e_j}Te_j, e_k\rangle = \delta_{k,j}\langle Te_k, e_k\rangle \\ &= \big\langle \langle Te_j, e_j\rangle\, e_j, e_k\big\rangle.\end{aligned}$$

Therefore, denoting $t_n = \langle Te_n, e_n\rangle$, we have $Te_n = t_n e_n$ and hence

$$T\xi = \sum_n \langle \xi, e_n\rangle Te_n = \sum_n t_n \langle \xi, e_n\rangle\, e_n.$$

That is, $T = M_a$ with $a = (t_n)$. And $a \in \ell^\infty(\mathbb{N})$ since

$$|t_n| = |\langle Te_n, e_n\rangle| \leq \|T\|.$$

So $T \in \mathcal{A}$.

(12.3.3) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Show that

$$\mathcal{Z}(\mathcal{M}') = \mathcal{Z}(\mathcal{M}).$$

Conclude that $\mathcal{M}$ is a factor if and only if $\mathcal{M}'$ is a factor.

Answer. We have

$$\mathcal{Z}(\mathcal{M}') = \mathcal{M}' \cap \mathcal{M}'' = \mathcal{M}' \cap \mathcal{M} = \mathcal{Z}(\mathcal{M}).$$

When either of $\mathcal{M}$ or $\mathcal{M}'$ is a factor, we have $\mathcal{Z}(\mathcal{M}) = \mathbb{C}\, I_h = \mathcal{Z}(\mathcal{M}')$, so the other one is a factor.

(12.3.4) Let $D \in M_n(\mathbb{C})$ be diagonal with all diagonal entries distinct. Show $\{D\}' = \{D\}'' = \{E \in M_n(\mathbb{C}) :\ \text{diagonal}\}$.

Answer. We have $D = \sum_{k=1}^n d_k E_{kk}$. Suppose that $TD = DT$. Writing $T = \sum_{k,j=1}^n t_{kj}E_{kj}$, we have

$$TD = \sum_{h,k,j=1}^n t_{kj}d_h E_{kj}E_{hh} = \sum_{k,j=1}^n d_j t_{kj}E_{kj}$$

and

$$DT = \sum_{h,k,j=1}^n t_{kj}d_h E_{hh}E_{kj} = \sum_{k,j=1}^n d_k t_{kj}E_{kj}.$$

Comparing entries we see that for each k, j we have

$$d_j t_{kj} = d_k t_{kj}.$$

If $k \neq j$, from $d_j \neq d_k$ we conclude that $t_{kj} = 0$. Thus the only nonzero entries of T are within those with $k = j$; that is, T is diagonal.

As for the double commutant, we have that

$$\{D\}' = \mathcal{A} = \{E \in M_n(\mathbb{C}) : \text{ diagonal } \}$$

is an abelian algebra, so $\mathcal{A} \subset \mathcal{A}'$. If $T \in \mathcal{A}'$, in particular $TD = DT$, so by the first part of the argument $T \in \mathcal{A}$. So $\mathcal{A}' \subset \mathcal{A}$, showing that $\{D\}'' = \mathcal{A}' = \mathcal{A}$.

(12.3.5) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a finite-dimensional C*-algebra. Show that $\mathcal{A}$ is a von Neumann algebra.

Answer. By Theorem 5.4.16 the restriction of the sot topology to $\mathcal{A}$ agrees with the norm topology. Thus $\mathcal{A}$ is sot-complete and therefore a von Neumann algebra.

(12.3.6) Prove Proposition 12.3.2.

Answer.

(i) For any $T \in \mathcal{M}$, $\lambda \in \mathbb{C}$, $T(\lambda I_{\mathcal{H}}) = \lambda T = (\lambda I_h)T$.

(ii) The fact that $\mathcal{A}$ is abelian means that each of its elements is in the commutant. Conversely, if $\mathcal{A} \subset \mathcal{A}'$ this implies that every element of $\mathcal{A}$ commutes with every element of $\mathcal{A}$, so $\mathcal{A}$ is abelian.

(iii) Suppose that $\{S_j\} \subset \mathcal{M}'$ is a net and $S_j \to S$ wot. For any $T \in \mathcal{M}$ and $\xi, \eta \in \mathcal{H}$,

$$\begin{aligned}\langle TS\xi, \eta\rangle &= \langle S\xi, T^*\eta\rangle = \lim_j \langle S_j\xi, T^*\eta\rangle \\ &= \lim_j \langle TS_j\xi, \eta\rangle = \lim_j \langle S_jT\xi, \eta\rangle = \langle ST\xi, \eta\rangle.\end{aligned}$$

As ξ, η were arbitrary, $TS = ST$. So $S \in \mathcal{M}'$ and $\mathcal{M}'$ is wot-closed. As wot is weaker than sot, it is also sot-closed. If $S_1, S_2 \in \mathcal{M}'$ and $T \in \mathcal{M}$, then $(S_1 + S_2)T = S_1T + S_2T = TS_1 + TS_2 = T(S_1 + S_2)$. Similarly, $S_1S_2T = S_1TS_2 = TS_1S_2$. So $\mathcal{M}'$ is an algebra.

(iv) If $T \in \mathcal{M}'$, this means that $TS = ST$ for all $S \in \mathcal{M}$. As $\mathcal{N} \subset \mathcal{M}$, we have $TS = ST$ for all $S \in \mathcal{N}$, so $T \in \mathcal{N}'$. That is, $\mathcal{M}' \subset \mathcal{N}'$.

(v) If $T \in \mathcal{M}'$ and $S \in \mathcal{M}$, then $S^* \in \mathcal{M}$ and we have $TS^* = S^*T$. Taking adjoints, $ST^* = T^*S$, so $T^* \in \mathcal{M}'$. That is, $\mathcal{M}'$ contains its adjoints.

(vi) Is $T \in \mathcal{M}$ and $S \in \mathcal{M}'$, then $ST = TS$, so $T \in \mathcal{M}''$.

(vii) By the above, $\mathcal{M}' \subset \mathcal{M}'''$. Now, if $T \in \mathcal{M}'''$ and $S \in \mathcal{M}$, then $S \in \mathcal{M}''$ and so $ST = ST$. So $\mathcal{M}''' \subset \mathcal{M}'$ and it follows that $\mathcal{M}''' = \mathcal{M}'$.

(12.3.7) Prove Proposition 12.3.4.

Answer. We identify $\mathcal{B}(\mathcal{H}^{(n)})$ with $M_n(\mathcal{B}(\mathcal{H}))$. We use the notation $\xi \otimes e_j$ to denote the element of $\mathcal{H}^{(n)} = \bigoplus_{k=1}^n \mathcal{H}$ that has ξ in the j^{th} entry and zeroes elsewhere. We use π_j to denote the projection onto the j^{th} entry. In particular $\pi_j(\xi \otimes e_j) = \xi$, and $\pi_k(\xi \otimes e_j) = 0$ if $k \neq j$.

Given $T \in \mathcal{B}(\mathcal{H}^{(n)})$ and $X \in \mathcal{M}^{(n)}$ (seen as a diagonal matrix in $M_n(\mathcal{B}(\mathcal{H}))$), $\xi \in \mathcal{H}$,

$$\pi_k TX(\xi \otimes e_j) = \pi_k T(X_j \xi \otimes e_j) = T_{kj} X_j \xi,$$

and

$$\pi_k XT(\xi \otimes e_j) = \pi_k X(T_{hj}\xi)_h = \pi_k(X_h T_{hj} \xi)_h = X_k T_{kj} \xi.$$

As the two equalities above can be obtained for any k, j, and any $\xi \in \mathcal{H}_j$, we get that $TX = XT$ if and only if $T_{kj} X_j = X_k T_{kj}$ for all k, j.

Next suppose that $S \in (\mathcal{M}^{(n)})''$. Given $T_1, \ldots, T_n \in \mathcal{M}'$ we can form $\tilde{T} = \bigoplus_k T_k \in \mathcal{B}(\mathcal{H}^{(n)})$; by the previous paragraph, $\tilde{T} \in (\mathcal{M}^{(n)})'$. Then $S\tilde{T} = \tilde{T}S$. Component wise, this is $S_{kj} T_j = T_k S_{kj}$. For $j \neq k$, we may take $T_j = I$, $T_k = 0$ to conclude that $S_{kj} = 0$; so S is diagonal. And $S_{kk} T_k = T_k S_{kk}$, so $S_{kk} \in (\mathcal{M}')' = \mathcal{M}''$. The converse is trivial to check.

(12.3.8) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{P} \subset \mathcal{M}$ a set of projections. Show that $\bigvee \mathcal{P}$ and $\bigwedge \mathcal{P}$ are respectively the supremum and infimum of $\mathcal{P}$.

Answer. For any $P \in \mathcal{P}$, since $\bigcup_{Q \in \mathcal{P}} Q\mathcal{H} \supset P\mathcal{H}$, we have that $\bigvee \mathcal{P} \geq P$. So $\bigvee \mathcal{P}$ is an upper bound. Now suppose that $Q \in \mathcal{M}$ is a projection and $Q \geq P$ for all $P \in \mathcal{P}$. Then $Q\mathcal{H} \supset \bigcup_{P \in \mathcal{P}} P\mathcal{H}$, so $Q \geq \bigvee \mathcal{P}$, showing that $\bigvee \mathcal{P}$ is the least upper bound of $\mathcal{P}$. The argument for the infimum is entirely similar.

(12.3.9) Given an alternative proof of Corollary 12.3.10 by using an approximate unit.

Answer. Since $\mathcal{M}$ is a C*-algebra it contains an approximate unit $\{E_j\}$ (Theorem 11.4.4). Since the approximate unit is monotone and bounded by definition, Proposition 12.1.10 shows that $E = \lim_{\text{sot}} E_j \in \mathcal{M}$ exists. For any $X \in \mathcal{M}$ and $\xi \in \mathcal{H}$,

$$\|(EX - X)\xi\| = \lim_j \|(E_jX - X)\xi\| \leq \|\xi\| \lim_j \|E_jX - X\| = 0.$$

Thus $EX = X$ for all $X \in \mathcal{M}$. The same argument works to show that $XE = X$ (or, we can use $EX = X$ for $X \geq 0$, take adjoints, and use that positive elements span the algebra). So $E = I_{\mathcal{M}}$.

(12.3.10) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate C*-algebra and $n \in \mathbb{N}$. Consider $M_n(\mathcal{A}) \subset \mathcal{B}(\mathcal{H}^n)$ and show that

$$M_n(\mathcal{A})' = \{X \otimes I_n : \ X \in \mathcal{A}'\}, \tag{12.4}$$

$$\{X \otimes I_n : \ X \in \mathcal{A}\}' = M_n(\mathcal{A}'), \tag{12.5}$$

$$M_n(\mathcal{A})'' = M_n(\mathcal{A}''). \tag{12.6}$$

Answer. Let $X \in \mathcal{A}'$ and $Y \in M_n(\mathcal{A})$. Then, writing $\tilde{X} = X \otimes I_n$,

$$(\tilde{X}Y)_{kj} = \sum_{h=1}^n (\tilde{X})_{kh} Y_{hj} = XY_{kj} = Y_{kj}X = \sum_{h=1}^n Y_{kh}(\tilde{X})_{hj} = (Y\tilde{X})_{kj}.$$

So $\tilde{X}Y = Y\tilde{X}$, and therefore $\{X \otimes I_n : \ X \in \mathcal{A}'\} \subset M_n(\mathcal{A})'$. Conversely, suppose that $\tilde{X} \in M_n(\mathcal{A})'$. Fix k, j and $A \in \mathcal{A}$, and consider the matrix $\tilde{A} \in M_n(\mathcal{A})$ that has $\tilde{A}_{kj} = A$ and zeros elsewhere. We have

$$(\tilde{X}\tilde{A})_{rs} = \sum_{h=1}^n X_{rh}\,\tilde{A}_{hs} = \delta_{j,s}\,X_{rk}A$$

and

$$(\tilde{A}\,\tilde{X})_{rs} = \sum_{h=1}^n \tilde{A}_{rh}\,X_{hs} = \delta_{r,k}\,AX_{js}.$$

These two expressions are be equal for any choice of r, s, k, j and A. If we take $r = k$ and $j = s$, we get that $X_{kk}A = AX_{jj}$ for all k, j. In particular $X_{kk} \in \mathcal{A}'$ for all k. If we take an approximate unit $\{E_\ell\}$ in $\mathcal{A}$, we have $I_{\mathcal{H}} = \lim_{\text{sot}} E_\ell$ by Corollary 12.3.10 and the fact that $\mathcal{A}$ is non-degenerate.

Then

$$X_{kk} = X_{kk} I_{\mathcal{H}} = \lim_{\text{sot}} X_{kk} E_\ell = \lim_{\text{sot}} E_\ell X_{jj} = X_{jj}.$$

Thus the diagonal of $\tilde{X}$ is constant, made out of elements of $\mathcal{A}'$. When $k \neq j$, choose $s = r = k$. Then the equality $\delta_{j,s} X_{rk} A = \delta_{r,k} A X_{js}$ becomes $0 = AX_{jk}$. Using again the approximate unit as above, we get that $X_{jk} = 0$. Thus $\tilde{X} = X_{11} \otimes I_n$, proving (12.4).

For (12.5), applying (12.4) to $\mathcal{A}''$, taking commutants, and using that $\mathcal{A}''' = \mathcal{A}'$,

$$\{X \otimes I_n : \ X \in \mathcal{A}''\}' = M_n(\mathcal{A}')''$$

By Exercise 12.1.21 $X_\alpha \otimes I_n \xrightarrow{\text{wot}} X \otimes I_n$ if and only if $X_\alpha \xrightarrow{\text{wot}} X$. Then, using the Double Commutant Theorem (12.3.5)

$$\begin{aligned}\{X \otimes I_n : \ X \in \mathcal{A}''\} &= \{X \otimes I_n : \ X \in \overline{\mathcal{A}}^{\text{wot}}\} = \overline{\{X \otimes I_n : \ X \in \mathcal{A}\}}^{\text{wot}} \\ &= \{X \otimes I_n : \ X \in \mathcal{A}\}''.\end{aligned}$$

We also know from Exercise 12.1.21 that if $X \in M_n(\mathcal{A}')$ then $X_\alpha \xrightarrow{\text{wot}} X$ if and only if $(X_\alpha)_{kj} \xrightarrow{\text{wot}} X_{kj}$ for all k, j. Therefore $M_n(\mathcal{A}')$ is a von Neumann algebra and

$$\begin{aligned}\{X \otimes I_n : \ X \in \mathcal{A}\}' &= (\{X \otimes I_n : \ X \in \mathcal{A}\}'')' = \{X \otimes I_n : \ X \in \mathcal{A}''\}' \\ &= M_n(\mathcal{A}')'' = M_n(\mathcal{A}'),\end{aligned}$$

which is (12.5).

Finally,

$$M_n(\mathcal{A})'' = \{X \otimes I_n : \ X \in \mathcal{A}'\}' = M_n((\mathcal{A}')') = M_n(\mathcal{A}'').$$

(12.3.11) Let $X \subset \mathcal{B}(\mathcal{H})$. Prove that $W^*(X) = (X \cup X^*)''$, and show by example that it is possible to have $X'' \subsetneq W^*(X)$.

Answer. We have $X \subset W^*(X)$ by definition and, since $W^*(X)$ is a C*-algebra, we also have $X^* \subset W^*(X)$. Thus $X \cup X^* \subset W^*(X)$ and so $(X \cup X^*)'' \subset W^*(X)'' = W^*(X)$ by Proposition 12.3.2 and Theorem 12.3.5. Conversely, using again Proposition 12.3.2 and Theorem 12.3.5 we have that $(X \cup X^*)''$ is a von Neumann algebra that contains X, hence $W^*(X) \subset (X \cup X^*)''$.

For an example, let $\mathcal{H} = \mathbb{C}^2$ and let X consist of the single element $S = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. It is easy to check that X' consists of the matrices of the form $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$. These in turn can be seen as $aI_2 + bS$. As everything commutes with the identity, it follows that $X'' = \{aI_2 + bS : a, b \in \mathbb{C}\}$. Which is not a $*$-algebra, since it does not contain S^*. A straightforward computation shows that $\{S, S^*\}' = \mathbb{C}I_2$. Then $\{S, S^*\}'' = (\mathbb{C}\, I_2)'' = M_2(\mathbb{C})$. Thus $X'' \subsetneq W^*(X) = M_2(\mathbb{C})$.

(12.3.12) Show that $\mathcal{S}$ and $\mathcal{S}^b$, as in Lemma 12.3.16, are closed under uniform limits.

Answer. Suppose that $\{f_n\} \subset \mathcal{S}$ and $f_n \to f$ uniformly. Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})^{\text{sa}}$ with $T_j \xrightarrow{\text{sot}} T$. Fix $\varepsilon > 0$. By hypothesis there exists n_0 such that $\|f_n - f\|_\infty < \varepsilon$ for all $n > n_0$. Fix $\xi \in \mathcal{H}$. Then, for $n \geq n_0$,

$$\begin{aligned}\|(f(T_j) - f(T))\xi\| &\leq \|(f(T_j) - f_n(T_j))\xi\| + \|(f_n(T_j) - f_n(T))\xi\| \\ &\qquad + \|(f_n(T) - f(T))\xi\| \\ &\leq 2\|f_n - f\|_\infty \|\xi\| + \|(f_n(T_j) - f_n(T))\xi\| \\ &\leq 2\varepsilon \|\xi\| + \|(f_n(T_j) - f_n(T))\xi\|.\end{aligned}$$

As $f_n \in \mathcal{S}$, we get that $\limsup_j \|(f(T_j) - f(T))\xi\| \leq 2\varepsilon \|\xi\|$. So by the Limsup Routine the limit exists and is zero. That is, $f(T_j) \xrightarrow{\text{sot}} f(T)$. The computation is $\mathcal{S}^b$ is exactly the same.

(12.3.13) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\mathcal{M}_0 \subset \mathcal{M}$ a subspace. Show that the following statements are equivalent:

(i) $\mathcal{M}_0$ is wot-dense in $\mathcal{M}$;

(ii) $\mathcal{M}_0$ is sot-dense in $\mathcal{M}$;

(iii) $\mathcal{M}_0$ is σ-weak dense in $\mathcal{M}$.

Answer. As $\mathcal{M}_0$ is convex we have $\overline{\mathcal{M}_0}^{\text{sot}} = \overline{\mathcal{M}_0}^{\text{wot}}$ by Corollary 12.1.3.

If $\mathcal{M}_0$ is σ-weak dense in $\mathcal{M}$, then it is wot dense as the wot is weaker. Conversely, suppose that $\mathcal{M}_0$ is sot dense in $\mathcal{M}$. Given $T \in \mathcal{M}$ there exists a net $\{T_j\} \subset \mathcal{M}_0$ with $T_j \xrightarrow{\text{sot}} T$. By Kaplansky's Density Theorem we may assume that the net $\{T_j\}$ is bounded. But then $T_j \xrightarrow{\text{wot}} T$ and bounded, so $T_j \xrightarrow{\sigma\text{-weak}} T$ by Lemma 12.1.21.

12.4. The Spectral Theorem

(12.4.1) Let $T \in \mathcal{K}(\mathcal{H})$ be normal. Show that Theorem 10.6.12 is a particular case of Theorem 12.4.4.

Answer. By Theorem 12.4.4 there exists a unique Borel measure μ_T on $\sigma(T)$ such that

$$T = \int_{\sigma(T)} \lambda \, d\mu_T(\lambda).$$

Since the identity function is 0 at 0 we can consider the integral over $\sigma(T)\setminus\{0\}$. From Theorem 9.6.13 we know that $\sigma(T)$ is either finite or a sequence $\{\lambda_k\}$ that converges to zero. So we can write $\sigma(T) \setminus \{0\}$ as a finite or countable disjoint union of singletons. There are no convergence issues for

$$\left\| \int_{\sigma(T)\cap[0,\varepsilon]} \lambda \, d\mu_T \right\| \leq \varepsilon \|\mu_T([0,\varepsilon])\| = \varepsilon.$$

As $\sigma(T) \cap (\varepsilon, \infty)$ is finite for all $\varepsilon > 0$, this shows that

$$T = \sum_k \int_{\{\lambda_k\}} \lambda \, d\mu_T(\lambda) = \sum_k \lambda_k \mu_T(\{\lambda_k\}).$$

The operators $P_k = \mu_T(\{\lambda_k\})$ are pairwise orthogonal projections. Since $TP_k = \lambda_k P_k$, each P_k is finite-rank for otherwise we would have an infinite-dimensional eigenspace, contradicting Theorem 9.6.13.

(12.4.2) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Show that there exists $S \in \mathcal{B}(\mathcal{H})$, selfadjoint, and a continuous $f : \sigma(S) \to \mathbb{C}$ such that $T = f(S)$. *(Hint: Proposition 7.6.7)*

Answer. By Theorem 7.6.5 there exists $f : \mathcal{C} \to \sigma(T)$, continuous and surjective. And by Proposition 7.6.7 there exists $g : \sigma(T) \to \mathcal{C}$ Borel with $f \circ g = \mathrm{id}_{\sigma(T)}$. By the Spectral Theorem (Theorem 12.4.4) there exists a spectral measure μ_T such that

$$T = \int_{\sigma(T)} \lambda \, d\mu_T(\lambda).$$

We define

$$S = \int_{\sigma(T)} g(\lambda)\, d\mu_T(\lambda).$$

This S is well-defined because g is bounded Borel. And S is selfadjoint because g is real-valued. We have

$$f(S) = \int_{\sigma(T)} (f \circ g)(\lambda)\, d\mu_T(\lambda) = \int_{\sigma(T)} \lambda\, d\mu_T(\lambda) = T.$$

(12.4.3) Expanding on the ideas of Example 12.4.7, show that if $g \in L^\infty[0,1]$, then $\mu_{M_g}(E) = M_{1_{g^{-1}(E)}}$.

Answer. If $\{f_r\}$ is a bounded sequence of polynomials in $L^\infty[0,1]$ with $f_r \to 1_E$ as in Example 12.4.7, then $f_r(M_g) = M_{f_r(g)}$ and

$$\begin{aligned}\langle \mu_{M_g}(E)h, h\rangle &= \lim_r \int_{[0,1]} f_r(g)\, |h|^2\, dm = \lim_r \int_{[0,1]} (f_r \circ g)\, |h|^2\, dm \\ &= \int_{[0,1]} (1_E \circ g)\, |h|^2\, dm\end{aligned}$$

So

$$\mu_{M_g}(E) = M_{1_E \circ g} = M_{1_{g^{-1}(E)}}.$$

(12.4.4) Show that the extreme points in the convex set $\mathcal{B}(\mathcal{H})_1^+$ of positive operators with norm at most 1, are precisely the projections.

Answer. We know that projections are extreme from Exercise 10.5.5. Now suppose that $T \geq 0$, $\|T\| \leq 1$, and T is not a projection. By Exercise 10.5.4 there exists $\lambda_0 \in (0,1) \cap \sigma(T)$. And by Corollary 12.4.14, $\mu_T(B_r(\lambda_0)) \neq 0$ for all $r > 0$. Fix $r = (1-\lambda_0)/3$ and let $E = \mathbb{R} \setminus B_r(\lambda_0)$. We have

$$\begin{aligned}\|T\, 1_{E^c}\| &= \sup\{|\lambda| :\ \lambda \in E^c\} \\ &= \sup\{|\lambda| :\ \lambda \in B_r(\lambda_0)\} \\ &\leq \lambda_0 + r < 1 - r.\end{aligned}$$

Let

$$T_1 = T + r\, 1_{E^c}(T), \qquad T_2 = T - r\, 1_{E^c}(T).$$

We have $\|T\,1_{E^c}(T) + r\,1_{E^c}(T)\| \leq (1-r) + r = 1$. Then

$$\|T_1\| = \max\left\{\|T\,1_E(T)\|, \|T\,1_{E^c}(T) + r\,1_{E^c}(T)\|\right\} \leq 1.$$

Similarly $\|T_2\| \leq 1$, and then $T = \frac{1}{2}(T_1 + T_2)$ is not extreme.

(12.4.5) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\lambda_0 \in \mathbb{C}$. Consider the extension of μ_T to all of $\mathbb{C}$ as in Exercise 2.3.8. Show that the following statements are equivalent:

(i) T is compact;

(ii) for every $\lambda_0 \in \sigma(T) \setminus \{0\}$ there exists $r > 0$ such that $\mu_T(B_r(\lambda_0))$ is a finite-rank projection.

Answer. (i) $\Longrightarrow$ (ii) We have that T is compact, and $\lambda_0 \in \sigma(T) \setminus \{0\}$. By Theorem 9.6.13 there exists $r > 0$ with $B_r(\lambda_0) \cap \sigma(T) = \{\lambda_0\}$. Recall that $\mu_T(B_r(\lambda_0)) = 1_{B_r(\lambda_0)}(T)$. If $\xi \in 1_{B_r(\lambda_0)}(T)\mathcal{H}$, as $t\,1_{B_r(\lambda_0)}(t) = \lambda_0\,1_{B_r(\lambda_0)}(t)$ on $\sigma(T)$, by functional calculus $T\,1_{B_r(\lambda_0)}(T) = \lambda_0\,1_{B_r(\lambda_0)}(T)$. Then

$$T\xi = T\,1_{B_r(\lambda_0)}(T)\xi = \lambda_0\,1_{B_r(\lambda_0)}(T)\xi = \lambda_0\xi.$$

Therefore $\mu_T(B_r(\lambda_0))\mathcal{H} \subset \ker(T - \lambda_0 I)$, which is finite-dimensional.

(ii) $\Longrightarrow$ (i) Combining that $A_n = \{\lambda \in \sigma(T) : \ \lambda \geq \frac{1}{n}\}$ is compact with Corollary 12.4.14 we get $\lambda_1, \ldots, \lambda_m \in A_n$ and $r_1, \ldots, r_m > 0$ with $A_n \subset \bigcup_{j=1}^m B_{r_j}(\lambda_j)$ and $\mu_T(B_{r_j}(\lambda_j)) \neq 0$ for $j = 1, \ldots, n$. Let $B_1, \ldots, B_m$ with $B_j \subset B_{r_j}(\lambda_j)$, pairwise disjoint and with union $\bigcup_{j=1}^m B_{r_j}(\lambda_j)$. Namely, we take $B_1 = B_{r_1}(\lambda_1)$, $B_2 = B_{r_2}(\lambda_2) \setminus B_1$, etc. Then

$$\mu_T(A_n) \leq \mu_T\Big(\bigcup_j B_{r_j}(\lambda_j)\Big) = \mu_T\Big(\bigcup_j B_j\Big) = \sum_{j=1}^m \mu_T(B_j).$$

As the B_j are pairwise disjoint the projections $\mu_T(B_j)$ are pairwise orthogonal, and each is below the corresponding $\mu_T(B_{r_j}(\lambda_j))$, so finite-rank. Therefore $\mu_T(A_n)$ is finite-rank. So $T\,\mu_T(A_n)$ is finite-rank. And since $|t - t\,1_{A_n}(t)| < \frac{1}{n}$ on $\sigma(T)$, $\|T - T\,\mu_T(A_n)\| < \frac{1}{n}$, so $T\,\mu_T(A_n) \to T$. Hence T is a norm-limit of finite-rank operators, and thus compact by Proposition 10.6.4.

(12.4.6) Let $\mathcal{M}$ be a von Neumann algebra and $T \in \mathcal{M}^+$ nonzero. Show that there exists a nonzero projection $P \in \mathcal{M}$ that commutes with T and $\lambda > 0$ such that $TP \geq \lambda P$.

Answer. Since $T \geq 0$ and nonzero, there exists nonzero λ with $2\lambda \in \sigma(T)$. So $\lambda > 0$. Let $P = \mu_T((\lambda, \infty))$. We have $P \neq 0$, for otherwise $T = T(I - P)$ and then

$$\|T\| = \|T(I - P)\| = \left\| \int_{|t| \leq \lambda} t \, d\mu_T(t) \right\| \leq \lambda,$$

contradicting that $2\lambda \in \sigma(T)$. The inequality $t\,1_{(\lambda,\infty)}(t) \geq \lambda 1_{(\lambda,\infty)}$ and functional calculus then give us

$$TP \geq \lambda P.$$

(12.4.7) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Show that the construction of μ_T in the proof of the Spectral Theorem gives $\mu_T(\sigma(T)) = I_{\mathcal{H}}$.

Answer. The function $1_{\sigma(T)}$ equals 1 on $\sigma(T)$ (this not deep!). In (12.9), we can take $f_r = 1$ for all r, and $f_r(T) = I_{\mathcal{H}}$. Hence $\langle \mu_T(\sigma(T))\xi, \xi\rangle = \langle \xi, \xi \rangle$ for all $\xi \in \mathcal{H}$, and using polarization we get $\mu_T(\sigma(T)) = I_{\mathcal{H}}$.

(12.4.8) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Show that $1_{\{0\}}(T)$ is the projection onto $\ker T$, and that $1_{\sigma(T)\setminus\{0\}}(T)$ is the projection onto $\overline{\operatorname{ran}}\, T$.

Answer. We know from Exercise 12.4.7 that $1_{\sigma(T)}(T) = I_{\mathcal{H}}$. Hence $1_{\{0\}}(T) + 1_{\sigma(T)\setminus\{0\}}(T) = I_{\mathcal{H}}$.

Let $\xi \in 1_{\{0\}}(T)\mathcal{H}$. As $T\,1_{\{0\}}(T) = 0$ (from the equality of functions $t\,1_{\{0\}}(t) = 0$), we get $T\xi = T\,1_{\{0\}}(T)\xi = 0$, so $\xi \in \ker T$. That is, $1_{\{0\}}(T)\mathcal{H} \subset \ker T$. Conversely, let $\xi \in \ker T$. Since $T\xi = 0$ we have $T^k\xi = 0$ for all $k \in \mathbb{N}$, so $p(T)\xi = 0$ for all polynomials p with $p(0) = 0$. As was done on page 859 of the Book we can get a sequence $\{p_j\}$ of polynomials with $p_j \to 1_{\{0\}}(t)$ pointwise. Since the limit takes the value 1 at 0, we may assume without loss of generality that $p_j(0) = 1$ for all j. We can write $p_j = 1 + q_j$, with q_j a polynomial with $q_j(0) = 0$. Hence $p_j(T) = I_{\mathcal{H}} + q_j(T)$ and $p_j(T)\xi = \xi + q_j(T)\xi = \xi$. This gives us, as in (12.9), $1_{\{0\}}(T)\xi = \xi$. Thus $\ker T \subset 1_{\{0\}}(T)\mathcal{H}$ and the equality $\ker T = 1_{\{0\}}(T)\mathcal{H}$ is proven.

Now $1_{\sigma(T)\setminus\{0\}}(T) = I_{\mathcal{H}} - 1_{\{0\}}(T)$ is the projection onto $(\ker T)^\perp = \overline{\operatorname{ran}} T^* = \overline{\operatorname{ran}} T$.

(12.4.9) Let $\mathcal{M} = \ell^\infty[0,1] \subset \mathcal{B}(L^2[0,1])$, seen as multiplication operators. Fix $t \in [0,1]$ and for $\delta > 0$ let $P_\delta = M_{1_{[t-\delta,t+\delta]}}$.

(i) Show that $P_\delta \xrightarrow{\text{sot}} 0$ as $\delta \to 0$.

(ii) Show that if $K \in \mathcal{K}(L^2[0,1])$ then $\|P_\delta K\| \to 0$.

(iii) Show that there exists $T \in \mathcal{B}(L^2[0,1])$ such that $P_\delta T$ does not converge in norm to 0.

Answer.

(i) Fix $f \in L^2[0,1]$. We have

$$\|P_{1/n}f\|_2^2 = \int_{[0,1]} 1_{[t-1/n,t+1/n]}|f|^2\,dm \to 0$$

by Dominated Convergence. For arbitrary δ, given $\varepsilon > 0$ there exists n such that $\|P_{1/n}f\|_2 < \varepsilon$; if $\delta < 1/n$, then

$$\|P_\delta f\|_2 = \|P_\delta P_{1/n}f\|_2 \leq \|P_{1/n}f\|_2 < \varepsilon.$$

Thus $P_\delta f \to 0$ for all f, which is to soay that $P_\delta \xrightarrow{\text{sot}} 0$.

(ii) This is Exercise 12.1.8, since the net is bounded.

(iii) We can take $T = I_{\mathcal{H}}$, then $\|P_\delta T\| = \|P_\delta\| = 1$ for all δ.

(12.4.10) Show that in Corollary 12.4.18, if a sot-dense separable C*-subalgebra $\mathcal{A}_0 \subset \mathcal{A}$ is prescribed, the operator T can be chosen so that $C^*(T) \supset \mathcal{A}_0$. *(Attention: the word "separable" has different meanings when referring to C* and von Neumann algebras, see page 903 of the Book)*

Answer. Since $\mathcal{A}_0$ is separable, it has a countable dense subset $\{T_n\}$. By considering the real and imaginary parts of each T_n, we may assume without loss of generality that $T_n = T_n^*$ for all n. For each $n, k \in \mathbb{N}$ by the Spectral Theorem there exist projections $P_{n,k,1}, \ldots, P_{n,k,r_{n,k}}$ such that

$$\operatorname{dist}(T_n, \operatorname{span}\{P_{n,k,1}, \ldots, P_{n,k,r_{n,k}}\}) < 1/k.$$

If we now bunch all the countably many projections $\{P_{n,k}\}$ with the projections in the proof of Corollary 12.4.18, we get that $P_{n,k,s} \in C^*(T)$ for all s, so $\mathcal{A}_0 \subset C^*(T)$ (note that the proof of Corollary 12.4.18 only uses continuous functional calculus).

(12.4.11) Let $\mathcal{M}$ be a von Neumann algebra and $T \in \mathcal{M}$ normal. Give an alternative proof of Corollary 12.3.8 by using Corollary 12.4.15 and the fact that $UTU^* = T$ for all unitaries $U \in \mathcal{M}'$.

Answer. Fix $U \in \mathcal{M}'$ a unitary. From the equality $UTU^* = T$ and the fact that $f(UTU^*) = Uf(T)U^*$ for all continuous f (or the uniqueness of the positive square root), $|UTU^*| = |T| = U|T|U^*$. Then

$$T = UTU^* = UV|T|U^* = (UVU^*)U|T|U^* = (UVU^*)|T|.$$

Since V^*V is the range projection of T^*, we have $V^*V, VV^* \in \mathcal{M}$ by Corollary 12.4.15. Let $W = UVU^*$. We have $W^*W = UV^*U^*UVU^* = UV^*VU^* = V^*V$ and similarly $WW^* = VV^*$. Then the uniqueness in the polar decomposition guarantees that $UVU^* = V$. That is, $UV = VU$. As we can do this for any unitary in $\mathcal{M}'$ and unitaries span $\mathcal{M}'$, we have that $V \in \mathcal{M}'' = \mathcal{M}$.

12.5. Cyclic and Separating Vectors

(12.5.1) In Proposition 12.5.6, where is σ-finiteness used?

Answer. The direct sum and the series of projections can be done in any dimension. But if $\mathcal{A}$ is not σ-finite, we cannot construct the vector ξ, as any vector in any Hilbert space has at most countably many nonzero coefficients. For instance consider $\mathcal{A} = \ell^\infty[0,1]$ acting on $\mathcal{H} = \ell^2[0,1]$. As any $\xi \in \mathcal{H}$ has only countably many non-zero entries, there exists s such that $\xi(s) = 0$. Then $\delta_s\xi = 0$, even though $\delta_s \neq 0$. So ξ cannot be separating for $\mathcal{A}$.

(12.5.2) In the context of Proposition 12.5.6, show an example of $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$, abelian and without a separating vector.

Answer. The result in Proposition 12.5.6 tells us that we need to look at an uncountably-dimensional $\mathcal{H}$. Let $\mathcal{H} = \ell^2[0,1]$ and take $\mathcal{A}$ to be the diagonal masa, that is

$$\mathcal{A} = \{E_{tt} : \ t \in [0,1]\}'',$$

where E_{tt} is the orthogonal projection onto $\mathbb{C}\delta_t$. Let $\xi \in \mathcal{H}$. So $\xi = \sum_t c_t\delta_t$, with $\sum_t |c_t|^2 = \|\xi\|^2$. Since the series is convergent, only countably c_t are nonzero. Let $s \in [0,1]$ such that $c_s = 0$. Then $E_{ss}\xi = 0$, and so ξ is not separating.

(12.5.3) Verify the facts about the atomic masa stated after its definition (12.18). The only nontrivial part is that $\mathcal{A}_a \not\simeq \mathcal{A}_m$, which can be seen by looking at the existence or not of minimal projections in both algebras.

Answer. For each $T \in \mathcal{A}_a$ there exists a sequence $\{t_n\}$ such that $T\xi_n = t_n\xi_n$ for all n. Since $|t_n| = \|T\xi_n\| \leq \|T\|$, this allows us to define $\gamma : \mathcal{A}_a \to \ell^\infty(\mathbb{N})$ by $\gamma(T) = \{t_n\}$. It is clear that γ is linear. It is also multiplicative: $ST\xi_n = t_nS\xi_n = s_nt_n\xi_n$, so $\gamma(ST) = \gamma(S)\gamma(T)$. Also,

$$\langle T^*\xi_n, \xi_m\rangle = \langle \xi_n, T\xi_m\rangle = \langle \overline{t_n}\xi_n, \xi_m\rangle,$$

so $\gamma(T^*) = \{\overline{t_n}\} = \{t_n\}^*$. If $\gamma(T) = 0$, then $T\xi_n = 0$ for all n and therefore $T = 0$; so γ is injective. And if $\{t_n\} \in \ell^\infty(\mathbb{N})$, we can define $T \in \mathcal{B}(\mathcal{H})$ by $T\xi_n = t_n\xi_n$ and extend by linearity. We have

$$\begin{aligned}\Big\|T\sum_{j=1}^m \alpha_j\xi_j\Big\|^2 &= \Big\|\sum_{j=1}^m \alpha_jt_j\xi_j\Big\|^2 = \sum_{j=1}^m |\alpha_j|^2\,|t_j|^2 \\ &\leq \|t\|_\infty^2 \sum_{j=1}^m |\alpha_j|^2 = \|t\|_\infty^2 \Big\|\sum_{j=1}^m \alpha_j\xi_j\Big\|.\end{aligned}$$

Thus T is bounded on a dense subspace, and being linear it extends to all of $\mathcal{H}$, bounded with the same norm.

We have in particular that $E_{kk} \in \mathcal{A}$ for all k. Also, since $\ell^\infty(\mathbb{N})$ is abelian, we get that $\mathcal{A}_a$ is abelian. If $S \in \mathcal{A}'$, then for each n we have $S\xi_n = SE_{nn}\xi_n = E_{nn}S\xi_n \in \mathbb{C}\xi_n$. So $S \in \mathcal{A}_a$, and $\mathcal{A}_a' = \mathcal{A}_a$, showing that it is a masa.

We can construct an easy $*$-monomorphism $\mathcal{A}_a \to \mathcal{A}_m$ by choosing an infinite partition $\{E_n\}$ of $[0,1]$ (say, $E_n = (\frac{1}{n+1}, \frac{1}{n}]$, missing 0 does not matter because it is a nullset) and mapping $\{t_n\} \subset \ell^\infty(\mathbb{N})$ to $\sum_n t_n 1_{E_n}$. This is a $*$-monomorphism which of course is not surjective.

We cannot have $\mathcal{A}_a \simeq \mathcal{A}_m$ because $\mathcal{A}$ has minimal projections (namely, E_{nn} for each n) while $L^\infty[0,1]$ does not. Every projection in $L^\infty[0,1]$ is 1_E for some Lebesgue measurable set E of positive measure; and these can always be divided to obtain proper subprojections (for instance, via Exercise 2.3.17).

(12.5.4) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a maximal abelian von Neumann algebra (a masa). Show that if $\mathcal{A}$ has a cyclic (equivalently, separating) vector then $\mathcal{A}$ is σ-finite.

Answer. Suppose that $\mathcal{A}$ is not σ-finite. Then there exist uncountably many pairwise orthogonal projections $\{P_j\} \subset \mathcal{A}$. By extending the family if needed we may assume that $\sum_j P_j = I_{\mathcal{H}}$. We have

$$\|\xi\|^2 = \Big\| \sum_j P_j\xi \Big\|^2 = \sum_j \|P_j\xi\|^2.$$

As this is finite, only finitely many $P_j\xi$ can be nonzero. That is, there exists some j such that $P_j\xi = 0$. As $P_j \neq 0$, this contradicts the fact that ξ is separating.

(12.5.5) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra such that $\xi \in \mathcal{H}$ is separating for $\mathcal{M}$. Show that $\mathcal{M}$ is σ-finite.

Answer. As in the argument from Exercise 12.5.4 the fact that a vector in a Hilbert space can only admit countably many nonzero components forces, if $\mathcal{M}$ is not σ-finite, the existence of a nonzero projection $P \in \mathcal{A}$ with $P\xi = 0$. Then ξ is not separating.

12.6. Normal Functionals

(12.6.1) Let $\varphi \in \mathcal{M}$ be a positive normal functional with support projection F_φ. Show that $\varphi(TF_\varphi) = \varphi(T)$ for all $T \in \mathcal{M}$.

Answer. Suppose first that $T \geq 0$. Then

$$\varphi(T) = \overline{\varphi(T)} = \overline{\varphi(F_\varphi T)} = \varphi(TF_\varphi).$$

For arbitrary $T \in \mathcal{M}$, we can write T as a linear combination of four positive elements, and so the equality $\varphi(TF_\varphi) = \varphi(T)$ follows by linearity.

(12.6.2) Let φ be a state on $\mathcal{M}$ and F a support projection for φ. Show that φ is faithful when restricted to $F\mathcal{M}F$.

Answer. Suppose that $FTF \geq 0$ and $\varphi(FTF) = 0$. If $FTF \neq 0$, by Exercise 12.4.6 there exists $\lambda > 0$ and a nonzero projection $Q \in F\mathcal{M}F$ that commutes with FTF and such that $\lambda Q \leq QFTF$. Then

$$\begin{aligned}\lambda\varphi(Q) &= \varphi(\lambda Q) \leq \varphi(QFTF)\\ &= \varphi\big((FTF)^{1/2}Q(FTF)^{1/2}\big) \leq \varphi(FTF) = 0.\end{aligned}$$

Hence $Q \leq I_{\mathcal{M}} - F$ by definition of support projection. But this gives $Q = QF(I_{\mathcal{M}} - F) = 0$. The contradiction shows that $FTF = 0$ and φ is faithful on $F\mathcal{M}F$.

(12.6.3) Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$ such that there exists a projection $P \in \mathcal{M}$ with $\varphi(P) = 0$. Show that there exists a pairwise orthogonal family $\{P_j\} \subset \mathcal{M}$, maximal with respect to the property that $\varphi(P_j) = 0$ for all j.

Answer. Let

$$\mathcal{F} = \big\{\{P_j\} \subset \mathcal{M} : \text{pairwise orthogonal proj. with } \varphi(P_j) = 0 \ \text{ for all } j\big\},$$

ordered by inclusion. The family is nonempty because $\{P\} \in \mathcal{F}$. Given a chain $\big\{\{P_j\}_{j\in J_k}\big\}_k \subset \mathcal{F}$, with $J_{k_1} \subset J_{k_2}$ if $k_1 \leq k_2$, the union $\{P_j\}_{j\in\bigcup_k J_k}$ is in $\mathcal{F}$ and is an upper bound for the chain. By Zorn's Lemma there exists a maximal $\{P_j\} \in \mathcal{F}$.

(12.6.4) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ be von Neumann algebras, and $U : \mathcal{H} \to \mathcal{K}$ a unitary such that $U\mathcal{M}U^* \subset \mathcal{N}$. Show that $\Gamma(T) = UTU^*$ is a σ-weak continuous $*$-monomorphism.

Answer. With U a unitary, that Γ is a $*$-monomorphism is straightforward. Now suppose that $\{T_j\} \subset \mathcal{M}$ with $T_j \xrightarrow{\sigma\text{-weak}} 0$. This means that $\mathrm{Tr}(AT_j) \to 0$ for all $A \in \mathcal{T}(\mathcal{H})$. Given $B \in \mathcal{T}(\mathcal{K})$, by Exercise 10.7.16 we have $U^*BU \in \mathcal{T}(\mathcal{H})$. Then (using Exercise 10.7.8)

$$\mathrm{Tr}(B\Gamma(T)) = \mathrm{Tr}(BUTU^*) = \mathrm{Tr}\,\big((U^*BU)T\big) \to 0.$$

So $\Gamma(T_j) \xrightarrow{\sigma\text{-weak}} 0$.

(12.6.5) Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$. Show that φ is normal if and only if its GNS representation π_φ is normal.

Answer. We have $\varphi(T) = \langle \pi_\varphi(T)\xi_\varphi, \xi_\varphi\rangle$ for all $T \in \mathcal{M}$. If π_φ is normal and $T_j \xrightarrow{\sigma\text{-weak}} T$, then

$$\begin{aligned}\varphi(T) &= \langle \pi_\varphi(\lim_j T_j)\xi_\varphi, \xi_\varphi\rangle \overset{(\pi_\varphi \text{ normal})}{=} \langle \lim_j \pi_\varphi(T_j)\xi_\varphi, \xi_\varphi\rangle \\ &= \lim_j \langle \pi_\varphi(T_j)\xi_\varphi, \xi_\varphi\rangle = \lim_j \varphi(T_j),\end{aligned}$$

the last equality because point functionals are σ-weak continuous.

Conversely, suppose that φ is normal. Then if $\{T_j\} \subset \mathcal{M}$ is an increasing net of selfadjoints with $T_j \nearrow T$, then $S^*T_jS \nearrow S^*TS$ for all $S \in \mathcal{M}$ and hence

$$\begin{aligned}\langle \pi_\varphi(T_j)\pi_\varphi(S)\xi_\varphi, \pi_\varphi(S)\xi_\varphi\rangle &= \varphi(S^*T_jS) \nearrow \varphi(S^*TS) \\ &= \langle \pi_\varphi(T)\pi_\varphi(S)\xi_\varphi, \pi_\varphi(S)\xi_\varphi\rangle.\end{aligned}$$

It follows that $\pi_\varphi(T_j) \nearrow \pi_\varphi(T)$. Composing with normal functionals of $\pi_\varphi(\mathcal{M})''$ and using Proposition 12.6.11, we get that π_φ is normal.

(12.6.6) Let $\mathcal{M}$ be a von Neumann algebra and $\psi \in \mathcal{M}^*$ normal and faithful. Show that $\pi_\psi(\mathcal{M})'' = \pi_\psi(\mathcal{M})$. *(This is a direct consequence of Corollary 12.6.12, but it is needed earlier in the text)*

Answer. Since π_ψ is injective, it is isometric. So it maps $\overline{B_1^{\mathcal{M}}(0)}$ onto $\overline{B_1^{\pi_\psi(\mathcal{M})}}$. As ψ is normal, π_φ is normal by Exercise 12.6.5; and the closed unit ball is wot-compact, so its image through π_φ is wot-closed, hence sot-closed (Corollary 12.1.3). Kaplansky then implies that $\pi_\psi(\mathcal{M})$ is sot-closed, and by the Double Commutant Theorem $\pi_\psi(\mathcal{M})'' = \pi_\psi(\mathcal{M})$.

12.7. Preduals and the Enveloping von Neumann Algebra

(12.7.1) Let $\mathcal{M}$ be a von Neumann algebra. Show that $\mathcal{M}_*$ is norm-closed.

Answer. Let $\{\varphi_n\} \subset \mathcal{M}_*$ be a Cauchy sequence. By being Cauchy in a metric space the sequence is bounded, so there exists $c > 0$ such that $\|\varphi_n\| \leq c$ for all n. Since $\mathcal{M}^*$ is complete, $\varphi = \lim \varphi_n \in \mathcal{M}^*$ exists. We need to show that φ is wot-continuous on bounded sets.

Let $\{T_j\} \subset \mathcal{M}$ be a bounded net such that $T_j \xrightarrow{\text{wot}} 0$. By enlarging c is needed, we may assume that $\|T_j\| \leq c$ for all j. Then

$$|\varphi(T_j)| \leq |(\varphi - \varphi_n)(T_j)| + |\varphi_n(T_j)| \leq c\,\|\varphi - \varphi_n\| + |\varphi_n(T_j)|.$$

Hence $\limsup_j |\varphi(T_j)| \leq c\,\|\varphi - \varphi_n\|$. As we are free to choose φ_n and $\varphi - \varphi_n \to 0$, the limit exists and is zero by the Limsup Routine; it follows that φ is wot-continuous on bounded sets, so $\varphi \in \mathcal{M}_*$ by Proposition 12.6.3.

(12.7.2) Provide an alternative proof to the fact that $\mathcal{M}_*$ is a predual for the von Neumann algebra $\mathcal{M}$ by using Corollary 7.3.8 to see that a predual for $\mathcal{M}$ is given by $\mathcal{T}(\mathcal{H})/\mathcal{M}_o$. This means identifying all normal functionals that agree on $\mathcal{M}$, so we have precisely the normal functionals of $\mathcal{M}$.

Answer. We know that $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is σ-weak-closed by Proposition 12.3.19. From Theorem 10.7.11 we know that $\mathcal{B}(\mathcal{H}) = \mathcal{T}(\mathcal{H})^*$. So the proof of Corollary 7.3.8 says that $\mathcal{T}(\mathcal{H})/\mathcal{M}_o$ is a predual for $\mathcal{M}$. For each class $\varphi = S + \mathcal{M}_o$ with $S \in \mathcal{T}(\mathcal{H})$, we are interpreting this as the functional $\varphi(T) = \operatorname{Tr}(ST)$. This is well-defined on $\mathcal{M}$ for if $S - S' \in \mathcal{M}_o$, this means that $\operatorname{Tr}(ST) = \operatorname{Tr}(S'T)$ for all $T \in \mathcal{M}$, and so they define the same linear functional. Said functional is normal by Proposition 12.6.3. So $\mathcal{T}(\mathcal{H})/\mathcal{M}_o \subset \mathcal{M}_*$. Now, given $\varphi \in \mathcal{M}_*$, by Proposition 12.6.3 there exists $S \in \mathcal{T}(\mathcal{H})$ with $\varphi(T) = \operatorname{Tr}(ST)$; that is, $\varphi = S + \mathcal{M}_o \in \mathcal{T}(\mathcal{H})/\mathcal{M}_o$. Hence $\mathcal{M}_* = \mathcal{T}(\mathcal{H})/\mathcal{M}_o$ is a predual for $\mathcal{M}$.

(12.7.3) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\gamma : \mathcal{A} \to \mathcal{B}$ a $*$-isomorphism. Show that $\mathcal{A}^{**} \simeq \mathcal{B}^{**}$ as von Neumann algebras.

Answer. We have a canonical isometric linear isomorphism $\gamma^{**} : \mathcal{A}^{**} \to \mathcal{B}^{**}$. But we need to account for the multiplication, and for this we look at the enveloping von Neumann algebras. Let $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be a universal representation. Then there exists, by Theorem 12.7.8, a unique linear surjective isometry $\tilde{\pi}_{\mathcal{A}} : \mathcal{A}^{**} \to \pi(\mathcal{A})''$ that extends $\pi_{\mathcal{A}}$. We can define a representation $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$ by $\pi_{\mathcal{B}} = \pi_{\mathcal{A}} \circ \gamma^{-1}$. This is universal: if $\rho : \mathcal{B} \to \mathcal{B}(\mathcal{K})$ is a representation, then $\rho \circ \gamma : \mathcal{A} \to \mathcal{B}(\mathcal{K})$ is a representation. By universality there exists $\rho' : \pi_{\mathcal{A}}(\mathcal{A})'' \to \rho(\gamma(\mathcal{A}))'' = \rho(\mathcal{B})''$ such that $\rho' \circ \pi_{\mathcal{A}} = \rho \circ \gamma$. Then $\rho' \circ \pi_{\mathcal{B}} = \rho$, showing that $\pi_{\mathcal{B}}$ is universal. Now $\mathcal{B}^{**} \simeq \pi_{\mathcal{B}}(\mathcal{B})'' = \pi(\mathcal{A})'' \simeq \mathcal{A}^{**}$, where the two isomorphisms are canonical.

We have $\pi_{\mathcal{B}} \circ \gamma = \pi_{\mathcal{A}}$ by construction, so the extensions from Theorem 12.7.8 satisfy $\tilde{\pi}_{\mathcal{B}} \circ \gamma^{**} = \tilde{\pi}_{\mathcal{A}}$.

(12.7.4) Let $\mathcal{A}$ be a unital C*-algebra. Show that the unit of $\mathcal{A}^{**}$ is the unit of $\mathcal{A}$; that is, show that $\hat{I}_{\mathcal{A}} = I_{\mathcal{A}^{**}}$.

Answer. Given $\Psi \in \mathcal{A}^{**}$, by Theorem 7.2.14 there exists a net $\{a_j\} \subset \mathcal{A}$ with $\hat{a}_j \xrightarrow{\text{weak}^*} \Psi$. Then, using the universal representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and its associated homeomorphism $\tilde{\pi} : \mathcal{A}^{**} \to \pi(\mathcal{A})''$ as in Theorem 12.7.8,

$$\begin{aligned}\tilde{\pi}(\hat{I}_{\mathcal{A}})\tilde{\pi}(\Psi) &= \lim_j \tilde{\pi}(\hat{I}_{\mathcal{A}})\tilde{\pi}(\hat{a}_j) = \lim_j \pi(I_{\mathcal{A}})\pi(a_j) = \lim_j \pi(a_j) \\ &= \lim_j \tilde{\pi}(\hat{a}_j) = \tilde{\pi}(\Psi).\end{aligned}$$

The same computation can be done on the right, so $\tilde{\pi}(\hat{I}_{\mathcal{A}}) = I_{\pi(\mathcal{A})''}$ (note that we can always have π non-degenerate by shrinking $\mathcal{H}$ if needed). And we are done, because $I_{\pi(\mathcal{A})''}$ is what we mean when we write $I_{\mathcal{A}^{**}}$, as we only see $\mathcal{A}^{**}$ as an algebra via $\tilde{\pi}$.

(12.7.5) Let $\mathcal{A}$ be a C*-algebra and $Z \in \mathcal{Z}(\mathcal{A}^{**})$ a central projection. Show that $(Z\mathcal{A})^{**} = Z\mathcal{A}^{**}$.

Answer. Given $\psi \in \mathcal{A}^{**}$, there exists a net $\{a_k\} \subset \mathcal{A}$ such that $\hat{a}_k \xrightarrow{\text{weak}^*} \psi$. When we see this in the enveloping von Neumann algebra, we have $a_k \xrightarrow{\sigma\text{-weak}} \psi$. Then $Za_k \xrightarrow{\sigma\text{-weak}} Z\psi$. So $Z\mathcal{A}^{**} \subset (Z\mathcal{A})^{**}$. For the reverse inclusion, we have $Z\mathcal{A} \subset Z\mathcal{A}^{**}$; thinking of the enveloping von Neumann algebra as the double commutant of the image through the universal representation, $(Z\mathcal{A})^{**} \subset Z\mathcal{A}^{**}$.

(12.7.6) Let $\mathcal{A}$ be a C*-algebra and $\mathcal{J} \subset \mathcal{A}$ a proper ideal. Show that $\mathcal{A}^{**} \simeq \mathcal{J}^{**} \oplus (\mathcal{A}/\mathcal{J})^{**}$ as C*-algebras.

Answer. Because $\mathcal{J}$ is a proper ideal of $\mathcal{A}$, by Hahn–Banach (Corollary 5.7.19) there exists nonzero $\varphi \in \mathcal{A}^*$ with $\varphi|_{\mathcal{J}} = 0$. When we look at $\mathcal{J}^{**} \subset \mathcal{A}^{**}$, the functional φ becomes normal and so $\varphi|_{\mathcal{J}^{**}} = 0$, which guarantees that $\mathcal{J}^{**}$ is a proper ideal of $\mathcal{A}^{**}$. By Corollary 12.3.12 there exists $Z \in \mathcal{Z}(\mathcal{A}^{**})$ with $\mathcal{J}^{**} = Z\mathcal{A}^{**}$. So

$$\mathcal{A}^{**} = \mathcal{J}^{**} \oplus (I_{\mathcal{A}^{**}} - Z)\mathcal{A}^{**}.$$

Consider now the map $\gamma : \mathcal{A}/\mathcal{J} \to (I_{\mathcal{A}^{**}} - Z)\mathcal{A}$ given by $\gamma(A+\mathcal{J}) = (I_{\mathcal{A}^{**}} - Z)A$. This is well-defined: it is linear and if $A \in \mathcal{J}$ then $A = ZA$ and so $(I_{\mathcal{A}^{**}} - Z)A = 0$. It is clearly surjective and if $(I_{\mathcal{A}^{**}} - Z)A = 0$ then $A = ZA$ so $A \in \mathcal{J}$, so it is injective. It is also straightforward by construction that γ is a $*$-homomorphism. Then γ^{**} is a $*$-isomorphism $(\mathcal{A}/\mathcal{J})^{**} \to (I_{\mathcal{A}^{**}} - Z)\mathcal{A}^{**}$, via Exercise 12.7.5. By Corollary 12.6.12, γ^{**} is normal. Now $\mathrm{id} \oplus \gamma^{**}$ is the desired isomorphism, where Exercise 12.7.3 confirms the multiplicativity.

(12.7.7) Let A be the closed unit ball of $\ell^\infty(\mathbb{R}) \subset \mathcal{B}(\ell^2(\mathbb{R}))$. Show that on A the σ-weak topology is precisely pointwise convergence.

Answer. By Lemma 12.1.21 the σ-weak topology agrees with the wot on A. Suppose that $\{f_j\} \subset A$ and $f_j \xrightarrow{\text{wot}} 0$. This means in particular that $f_j(t) = \langle f_j e_t, e_t\rangle \to 0$. Conversely, suppose that $f_j(t) \to 0$ for all t. Fix $g \in \ell^2(\mathbb{R})$. Let $\varepsilon > 0$. Choose t_0 such that $\sum_{|t|>t_0} |g(t)|^2 < \varepsilon$. Then

$$\begin{aligned} |\langle f_j g, g\rangle| &\leq \sum_{|t|\leq t_0} |f_j(t)|\,|g(t)|^2 + \sum_{|t|>t_0} |f_j(t)|\,|g(t)|^2 \\ &\leq \sum_{|t|\leq t_0} |f_j(t)|\,|g(t)|^2 + \sum_{|t|>t_0} |g(t)|^2 \\ &\leq \sum_{|t|\leq t_0} |f_j(t)|\,|g(t)|^2 + \varepsilon \end{aligned}$$

Then $\limsup_j |\langle f_j g, g\rangle| = 0$. By the Limsup Routine, the limit exists and is 0. Applying polarization we obtain $f_j \xrightarrow{\text{wot}} 0$.

(12.7.8) Show that $\ell^\infty(\mathbb{R}) \subset \mathcal{B}(\ell^2(\mathbb{R}))$ is σ-weak separable.

Answer. Since the whole space is a countable union of balls, it is enough to show that the unit ball has a countable dense subset. By Exercise 12.7.7 we consider pointwise convergence. Our countable dense set will be

$$C = \Big\{ \sum_{k=1}^{n} q_k \, 1_{E_k}, \ n \in \mathbb{N}, \ q_k \in \mathbb{Q} + i\mathbb{Q}, \ |q_k| \leq 1, \ E_1, \ldots, E_n$$
$$\text{partition of } \mathbb{R} \text{ with endpoints in } \mathbb{Q} \Big\}.$$

Fix $f \in \ell^\infty(\mathbb{R})$ and $\varepsilon > 0$. For each $F = \{t_1, \ldots, t_n\} \subset \mathbb{R}$ finite, given $t_k \in F$ choose $q_k \in \mathbb{Q} + i\mathbb{Q}$ with $|t - q_k| < \varepsilon$. Let $r_1, \ldots, r_{n+1} \in \mathbb{Q}$ with $r_k < t_k < r_{k+1}$ for all $k = 1, \ldots, n$. Take $E_1 = (-\infty, r_1)$, $E_{n+1} = (r_{n+1}, \infty)$, and $E_k = (r_{k-1}, r_k)$. Then $g_F = \sum_k q_k 1_{E_k} \in C$ and $|f(t) - g_F(t)| < \varepsilon$ for all $t \in F$. So if we consider the seminorms p_t as in Example 5.4.14 and $N_{F,\varepsilon} = \{g : \ p_t(f - g) < \varepsilon, \ t \in F\}$ is a basic neighbourhood around f, for any finite $F' \supset F$ we will have $g_{F'} \in N_{F,\varepsilon}$. This shows that $g_F \to f$ pointwise, when we consider the finite subsets of $\mathbb{R}$ ordered by inclusion.

(12.7.9) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{X}$ a Banach space such that $\mathcal{M} = \mathcal{X}^*$. Put $\mathcal{P} = \{\varphi_1 - \varphi_2 + i(\varphi_3 - \varphi_4) : \ \varphi_j \in \mathcal{X}^+, \ j = 1, \ldots, 4\}$. Show that $\mathcal{P}$ is a subspace.

Answer. For the subspace part, that $\mathcal{P}$ is closed under addition is just the fact that sums of positive functionals are positive, and $\mathcal{X}$ is a vector space. As for the multiplication by scalars, by writing $\lambda \in \mathbb{C}$ in the form $\lambda = a_1 - a_2 + i(a_3 - a_4)$ with $a_1, a_2, a_3, a_4 \geq 0$ we get

$$\begin{aligned} \lambda\big(\varphi_1 - \varphi_2 + i(\varphi_3 - \varphi_4)\big) = (a_1\varphi_1 + a_2\varphi_2 + a_3\varphi_4 + a_4\varphi_3) \\ - (a_1\varphi_2 + a_2\varphi_1 + a_3\varphi_3 + a_4\varphi_4) \\ + i(a_1\varphi_3 + a_2\varphi_4 + a_3\varphi_1 + a_4\varphi_2) \\ - i(a_1\varphi_4 + a_2\varphi_3 + a_3\varphi_2 + a_4\varphi_1) \\ \in \mathcal{P}. \end{aligned}$$

(12.7.10) Let $\mathcal{A}$ be a C*-algebra and $\mathcal{X}, \mathcal{Y}$ Banach spaces such that $\mathcal{X}^* = \mathcal{Y}^* = \mathcal{A}$. Show that $\mathcal{X} \simeq \mathcal{Y}$ isometrically.

Answer. By the existence of the predual we get from Theorem 12.7.5 that there exists a faithful representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ such that $\pi(\mathcal{A})$ is a von Neumann algebra. If $\gamma : \mathcal{X}^* \to \mathcal{A}$ is an isometric isomorphism, we get that $\pi \circ \gamma : X^* \to \pi(\mathcal{A})$ is an isometric isomorphism of Banach spaces. Now

Theorem 12.7.2 implies that $\mathcal{X} \simeq \pi(\mathcal{A})_*$ isometrically. As the same can be done for $\mathcal{Y}$, we get that $\mathcal{Y} \simeq \mathcal{X}$ isometrically.

CHAPTER

13

Constructions with C*-Algebras

13.1. Algebraic Tensor Products

(13.1.1) Prove Proposition 13.1.2.

Answer. This follows rather directly from the definition. Indeed, for ϕ bilinear

$$\big(\lambda(x \otimes y)\big)(\phi) = \lambda(x \otimes y)(\phi) = \lambda\phi(x, y) = \phi(\lambda x, y) = (\lambda x \otimes y)(\phi),$$

so $\lambda(x \otimes y) = (\lambda x) \otimes y$. The other scalar multiplication is similar.

For the sum, for any bilinear ϕ

$$\begin{aligned}\big((x + z) \otimes y\big)(\phi) &= \phi(x + z, y) = \phi(x, y) + \phi(z, y) = (x \otimes y)(\phi) + (z \otimes y)(\phi)\\ &= \big((x \otimes y) + (z \otimes y)\big)(\phi),\end{aligned}$$

so $(x + z) \otimes y = x \otimes y + z \otimes y$. The other distributivity is proven similarly.

(13.1.2) Let $\{e_k\}$ be a basis for $\mathcal{X}$ and $\{f_j\}$ be a basis for $\mathcal{Y}$. Show that $B = \{e_k \otimes f_j\}_{k,j}$ is a basis for $\mathcal{X} \otimes \mathcal{Y}$.

Answer. The linear independence follows from Proposition 13.1.3. And given $z \in \mathcal{X} \otimes \mathcal{Y}$, by definition there exist $x_1, \ldots, x_r \in \mathcal{X}$ and $y_1, \ldots, y_r \in \mathcal{Y}$ with $z = \sum_{s=1}^r x_s \otimes y_s$. Expressing each x_s and y_s in its respective basis we have for each s

$$x_s = \sum_k \alpha_{s,k} e_k, \qquad y_s = \sum_j \beta_{s,j} f_j.$$

Then

$$z = \sum_{s=1}^r \sum_{k,j} \alpha_{s,k}\beta_{s,j}\, e_k \otimes f_j = \sum_{k,j} \left(\sum_{s=1}^r \alpha_{s,k}\beta_{s,j} \right) e_k \otimes f_j.$$

Thus $\mathcal{X} \otimes \mathcal{Y} = \operatorname{span} B$, and B is a basis.

(13.1.3) Let X, Y be complex vector spaces and $X \otimes' Y$ a tensor product defined in some way other than via our bilinear maps. This means that $X \otimes' Y$ is a vector space, spanned by elements of the form $x \otimes' y$ which satisfy the properties in Propositions 13.1.2 and 13.1.3. Show that $X \otimes' Y \simeq X \otimes Y$ canonically.

Answer. Consider the bilinear map $\phi : X \times Y \to X \otimes' Y$ given by $\phi(x, y) = x \otimes' y$. By Theorem 13.1.6 there exists a linear map $\Psi : X \otimes Y \to X \otimes' Y$ that satisfies $\Psi(x \otimes y) = x \otimes' y$ for all $x \in X$ and $y \in Y$. As Ψ is linear, it follows from Exercise 13.1.2 that it is bijective, since it maps a basis to a basis. Therefore Ψ is a linear bijection and $X \otimes' Y \simeq X \otimes Y$ via a map that sends $x \otimes' y \longmapsto x \otimes y$.

(13.1.4) Use Theorem 13.1.6 to show that the map $\mathcal{X} \otimes \mathcal{Y} \to \mathcal{Y} \otimes \mathcal{X}$ induced by $x \otimes y \longmapsto y \otimes x$ is an isomorphism.

Answer. Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathcal{Y} \otimes \mathcal{X}$ be given by $\phi(x, y) = y \otimes x$. By Theorem 13.1.6, there exists linear $T_\phi : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{Y} \otimes \mathcal{X}$ with

$$T_\phi(x \otimes y) = \phi(y, x) = y \otimes x.$$

Similarly we can get linear $S : \mathcal{Y} \otimes \mathcal{X} \to \mathcal{X} \otimes \mathcal{Y}$ with $S(y \otimes x) = x \otimes y$. As $S \circ T_\phi(x \otimes y) = x \otimes y$ for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, it follows by linearity that $S \circ T_\phi = \mathrm{id}_{\mathcal{X} \otimes \mathcal{Y}}$. We also get that $T_\phi \circ S = \mathrm{id}_{\mathcal{Y} \otimes \mathcal{X}}$. Thus T_ϕ is a vector space isomorphism of $\mathcal{Y} \otimes \mathcal{X}$ onto $\mathcal{X} \otimes \mathcal{Y}$.

(13.1.5) Using Theorem 13.1.6 as in Exercise 13.1.4, show that there are canonical isomorphisms (in that they do the obvious thing to the elementary tensors) as follows:

(i) $\mathbb{C} \otimes \mathcal{X} \simeq \mathcal{X}$;

(ii) $\mathbb{C}^n \otimes \mathcal{X} \simeq \mathcal{X}^n$;

(iii) $\mathcal{X} \otimes \mathcal{Y}^* \simeq B(\mathcal{Y}, \mathcal{X})$, if $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional;

(iv) $M_n(\mathbb{C}) \otimes \mathcal{X} \simeq M_n(\mathcal{X})$.

Note that when $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional one could establish the existence of isomorphisms as above by dimension considerations. But we do not always require finite-dimension, and we want our isomorphisms to be canonical.

Answer.

(i) Consider the bilinear map $\phi : \mathbb{C} \times \mathcal{X} \to \mathcal{X}$ given by $\phi(\lambda, x) = \lambda x$. By Theorem 13.1.6, there exists a linear map $T : \mathbb{C} \otimes \mathcal{X} \to \mathcal{X}$ such that $T(\lambda \otimes x) = \lambda x$. It is obvious that T is onto, so we only need to show that T is on-to-one. Suppose that $T(\sum_j \lambda_j \otimes x_j) = 0$. This means that $\sum_j \lambda_j x_j = 0$. Then

$$\sum_j \lambda_j \otimes x_j = \sum_j 1 \otimes \lambda_j x_j = 1 \otimes \sum_j \lambda_j x_j = 0,$$

so T is one-to-one.

(ii) Now consider the bilinear map $\phi : \mathbb{C}^n \times \mathcal{X} \to \mathcal{X}^n$ given by

$$\phi((\lambda_1, \ldots, \lambda_n), x) = (\lambda_1 x, \ldots, \lambda_n x).$$

By Theorem 13.1.6 there exists a linear map $T : \mathbb{C}^n \otimes \mathcal{X} \to \mathcal{X}^n$ such that $T((\lambda_1, \ldots, \lambda_n) \otimes x) = (\lambda_1 x, \ldots, \lambda_n x)$. We can define an inverse for T explicitly by

$$T^{-1}(x_1 \ldots, x_n) = \sum_{k=1}^{n} e_k \otimes x_k \in \mathbb{C}^n \otimes \mathcal{X}.$$

(iii) This time the bilinear map is $\phi(x, f) = f(\cdot)x \in B(\mathcal{Y}, \mathcal{X})$. By Theorem 13.1.6 there exists a linear map $T : \mathcal{X} \otimes \mathcal{Y}^* \to B(\mathcal{Y}, \mathcal{X})$ with $T(x \otimes f) = f(\cdot)x$.

If $\mathcal{X}$ is finite-dimensional, then given a basis $x_1, \ldots, x_n$ of $\mathcal{X}$ we can write any $A \in B(\mathcal{Y}, \mathcal{X})$ as $Ay = \sum_{k=1}^n f_j(y)x_j$, where $f_j \in \mathcal{Y}^*$ is determined uniquely by the fact that $\{x_j\}$ is a basis. Thus

$$T^{-1}\Big(\sum_{k=1}^n f_j(\cdot)x_j\Big) = \sum_{k=1}^n x_j \otimes f_j$$

gives an inverse for T and so T is an isomorphism.

(iv) Consider the bilinear map $\phi : M_n(\mathbb{C}) \times \mathcal{X} \to M_n(\mathcal{X})$ given by $\phi(A, x) = [a_{kj}x]_{k,j}$. By Theorem 13.1.6 there exists a linear map $T : M_n(\mathbb{C}) \otimes \mathcal{X} \to M_n(\mathcal{X})$ with $T(A \otimes x) = [a_{kj}x]_{k,j}$. Given any $X \in M_n(\mathcal{X})$, we can define an inverse for T by

$$T^{-1}(X) = \sum_{k,j} E_{kj} \otimes x_{kj}.$$

(13.1.6) Prove Corollary 13.1.9.

Answer. Via Theorem 13.1.6 we define $\varphi \times \psi = T_b$, where $b : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{A}$ is the bilinear form $b(x, y) = \varphi(x)\psi(y)$. If $L : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{A}$ is linear and $L(x \otimes y) = \varphi(x)\psi(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, the bilinear form induced by L is $b_L(x, y) = L(x \otimes y) = \varphi(x)\psi(y)$, which agrees with b; and so $L = T_{b_L} = T_b = \varphi \times \psi$.

In the case where $\mathcal{X}, \mathcal{Y}$ are algebras and φ, ψ homomorphisms with commuting ranges, due to the linearity we only have to show multiplicativity on elementary tensors. We have

$$\begin{aligned}
(\varphi \times \psi)\big((x_1 \otimes y_1)(x_2 \otimes y_2)\big) &= (\varphi \times \psi)(x_1x_2 \otimes y_1y_2) \\
&= \varphi(x_1x_2)\psi(y_1y_2) \\
&= \varphi(x_1)\varphi(x_2)\psi(y_1)\psi(y_2) \\
&= \varphi(x_1)\psi(y_1)\varphi(x_2)\psi(y_2) \\
&= (\varphi \times \psi)(x_1 \otimes y_1)\,(\varphi \times \psi)(s_2 \otimes y_2).
\end{aligned}$$

(13.1.7) In the situation of Corollary 13.1.9 where $\mathcal{X}, \mathcal{Y}$ are algebras and φ, ψ homomorphisms, show an example where both φ and ψ are injective but $\varphi \times \psi$ is not (*Hint: abelian algebras and finite-dimension are enough*).

Answer. Let $\mathcal{X} = \mathcal{Y} = \mathcal{A} = \mathbb{C}^2$, with pointwise addition and multiplication, and let $\varphi = \psi = \text{id}$. Then φ and ψ are injective homomorphisms with commuting ranges. But

$$(\varphi \times \psi)(e_1 \otimes e_2) = e_1 e_2 = 0.$$

13.2. Completely Positive Maps

(13.2.1) Prove that if $\mathcal{S} \subset \mathcal{A}$ is an operator system, then $M_n(\mathcal{S}) \subset M_n(\mathcal{A})$ is an operator system.

Answer. This of course assumes that $\mathcal{A}$ is unital. In that case, $M_n(\mathcal{S})$ contains the identity matrix $\sum_k I_{\mathcal{A}} \otimes E_{kk}$. It is obvious that $M_n(\mathcal{S})$ is a subspace of $M_n(\mathcal{A})$, so all that remains is to check that $M_n(\mathcal{S})$ is closed under taking adjoints. Given $S = \sum_{k,j} s_{kj} \otimes E_{kj} \in M_n(\mathcal{S})$, we have

$$S^* = \sum_{k,j} s_{kj}^* \otimes E_{jk}.$$

As each $s_{kj}^* \in \mathcal{S}$—since $\mathcal{S}$ is an operator system—we get that $S^* \in M_n(\mathcal{S})$.

(13.2.2) Show that that if $n, m \in \mathbb{N}$ with $n < m$, and $\phi : \mathcal{S} \to \mathcal{B}$, then $\|\phi^{(m)}\| \geq \|\phi^{(n)}\|$, and $\phi^{(m)} \geq 0$ implies $\phi^{(n)} \geq 0$.

Answer. We may assume without loss of generality that $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$. We have

$$\|\phi^{(n)}\| = \sup\{\|\phi^{(n)}(A)\| : \ A \in M_n(S), \ \|A\| = 1\}.$$

Since

$$\phi^{(n)}(A) = \phi^{(m)}(A \oplus 0_{m-n}),$$

every number $\|\phi^{(n)}(A)\|$ can be written as $\|\phi^{(m)}(\tilde{A})\|$ with $\|\tilde{A}\| = 1$. Thus $\|\phi^{(n)}\| \leq \|\phi^{(m)}\|$.

Similarly, if $\phi^{(m)} \geq 0$ and $A \in M_n(S)$ is positive, then $\tilde{A} = A \oplus 0 \geq 0$ and

$$\langle \phi^{(n)}(A)\xi, \xi\rangle = \langle \phi^{(m)}(\tilde{A})\tilde{\xi}, \tilde{\xi}\rangle \geq 0$$

for any $\xi \in \mathcal{H}^n$, with $\tilde{\xi} = \xi \oplus 0$.

(13.2.3) Let $\mathcal{S} = \mathcal{A}$ be a C*-algebra, and $\phi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Show that ϕ is completely positive.

Answer. If $A \in M_n(\mathcal{A})$ is positive, we can write $A = B^*B$ for some $B \in M_n(\mathcal{A})$. Then

$$\begin{aligned}
\phi^{(n)}(A) = \phi^{(n)}(B^*B) &= \sum_{k,j} \phi^{(n)}\big((B^*B)_{kj}\big) \otimes E_{kj} \\
&= \sum_{k,j}\sum_{h} \phi(B_{hk}^* B_{hj}) \otimes E_{kj} = \sum_{k,j}\sum_{h} \phi(B_{hk})^*\phi(B_{hj}) \otimes E_{kj} \\
&= \sum_{k,j}\sum_{h} \big(\phi(B_{hk}) \otimes E_{hk}\big)^*\big(\phi(B_{hj}) \otimes E_{hj}\big) \\
&= \sum_{h} \Big(\sum_{k} \phi(B_{hk} \otimes E_{hk}\Big)^* \Big(\sum_{k} \phi(B_{hk} \otimes E_{hk}\Big) \geq 0.
\end{aligned}$$

(13.2.4) Show an example of a C*-algebra $\mathcal{A}$ with a dense subalgebra $\mathcal{A}_0$ and a $*$-homomorphism $\rho : \mathcal{A}_0 \to \mathcal{A}_0$ that is unbounded (so, in particular, it doesn't extend to $\mathcal{A}$).

Answer. Let $\mathcal{A} = C[0,1]$, $\mathcal{A}_0 = \mathbb{C}[x]$, and $\rho : \mathcal{A}_0 \to \mathcal{A}_0$ given by $\rho(p) = p(2)$. Then ρ is a $*$-homomorphism. If $p_n(x) = x^n$, then $\|p_n\| = 1$ for all n but $\rho(p_n) = 2^n$, so ρ is unbounded.

(13.2.5) Given a compact Hausdorff space T, show that the C*-algebras $\mathcal{A} = M_n(C(T))$ and $\mathcal{B} = C(T, M_n(\mathbb{C}))$ are canonically isomorphic, where the norm in $\mathcal{B}$ is given by

$$\|y\|_{\mathcal{B}} = \sup\{\|y(t)\| : \ t \in T\}.$$

Answer. Let $\pi : \mathcal{A} \to \mathcal{B}$ be given by

$$\pi(A)(t) = \sum_{k,j} A_{kj}(t) \otimes E_{kj}.$$

Because the algebraic operations between functions are defined pointwise, it is clear that π is a *-homomorphism. It is injective, for if $A_{kj}(t) = 0$ for all

t and all k,j then $A_{kj} = 0$ for all k,j, and therefore $A = 0$. To see that π is surjective, if $g : T \to M_n(\mathbb{C})$ is continuous, let $A = \sum_{k,j} g_{kj} \otimes E_{kj}$, where $g_{k,j}(t) = e_k^* g(t) e_j$; then $\pi(A) = g$.

(13.2.6) Prove that a positive map $\phi : \mathcal{S} \to \mathcal{B}$ maps selfadjoint elements to selfadjoint elements.

Answer. Let $a \in \mathcal{S}$ be selfadjoint. Then $a + \|a\|\, I_{\mathcal{A}}$ is positive, so $z = \phi(a) + \|a\|\, \phi(I_{\mathcal{A}})$ is positive. Therefore $\phi(a) = z - \|a\|\, \phi(I_{\mathcal{A}})$ is a linear combination of positive elements and hence selfadjoint.

(13.2.7) Show that

$$\alpha + \beta z + \gamma \bar{z} \geq 0,\ z \in \overline{\mathbb{D}} \iff \alpha \geq 0,\ \gamma = \bar{\beta},\ 2|\beta| \leq \alpha$$
$$\iff \begin{bmatrix} \alpha & 2\beta \\ 2\gamma & \alpha \end{bmatrix} \geq 0.$$

Answer. Suppose first that $\alpha + \beta z + \gamma \bar{z} \geq 0$ for all z in the disk. From $z = 0$ we get that $\alpha \geq 0$. It follows that $\beta z + \gamma \bar{z} \in \mathbb{R}$ for all z. With $z = 1$ and $z = i$ we obtain $\beta + \gamma \in \mathbb{R}$, $(\beta - \gamma)i \in \mathbb{R}$. We get from these that $\operatorname{Im}\gamma = -\operatorname{Im}\beta$, and $\operatorname{Re}\beta = \operatorname{Re}\gamma$; so $\gamma = \bar{\beta}$. Now we have that $\alpha + 2\operatorname{Re}\beta z \geq 0$ for all z. Choosing z such that $\beta z = -|\beta|$ we get $2|\beta| \leq \alpha$.

Now assume that $\alpha \geq 0$, $\gamma = \bar{\beta}$, and $2|\beta| \leq \alpha$. We have

$$\begin{aligned}
\left\langle \begin{bmatrix} \alpha & 2\beta \\ 2\gamma & \alpha \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \right\rangle &= \alpha|z_1|^2 + \alpha|z_2|^2 + 4\operatorname{Re}\beta z_1 \overline{z_2} \\
&\geq \alpha|z_1|^2 + \alpha|z_2|^2 - 4|\beta|\,|z_1|\,|z_2| \\
&\geq \alpha|z_1|^2 + \alpha|z_2|^2 - 2\alpha\,|z_1|\,|z_2| \\
&= \alpha\left(|z_1|^2 - |z_2|^2\right) \geq 0.
\end{aligned}$$

So the matrix is positive.

Finally, suppose that the matrix is positive; that is,

$$\alpha|z_1|^2 + \alpha|z_2|^2 + 2\beta \overline{z_1} z_2 + 2\gamma z_1 \overline{z_2} \geq 0$$

for all z_1, z_2. Given $z = re^{i\theta} \in \overline{\mathbb{D}}$, let

$$z_1 = \sqrt{\frac{1 + \sqrt{1 - r^2}}{2}}\, e^{-i\theta}, \qquad z_2 = \sqrt{\frac{1 - \sqrt{1 - r^2}}{2}}.$$

Then $|z_1|^2 + |z_2|^2 = 1$ and $2\overline{z_1} z_2 = z$, giving us $\alpha + \beta z + \gamma \bar{z} \geq 0$.

(13.2.8) Prove equation (13.4), i.e.

$$\sum_{k,j=1}^{n} \langle \phi(a_j^* a_k)\xi_k, \xi_j\rangle = \left\langle \phi^{(n)}(A^*A)\xi, \xi \right\rangle,$$

where $\xi = (\xi_1, \ldots, \xi_n)^\top \in \mathcal{H}^n$ and $A \in M_n(\mathcal{A})$ is the matrix with some row $a_1, \ldots, a_n$ and zeroes elsewhere.

Answer. If we write $\xi = \begin{bmatrix} \xi_1 & \cdots & \xi_n \end{bmatrix}^\top$ and $A = \sum_j a_j \otimes E_{rj}$, then

$$\begin{aligned} A^*A &= \sum_{k,j} (a_k \otimes E_{rk})^*(a_j \otimes E_{rj}) \\ &= \sum_{k,j} (a_k^* \otimes E_{kr})(a_j \otimes E_{rj}) \\ &= \sum_{k,j} a_k^* a_j \otimes E_{kj}. \end{aligned}$$

Then

$$\begin{aligned} \langle \phi^{(n)}(A^*A)\xi, \xi\rangle &= \sum_{k,j}\sum_{r,s} \langle (\phi(a_k^* a_j) \otimes E_{kj})(\xi_r \otimes e_r), \xi_s \otimes e_s\rangle \\ &= \sum_{k,j}\sum_{s} \phi(a_k^* a_j)\xi_k, \xi_j\rangle. \end{aligned}$$

(13.2.9) Show that if $X \in M_n(\mathcal{A})$ then we can write $X = X_1 + \cdots + X_n$—where X_k is the matrix such that its k^{th} row is that of X, and the rest of the rows are zero—and then $X^*X = \sum_k X_k^* X_k$.

Answer. We have $X_k = \sum_j X_{kj} \otimes E_{kj}$, so $X = \sum_k X_k$. And

$$\begin{aligned} X^*X = \sum_{r,s} X_r^* X_s &= \sum_{r,s}\sum_{j,h} (X_{rh} \otimes E_{rh})^*(X_{sj} \otimes E_{sj}) \\ &= \sum_{r,s}\sum_{j,h} X_{rh}^* X_{sj} \otimes E_{hr}E_{sj} \\ &= \sum_{r}\sum_{j,h} (X_{rh} \otimes E_{rh})^*(X_{rj} \otimes E_{rj}) \\ &= \sum_{r} X_r^* X_r. \end{aligned}$$

(13.2.10) In the proof of Proposition 13.2.12, show that $\|X\| = \|\xi\|$.

Answer. We have

$$\|X\|^2 = \|X^*X\| = \left\| \begin{bmatrix} \sum_j |c_j|^2 & 0 \\ 0 & 0 \end{bmatrix} \right\| = \sum_j |c_j|^2 = \|\xi\|^2,$$

where we are using that $\|E_{11}\| = 1$ (since it is a projection).

(13.2.11) Show that if $A \in M_n(\mathbb{C})$ and $B \in M_n(\mathcal{S})$, then $AB \in M_n(\mathcal{S})$. Show an example where $A, B \in M_n(\mathcal{S})$ and $AB \notin M_n(\mathcal{S})$.

Answer. Because A is scalar, the k, j entry in AB is $(AB)_{kj} = \sum_j A_{kh}B_{hj}$, a linear combination of elements in $\mathcal{S}$. Hence $AB \in M_n(\mathcal{S})$.

For the example, consider the operator system

$$\mathcal{S} = \{\alpha + \beta t : \ \alpha, \beta \in \mathbb{C}\} \subset C[0, 1]\}.$$

Already for $n = 1$ we have, with $g(t) = t$, $g \in \mathcal{S}$ but $g^2 \notin \mathcal{S}$. We can make this look like matrices with

$$A = \begin{bmatrix} g & 0 \\ 0 & 0 \end{bmatrix} \in M_2(\mathcal{S}), \qquad \text{while} \qquad A^2 = \begin{bmatrix} g^2 & 0 \\ 0 & 0 \end{bmatrix} \notin M_2(\mathcal{S}).$$

(13.2.12) Show that the set K_{00} in the proof of Stinespring's Theorem 13.2.15 is actually a subspace. *(Hint: Cauchy-Schwarz)*

Answer. Suppose that $\bar{\xi}, \bar{\eta} \in \mathcal{K}_{00}$ and $\alpha \in \mathbb{C}$. It was established in (13.7) that $\langle \cdot, \cdot \rangle_{\mathcal{K}}$ is a positive sesquilinear form, so Cauchy–Schwarz applies. Then

$$|\langle \bar{\xi}, \bar{\eta} \rangle|^2 \leq \langle \bar{\xi}, \bar{\xi} \rangle \, \langle \bar{\eta}, \bar{\eta} \rangle = 0.$$

Thus

$$\langle \bar{\xi} + \alpha\bar{\eta}, \bar{\xi} + \alpha\bar{\eta} \rangle = \langle \bar{\xi}, \bar{\xi} \rangle + |\alpha|^2 \langle \bar{\eta}, \bar{\eta} \rangle + 2\mathrm{Re}\, \bar{\alpha} \, \langle \bar{\xi}, \bar{\eta} \rangle = 0.$$

This shows that $\tilde{\xi} + \alpha\tilde{\eta} \in K_{00}$. The argument shows also that

$$K_{00} = \{\tilde{\xi} : \ \langle \tilde{\xi}, \tilde{\eta} \rangle = 0, \quad \text{for all } \tilde{\eta} \in \mathcal{A} \otimes \mathcal{H}\}.$$

(13.2.13) Prove the non-unital case and the uniqueness, up to unitary conjugation, of the Stinespring's Dilation (Theorem 13.2.15).

Answer. Suppose that $(\mathcal{K}_1, \pi_1, V_1)$ is another minimal Stinespring triple for ϕ. So $\mathcal{K}_1 = \overline{\pi_1(\mathcal{A})V_1\mathcal{H}}$, $\mathcal{K} = \overline{\pi(\mathcal{A})V\mathcal{H}}$. Define a map $W : \mathcal{K} \to \mathcal{K}_1$ by

$$W\pi(a)V\xi = \pi_1(a)V_1\xi.$$

We first need to check this is well defined: if $\pi(a)V\xi = \pi(b)V\eta$, then

$$\begin{aligned}
\|\pi_1(a)V_1\xi - \pi_1(b)V_1\eta\|^2 &= \langle \pi_1(a)V_1\xi, \pi_1(a)V_1\xi\rangle + \langle \pi_1(b)V_1\eta, \pi_1(b)V_1\eta\rangle \\
&\quad - 2\mathrm{Re}\,\langle \pi_1(a)V_1\xi, \pi_1(b)V_1\eta\rangle \\
&= \langle V_1^*\pi_1(a^*a)V_1\xi, \xi\rangle + \langle V_1^*\pi_1(b^*b)V_1\eta, \eta\rangle \\
&\quad - 2\mathrm{Re}\,\langle V_1^*\pi_1(b^*a)V_1\xi, \eta\rangle \\
&= \langle \phi(a^*a)\xi, \xi\rangle + \langle \phi(b^*b)\eta, \eta\rangle - 2\mathrm{Re}\,\langle \phi(b^*a)\xi, \eta\rangle \\
&= \langle V^*\pi(a^*a)V\xi, \xi\rangle + \langle V^*\pi(b^*b)V\eta, \eta\rangle \\
&\quad - 2\mathrm{Re}\,\langle V^*\pi(b^*a)V\xi, \eta\rangle \\
&= \langle \pi(a)V\xi, \pi(a)V\xi\rangle + \langle \pi(b)V\eta, \pi(b)V\eta\rangle \\
&\quad - 2\mathrm{Re}\,\langle \pi(a)V\xi, \pi(b)V\eta\rangle \\
&= \|\pi(a)V\xi - \pi(b)V\eta\|^2 = 0,
\end{aligned}$$

so $\pi_1(a)V_1\xi = \pi_1(b)V_1\eta$. Also, the same computation but with $b = 0$ shows that

$$\|W\pi(a)V\xi\| = \|\pi_1(a)V_1\xi\| = \overline{\|\phi(a^*a)\xi\|} = \|\pi(a)V\xi\|,$$

and hence W is an isometry. From $\mathcal{K}_1 = \overline{\pi_1(\mathcal{A})V_1\mathcal{H}}$ the range of W is dense; as W is an isometry its range is also closed, and thus W is a unitary. Finally, we show that the unitary W conjugates one Stinespring triple into the other: by construction, $W\mathcal{K} = \mathcal{K}_1$, and

$$\begin{aligned}
W\pi(a)\,\pi(b)V\xi &= W\pi(ab)V\xi = \pi_1(ab)V_1\xi \\
&= \pi_1(a)\,\pi_1(b)V_1\xi = \pi_1(a)W\pi(b)V\xi;
\end{aligned}$$

as the elements of the form $\pi(b)V\xi$ are dense in $\mathcal{K}$, we get that $W\pi(a) = \pi_1(a)W$ and hence $W\pi(a)W^* = \pi_1(a)$ for all $a \in \mathcal{A}$. Assuming that $\mathcal{A}$ is unital, for all $\xi \in \mathcal{H}$ we have

$$WV_1\xi = W\pi_1(I_{\mathcal{A}})V_1\xi = \pi(I_{\mathcal{A}})V\xi = V\xi.$$

So $WV_1 = V$. If $\mathcal{A}$ is non-unital we use that $\pi(e_j) \xrightarrow{\text{sot}} I_{\mathcal{H}}$ (Exercise 12.1.23). Then

$$WV_1\xi = \lim_j WV_1\pi_1(e_j)\xi = \lim_j V\pi(e_j)\xi = V\xi;$$

so again we get $WV_1 = V$.

When $\mathcal{A}$ is not unital, the argument in the proof of Theorem 13.2.15 can still be carried by using an approximate unit. Indeed, if $\{e_j\} \subset \mathcal{A}$ is an approximate unit, then the net $\{e_j \otimes \xi\}$ is weakly convergent in $\mathcal{K}$; for

$$\langle e_j \otimes \xi - e_k \otimes \xi, b \otimes \xi\rangle = \langle \phi(b^*(e_j - e_k))\xi, \xi\rangle \to \langle \phi(b^* - b^*)\xi, \xi\rangle = 0.$$

And for arbitrary $\eta \in \mathcal{K}$, given $\varepsilon > 0$ there exists $\eta_0 = \sum_\ell b_\ell \otimes \eta_\ell$ with $\|\eta - \eta_0\| < \varepsilon$. Then

$$\begin{aligned}
|\langle e_j \otimes \xi - e_k \otimes \xi, \eta\rangle| &\leq |\langle e_j \otimes \xi - e_k \otimes \xi, \eta_0\rangle| + |\langle e_j \otimes \xi - e_k \otimes \xi, \eta - \eta_0\rangle| \\
&\leq |\langle e_j \otimes \xi - e_k \otimes \xi, \eta_0\rangle| + \|\eta - \eta_0\| \, \|(e_k - e_j) \otimes \xi\| \\
&\leq |\langle e_j \otimes \xi - e_k \otimes \xi, \eta_0\rangle| + \varepsilon \, \langle \phi((e_k - e_j)^2)\xi, \xi\rangle^{1/2} \\
&\leq |\langle e_j \otimes \xi - e_k \otimes \xi, \eta_0\rangle| + 2\varepsilon \, \|\varphi\|^{1/2} \, \|\xi\|.
\end{aligned}$$

If now $\eta_1, \ldots, \eta_r \in \mathcal{K}$ are given, the above computation shows that the net $\{\langle e_j \otimes \xi, \eta\rangle\}$ is Cauchy in $\mathbb{C}$. So a limit in $\mathbb{C}$ exists. Doing this for every η we deduce that $\gamma(\eta) = \lim_j \langle \eta, e_j \otimes \xi\rangle$ defines a linear functional on $\mathcal{H}$. It is clearly bounded, for $\|e_j \otimes \xi\| \leq \|\xi\|$ for all j, so $\|\gamma\| \leq \|\xi\|$. By Riesz Representation (Theorem 4.5.4) there exists $\tilde{\xi} \in \mathcal{K}$ with $\tilde{\xi} = \lim_{\text{weak}} e_j \otimes \xi$. We define $V\xi = \tilde{\xi}$. The only moment where we used the definition of V was to check the formula $V^*\pi(a)V = \phi(a)$. In this case we can do (note that there is no double limit below, just two limits applied one after the other)

$$\begin{aligned}
\langle V^*\pi(a)V\xi, \eta\rangle = \langle \pi(a)V\xi, V\eta\rangle = \langle \pi(a)\tilde{\xi}, \tilde{\eta}\rangle &= \lim_j \lim_k \langle \pi(a)(e_j \otimes \xi), e_k \otimes \eta\rangle \\
&= \lim_j \lim_k \langle \pi(e_k a e_j)\xi, \eta\rangle = \lim_j \langle \phi(a e_j)\xi, \eta\rangle = \langle \phi(a)\xi, \eta\rangle.
\end{aligned}$$

For the norm of V, we have

$$\begin{aligned}
\|V\xi\|^2 = \langle V\xi, V\xi\rangle &= \lim_j \lim_k \langle e_j \otimes \xi, e_k \otimes \xi\rangle \\
&= \lim_j \lim_k \langle \phi(e_k e_j)\xi, \xi\rangle \leq \|\phi\| \, \|\xi\|^2
\end{aligned}$$

for all $\xi \in \mathcal{H}$, so $\|V\|^2 \leq \|\phi\|$. Conversely, given $\varepsilon > 0$ choose $a \in \mathcal{A}$ with $\|a\| = 1$ and $\|\phi\| \leq \varepsilon + \|\phi(a)\|$. Now choose $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ and $\|\phi(a)\| \leq \varepsilon + \|\phi(a)\xi\|$. Then

$$\begin{aligned}
\|\phi\| &\leq 2\varepsilon + \|\phi(a)\xi\| = 2\varepsilon + \|V^*\pi(a)V\xi\| = 2\varepsilon + \langle V^*\pi(a^*)VV^*\pi(a)V\xi, \xi\rangle^{1/2} \\
&\leq 2\varepsilon + \|V\| \, \langle V^*\pi(a^*a)V\xi, \xi\rangle^{1/2} = 2\varepsilon + \|V\| \, \langle \pi(a^*a)V\xi, V\xi\rangle^{1/2} \\
&\leq 2\varepsilon + \|V\| \, \|\pi(a^*a)\|^{1/2} \, \|V\xi\| \leq 2\varepsilon + \|V\|^2.
\end{aligned}$$

As ε was arbitrary we get the reverse inequality and then $\|V\|^2 = \|\phi\|$.

(13.2.14) Let $\Phi : \mathcal{A} \to \mathcal{B}$ be a linear map. Let $\gamma : \mathcal{B} \to \mathcal{B}(\mathcal{H})$ be a faithful representation and $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ a decomposition of $\mathcal{H}$ such that $\gamma(\mathcal{B})\mathcal{H}_j \subset \mathcal{H}_j$. Show that Φ is completely positive if and only if each restriction $P_j(\gamma \circ \Phi)P_j$ is completely positive, where $P_j : \mathcal{H} \to \mathcal{H}_j$ is the canonical projection.

Answer. If Φ is completely positive, then $P_j(\gamma \circ \Phi)P_j$ is completely positive since it is a composition of completely positive maps.

Conversely, suppose that $P_j(\gamma \circ \Phi)P_j$ is completely positive for each j. Fix $a_1, \ldots, a_n \in \mathcal{A}$ and $\xi_1, \ldots, \xi_n \in \mathcal{H}$. For each ξ_k we have (since the P_j are pairwise orthogonal projections) a decomposition $\xi_k = \sum_j P_j \xi_k$ (Exercise 12.1.22). Then (since $\gamma(\Phi(a))P_j = P_j\gamma(\Phi(a))$ for all $a \in \mathcal{A}$)

$$\begin{aligned}
\sum_{k,h=1}^{n} \langle \gamma(\Phi(a_h^* a_k))\xi_k, \xi_h \rangle &= \sum_{g,j} \sum_{k,h=1}^{n} \langle \gamma(\Phi(a_h^* a_k))P_j\xi_k, P_g\xi_h \rangle \\
&= \sum_{g,j} \sum_{k,h=1}^{n} \langle P_g\gamma(\Phi(a_h^* a_k))P_j\xi_k, \xi_h \rangle \\
&= \sum_{j} \sum_{k,h=1}^{n} \langle P_j\gamma(\Phi(a_h^* a_k))P_j\xi_k, \xi_h \rangle \\
&= \sum_{j} \sum_{k,h=1}^{n} \langle P_j\gamma(\Phi(a_h^* a_k))P_j\xi_k, \xi_j \rangle \geq 0.
\end{aligned}$$

By Lemma 13.2.10 $\gamma \circ \Phi$ is completely positive, and then Φ is completely positive since γ^{-1} is.

(13.2.15) Show that the transpose map $\phi_T : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ is positive, unital, and contractive.

Answer. We have $\phi_T(A^*A) = (A^*A)^\top = A^\top(A^\top)^* \geq 0$, so ϕ_T is positive.

Also, $\phi_T(I_2) = I_2^\top = I_2$, so ϕ_T is unital. As for the norm,

$$\|\phi_T(A)\|^2 = \|A^\top\|^2 = \|(A^\top)^*(A^\top)\| = \|(AA^*)^\top\| = \|AA^*\| = \|A\|^2.$$

(13.2.16) Work out the missing details in Example 13.2.20. Namely, show that the matrix inside $\phi_T^{(2)}$ is positive, while its image is not positive.

Answer. Let
$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$
As A is real and symmetric, it is selfadjoint. Also, $A^2 = 2A$, so $A = (A/\sqrt{2})^2 \geq 0$. Or, we deduce from $A^2 = 2A$ that $\sigma(A) = \{0, 2\}$.

The other matrix is
$$B = \phi_T^{(2)}(A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$
One can compute directly that $B(e_2 - e_3) = -(e_2 - e_3)$, which shows that $-1 \in \sigma(B)$. Hence B is not positive.

(13.2.17) Let $\varphi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be contractive and completely positive with minimal Stinespring triple $(\pi, \mathcal{K}, V)$. Use ideas from the proof of Lemma 13.2.37 to show there exists a $*$-homomorphism $\rho : \varphi(\mathcal{A})' \to \pi(\mathcal{A})' \subset \mathcal{B}(\mathcal{K})$ that satisfies
$$\varphi(a)T = V^*\pi(a)\rho(T)V, \qquad a \in \mathcal{A},\ T \in \varphi(A)'.$$

Answer. Given $T \in \varphi(\mathcal{A})'$ we define for $a \in \mathcal{A}$ and $\xi \in \mathcal{H}$
$$\rho(T)\pi(a)V\xi = \pi(a)VT\xi$$
and extended by linearity. If we get that $\rho(T)$ is well-defined and bounded, then
$$\rho(T_1T_2)\pi(a)V\xi = \pi(a)VT_1T_2\xi = \rho(T_1)\pi(a)VT_2\xi = \rho(T_1)\rho(T_2)\pi(a)V\xi,$$
and
$$\begin{aligned} \langle \rho(T^*)\pi(a)V\xi, \pi(b)V\eta\rangle &= \langle \pi(a)VT^*\xi, \pi(b)V\eta\rangle = \langle \xi, TV^*\pi(a)^*\pi(b)V\eta\rangle \\ &= \langle \xi, T\varphi(a^*b)\eta\rangle = \langle \xi, \varphi(a^*b)T\eta\rangle \\ &= \langle \pi(a)V\xi, \pi(b)VT\eta\rangle = \langle \pi(a)V\xi, \rho(T)\pi(b)V\eta\rangle \\ &= \langle \rho(T)^*\pi(a)V\xi, \pi(b)V\eta\rangle. \end{aligned}$$

As this can be done for all $a, b \in \mathcal{A}$ and all $\xi, \eta \in \mathcal{H}$ these equalities survive sums, and so $\rho(T^*) = \rho(T)$. Now we check that $\rho(T)$ is well-defined and bounded. We have

$$\begin{aligned}
\left\|\rho(T)\left(\sum_j \pi(a_j)VT\xi_j\right)\right\|^2 &= \left\|\sum_j \pi(a_j)VT\xi_j\right\|^2 \\
&= \sum_{k,j} \langle T^*V^*\pi(a_j^*a_k)VT\xi_k, \xi_j\rangle \\
&= \sum_{k,j} \langle T^*\varphi(a_j^*a_k)T\xi_k, \xi_j\rangle \\
&= \langle (\tilde{T})^*\varphi^{(m)}(A^*A)\tilde{T}\tilde{\xi}, \tilde{\xi}\rangle \\
&= \langle \varphi^{(m)}(A^*A)^{1/2}\,(\tilde{T})^*\tilde{T}\,\varphi^{(m)}(A^*A)^{1/2}\tilde{\xi}, \tilde{\xi}\rangle \\
&\leq \|T\|^2\,\langle \varphi^{(m)}(A^*A)\tilde{\xi}, \tilde{\xi}\rangle \\
&= \|T\|^2 \left\|\sum_j \pi(a_j)V\xi_j\right\|^2.
\end{aligned}$$

where $A = \begin{bmatrix} a_1 & \cdots & a_m\end{bmatrix}$, $\tilde{\xi} = \begin{bmatrix}\xi_1 & \cdots & \xi_m\end{bmatrix}^\top$, and $\tilde{T} \in M_m(\varphi(\mathcal{A})')$ is the matrix with T in the diagonal and zeros elsewhere. This shows that $\rho(T)$ is bounded; it also shows that it is well-defined, for if an element in $\pi(\mathcal{A})V\mathcal{H}$ is represented in two different ways we can take the difference and then the above computation shows that $\rho(T)$ is zero on the difference.

Being bounded on a dense subspace, ρ extends uniquely to an operator $\rho(T) \in \mathcal{B}(\mathcal{K})$, and from

$$\begin{aligned}
\rho(T)\pi(a)\,\pi(b)V\xi = \rho(T)\pi(ab)V\xi = \pi(ab)VT\xi = \pi(a)\,\pi(b)VT\xi \\
= \pi(a)\rho(T)\,\pi(b)V\xi,
\end{aligned}$$

we conclude that $\rho(T) \in \pi(\mathcal{A})'$.

(13.2.18) Let $X \in M_n(\mathcal{A})^+$. Show that if $X_{kk} = 0$ for some k, then $X_{kj} = X_{jk} = 0$ for all $j = 1, \ldots, n$.

Answer. Since X is positive, we have $X = Y^*Y$ for some $Y \in M_n(\mathcal{A})$. Then

$$0 = X_{kk} = \sum_{j=1}^n (Y^*)_{kj}Y_{jk} = \sum_{j=1}^n \overline{Y_{jk}}Y_{jk} = \sum_{j=1}^n |Y_{jk}|^2$$

It follows that $Y_{jk} = 0$ for all $j = 1, \ldots, n$. Now, for any j,

$$X_{kj} = \sum_{\ell=1}^{n} \overline{Y_{\ell k}} Y_{jk} = 0.$$

And here is a different argument:

$$|X_{jk}| = |\langle Xe_k, e_j|\rangle \leq \|Xe_k\| \, \|e_j\| \leq \|X^{1/2}\| \, \|X^{1/2}e_k\|$$
$$= \|X^{1/2}\| \, \langle Xe_k, e_k\rangle = \|X^{1/2}\| \, X_{kk} = 0.$$

In both cases, we have $X_{kj} = \overline{X_{jk}}$, so it is enough to show that one of them is zero.

(13.2.19) Let $\begin{bmatrix} x & y \\ y^* & z \end{bmatrix} \in M_2(\mathcal{A})$ with $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} x & y \\ y^* & z \end{bmatrix} \leq \begin{bmatrix} I_{\mathcal{A}} & 0 \\ 0 & 0 \end{bmatrix}$. Show that $y = z = 0$. Use the same idea to conclude that if $X \in M_n(\mathcal{A})$ and $0 \leq X \leq I_{\mathcal{A}} \otimes E_{kk}$, then $X = a \otimes E_{kk}$ for some $a \in \mathcal{A}$.

Answer. The second inequality is

$$\begin{bmatrix} I_{\mathcal{A}} & -y \\ -y^* & -z \end{bmatrix} \geq 0.$$

The 2,2 entry gives that $-z \geq 0$, while the first inequality $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} x & y \\ y^* & z \end{bmatrix}$ gives $z \geq 0$. Thus $z = 0$ and now Exercise 13.2.18 gives us $y = 0$.

For $X \in M_n(\mathcal{A})$, the idea is the same. We write $X = \sum_{k,j} x_{kj} \otimes E_{kj}$. Then, for any h,

$$\langle x_{hh}\xi, \xi\rangle = \langle X(\xi \otimes \xi_h), \xi \otimes \xi_h\rangle \geq 0,$$

so $x_{hh} \geq 0$. We also have, for any $\xi \in \mathcal{H}$ and $h \neq k$,

$$0 \leq \langle (I_{\mathcal{A}} \otimes E_{kk} - X)(\xi \otimes \xi_h), (\xi \otimes \xi_h\rangle = -\langle x_{hh}\xi, \xi\rangle,$$

so $-x_{hh} \geq 0$. It follows that $x_{hh} = 0$ for all $h \neq k$. Now we can repeat the argument as in Exercise 13.2.18: for $\xi \in H$ with $\|\xi\| = 1$,

$$\langle x_{jh}\xi, \eta\rangle = \langle X(\xi \otimes \xi_h), \xi \otimes \xi_j\rangle \leq \|X(\xi \otimes \xi_h\|^2$$
$$\leq \|X^{1/2}\|^2 \, \|X^{1/2}(\xi \otimes \xi_h)\|^2$$
$$= \|X\| \, \langle X(\xi \otimes \xi_h), \xi \otimes \xi_h\rangle$$
$$= \|X\| \, \langle x_{hh}\xi, \xi\rangle = 0.$$

Then $x_{jh} = 0$ for all $h \neq k$ and all j. Since $X^* = X$, we also get $x_{hj} = 0$. Thus only x_{kk} is possibly nonzero, and then $X = x_{kk} \otimes E_{kk}$.

(13.2.20) Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be unital and 2-positive. Show that ϕ is bounded and $\|\phi\| = 1$.

Answer. Fix $a \in \mathcal{A}$ with $\|a\| = 1$. Then $a^*a \leq I_{\mathcal{A}}$ and, using Kadison's Schwarz inequality (13.8),

$$\|\phi(a)\|^2 = \|\phi(a)^*\phi(a)\| \leq \|\phi(a^*a)\| \leq \|\phi(I_{\mathcal{A}})\| = \|I_h\| = 1.$$

Thus, ϕ is bounded and $\|\phi\| \leq 1$. As $I_{\mathcal{H}} = \phi(I_{\mathcal{A}})$, we have $\|\phi\| = 1$.

(13.2.21) Let $\mathcal{A}$ be a non-unital C*-algebra and $f \in S(\mathcal{A})$ pure. Show that the unique extension $\tilde{f}$ of f to $\tilde{\mathcal{A}}$ (which exists by Proposition 11.5.6) is pure.

Answer. By Proposition 13.2.41 it is enough to show that $\tilde{f}$ is extreme. Suppose that $\tilde{f} = tg + (1-t)h$ for $g, h \in S(\tilde{\mathcal{A}})$ and $t \in [0,1]$. By restriction to $\mathcal{A}$ and the fact that f is pure, we get that $g|_{\mathcal{A}} = h|_{\mathcal{A}} = f$. Then $g(a, \lambda) = g(a, 0) + \lambda = f(a) + \lambda = \tilde{f}(a, \lambda)$, and similarly for h. Thus $g = h = \tilde{f}$, and so $\tilde{f}$ is pure.

(13.2.22) Show that, using the matrix units $\{E_{kj}\}$ as the basis of $M_n(\mathbb{C})$, the basis can be ordered in such a way that the matrix representation of the multiplication operator $M_X : Y \longmapsto XY$ is $X \otimes I_n$.

Answer. The $(n^2 \times n^2)$ matrix of M_X is obtained via the equation

$$XE_{k,j} = \sum_{h,\ell} X_{k,j,h,\ell} E_{h\ell}. \tag{AB.13.1}$$

As $X = \sum_{h,\ell} X_{h,\ell} E_{h\ell}$, we obtain

$$XE_{kj} = \sum_{h} X_{h,k} E_{hj}. \tag{AB.13.2}$$

Comparing (AB.13.1) with (AB.13.2), we get

$$X_{k,j,h,\ell} = \delta_{j,\ell}\, X_{h,k}.$$

So if we see the matrix of M_X as a block matrix with the blocks indexed by j, ℓ, we will have a copy of X in each diagonal block: $M_X \simeq X \otimes I_n$.

To visualize this more concretely, consider the case $n = 2$. We identify $M_2(\mathbb{C})$ with $\mathbb{C}^4$ by

$$Y \longmapsto \begin{bmatrix} Y_{11} \\ Y_{21} \\ Y_{12} \\ Y_{22} \end{bmatrix},$$

and then $Y \longmapsto XY$ is achieved by

$$\begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} Y_{11} \\ Y_{21} \\ Y_{12} \\ Y_{22} \end{bmatrix}.$$

(13.2.23) Let $\{\phi_j\} \subset CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$ be a bounded net. Show that $\phi_j \xrightarrow{BW} \phi$ if and only if $\langle \phi_j(x)\xi, \xi\rangle \to \langle \phi(x)\xi, \xi\rangle$ for all $\xi \in \mathcal{H}$ and $x \in \mathcal{S}$.

Answer. Let $c > 0$ with $\|\varphi_j\| \leq c$ for all j.

Suppose first that $\phi_j \xrightarrow{BW} \phi$ and $\xi \in \mathcal{H}$. Let P be the rank-one projection with $P\xi = \xi$. Then

$$\langle \varphi_j(x)\xi, \xi\rangle = \operatorname{Tr}(P\varphi_j(x)P) = \operatorname{Tr}(P\varphi_j(x)) \to \operatorname{Tr}(P\varphi(x)) = \langle \varphi(x)\xi, \xi\rangle.$$

Conversely, if $\langle \phi_j(x)\xi, \xi\rangle \to \langle \phi(x)\xi, \xi\rangle$ for all ξ then using polarization we get that $\langle \phi_j(x)\xi, \eta\rangle \to \langle \phi(x)\xi, \eta\rangle$ for all $\xi, \eta \in \mathcal{H}$. From this we get that $\|\varphi(x)\| \leq c\|x|$ and that $\operatorname{Tr}(S\varphi_j(x)) \to \operatorname{Tr}(S\varphi(x))$ for all finite-rank S. Given $S \in \mathcal{T}(\mathcal{H})$, by Proposition 10.7.9 there exist finite-rank operators $\{S_n\}$ with $\|S - S_n\|_1 \to 0$. Then

$$\begin{aligned} |\operatorname{Tr}(S(\varphi_j(x) - \varphi(x)))| &\leq |\operatorname{Tr}(S_n(\varphi_j(x) - \varphi(x)))| + \|S - S_n\|_1 \, \|\varphi_j(x) - \varphi(x)\| \\ &\leq |\operatorname{Tr}(S_n(\varphi_j(x) - \varphi(x)))| + 2c\|x\| \, \|S - S_n\|_1. \end{aligned}$$

Then $\limsup_j |\operatorname{Tr}(S(\varphi_j(x) - \varphi(x)))| \leq 2c\|x\| \, \|S - S_n\|_1$. As we can do this for all n, the the limit exists and is zero by the Limsup Routine.

(13.2.24) Let $P \in \mathcal{B}(\mathcal{H})$ be a finite-rank projection, $n = \operatorname{Tr}(P)$. Show that $P\mathcal{B}(\mathcal{H})P \simeq M_n(\mathbb{C})$ as C*-algebras.

Answer. Since P has rank n, there exist n pairwise orthogonal rank-one projections $P_1, \ldots, P_n$ with $\sum_j P_j = P$. Let $\{\xi_j\} \subset \mathcal{H}$ unit vectors with $P_j\xi_j = \xi_j$ (hence orthonormal) and put $V_{1j}\xi = \langle \xi, \xi_j\rangle\xi_1$. Then $(V_{1j})^*V_{1j} =$

P_j, $V_{1j}V_{1j}^* = P_1$. If we define

$$V_{kj} = V_{1k}^* V_{1j},$$

we get a system of matrix units. Let $\rho : M_n(\mathbb{C}) \to P\mathcal{B}(\mathcal{H})P$ be given by $\rho(E_{kj}) = V_{kj}$ and extended by linearity. The matrix unit properties guarantee that ρ is a $*$-homomorphism. As $M_n(\mathbb{C})$ is simple, ρ is injective. Now given $T \in P\mathcal{B}(\mathcal{H})P$, we have

$$V_{1k}TV_{j1}\xi = \langle \xi, \xi_1\rangle \langle T\xi_j, \xi_k\rangle \xi_1 = \langle T\xi_j, \xi_k\rangle P_1\xi.$$

Then

$$\begin{aligned} T = PTP = \sum_{k,j} P_k T P_j &= \sum_{k,j} V_{k1}(V_{1k}TV_{j1})V_{1j} \\ &= \sum_{k,j} \langle T\xi_j, \xi_k\rangle V_{k1}P_1V_{1j} \\ &= \sum_{k,j} \langle T\xi_j, \xi_k\rangle V_{kj} = \rho\Big(\sum_{k,j} \langle T\xi_j, \xi_k\rangle E_{kj}\Big). \end{aligned}$$

So ρ is surjective, and therefore a $*$-isomorphism.

(13.2.25) Show that the composition of cp maps is cp.

Answer. This is simply the observation that $(\varphi \circ \psi)^{(n)} = \varphi^{(n)} \circ \psi^{(n)}$ and that a composition of positive maps is positive.

(13.2.26) Show that if $A \in M_n(\mathcal{S})$, i.e. $A = \sum_{k,j=1}^n a_{kj} \otimes E_{kj}$ with $a_{kj} \in \mathcal{S}$ for all k, j, then $\|A\| \le \sum_{k,j} \|a_{kj}\|$.

Answer. We have $\|A\| \le \sum_{k,j=1}^n \|a_{kj} \otimes E_{kj}\|$. And, given any $\xi \in \mathcal{H}^n$,

$$\|(a_{kj} \otimes E_{kj})\xi\| = \|a_{kj}\xi_j \otimes e_k\| \le \|a_{kj}\| \, \|\xi_j\| \le \|a_{kj}\| \, \|\xi\|.$$

(13.2.27) Let $\mathcal{A}$ be a C*-algebra and $\varphi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a contractive completely positive map. Show that φ admits a unique ucp extension to the unitization $\tilde{\mathcal{A}}$.

Answer. The extension should necessarily be $\tilde\varphi(a, \lambda) = \varphi(a) + \lambda\varphi(0, 1) = \varphi(a) + \lambda I_{\mathcal{H}}$ since $\tilde\varphi$ is required to be unital. So the only task ahead is to show that $\tilde{\varphi}$ is cp. Let $\varphi = V^*\pi V$ be a minimal Stinespring dilation, with $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{K})$ a representation and $V : \mathcal{K} \to \mathcal{H}$ linear and bounded. Since φ is contractive, we have $\|V\| = \|\phi\|^{1/2} \leq 1$. Let $Z = I_{\mathcal{H}} - V^*V \geq 0$. We have

$$\tilde{\varphi}(a, \lambda) = V^*\pi(a)V + \lambda V^*V + \lambda Z = V^*(\pi(a) + \lambda I_{\mathcal{H}})V + \lambda Z.$$

As conjugating with V is cp, the map $(a, \lambda) \longmapsto \lambda Z$ is cp (being of the form "state times fixed operator"), and the sum of cp maps is cp, it only remains to show that we can extend representations to the unitization. This was done in Exercise 11.6.10.

(13.2.28) Let $\mathcal{A} = M_n(\mathbb{C})$, and $\mathcal{B}$ the diagonal subalgebra. Show that the map $A \longmapsto \operatorname{diag}(A_{11}, \ldots, A_{nn})$ is a conditional expectation.

Answer. Denoting the map by $\mathcal{E}$, we have

$$\begin{aligned}\mathcal{E}(A + \lambda B) &= \operatorname{diag}(A_{11} + \lambda B_{11}, \ldots, A_{nn} + \lambda B_{nn}) \\ &= \operatorname{diag}(A_{11}, \ldots, A_{nn}) + \lambda \operatorname{diag}(B_{11}, \ldots, B_{nn}).\end{aligned}$$

If B is already diagonal, the diagonal of AB is $\operatorname{diag}(A_{11}B_{11}, \ldots, A_{nn}B_{nn})$. So $\mathcal{E}(AB) = \mathcal{E}(A)B$, and the other side is similar. The positivity of $\mathcal{E}$ is the fact that the diagonal of a positive matrix is positive, namely $A_{kk} = e_k^* A e_k \geq 0$.

(13.2.29) Let $\mathcal{A}$ be a unital C*-algebra and $\varphi \in \mathcal{A}^*$ a positive linear functional. Show that $a \longmapsto \varphi(a)\, I_{\mathcal{A}}$ is a conditional expectation onto $\mathbb{C}\, I_{\mathcal{A}}$.

Answer. The linearity and positivity are those of φ. If $b = \lambda I_{\mathcal{A}} \in \mathcal{B}$, then $\varphi(ab)I_{\mathcal{A}} = (\varphi(a)I_{\mathcal{A}})\lambda = (\varphi(a)I_{\mathcal{A}})b$. So the map is $\mathcal{B}$-linear, and therefore it is a conditional expectation.

(13.2.30) Write a direct proof of (ii) $\Longrightarrow$ (i) in Proposition 13.2.68.

Answer. Since $\mathcal{E}$ satisfies (13.8), it is positive. It remains to show that $\mathcal{E}(ba) = b\mathcal{E}(a)$ for all $a \in \mathcal{A}$, $b \in \mathcal{B}$ (because $\mathcal{E}$ is positive, the equality to the other side can be obtained by taking adjoints). We have $\mathcal{E}(b)^*\mathcal{E}(b) = b^*b = \mathcal{E}(b^*b)$. The proof of Theorem 13.2.29 only uses the 2-positivity to have access to Kadison's Schwarz inequality, which is a hypothesis here. Hence the proof applies, and b is in the multiplicative domain of $\mathcal{E}$. This means that $\mathcal{E}(ba) = \mathcal{E}(b)\mathcal{E}(a) = b\mathcal{E}(a)$.

(13.2.31) Show that the map $\mathcal{E}$ from Example 13.2.70 is a faithful normal conditional expectation.

Answer. For each $T \in \mathcal{B}(\mathcal{H})$, we have $F_jTF_j = \lambda_{T,j}F_j$ for some $\lambda_{T,j} \in \mathbb{C}$ (due to the minimality of F_j), and $|\lambda_{T,j}| \leq \|T\|$. The series $\sum_j F_jTF_j$ converges sot since the F_j are pairwise orthogonal (Exercise 12.1.10). So $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{A}$ is a linear map, and $\|\mathcal{E}(T)\| = \sup\{\|F_jTF_j\| : \ j\} \leq \|T\|$. If $A \in \mathcal{A}$, then

$$\mathcal{E}(A) = \sum_j F_jAF_j = \sum_j AF_j = A\,I_{\mathcal{A}} = A,$$

showing that $\mathcal{E}$ is a projection of norm 1 and thus a conditional expectation by Proposition 13.2.68.

If $\mathcal{E}(T^*T) = 0$, then $\sum_j F_jT^*TF_j = 0$. Compressing with a single F_j we get

$$0 = F_jT^*TF_j = (TF_j)^*TF_j,$$

so $TF_j = 0$. Then $T = T\,I_{\mathcal{A}} = \sum_j TF_j = 0$ and $\mathcal{E}$ is faithful.

Finally, normality. Normality in this case means that if $T_\alpha \xrightarrow{\sigma\text{-weak}} 0$ then $\mathcal{E}(T_\alpha) \xrightarrow{\sigma\text{-weak}} 0$.

Fix $\varphi \in S(\mathcal{A})$ a normal state. We want to check that $\varphi\circ\mathcal{E}$ is normal. By Proposition 12.6.3 for a linear functional it is enough to check sot-continuity on bounded sets. Suppose that $T_\alpha \xrightarrow{\text{sot}} 0$. Fix $\xi \in \mathcal{H}$ and $\varepsilon > 0$. By Exercise 12.1.10 there exists a finite set of indices F_0 such that $\sum_{j\notin F_0} \|F_j\xi\|^2 < \varepsilon$. Then

$$\begin{aligned}\|\mathcal{E}(T_\alpha)\xi\|^2 &= \Big\|\sum_j F_jT_\alpha F_j\xi\Big\|^2 = \sum_j \|F_jT_\alpha F_j\xi\|^2 \\ &\leq \sum_{j\in F_0} \|F_jT_\alpha F_j\xi\|^2 + \varepsilon.\end{aligned}$$

Then $\limsup_\alpha \|\mathcal{E}(T_\alpha)\| \leq \varepsilon$, and the Limsup Routine guarantees $\mathcal{E}(T_\alpha) \xrightarrow{\text{sot}} 0$. Then $\varphi\circ\mathcal{E}(T_\alpha) \to 0$. Hence $\varphi\circ\mathcal{E}$ is normal. Using again Proposition 12.6.3, this means that if $T_\alpha \xrightarrow{\sigma\text{-weak}} 0$, then $\varphi(\mathcal{E}(T_\alpha)) \to 0$. Gives any $S \in \mathcal{T}(\mathcal{H})$ we can take $\varphi = \mathrm{Tr}(S\cdot)$, and so we have shown that $\mathrm{Tr}(S\mathcal{E}(T_\alpha)) \to 0$ for all $S \in \mathcal{T}(\mathcal{H})$, which is saying precisely that $\mathcal{E}(T_\alpha) \xrightarrow{\sigma\text{-weak}} 0$.

(13.2.32) Show that if on $\ell^\infty(\mathbb{Z})$ we consider the states

$$\varphi_N(a) = \frac{1}{2N+1}\sum_{n=-N}^{N} a(n).$$

and φ is a weak*-accumulation point for $\{\varphi_N\}_N$, then φ is an invariant mean. That is, if $b(n) = a(n+m)$ for all n (that is, b is a translate of a), show that $\varphi(b) = \varphi(a)$.

Answer. Let $\{N_j\}$ be a net of integers such that $\varphi_{N_j} \to \varphi$ pointwise. We have

$$\begin{aligned}|\varphi_{N_j}(b) - \varphi_{N_j}(a)| &= \frac{1}{2N_j+1}\left|\sum_{n=-N_j}^{N_j} a(n+m) - a(n)\right| \\ &\leq \frac{1}{2N_j+1}\sum_{n=-N_j}^{N_j} |a(n+m) - a(n)| \\ &= \frac{1}{2N_j+1}\left(\sum_{n=-N_j-m}^{-N_j-1} |a(n)| + \sum_{n=N_j-m+1}^{N_j} |a(n)|\right) \\ &\leq \frac{2m\|a\|_\infty}{2N_j+1} \to 0.\end{aligned}$$

Therefore $\varphi(b) = \varphi(a)$.

13.3. Group Algebras

(13.3.1) Prove Proposition 13.3.2.

Answer. Given $\gamma \in \mathcal{H}_G$, since G is a group

$$\|\lambda(g)\gamma\|^2 = \sum_{h\in G} |\gamma(g^{-1}h)|^2 = \sum_{h\in G} |\gamma(h)|^2 = \|\gamma\|^2.$$

This shows that $\lambda(g)$ is an isometry. The linearity follows from

$$\big(\lambda(g)(\alpha\gamma_1+\gamma_2)\big)(h) = (\alpha\gamma_1+\gamma_2)(g^{-1}h) = \alpha\gamma_1(g^{-1}h)+\gamma_2(g^{-1}h)$$
$$= (\alpha\lambda(g)\gamma_1)(h) + (\lambda(g)\gamma_2)(h).$$

So $\lambda(g)\in\mathcal{B}(\mathcal{H}_G)$. We have

$$(\lambda(gh)\gamma)(k) = \gamma((gh)^{-1}k) = \gamma(h^{-1}g^{-1}k) = (\lambda(h)\gamma)(g^{-1}k) = (\lambda(g)\lambda(h)\gamma)(k)$$

for all $g,h,k\in G$, so $\lambda(gh)=\lambda(g)\lambda(h)$. As $\lambda(e)=I_{\mathcal{H}_G}$ and every g is invertible in G, we get that $\lambda(g)$ is invertible with inverse $\lambda(g^{-1})$ (no need for theorems here, as $\lambda(g^{-1})$ is a bounded operator). So $\lambda(g)$ is an invertible isometry and hence a unitary.

We have, for all $g,h,k\in G$,

$$(\lambda(g)\delta_h)(k) = \delta_h(g^{-1}k) = \delta_{gh}(k).$$

Hence $\lambda(g)\delta_h = \delta_{gh}$.

(13.3.2) Prove Proposition 13.3.4.

Answer. Given $\gamma\in\mathcal{H}_G$, since G is a group

$$\|\rho(g)\gamma\|^2 = \sum_{h\in G}|\gamma(hg)|^2 = \sum_{h\in G}|\gamma(h)|^2 = \|\gamma\|^2.$$

This shows that $\rho(g)$ is an isometry. The linearity follows from

$$\big(\rho(g)(\alpha\gamma_1+\gamma_2)\big)(h) = (\alpha\gamma_1+\gamma_2)(hg) = \alpha\gamma_1(hg)+\gamma_2(hg)$$
$$= (\alpha\rho(g)\gamma_1)(h) + (\rho(g)\gamma_2)(h).$$

So $\rho(g)\in\mathcal{B}(\mathcal{H}_G)$. We have

$$(\rho(gh)\gamma)(k) = \gamma(kgh) = (\rho(h)\gamma)(kg) = (\rho(g)\rho(h)\gamma)(k)$$

for all $g,h,k\in G$, so $\rho(gh)=\rho(g)\rho(h)$. As $\rho(e)=I_{\mathcal{H}_G}$ and every g is invertible in G, we get that $\rho(g)$ is invertible with inverse $\rho(g^{-1})$ (no need for theorems here, as $\rho(g^{-1})$ is a bounded operator). So $\rho(g)$ is an invertible isometry and hence a unitary.

We have, for all $g,h,k\in G$,

$$(\rho(g)\delta_h)(k) = \delta_h(kg) = \delta_{hg^{-1}}(k).$$

Hence $\rho(g)\delta_h = \delta_{hg^{-1}}$.

(13.3.3) Let $x\in C^*_\lambda(G)$ and $y\in C^*_\rho(G)$. Show that $xy=yx$.

Answer. We have

$$\lambda(g)\rho(h)\delta_k = \delta_{gkh^{-1}} = \rho(h)\lambda(g)\delta_k.$$

By linearity and continuity, $xy\delta_k = yx\delta_k$ for all $k \in G$. Then linearity and continuity again give us $xy = yx$.

(13.3.4) Let J be given by $(J\xi)(g) = \xi(g^{-1})$. Show that $J \in \mathcal{B}(\mathcal{H}_G)$ is a unitary and $J\lambda(g)J = \rho(g)$ for all $g \in G$.

Answer. Given $\xi \in \mathcal{H}_G$,

$$(J\lambda(g)J\xi)(h) = (\lambda(g)J\xi)(h^{-1}) = (J\xi)(g^{-1}h^{-1}) = \xi(hg) = (\rho(g)\xi)(h)$$

for all $g, h \in G$. Then $J\lambda(g)J = \rho(g)$. The fact that J is a unitary follows from $J^2 = I_{\mathcal{H}_G}$ and $\|J\xi\| = \|\xi\|$, since $g \longmapsto g^{-1}$ is a bijection on G.

(13.3.5) Let G be a discrete group, $c \in \ell^2(G)$ with the property that $c * \eta \in \ell^2(G)$ for all $\eta \in \ell^2(G)$, and $T : \ell^2(G) \to \ell^2(G)$ the operator $T\eta = c * \eta$. Use the Closed Graph Theorem to show that T is bounded.

Answer. We want to use the Closed Graph Theorem (6.3.12) and Exercise 6.3.9.

Suppose that $\eta_n \to 0$ and $c*\eta_n \to \xi$. Since $\|c*\eta_n\|_\infty \leq \|c\|_2\|\eta_n\|_2 \to 0$, we have that $c * \eta_n \to 0$ pointwise. Now Proposition 7.1.20 implies that $c * \eta_n \xrightarrow{\text{weak}} 0$, so $\xi = 0$ since it's the norm limit of the sequence.

(13.3.6) Let $c, \eta \in \ell^2(G)$, $F \subset G$ and $P_F \in \mathcal{B}(\ell^2(G))$ the projection $(P_F\xi)(g) = 1_F(g)\xi(g)$. Show that $(P_Fc) * \eta = c * (P_{F^{-1}g}\eta)$.

Answer. We have, for each $g \in G$,

$$\begin{aligned}\big[(P_Fc) * \eta\big](g) &= \sum_{h\in G} (P_Fc)(h)\,\eta(h^{-1}g) = \sum_{h\in F} c(h)\,\eta(h^{-1}g)\\ &= \sum_{k\in F^{-1}g} c(gk^{-1})\,\eta(k) = \sum_{k\in G} c(gk^{-1})\,(P_{F^{-1}g}\eta)(k)\\ &= \sum_{h\in G} c(h)\,(P_{F^{-1}g}\eta)(h^{-1}g) = \big[c * (P_{F^{-1}g}\eta)\big](g).\end{aligned}$$

(13.3.7) For each $g \in G$ let P_g be the orthogonal projection onto $\mathbb{C}\,\delta_g$. Show that

$$\lambda(g)P_e\lambda(g)^* = P_g, \qquad \text{and} \qquad \rho(g)\,P_e\,\rho(g)^* = P_{g^{-1}}.$$

Answer. To avoid confusion with the canonical basis, we use the notation

$$\delta(a,b) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}$$

We have

$$\begin{aligned} \rho(g)\,P_e\,\rho(g)^*\,\delta_s &= \rho(g)\,P_e\,\rho(g^{-1})\delta_s = \rho(g)\,P_e\,\delta_{sg} \\ &= \langle \delta_s, \delta_{g^{-1}} \rangle\,\delta_{g^{-1}} = P_{g^{-1}}\delta_s. \end{aligned}$$

So P_g and $\rho(g)P_e\rho(g)^*$ agree on each element of the canonical basis and are thus equal. The other equality is similar.

(13.3.8) Show that the tracial state τ is faithful on $L(G)$.

Answer. We need to show that $\tau(\xi^* * \xi) = 0$ implies that $\xi = 0$. The adjoint of $\xi = \sum_g c_g\delta_g$ as an element of $L(G)$ is given by

$$\xi^* = \sum_g \overline{c_g}\delta_g^* = \sum_g \overline{c_g}\delta_{g^{-1}} = \sum_g \overline{c_{g^{-1}}}\delta_g.$$

Then $\xi^*(h) = \overline{c_{h^{-1}}} = \overline{\xi(h^{-1})}$. So

$$(\xi^* * \xi)(g) = \sum_h \xi^*(h)\xi(h^{-1}g) = \sum_h \overline{\xi(h^{-1})}\xi(h^{-1}g),$$

and

$$\begin{aligned} \tau(\xi^* * \xi) &= \langle (\xi^* * \xi)\delta_e, \delta_e \rangle = (\xi^* * \xi)(e) \\ &= \sum_h \overline{\xi(h^{-1})}\xi(h^{-1}) = \sum_h |\xi(h)|^2. \end{aligned}$$

Therefore, if $\tau(\xi^* * \xi) = 0$, then $\xi(h) = 0$ for all h; that is, $\xi = 0$ and τ is faithful.

(13.3.9) Write an alternative proof of (13.17) by writing each coordinate of $T\eta$.

Answer. For $g \in G$, we have

$$(T\eta)(g) = \langle T\eta, \delta_g \rangle = \sum_h \alpha_h \langle T\delta_h, \delta_g \rangle = \sum_h \alpha_h \sum_k c_k \langle \delta_{kh}, \delta_g \rangle$$
$$= \sum_h \alpha_h c_{gh^{-1}} = \sum_h c_h \alpha_{h^{-1}g}.$$

(13.3.10) Let G be a discrete group. The **full C*-algebra** of G is the completion $C^*(G)$ of $\mathbb{C}G$ via the norm

$$\Big\| \sum_{g\in G} a_g\, g \Big\| = \sup\Big\{ \Big\| \sigma\Big(\sum_{g\in G} a_g\, g \Big) \Big\| : \ \sigma \in R_G \Big\},$$

where R_G is the set of $*$-representations $\sigma : \mathbb{C}G \to \ell^2(S_n)$ where each S_n is a set of cardinality n for each $n \leq |G|$ (these convoluted choice guarantees that the set of representations is actually a set). Show that $C^*_\lambda(G)$ and $C^*_\rho(G)$ are quotients of $C^*(G)$.

Answer. By definition of the norm on $C^*(G)$, we have $\|x\|_\lambda \leq \|x\|$ for all $x \in \mathbb{C}G$. So the identity map $\mathbb{C}G \to \mathbb{C}G$, with $\|\cdot\|$ in the domain and $\|\cdot\|_\lambda$ in the codomain, is bounded. So we get a $*$-homomorphism $\beta : C^*(G) \to C^*_\lambda(G)$ with dense range (hence surjective).

The argument for $C^*_\rho(G)$ runs the same.

(13.3.11) Let G be a discrete group and $T \in \mathcal{B}(\ell^2(G))$. Show that $T \in L(G)$ if and only if T is "Toeplitz", in the sense that diagonals are constant, meaning that for $g, k, h \in G$,

$$\langle T\delta_{hg}, \delta_g \rangle = T_{g,hg} = T_{k,hk} = \langle T\delta_{hk}, \delta_h \rangle. \qquad (13.18)$$

Answer. We know that $L(G) = \lambda(G)''$. So there exists a net $\{T_j\} \subset \operatorname{span}\lambda(G)$ with $T_j \xrightarrow{\text{sot}} T$. As the equality (13.18) survives wot limits, it is enough to show that $T \in \operatorname{span}\lambda(G)$ has the property, and by linearity it is enough to show it for $\lambda(r)$ for a fixed $r \in G$. We have

$$\langle \lambda(r)\delta_{hg}, \delta_g \rangle = \langle \delta_{rhg}, \delta_g \rangle = \langle \delta_{rhk}, \delta_k \rangle,$$

since the left inner product will be 1 or 0 depending on whether $rh = e$, and the same with the right one.

For the converse, suppose that T satisfies (13.18). Fix $k \in G$. We have

$$\begin{aligned}\langle T\rho(k)\delta_g, \delta_h\rangle = \langle T\delta_{gk^{-1}}, \delta_h\rangle = \langle T\delta_{gk^{-1}}, \delta_{hkk^{-1}}\rangle = \langle T\delta_g, \delta_{(hk}\rangle \\ = \langle T\delta_g, \rho(k^{-1})\delta_h\rangle = \langle \rho(k)T\delta_g, \delta_h\rangle.\end{aligned}$$

It follows that $T \in \rho(G)' = L(G)$.

13.4. Topological Tensor Products

(13.4.1) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces with orthonormal bases $\{\xi_j\}$ and $\{\eta_k\}$, respectively. Show that $\{\xi_j \otimes \eta_k\}$ is an orthonormal basis for $\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$.

Answer. We have

$$\langle \xi_j \otimes \eta_k, \xi_r \otimes \eta_s\rangle = \langle \xi_j, \xi_r\rangle\langle \eta_k, \eta_s\rangle = \delta_{j,r}\delta_{k,s},$$

so the family $\{\xi_j \otimes \eta_k\}$ is orthonormal. We also have $\operatorname{span}\{\xi_j \otimes \eta_k : \ k, j\} = \mathcal{H} \otimes \mathcal{K}$ is dense in $\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$, so $\{\xi_j \otimes \eta_k\}$ is an orthonormal basis.

(13.4.2) Prove the isomorphisms (13.20).

Answer. With the notation from Lemma 13.4.3, let $\Gamma : \mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K} \to \bigoplus_{k\in K} \mathcal{H}$ be $\Gamma(\xi) = \bigoplus_k \zeta_k$. Let uniqueness in the lemma gives us that Γ is well-defined and injective. The linearity of Γ can be obtained either by looking at the definition of ζ_k, or out of the uniqueness of the ζ_k. Given $\tilde{\zeta} = \bigoplus_k \zeta_k \in \bigoplus_k \mathcal{H}$ we have $\sum_k \|\zeta_k\|^2 < \infty$. Then

$$\tilde{\zeta} = \Gamma\Big(\sum_k \zeta_k \otimes \eta_k\Big).$$

Thus Γ is a linear bijection. It remains to check that it preserves the inner product. For this by polarization it is enough to check that it preserves norms;

and

$$\Big\langle \Gamma\Big(\sum_k \zeta_k \otimes \eta_k\Big), \Gamma\Big(\sum_k \zeta_k \otimes \eta_k\Big)\Big\rangle = \Big\langle \bigoplus_k \zeta_k, \bigoplus_k \zeta_k \Big\rangle = \sum_k \|\zeta_k\|^2$$
$$= \Big\langle \sum_k \zeta_k \otimes \eta_k, \sum_k \zeta_k \otimes \eta_k \Big\rangle.$$

The second isomorphism is proven in the same manner, with the roles of $\mathcal{H}$ and $\mathcal{K}$ exchanged.

(13.4.3) Show that the tensor product of operators obeys the same arithmetic rules as the elementary tensors of vectors, as in Proposition 13.1.2.

Answer. The properties follow directly from the corresponding properties of vectors. For instance,

$$\big((T_1+T_2)\otimes S\big)(\xi\otimes\eta) = (T_1+T_2)\xi\otimes S\eta = T_1\xi\otimes S\eta + T_2\xi\otimes S\eta$$
$$= \big(T_1\otimes S + T_2\otimes S\big)(\xi\otimes\eta).$$

By linearity and continuity we get $(T_1+T_2)\otimes S = T_1\otimes S + T_2\otimes S$.

(13.4.4) Let $\mathcal{M}\subset\mathcal{B}(\mathcal{K})$ be a von Neumann algebra and $\mathcal{H}$ a Hilbert space. Fix an orthonormal basis $\{\eta_j\}_{j\in J}$ for $\mathcal{H}$ and consider the associated matrix units $\{E_{kj}\}$. Show that for each $\tilde{T}\in \mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H})$ there exist unique operators $\{T_{kj}\}\subset\mathcal{M}$ such that

$$\tilde{T} = \sum_{k,j} T_{kj}\otimes E_{kj},$$

where the series converges sot.

Answer.

Let $\gamma:\mathcal{M}\to(I_\mathcal{K}\otimes E_{11})(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H}))(I_\mathcal{K}\otimes E_{11})$ be given by $\gamma(T) = T\otimes E_{11}$. Since E_{11} is a projection, it is straightforward to check that γ is a $*$-homomorphism. It is injective by Proposition 13.1.3. And if $\tilde{T}\in (I_\mathcal{K}\otimes E_{11})(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H}))(I_\mathcal{K}\otimes E_{11})$, positive, define a form on $\mathcal{K}$ by

$$[\xi,\nu] = \langle \tilde{T}(\xi\otimes\eta_1), \nu\otimes\eta_1\rangle.$$

This form is sesquilinear and positive, so by Proposition 10.1.5 there exists $T\in\mathcal{B}(\mathcal{H})$ with $[\xi,\nu]=\langle T\xi,\nu\rangle$. Let $S\in\mathcal{M}'$. Then $S\otimes I_\mathcal{H}\in(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H}))'$.

We have

$$\begin{aligned}\langle TS\xi, \nu\rangle &= [S\xi, \nu] = \langle \tilde{T}(S \otimes I_{\mathcal{H}})(\xi \otimes \eta_1), \nu \otimes \eta_1\rangle \\ &= \langle (S \otimes I_{\mathcal{H}})\tilde{T}(\xi \otimes \eta_1), \nu \otimes \eta_1\rangle \\ &= \langle \tilde{T}(\xi \otimes \eta_1), S^*\nu \otimes \eta_1\rangle \\ &= [\xi, S^*\nu] = \langle T\xi, S^*\nu\rangle = \langle ST\xi, \nu\rangle.\end{aligned}$$

As this can be done for all $\xi, \nu \in \mathcal{K}$, we have that $ST = TS$. Thus $T \in \mathcal{M}'' = \mathcal{M}$. By definition of T we have

$$\begin{aligned}\langle \tilde{T}(\xi \otimes \eta_j), \xi \otimes \eta_j\rangle &= \langle (I_{\mathcal{K}} \otimes E_{11})\tilde{T}(I_{\mathcal{K}} \otimes E_{11})(\xi \otimes \eta_j), \xi \otimes \eta_j\rangle \\ &= \langle \tilde{T}(I_{\mathcal{K}} \otimes E_{11})(\xi \otimes \eta_j), (I_{\mathcal{K}} \otimes E_{11})(\xi \otimes \eta_j)\rangle \\ &= \delta_{j,1}\langle \tilde{T}(\xi \otimes \eta_1), \xi \otimes \eta_1\rangle = \delta_{j,1}\langle T\xi, \xi\rangle \\ &= \langle (T \otimes E_{11})(\xi \otimes \eta_j), \xi \otimes \eta_j\rangle.\end{aligned}$$

It follows by polarization that $\tilde{T} = T \otimes E_{11} = \gamma(T)$. As any C*-algebra is spanned by its positive elements, we get that γ is surjective and hence a bijection.

Fix $\tilde{T} \in \mathcal{M} \,\overline{\otimes}\, \mathcal{B}(\mathcal{H})$. For each $k, j \in J$, let $T_{kj} = \gamma^{-1}\big((I_{\mathcal{K}} \otimes E_{1k})\tilde{T}(I_{\mathcal{K}} \otimes E_{j1})\big)$. By Lemma 13.4.3 for each $\xi \in \mathcal{K}$ there exist unique vectors $\{\zeta_k\} \subset \mathcal{K}$ such that

$$\xi = \sum_k \zeta_k \otimes \eta_k$$

and $\sum_k \|\zeta_k\|^2 = \|\xi\|^2$. Fix $F \subset J$, finite. Then

$$\begin{aligned}
\sum_{k,j\in F} (T_{kj}\otimes E_{kj})\xi &= \sum_r \sum_{k,j\in F} T_{kj}\zeta_r \otimes E_{kj}\eta_r = \sum_{k,j\in F} T_{kj}\zeta_j\otimes \eta_k \\
&= \sum_{k,j\in F} (T_{kj}\otimes E_{kk})(\zeta_j\otimes\eta_k) \\
&= \sum_{k,j\in F} (I_{\mathcal{K}}\otimes E_{k1})(T_{kj}\otimes E_{11})(I_{\mathcal{K}}\otimes E_{1k})(\zeta_j\otimes\eta_k) \\
&= \sum_{k,j\in F} (I_{\mathcal{K}}\otimes E_{k1})(I_{\mathcal{K}}\otimes E_{1k})\tilde{T}(I_{\mathcal{K}}\otimes E_{j1})(\zeta_j\otimes\eta_1) \\
&= \sum_{k,j\in F} (I_{\mathcal{K}}\otimes E_{kk})\tilde{T}(\zeta_j\otimes\eta_j) \\
&= \sum_{k,j\in F} (I_{\mathcal{K}}\otimes E_{kk})\tilde{T}(I_{\mathcal{K}}\otimes E_{jj}))(\zeta_j\otimes\eta_j) \\
&= \sum_{k,j\in F} \sum_r (I_{\mathcal{K}}\otimes E_{kk})\tilde{T}(I_{\mathcal{K}}\otimes E_{jj}))(\zeta_r\otimes\eta_r) \\
&= (I_{\mathcal{K}}\otimes P_F)\tilde{T}(I_{\mathcal{K}}\otimes P_F)\xi.
\end{aligned}$$

Now, with a similar argument as that in Exercise 10.6.9, since the projections $I_{\mathcal{K}}\otimes P_F$ are not finite-rank but they do converge sot to $I_{\mathcal{K}}\otimes I_{\mathcal{H}}$,

$$\sum_{k,j\in F} (T_{kj}\otimes E_{kj}) \xrightarrow{\text{sot}} \tilde{T}.$$

(13.4.5) Let $\mathcal{M}\subset\mathcal{B}(\mathcal{H})$ be a non-degenerate von Neumann algebra and consider matrix units $\{E_{kj}\}_{k,j\in J}\subset\mathcal{M}$ such that $\sum_k E_{kk} = I_{\mathcal{M}}$ (with the series converging sot). Fix $j_0\in J$ and let $P = E_{j_0,j_0}$. Show that there exists a Hilbert space $\mathcal{K}$ such that $\mathcal{M}\simeq P\mathcal{M}P\,\overline{\otimes}\,\mathcal{B}(\mathcal{K})$ (we take as assumed knowledge that $P\mathcal{M}P\subset\mathcal{B}(P\mathcal{H})$ is a von Neumann algebra; this will be proven in Proposition 14.1.1).

Answer. Let $\mathcal{N} = P\mathcal{M}P$, let $\mathcal{K} = \ell^2(J)$, and let $\{G_{kj}\}_{k,j\in J}\subset\mathcal{B}(\mathcal{K})$ be the matrix units corresponding to the canonical basis. By Exercise 13.4.4 we know that any $\tilde{T}\in\mathcal{N}\,\overline{\otimes}\,\mathcal{B}(\mathcal{K})$ can be written in the form

$$\tilde{T} = \sum_{k,j} T_{kj}\otimes G_{kj}$$

for certain operators $\{T_{kj}\} \subset \mathcal{N}$. Let $W : P\mathcal{H} \overline{\otimes} \ell^2(J) \to \mathcal{H}$ be given by

$$W : \sum_j \zeta_j \otimes \delta_j \longmapsto \sum_j E_{j,j_0}\zeta_j.$$

where we are using Lemma 13.4.3 to write the elements of $P\mathcal{H} \overline{\otimes} \ell^2(J)$. The uniqueness in the lemma makes W well-defined. It is clearly linear, for all the algebra occurs on the side of the ζ_j. It is surjective, for given $\xi \in \mathcal{H}$ we have

$$\xi = \sum_j E_{jj}\xi = \sum_j E_{j,j_0}\, E_{j_0,j}\xi = W\Big(\sum_j E_{j_0,j}\xi \otimes \delta_j\Big).$$

And (using the continuity of the inner product to exchange with the series)

$$\begin{aligned}\Big\|W\Big(\sum_j \zeta_j \otimes \delta_j\Big)\Big\|^2 &= \Big\|\sum_j E_{j,j_0}\zeta_j\Big\|^2 = \sum_{k,j}\langle E_{j,j_0}\zeta_j, E_{k,j_0}\zeta_k\rangle \\ &= \sum_{k,j}\langle E_{j_0,k}E_{j,j_0}\zeta_j, \zeta_k\rangle \\ &= \sum_j \langle E_{j_0,j_0}\zeta_j, \zeta_j\rangle = \sum_j \|\zeta_j\|^2 \\ &= \Big\|\sum_j \zeta_j \otimes \delta_j\Big\|^2.\end{aligned}$$

So W is a unitary. Let $\Gamma : \mathcal{N} \otimes \mathcal{B}(\mathcal{K}) \to \mathcal{M}$ be given by

$$\Gamma\Big(\sum_{k,j\in F} T_{kj} \otimes G_{kj}\Big) = \sum_{k,j\in F} E_{k,j_0}T_{kj}E_{j_0,j},$$

where $F \in J$ is some finite subset.

Let $\tilde{T} = \sum_{k,j\in F} T_{kj} \otimes G_{kj} \in \mathcal{N} \otimes \mathcal{B}(\mathcal{K})$. For $\xi \in \mathcal{H}$,

$$\begin{aligned}\Gamma(\tilde{T})\,\xi &= \sum_{k,j\in F} E_{k,j_0}T_{kj}E_{j_0,j}\xi = \sum_{k\in F} E_{k,j_0}\Big(\sum_{j\in F} T_{kj}\, E_{j_0,j}\xi\Big) \\ &= W\Big(\sum_{k\in F}\Big(\sum_{j\in F} T_{kj}\, E_{j_0,j}\xi\Big) \otimes \delta_k\Big) \\ &= W\Big(\sum_{k\in F}\Big(\sum_{j\in F} T_{kj}\, E_{j_0,j}\xi\Big) \otimes G_{kj}\delta_j\Big) \\ &= W\Big(\sum_{k,j\in F} T_{kj} \otimes G_{kj}\ \sum_r E_{j_0,r}\xi \otimes \delta_r\Big) \\ &= W\Big(\sum_{k,j\in F} T_{kj} \otimes G_{kj}\Big)W^*\xi \\ &= W\tilde{T}W^*\xi.\end{aligned}$$

This shows that we can define $\Gamma(\tilde{T}) = W\tilde{T}W^*$ on all of $\mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$, and that $\Gamma(\tilde{T}) \in \mathcal{M}$. It remains to see that Γ is surjective. Given $T \in \mathcal{M}$, we have

$$T = \sum_{k,j} E_{kk} T E_{jj} = \sum_{k,j} E_{k,j_0} \left(E_{j_0,k} T E_{j,j_0}\right) E_{j_0,j}$$
$$= \Gamma\Big(\sum_{k,j} E_{j_0,k} T E_{j,j_0} \otimes G_{kj}\Big).$$

Hence Γ is surjective if we show that $\sum_{k,j} E_{j_0,k} T E_{j,j_0} \otimes G_{kj}$ is bounded, since $E_{j_0,k} T E_{j,j_0} \in \mathcal{N}$ for all k,j. This uses the same ideas as above. Given $\tilde{\xi} = \sum_r \xi_r \otimes \delta_r$,

$$\begin{aligned}
\Big\|\Big(\sum_{k,j} E_{j_0,k} T E_{j,j_0} \otimes G_{kj}\Big)\tilde{\xi}\Big\|^2 &= \Big\|\sum_{k,j} E_{j_0,k} T E_{j,j_0}\xi_j \otimes \delta_k\Big\|^2 \\
&= \Big\|\sum_{k} E_{j_0,k}\Big(\sum_j T E_{j,j_0}\xi_j\Big) \otimes \delta_k\Big\|^2 \\
&= \sum_k \|E_{j_0,k} T W\tilde{\xi}\|^2 \\
&= \sum_k \langle W^*T^*E_{kk}TW\tilde{\xi}, \tilde{\xi}\rangle \\
&= \langle W^*T^*TW\tilde{\xi}, \tilde{\xi}\rangle = \|TW\tilde{\xi}\|^2 \\
&\leq \|T\|^2\, \|\tilde{\xi}\|^2.
\end{aligned}$$

(13.4.6) Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Show that

$$\mathcal{K}(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2) = \overline{\mathcal{K}(\mathcal{H}_1) \otimes \mathcal{K}(\mathcal{H}_2)}.$$

Answer. Since $\mathcal{F}(\mathcal{H}_1) \otimes \mathcal{F}(\mathcal{H}_2) \subset \mathcal{F}(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2)$ we get $\overline{\mathcal{K}(\mathcal{H}_1) \otimes \mathcal{K}(\mathcal{H}_2)} \subset \mathcal{K}(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2)$ by Proposition 10.6.4.

For the reverse inclusion, by Propositions 10.6.1 and 10.6.4 it is enough to show that if $T \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is rank-one, then $T \in \overline{\mathcal{K}(\mathcal{H}_1) \otimes \mathcal{K}(\mathcal{H}_2)}$. Suppose then that $T = \tilde{\xi}\tilde{\eta}^*$ for $\tilde{\xi}, \tilde{\eta} \in \mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2$ and fix $\varepsilon > 0$. There exist $\xi, \eta \in \mathcal{H}_1 \otimes \mathcal{H}_2$ with $\|\tilde{\xi} - \xi\| < \varepsilon$ and $\|\tilde{\eta} - \eta\| < \varepsilon$. Then

$$\|(T - \xi\eta^*)\nu\| \leq \|(\tilde{\xi} - \xi)\tilde{\eta}^*\nu\| + \|\xi(\tilde{\eta} - \eta)^*\nu\| \leq \varepsilon(\|\tilde{\eta}\| + \|\tilde{\xi}\| + \varepsilon)\, \|\nu\|.$$

Thus we may assume without loss of generality that $T = \xi\eta^*$ for $\xi, \eta \in \mathcal{H}_1 \otimes \mathcal{H}_2$. And now

$$\begin{aligned} T &= \Big(\sum_{k=1}^m \xi_{1,k} \otimes \xi_{2,k}\Big)\Big(\sum_{k=1}^n \eta_{1,k} \otimes \eta_{2,k}\Big)^* \\ &= \sum_{k,j} \xi_{1,k}\eta_{1,j}^* \otimes \xi_{2,k}\eta_{2,j}^* \in \mathcal{K}(\mathcal{H}_1) \otimes \mathcal{K}(\mathcal{H}_2). \end{aligned}$$

(13.4.7) Let $S \in \mathcal{T}(\mathcal{H})$, $T \in \mathcal{T}(\mathcal{K})$. Show that $S \otimes T \in \mathcal{T}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$.

Answer. Fix orthonormal bases $\{\xi_j\}$ and $\{\eta_k\}$ for $\mathcal{H}$ and $\mathcal{K}$. We know from Exercise 13.4.1 that $\{\xi_j \otimes \eta_k\}$ is an orthonormal basis for $\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$. We have $|S \otimes T| = |S| \otimes |T|$, since $|S| \otimes |T| \geq 0$ and

$$(|S| \otimes |T|)^2 = |S|^2 \otimes |T|^2 = S^*S \otimes T^*T = (S \otimes T)^*(S \otimes T).$$

Then

$$\begin{aligned} \operatorname{Tr}(|S \otimes T|) &= \sum_{k,j} \langle (|S| \otimes |T|)(\xi_j \otimes \eta_k), \xi_j \otimes \eta_k\rangle \\ &= \sum_{k,j} \langle\langle |S|\xi_j, \xi_j\rangle \, \langle |T|\eta_k, \eta_k\rangle = \operatorname{Tr}(|S|)\operatorname{Tr}(|T|) < \infty \end{aligned}$$

(no issues with summation order by Tonelli, since everything is non-negative). So $S \otimes T \in \mathcal{T}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$.

(13.4.8) Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ be Hilbert spaces. Show that

$$(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2) \mathbin{\overline{\otimes}} \mathcal{H}_3 \simeq \mathcal{H}_1 \mathbin{\overline{\otimes}} (\mathcal{H}_2 \mathbin{\overline{\otimes}} \mathcal{H}_3)$$

canonically.

Answer. We know from Proposition 13.1.7 that the linear map $U : (\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}3 \simeq \mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3)$ induced by

$$U : (\xi \otimes \eta) \otimes \nu \longmapsto \xi \otimes (\eta \otimes \nu)$$

is well-defined. And since

$$\begin{aligned} \langle U[(\xi_1 \otimes \eta_1) \otimes \nu_1], U[(\xi_2 \otimes \eta_2) \otimes \nu_2]\rangle &= \langle \xi_1 \otimes (\eta_1 \otimes \nu_1), \xi_2 \otimes (\eta_2 \otimes \nu_2)\rangle \\ &= \langle \xi_1, \xi_2\rangle \, \langle \eta_1 \otimes \nu_1, \eta_2 \otimes \nu_2\rangle \\ &= \langle \xi_1, \xi_2\rangle \, \langle \eta_1, \eta_2\rangle \, \langle \nu_1, \nu_2\rangle \\ &= \langle (\xi_1 \otimes \eta_1) \otimes \nu_1, (\xi_2 \otimes \eta_2) \otimes \nu_2\rangle, \end{aligned}$$

together with the linearity of U this shows that U is isometric. Taking limits, U extends first to an isometry

$$(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2) \otimes \mathcal{H}_3 \simeq \mathcal{H}_1 \otimes (\mathcal{H}_2 \mathbin{\overline{\otimes}} \mathcal{H}_3),$$

and then to an isometry

$$(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2) \mathbin{\overline{\otimes}} \mathcal{H}_3 \simeq \mathcal{H}_1 \mathbin{\overline{\otimes}} (\mathcal{H}_2 \mathbin{\overline{\otimes}} \mathcal{H}_3).$$

As it has dense range, U is a unitary.

(13.4.9) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space, and $n \in \mathbb{N}$. Show that $M_n(\mathbb{C}) \otimes \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C}) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H}) \simeq \mathcal{B}(\mathcal{H})$ as von Neumann algebras.

Answer. We discussed at the beginning of Section 11.7 how $M_n(\mathcal{M})$ is complete. And $M_n(\mathcal{M}) \simeq M_n(\mathbb{C}) \otimes \mathcal{M}$ canonically by Exercise 13.1.5. Thus $M_n(\mathbb{C}) \otimes \mathcal{B}(\mathcal{H})$ is complete and therefore equal to $M_n(\mathbb{C}) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$.

As $\dim \mathcal{H} = \infty$, we can split an orthonormal basis into n sets of equal cardinality, and this way we induce a unitary $U : \mathcal{H} \to \bigoplus_{k=1}^n \mathcal{H}$. Then $\mathcal{B}(\mathcal{H}) \simeq \mathcal{B}(\bigoplus_{k=1}^n \mathcal{H})$, so we have reduced the problem to showing that

$$\mathcal{B}\Big(\bigoplus_{k=1}^n \mathcal{H}\Big) \simeq M_n(\mathcal{B}(\mathcal{H})).$$

We can achieve this by naturally interpreting a matrix in $M_n(\mathcal{B}(\mathcal{H}))$ as an operator on $\bigoplus_k \mathcal{H}$; this was done at the beginning of Section 11.7.

(13.4.10) Let $\mathcal{M}_k \subset \mathcal{B}(\mathcal{H}_k)$, $k = 1, 2, 3$, be von Neumann algebras. Show that there is a canonical isomorphism

$$(\mathcal{M}_1 \mathbin{\overline{\otimes}} \mathcal{M}_2) \mathbin{\overline{\otimes}} \mathcal{M}_3 \simeq \mathcal{M}_1 \mathbin{\overline{\otimes}} (\mathcal{M}_2 \mathbin{\overline{\otimes}} \mathcal{M}_3).$$

Answer. We know from Exercise 13.4.8 that the underlying tensor product of Hilbert spaces behaves the right way. The unitary that implements the isomorphism at the level of Hilbert spaces then gives

$$U\big[(T_1 \otimes T_2) \otimes T_3\big]U^* = T_1 \otimes (T_2 \otimes T_3)$$

and the same for any linear combination of such operators. As unitary conjugation is as continuous as any map in a von Neumann algebra will ever be, it extends to the closures and so

$$U\big[(\mathcal{M}_1 \mathbin{\overline{\otimes}} \mathcal{M}_2) \mathbin{\overline{\otimes}} \mathcal{M}_3\big]U^* = \mathcal{M}_1 \mathbin{\overline{\otimes}} (\mathcal{M}_2 \mathbin{\overline{\otimes}} \mathcal{M}_3).$$

(13.4.11) Let $\mathcal{H},\mathcal{K}$ be separable infinite-dimensional Hilbert spaces. Show that the subalgebras

$$\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K}),\qquad \mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{B}(\mathcal{K}),$$

and

$$\mathcal{B}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})$$

are three distinct ideals of $\mathcal{B}(\mathcal{H})\otimes_{\min}\mathcal{B}(\mathcal{K})$.

Answer. Let $\varphi\in S(\mathcal{B}(\mathcal{H}))$ and $\psi\in S(\mathcal{B}(\mathcal{K}))$ such that $\varphi|_{\mathcal{K}(\mathcal{H})}=0$ and $\psi|_{\mathcal{K}(\mathcal{K})}=0$. These exist because we can apply Corollary 11.5.8 to the Calkin algebra $\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ to get a nonzero state φ', and then we define $\varphi=\varphi'\circ q$, where $q:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ is the quotient map; and we do the similar thing with ψ.

By Corollary 13.4.21 we can consider the state $\varphi\times\psi\in S(\mathcal{B}(\mathcal{H})\otimes_{\min}\mathcal{B}(\mathcal{K}))$. Let $T\in\mathcal{K}(\mathcal{H})$ and $S\in\mathcal{K}(\mathcal{K})$ be nonzero and positive. Then $T\otimes I_{\mathcal{K}}\in\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{B}(\mathcal{K})$, $I_{\mathcal{H}}\otimes S\in\mathcal{B}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})$. Let also $\nu\in S(\mathcal{B}(\mathcal{H}))$ and $\varrho\in S(\mathcal{B}(\mathcal{K}))$ be given by $\nu(X)=\operatorname{Tr}(WX)$, where $W\in\mathcal{T}(\mathcal{H})$ is positive, with $\operatorname{Tr}(W)=1$, and injective; and form ϱ similarly. We have $(\nu\times\psi)(R\otimes S)=0$ for all $R\in\mathcal{K}(\mathcal{H})$ and $S\in\mathcal{K}(\mathcal{K})$, so by linearity and continuity $\nu\times\psi|_{\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})}=0$, while $(\nu\times\psi)(T\otimes I_{\mathcal{K}})=\nu(T)>0$. This shows that $\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})\subsetneq\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{B}(\mathcal{K})$, and a similar argument with ϱ and φ shows that $\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})\subsetneq\mathcal{B}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})$. Finally, $\nu\times\psi$ is nonzero on $\mathcal{K}(\mathcal{H})\otimes_{\min}\mathcal{B}(\mathcal{K})$ but zero on $\mathcal{B}(\mathcal{H})\otimes_{\min}\mathcal{K}(\mathcal{K})$, so these two are also distinct.

(13.4.12) Let $\pi:\mathcal{A}\to\mathcal{B}(\mathcal{H})$ a representation. In Definition 12.6.15 we considered the **amplification** of π given by $\tilde\pi:\mathcal{A}\to\bigoplus_{k\in K}\mathcal{H}$.

Show that there is a unitary $U:\bigoplus_{k\in K}\mathcal{H}\to\mathcal{H}\,\overline{\otimes}\,\mathcal{K}$, where $\mathcal{K}$ is a Hilbert space with $\dim\mathcal{K}=|K|$, such that $\tilde\pi=U^*(\pi\otimes I_{\mathcal{K}})U$.

Answer. We write $\bigoplus_{k\in K}\mathcal{H}=\mathcal{H}^{|K|}$ for convenience. Let $\mathcal{K}=\ell^2(K)$. Given $\tilde\xi=\{\xi_k\}\in\mathcal{H}^{|K|}$, we define

$$U\tilde\xi=\sum_{k\in K}\xi_k\otimes e_k.$$

This is linear because tensor products are linear on each component, and

$$\|U\tilde{\xi}\|^2 = \langle U\tilde{\xi}, U\tilde{\xi}\rangle = \sum_{k,j\in K} \langle \xi_k \otimes e_k, \xi_j \otimes e_j\rangle$$

$$= \sum_{k\in K} \|\xi_k\|^2 = \|\tilde{\xi}\|^2.$$

So U is an isometry. Given $K_0 \subset K$ finite and $\sum_{k\in K_0} \xi_k \otimes e_{j_k} \in \mathcal{H}\otimes\mathcal{K}$, we have

$$\sum_{k\in K_0} \xi_k \otimes e_{j_k} = U\tilde{\xi},$$

where $\tilde{\xi}(j_k) = \xi_k$ for all $k \in K_0$. This shows that U has dense range; being an isometry, it is surjective and hence a unitary. Now

$$(\pi(a)\otimes I_{\mathcal{K}})U\tilde{\xi} = (\pi(a)\otimes I_{\mathcal{K}})\sum_k \xi_k \otimes e_k = \sum_k \pi(a)\xi_k \otimes e_k$$

$$= U(\{\pi(a)\xi_k\}) = U\tilde{\pi}(a)\tilde{\xi}.$$

This can be done for all $\tilde{\xi} \in \mathcal{H}^{|K|}$, so $(\pi(a)\otimes I_{\mathcal{K}})U = U\tilde{\pi}(a)$. With U a unitary, the equality $\tilde{\pi} = U^*(\pi\otimes I_{\mathcal{K}})U$ holds.

(13.4.13) Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces and $\{P_j\}_{j\in J}\subset\mathcal{B}(\mathcal{H})$, $\{Q_k\}_{k\in K}\subset\mathcal{B}(\mathcal{K})$ increasing nets of projections. Show that $\{P_j\otimes Q_k\}\subset\mathcal{B}(\mathcal{H}\otimes\mathcal{K})$ is an increasing net of projections. Show that if $P_j \xrightarrow{\text{sot}} I_{\mathcal{H}}$ and $Q_k \xrightarrow{\text{sot}} I_{\mathcal{K}}$ then $P_j\otimes Q_k \xrightarrow{\text{sot}} I_{\mathcal{H}\overline{\otimes}\mathcal{K}}$.

Answer. We have $(P_j\otimes Q_k)^*(P_j\otimes Q_k) = P_j^*P_j\otimes Q_k^*Q_k = P_j\otimes Q_k$, so they are projections.

We order $J\times K$ by saying that $(j_1,k_1)\leq(j_2,k_2)$ if $j_1\leq j_2$ and $k_1\leq k_2$.

If $j_1\leq j_2$ and $k_1\leq k_2$ then $P_{j_1}\leq P_{j_2}$ and $Q_{k_1}\leq Q_{k_2}$. Using Proposition 10.5.3 we have $(P_{j_2}\otimes Q_{k_2})(P_{j_1}\otimes Q_{k_1}) = P_{j_2}P_{j_1}\otimes Q_{k_2}Q_{k_1} = P_{j_1}\otimes Q_{k_1}$, so again by Proposition 10.5.3 we get $P_{j_2}\otimes Q_{k_2}\geq P_{j_1}\otimes Q_{k_1}$.

As for the limit,

$$\begin{aligned}\|(P_j\otimes Q_k)(\xi\otimes\eta)-\xi\otimes\eta\| &\leq \|(P_j\otimes Q_k)(\xi\otimes\eta)-\xi\otimes Q_k\eta\| \\ &\quad + \|\xi\otimes Q_k\eta-\xi\otimes\eta\| \\ &\leq \|P_j\xi-\xi\|\,\|Q_k\eta\| + \|\xi\|\,\|Q_k\eta-\eta\| \\ &\leq \|\eta\|\,\|P_j\xi-\xi\| + \|\xi\|\,\|Q_k\eta-\eta\| \xrightarrow[j,k]{} 0.\end{aligned}$$

When $\nu = \sum_{r=1}^m \xi_r\otimes\eta_r$, by linearity of the limit we obtain $(P_j\otimes Q_k)\nu\to\nu$. As $P_j\otimes Q_k\in\mathcal{B}(\mathcal{H}\otimes\mathcal{K})$ by Proposition 13.4.4, for arbitrary $\nu\in\mathcal{H}\,\overline{\otimes}\,\mathcal{K}$ given

$\varepsilon > 0$ there exists $\nu_0 \in \mathcal{H} \otimes \mathcal{K}$ with $\|\nu - \nu_0\| < \varepsilon$. Then

$$\begin{aligned}\|(P_j \otimes Q_k)\nu - \nu\| &\le \|(P_j \otimes Q_k)(\nu - \nu_0)\| + \|(P_j \otimes Q_k)\nu_0 - \nu_0\| + \|\nu_0 - \nu\| \\ &\le 2\varepsilon + \|(P_j \otimes Q_k)\nu_0 - \nu_0\|.\end{aligned}$$

Thus

$$\limsup_{j,k} \|(P_j \otimes Q_k)\nu - \nu\| \le 2\varepsilon,$$

and as ε was arbitrary the Limsup Routine shows that the limit exists and is zero.

(13.4.14) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras. Use Proposition 13.1.4 to show that the product and involution are well-defined on $\mathcal{A} \otimes \mathcal{B}$.

Answer. Suppose that

$$\sum_{j=1}^{n} a_j \otimes b_j = \sum_{r=1}^{m} a'_r \otimes b'_r. \tag{AB.13.3}$$

By relabelling $a_{n+j} = -a'_j$ and $b_{n+j} = b'_j$, we may write the above as

$$\sum_{j=1}^{n+m} a_j \otimes b_j = 0.$$

Then Proposition 13.1.4 gives us coefficients $\{\gamma_{kj}\}$ such that

$$\sum_{k=1}^{n+m} \gamma_{kj} a_k = 0, \qquad \sum_{j=1}^{n+m} \gamma_{kj} b_j = b_k. \tag{AB.13.4}$$

We want to show that

$$\Big(\sum_{j=1}^{n} a_j \otimes b_j\Big)\Big(\sum_{k=1}^{p} c_k \otimes d_k\Big) = \Big(\sum_{r=1}^{m} a'_r \otimes b'_r\Big)\Big(\sum_{k=1}^{p} c_k \otimes d_k\Big),$$

which expanded according to the definition of product in $\mathcal{A} \otimes \mathcal{B}$ amounts to

$$\sum_{j=1}^{n}\sum_{k=1}^{p} a_j c_k \otimes b_j d_k = \sum_{r=1}^{m}\sum_{k=1}^{p} a'_r c_k \otimes b'_r d_k.$$

Now

$$\begin{aligned}\sum_{j=1}^{n}\sum_{k=1}^{p} a_j c_k \otimes b_j d_k &= \sum_{j=1}^{n}\sum_{k=1}^{p}\sum_{s=1}^{n+m} \gamma_{js} a_j c_k \otimes b_s d_k \\ &= \sum_{s=1}^{n+m}\sum_{k=1}^{p}\Big(\sum_{j=1}^{n}\gamma_{js}a_j\Big)c_k \otimes b_s d_k \\ &= -\sum_{s=1}^{n+m}\sum_{k=1}^{p}\Big(\sum_{j=n+1}^{n+m}\gamma_{js}a_j\Big)c_k \otimes b_s d_k \\ &= -\sum_{j=n+1}^{n+m}\sum_{k=1}^{p} a_j c_k \otimes \Big(\sum_{s=1}^{n+m}\gamma_{js}b_s\Big)d_k \\ &= -\sum_{j=n+1}^{n+m}\sum_{k=1}^{p} a_j c_k \otimes b_j d_k \\ &= \sum_{r=1}^{m}\sum_{k=1}^{p} a'_r c_k \otimes b'_r d_k.\end{aligned}$$

The same argument can be used for different presentations of $\sum_k c_k \otimes d_k$, and therefore the product does not depend on the presentation.

With the involution we can use a similar idea. If we have the equality (AB.13.3), we can again relabel and use the relations (AB.13.4). We want to show that $\sum_{j=1}^{n+m} a_j \otimes b_j = 0$ implies

$$\sum_{j=1}^{n+m} a_j^* \otimes b_j^* = 0.$$

We have, by Proposition 13.1.4,

$$\sum_{k=1}^{n+m} \overline{\gamma_{kj}} a_k^* = 0, \qquad \sum_{j=1}^{n+m} \overline{\gamma_{kj}} b_j^* = b_k^*.$$

Then

$$\begin{aligned}\sum_{j=1}^{n+m} a_j^* \otimes b_j^* &= \sum_{j=1}^{n+m} a_j^* \otimes \sum_{s=1}^{n+m} \overline{\gamma_{js}} b_s^* \\ &= \sum_{s=1}^{n+m}\Big(\sum_{j=1}^{n+m}\overline{\gamma_{js}}a_j^*\Big)\otimes b_s^* = 0\end{aligned}$$

as desired.

(13.4.15) Let $\mathcal{A},\mathcal{B}$ be C*-algebras. Show that the product on $\mathcal{A}\otimes\mathcal{B}$ is well-defined (use Theorem 13.1.6 and Corollary 13.1.8).

Answer. For each $a\in\mathcal{A}$ consider the multiplication operator $L_a:\mathcal{A}\to\mathcal{A}$ given by left multiplication by a, and similarly we have $L_b:\mathcal{B}\to\mathcal{B}$. By Corollary 13.1.8 there exists a linear map $L_a\otimes L_b:\mathcal{A}\otimes\mathcal{B}\to\mathcal{A}\otimes\mathcal{B}$, with the property that

$$(L_a\otimes L_b)(c\otimes d)=ac\otimes bd.$$

If $\mathcal{L}$ is the space of linear maps $\mathcal{A}\otimes\mathcal{B}\to\mathcal{A}\otimes\mathcal{B}$, we can consider a bilinear form $\phi:\mathcal{A}\times\mathcal{B}\to\mathcal{L}$ given by $\phi(a,b)=L_a\otimes L_b$. By Theorem 13.1.6 there exists a linear map $M:\mathcal{A}\otimes\mathcal{B}\to\mathcal{L}$ such that $M(a\otimes b)=L_a\otimes L_b$. This allows us to define, for $x,y\in\mathcal{A}\otimes\mathcal{B}$,

$$xy=(Mx)(y).$$

This is bilinear, for M is linear and the map Mx is linear. And, on elementary tensors,

$$(M(a\otimes b))(c\otimes d)=(L_a\otimes L_b)(c\otimes d)=ac\otimes bd.$$

(13.4.16) Let $\mathcal{A},\mathcal{B},\mathcal{C}$ be unital C*-algebras and $\pi:\mathcal{A}\otimes\mathcal{B}\to\mathcal{C}$ a $*$-homomorphism. Show that there exist $*$-homomorphisms $\pi_\mathcal{A}:\mathcal{A}\to\mathcal{C}$ and $\pi_\mathcal{B}:\mathcal{B}\to\mathcal{C}$, with commuting ranges, such that $\pi=\pi_\mathcal{A}\times\pi_\mathcal{B}$.

Answer. Let $\pi_\mathcal{A}(a)=\pi(a\otimes I_\mathcal{B})$, $\pi_\mathcal{B}(b)=\pi(I_\mathcal{A}\otimes b)$. Then $\pi_\mathcal{A},\pi_\mathcal{B}$ are $*$-homomorphisms. We have

$$\begin{aligned}\pi_\mathcal{A}(a)\pi_\mathcal{B}(b)&=\pi(a\otimes I_\mathcal{B})\pi(I_\mathcal{A}\otimes b)=\pi(a\otimes b)\\&=\pi(I_\mathcal{A}\otimes b)\pi(a\otimes I_\mathcal{B})=\pi_\mathcal{B}(b)\pi_\mathcal{A}(a),\end{aligned}$$

so the ranges commute. Now Corollary 13.1.9 guarantees that $\pi=\pi_\mathcal{A}\times\pi_\mathcal{B}$.

(13.4.17) Let $\mathcal{A},\mathcal{B}$ be C*-algebras and $\pi:\mathcal{A}\otimes\mathcal{B}\to\mathcal{B}(\mathcal{H})$ a representation. Fix $a\in\mathcal{A}$. Show that the map $\rho:b\longmapsto\pi(a\otimes b)$ is linear and bounded. *(Hint: for the bounded part, assume that $a\geq 0$ so that $\varphi\circ\rho$ is positive for any state φ, and use the Closed Graph Theorem)*

Answer. We have

$$\rho(b_1 + \lambda b_2) = \pi(a \otimes (b_1 + \lambda b_2)) = \pi(a \otimes b_1 + \lambda a \otimes b_2)$$
$$= \pi(a \otimes b_1) + \lambda\pi(a \otimes b_2) = \rho(b_1) + \lambda\rho(b_2).$$

So ρ is linear. To show that ρ is bounded, we can assume without loss of generality that $a \geq 0$, for an arbitrary a is a linear combination of positives and then ρ will be a linear combination of bounded. Suppose that $b_n \to 0$ and $\rho(b_n) \to T$. Let $\varphi \in S(\mathcal{B}(\mathcal{H}))$. We have, for $b \in \mathcal{B}$,

$$\varphi(\rho(b^*b)) = \varphi(\pi(a \otimes b^*b)) = \varphi(\pi(a^{1/2} \otimes b)^*\pi(a^{1/2} \otimes b)) \geq 0.$$

So $\varphi \circ \rho$ is a positive linear functional on $\mathcal{B}$. By Proposition 11.5.4, $\varphi \circ \rho$ is bounded. Then

$$\varphi(T) = \varphi(\lim_n \rho(b_n)) \overset{(\varphi \text{ bounded})}{=} \lim_m \varphi(\rho(b_n) \underset{(\varphi\circ\rho \text{ bounded})}{=} \varphi \circ \rho(\lim_n b_n) = 0.$$

As φ can be any state, it follows that $T = 0$ (by Corollary 11.5.8) and so ρ is bounded by the Closed Graph Theorem (6.3.12).

(13.4.18) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a non-degenerate representation. Show that there exist non-degenerate representations $\pi_\mathcal{A} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\pi_\mathcal{B} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$, with commuting ranges, such that $\pi = \pi_\mathcal{A} \times \pi_\mathcal{B}$. When π is faithful, so are $\pi_\mathcal{A}$ and $\pi_\mathcal{B}$.

(As opposed to Exercise 13.4.16 the algebras are not required to be unital, so a different method is required; use the non-degeneracy to define $\pi_\mathcal{A}$ and $\pi_\mathcal{B}$ on a dense subspace of $\mathcal{H}$, and use Exercise 13.4.17 when needed)

Answer. Since $\pi(\mathcal{A} \otimes \mathcal{B})\mathcal{H}$ is dense in $\mathcal{H}$, we define

$$\pi_\mathcal{A}(a)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi = \sum_k \pi(aa_k \otimes b_k)\xi.$$

We need to check that this is well-defined and that $\pi_\mathcal{A}(a) \in \mathcal{B}(\mathcal{H})$. Suppose that

$$\sum_k \pi(a_k \otimes b_k)\xi = \sum_r \pi(c_r \otimes d_r)\eta.$$

Let $\{e_\ell\}$ be an approximate unit for $\mathcal{B}$. Using Exercise 13.4.17,

$$\begin{aligned}\pi_{\mathcal{A}}(a)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi &= \sum_k \pi(aa_k \otimes b_k)\xi = \lim_\ell \sum_k \pi(aa_k \otimes e_\ell b_k)\xi \\ &= \lim_\ell \pi(a \otimes e_\ell) \sum_k \pi(a_k \otimes b_k)\xi \\ &= \lim_\ell \pi(a \otimes e_\ell) \sum_r \pi(c_r \otimes d_k)\eta \\ &= \lim_\ell \sum_r \pi(ac_r \otimes e_\ell d_k)\eta = \sum_r \pi(ac_r \otimes d_k)\eta.\end{aligned}$$

Thus $\pi_{\mathcal{A}}(a)$ is well-defined and linear by construction. By Exercise 13.4.17 there exists $c > 0$ with $\|\pi(a \otimes e_\ell)\| \leq c$ for all ℓ. Then

$$\begin{aligned}\Big\|\pi_{\mathcal{A}}(a)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi\Big\| &= \lim_\ell \Big\|\pi(a \otimes e_\ell)\sum_k \pi(a_k \otimes b_k)\xi\Big\| \\ &\leq c\Big\|\pi\Big(\sum_k a_k \otimes b_k\Big)\xi\Big\|.\end{aligned}$$

Hence $\pi_{\mathcal{A}}(a)$ is bounded and by Proposition 6.1.9 it extends uniquely to an operator $\pi_{\mathcal{A}}(a) \in \mathcal{B}(\mathcal{H})$.

The definition and justification for $\pi_{\mathcal{B}}$ is entirely analogous. As for the commuting ranges, that's straightforward:

$$\begin{aligned}\pi_{\mathcal{A}}(a)\pi_{\mathcal{B}}(b)\Big(\sum_k a_k \otimes b_k\Big)\xi &= \Big(\sum_k \pi(aa_k \otimes bb_k)\Big)\xi \\ &= \pi_{\mathcal{B}}(b)\pi_{\mathcal{A}}(a)\Big(\sum_k a_k \otimes b_k\Big)\xi,\end{aligned}$$

so $\pi_{\mathcal{B}}(b)\pi_{\mathcal{A}}(a)$ and $\pi_{\mathcal{A}}(a)\pi_{\mathcal{B}}(b)$ agree on a dense subset and are therefore equal. It remains to check that the representations are non-degenerate. Given an approximate unit $\{e_\ell\}$ for $\mathcal{B}$,

$$\pi_{\mathcal{B}}(e_\ell)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi = \sum_k \pi(a_k \otimes e_\ell b_k) \underset{\ell}{\to} \pi\Big(\sum_k a_k \otimes b_k\Big)\xi.$$

Thus $\pi_{\mathcal{B}}(\mathcal{B})\mathcal{H}$ is dense in $\mathcal{H}$, and $\pi_{\mathcal{B}}$ is non-degenerate. A similar argument establishes that $\pi_{\mathcal{A}}$ is non-degenerate.

When π is faithful and $\pi_{\mathcal{A}}(a) = 0$, then $\pi(a \otimes b) = \pi_{\mathcal{A}}(a)\pi_{\mathcal{B}}(b) = 0$ for all $b \in \mathcal{B}$. This implies that $a \otimes b = 0$ for all $b \in \mathcal{B}$ and then $a = 0$. Hence $\pi_{\mathcal{A}}$ is faithful, and the argument for $\pi_{\mathcal{B}}$ is entirely analogous.

(13.4.19) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a non-degenerate representation. Show that if $\{e_\ell\}$ and $\{f_t\}$ are approximate units for $\mathcal{B}$ and $\mathcal{A}$ respectively, the representations $\pi_\mathcal{A}$ and $\pi_\mathcal{B}$ of Exercise 13.4.18 satisfy

$$\pi_\mathcal{A}(a) = \lim_\ell \pi(a \otimes e_\ell), \qquad \pi_\mathcal{B}(b) = \lim_t \pi(f_t \otimes b),$$

where the limits are sot.

Answer. This was done more or less explicitly in the answer to Exercise 13.4.18. Concretely, with $\rho_k(b) = \pi(aa_k \otimes b)$ and using Exercise 13.4.17,

$$\begin{aligned}\Big\|\pi_\mathcal{A}(a)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi - \pi(a \otimes e_\ell)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi\Big\| \\ &= \Big\|\sum_k aa_k \otimes (b_k - b_k e_\ell)\Big)\xi\Big\| \\ &= \Big\|\sum_k \rho_k(b_k - b_k e_\ell)\xi\Big\| \\ &\le \|\xi\| \sum_k c_k \|b_k - b_k e_\ell\| \underset{\ell}{\to} 0.\end{aligned}$$

For arbitrary $\eta \in \mathcal{H}$ fix $\varepsilon > 0$ and choose $\xi \in \mathcal{H}$ and $\{a_k\}_{k=1}^m \subset \mathcal{A}$, $\{b_k\}_{k=1}^m \subset \mathcal{B}$ with

$$\Big\|\eta - \Big(\sum_{k=1}^m a_k \otimes b_k\Big)\xi\Big\| < \varepsilon.$$

Then, using that $\|\pi(a \otimes e_\ell)\| \le c\|\varepsilon_\ell\| \le c$ by Exercise 13.4.17,

$$\begin{aligned}&\|\pi_\mathcal{A}(a)\eta - \pi(a \otimes e_\ell)\eta\| \\ &\quad\le \|\pi_\mathcal{A}(a)\| \Big\|\eta - \Big(\sum_k a_k \otimes b_k\Big)\xi\Big\| \\ &\qquad + \Big\|\pi_\mathcal{A}(a)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi - \pi(a \otimes e_\ell)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi\Big\| \\ &\qquad + \|\pi(a \otimes e_\ell)\| \Big\|\Big(\sum_k a_k \otimes b_k\Big)\xi - \eta\Big\| \\ &\quad\le (1+c)\|a\|\,\varepsilon \\ &\qquad + \Big\|\pi_\mathcal{A}(a)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi - \pi(a \otimes e_\ell)\pi\Big(\sum_k a_k \otimes b_k\Big)\xi\Big\|.\end{aligned}$$

Hence

$$\limsup_{\ell} \|\pi_{\mathcal{A}}(a)\eta - \pi(a\otimes e_\ell)\eta\| \leq (1+c)\|a\|\,\varepsilon.$$

As ε was arbitrary, the Limsup Routine shows that the limit exists and is zero. The computation for $\pi_{\mathcal{B}}$ is entirely analog.

(13.4.20) Let $\mathcal{A},\mathcal{B}$ be C*-algebras and $\pi_{\mathcal{A}}:\mathcal{A}\to\mathcal{H}_{\mathcal{A}}$ and $\pi_{\mathcal{B}}:\mathcal{B}\to\mathcal{H}_{\mathcal{B}}$ cyclic representations. Show that $\pi_{\mathcal{A}}\otimes\pi_{\mathcal{B}}$ is cyclic.

Answer. We have $\xi_{\mathcal{A}}\in\mathcal{H}_{\mathcal{A}}$ cyclic for $\pi_{\mathcal{A}}$ and $\xi_{\mathcal{B}}\in\mathcal{H}_{\mathcal{B}}$ cyclic for $\pi_{\mathcal{B}}$. We want to show that $(\pi_{\mathcal{A}}\otimes\pi_{\mathcal{B}})(\mathcal{A}\otimes\mathcal{B})(\xi_{\mathcal{A}}\,\overline{\otimes}$
$xi_{\mathcal{B}})$ is dense in $\mathcal{H}_{\mathcal{A}}\,\overline{\otimes}\,\mathcal{H}_{\mathcal{B}}$. Given $\xi\in\mathcal{H}_{\mathcal{A}}$ and $\eta\in\mathcal{H}_{\mathcal{B}}$ there exist sequences $\{a_n\}\subset\mathcal{A}$ and $\{b_n\}\subset\mathcal{B}$ with $\pi_{\mathcal{A}}(a_n)\xi_{\mathcal{A}}\to\xi$ and $\pi_{\mathcal{B}}(b_n)\xi_{\mathcal{B}}\to\eta$. Then

$$\begin{aligned}
\|(\pi_{\mathcal{A}}\otimes\pi_{\mathcal{B}})(a_n\otimes b_n)(\xi_{\mathcal{A}}\otimes\xi_{\mathcal{B}})-\xi\otimes\eta\| &= \|\pi_{\mathcal{A}}(a_n)\xi_{\mathcal{A}}\otimes\pi_{\mathcal{B}}(b_n)\xi_{\mathcal{B}}-\xi\otimes\eta\|\\
&\leq \|(\pi_{\mathcal{A}}(a_n)\xi_{\mathcal{A}}-\xi)\otimes\pi_{\mathcal{B}}(b_n)\xi_{\mathcal{B}}\|\\
&\qquad+\|\xi\otimes(\pi_{\mathcal{B}}(b_n)\xi_{\mathcal{B}}-\eta)\|\\
&= \|(\pi_{\mathcal{A}}(a_n)\xi_{\mathcal{A}}-\xi)\|\,\|\pi_{\mathcal{B}}(b_n)\xi_{\mathcal{B}}\|\\
&\qquad+\|\xi\|\,\|\pi_{\mathcal{B}}(b_n)\xi_{\mathcal{B}}-\eta\|\\
&\xrightarrow[n]{}0.
\end{aligned}$$

(13.4.21) Let $\mathcal{A},\mathcal{B}$ be C*-algebras with $\mathcal{A}$ non-unital, γ a C*-norm on $\tilde{\mathcal{A}}\otimes\mathcal{B}$, and $\psi\in S(\tilde{\mathcal{A}})$ the unique state with $\ker\psi=\mathcal{A}$. Show that for the map $\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}}:\tilde{\mathcal{A}}\otimes_\gamma\mathcal{B}\to\mathbb{C}\otimes\mathcal{B}$ we have $\ker(\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}})=\mathcal{A}\otimes_\gamma\mathcal{B}$.

Answer. If $\mathcal{B}$ is not unital, by Exercise 11.6.10 we have that $\mathrm{id}_{\tilde{\mathcal{B}}}$ is the unique $*$-homomorphism that extends $\mathrm{id}_{\mathcal{B}}$ to $\tilde{\mathcal{B}}$. Then Exercise 13.2.27 guarantees that $\psi\otimes_\gamma\mathrm{id}_{\tilde{\mathcal{B}}}$ is the unique ucp extension of $\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}}$ to $\tilde{\mathcal{A}}\otimes_\gamma\tilde{\mathcal{B}}\to\mathbb{C}\otimes\tilde{\mathcal{B}}$. In the end, we may assume without loss of generality that $\mathcal{B}$ is unital.

Let $x\in\tilde{\mathcal{A}}\otimes_\gamma\mathcal{B}$ with $\|x\|_\gamma=1$ and $(\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}})(x)=0$. Then there exists a sequence $\{x_n\}\subset\tilde{\mathcal{A}}\otimes\mathcal{B}$ with $\|x-x_n\|_\gamma\to0$. As $\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}}$ is γ-continuous, we have $0=(\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}})(x)=\lim_n(\psi\otimes_\gamma\mathrm{id}_{\mathcal{B}})(x_n)$. Each of these x_n is of the form $x_n=\sum_j a_{nj}\otimes b_{nj}$, where we might assume that the b_{nj} are linearly independent for each n (Remark 13.1.5). Let

$$z_n = \sum_j (a_{nj} - \psi(a_{nj})I_{\tilde{\mathcal{A}}}) \otimes b_{nj}.$$

By definition of ψ, $a_{nj} - \psi(a_{nj})I_{\tilde{\mathcal{A}}} \in \ker\psi = \mathcal{A}$ for all n, j. So $z_n \in \mathcal{A} \otimes \mathcal{B}$ for all n. Also

$$\|x_n - z_n\|_\gamma = \Big\| \sum_j \psi(a_{nj})I_{\tilde{\mathcal{A}}} \otimes b_{nj} \Big\|_\gamma = \Big\| I_{\tilde{\mathcal{A}}} \otimes \sum_j \psi(a_{nj})b_{nj} \Big\|$$

$$= \Big\| \sum_j \psi(a_{nj})b_{nj} \Big\| = \|(\psi \otimes_\gamma \mathrm{id}_{\mathcal{B}})(x_n)\| \to 0,$$

since $\|x - x_n\|_\gamma \to 0$. It follows that $x = \lim_n z_n \in \mathcal{A} \otimes_\gamma \mathcal{B}$. As we also have $\mathcal{A} \otimes_\gamma \mathcal{B} \subset \ker(\psi \otimes_\gamma \mathrm{id}_{\mathcal{B}})$ by the γ-continuity of $\psi \otimes_\gamma \mathrm{id}_{\mathcal{B}}$, the equality is proven.

(13.4.22) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras, $x \in \mathcal{A} \otimes_{\min} \mathcal{B}$. Show that if $(\varphi \otimes \psi)(x^*x) = 0$ for all $\varphi \in S(\mathcal{A})$ and $\psi \in S(\mathcal{B})$, then $x = 0$.

Answer. Let $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H}_{\mathcal{A}})$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{H}_{\mathcal{B}})$ be faithful representations. By Proposition 13.4.8 and Corollary 13.4.24 we have a faithful representation $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}} : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{B}(\mathcal{H}_{\mathcal{A}} \bar{\otimes} \mathcal{H}_{\mathcal{B}})$. If $x \neq 0$ then $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(x^*x) \neq 0$. So there exists an elementary tensor $\xi \otimes \eta \in \mathcal{H}_{\mathcal{A}} \bar{\otimes} \mathcal{H}_{\mathcal{B}}$ with $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(|x|)(\xi \otimes \eta) \neq 0$ (otherwise, the equality to zero would extend to all vectors implying that $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(|x|) = 0$ and this forces $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(|x|^2) = 0$). Let

$$\varphi(a) = \langle \pi_{\mathcal{A}}(a)\xi, \xi \rangle, \qquad \psi(b) = \langle \pi_{\mathcal{B}}(b)\eta, \eta \rangle.$$

For any $z \in \mathcal{A} \otimes \mathcal{B}$ we have $z = \sum_j a_j \otimes b_j$ and

$$(\varphi \otimes \psi)(z) = \sum_j \varphi(a_j)\psi(b_j) = \sum_j \langle (\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(z)(\xi \otimes \eta), \xi \otimes \eta \rangle.$$

By continuity the equality holds for all $z \in \mathcal{A} \otimes_{\min} \mathcal{B}$. Then

$$(\varphi \otimes \psi)(x^*x) = \langle (\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(x^*x)\, \xi \otimes \eta, \xi \otimes \eta \rangle$$

$$= \|(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(|x|)(\xi \otimes \eta)\|^2 > 0.$$

(13.4.23) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H}_{\mathcal{A}})$, $\mathcal{B} \subset \mathcal{B}(\mathcal{H}_{\mathcal{B}})$ be separable C*-algebras. Show that if $\mathcal{A} \otimes_{\min} \mathcal{B} \subset \mathcal{B}(\mathcal{H}_{\mathcal{A}} \bar{\otimes} \mathcal{H}_{\mathcal{B}})$ contains a nonzero compact operator, then $\mathcal{A}$ contains a nonzero compact operator.

Answer. Suppose that the only compact operator in $\mathcal{A}$ is 0. Since $\mathcal{A}$ is separable, any closed subset is separable by Proposition 1.8.5. Let $\{A_n\} \subset \mathcal{A}^+$ be a countable dense subset for $\overline{B_1^{\mathcal{A}^+}(0)}$. Choose $\varphi_n \in S(\mathcal{B}(\mathcal{H}_\mathcal{A}))$ with $\varphi(A_n) = 1$ and $\varphi|_{\mathcal{K}(\mathcal{H}_\mathcal{A})} = 0$; this can always be done (since A_n is not compact) by applying Corollary 11.5.8 in the Calkin algebra $\mathcal{B}(\mathcal{H}_\mathcal{A})/\mathcal{K}(\mathcal{H}_\mathcal{A})$. Let $\varphi \in S(\mathcal{B}(\mathcal{H}_\mathcal{A}))$ be given by

$$\varphi(A) = \sum_{n=1}^\infty 2^{-n}\varphi_n(A).$$

This is clearly a state since it is positive and $\varphi(I_{\mathcal{H}_\mathcal{A}}) = 1$. We also have that φ is faithful on $\mathcal{A}$. Indeed, if $A \geq 0$ and $\varphi(A) = 0$, then $\varphi_n(A) = 0$ for all n. Let $A' = A/\|A\|$. Choose n such that $\|A' - A_n\| < 1/2$. Then

$$1 = \varphi(A_n) = \varphi(A_n - A') \leq \|A_n - A'\| < 1/2,$$

a contradiction. Hence $A = 0$ and φ is faithful on $\mathcal{A}$. As $\mathcal{B}$ is also separable, there exists $\psi \in S(\mathcal{B})$ faithful (Proposition 11.5.12). The state $\varphi \otimes \psi$ is faithful on $\mathcal{A} \otimes_{\min} \mathcal{B}$ by Lemma 13.4.26. Now suppose that $T \in \mathcal{A} \otimes_{\min} \mathcal{B}$ is compact. Then T^*T is compact. By the proof of Proposition 9.8.5 there exist finite-rank projections $\{P_F\}$ such that $\|T - P_F T\| \to 0$. Let $\{T_n\} \subset \mathcal{A} \otimes \mathcal{B}$ with $T_n \to T$. As everything is bounded, $P_F T_n \to T$ if we move on both indices. The state $\varphi \otimes \psi$ is zero on $\mathcal{K}(\mathcal{H}_\mathcal{A} \overline{\otimes} \mathcal{H}_\mathcal{B})$ by Exercise 13.4.6 and continuity, so $(\varphi \otimes \psi)(T_n^* P_F T_n) = 0$ for all n and F. Hence

$$(\varphi \otimes \psi)(T^*T) = \lim_{n,F}(\varphi \otimes \psi)(T_n^* P_F T_n) = 0.$$

By the faithfulness of $\varphi \otimes \psi$ we get that $T = 0$. So $\mathcal{A} \otimes_{\min} \mathcal{B}$ contains no nonzero compact operator.

(13.4.24) Let X be set, $\mathcal{A}$ a C*-algebra, and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. Let $\delta : c_0(X) \to \mathcal{B}(\ell^2(X))$ be given by $\delta(f) = M_f$. Given $Z = \sum_k f_k \otimes a_k \in c_0(X) \otimes \mathcal{A}$, show that

$$\Big\|(\delta \otimes \pi)(Z)\Big\|_{\mathcal{B}(\ell^2(X)\overline{\otimes}\mathcal{H})} = \sup\Big\{\Big\|(\delta_x \otimes \pi)(Z)\Big\|_{\mathcal{B}(\mathcal{H})} : x \in X\Big\}.$$

Answer. Let $Z = \sum_k f_k \otimes a_k \in C_0(X) \otimes \mathcal{A}$ and $\tilde{\xi} = \sum_j g_j \otimes \xi_j, \tilde{\eta} = \sum_j h_j \otimes \xi_j \in \ell^2(X) \overline{\otimes} \mathcal{H}$ (where $\{\xi_j\}_j$ is an orthonormal basis, as in Lemma 13.4.3). Write

$$D_Z = \sup\Big\{\Big\|(\delta_x \otimes \pi)\Big(\sum_k f_k \otimes a_k\Big)\Big\|_{\mathcal{B}(\mathcal{H})} : x \in X\Big\}$$

We have

$$\begin{aligned}
\|(\delta\otimes\pi)(Z)\tilde{\xi}\|^2 &= \Big\|(\delta\otimes\pi)\Big(\sum_k f_k\otimes a_k\Big)\Big(\sum_j g_j\otimes\xi_j\Big)\Big\|^2_{\ell^2(X)\overline{\otimes}\mathcal{H}} \\
&= \Big\|\sum_k\sum_j f_k g_j\otimes\pi(a_k)\xi_j\Big\|^2_{\ell^2(X)\overline{\otimes}\mathcal{H}} \\
&= \sum_{k_1,k_2}\sum_{j_1,j_2}\langle f_{k_1}g_{j_1}, f_{k_2}g_{j_2}\rangle\,\langle\pi(a_{k_1})\xi_{j_1},\pi(a_{k_2})\xi_{j_2}\rangle \\
&= \sum_{k_1,k_2}\sum_{j_1,j_2}\sum_x f_{k_1}(x)g_{j_1}(x)\,\overline{f_{k_2}(x)g_{j_2}(x)}\,\langle\pi(a_{k_1})\xi_{j_1},\pi(a_{k_2})\xi_{j_2}\rangle \\
&= \sum_x\sum_{k_1,k_2}\sum_{j_1,j_2} f_{k_1}(x)g_{j_1}(x)\,\overline{f_{k_2}(x)g_{j_2}(x)}\,\langle\pi(a_{k_1})\xi_{j_1},\pi(a_{k_2})\xi_{j_2}\rangle \\
&= \sum_x\Big\|(\delta_x\otimes\pi)\Big(\sum_k f_k\otimes a_k\Big)\Big(\sum_j g_j(x)\xi_j\Big)\Big\|^2_{\mathcal{H}} \\
&\le D_Z\sum_x\Big\|\sum_j g_j(x)\xi_j\Big\|^2 \\
&= D_Z\,\|\tilde{\xi}\|^2.
\end{aligned}$$

(for the exchanging of series, the sums on k_1, k_2 are actual sums, and the series for j_1, j_2, and x converge absolutely by Cauchy-Schwarz). Thus $\|(\delta\otimes\pi)(Z)\|\le\sup\{(\delta_x\otimes\pi)(Z)\|_{\mathcal{B}(\mathcal{H})}$.

Conversely, fix $\varepsilon>0$ and $y\in X$ such that

$$\Big\|\sum_k f_k(y)\pi(a_k)\Big\|^2_{\mathcal{B}(\mathcal{H})}\ge\sup\Big\{\Big\|\sum_k f_k(x)\pi(a_k)\Big\|:\ x\in X\Big\}^2-\varepsilon.$$

Fix $\xi\in\mathcal{H}$ with $\|\xi\|=1$ and

$$\Big\|\sum_k f_k(y)\pi(a_k)\,\xi\Big\|^2\ge\Big\|\sum_k f_k(y)\pi(a_k)\Big\|^2-\varepsilon.$$

Let $\tilde{\xi} = \delta_y \otimes \xi \in \ell^2(X) \otimes \mathcal{H}$. Then

$$\begin{aligned}
\|(\delta \otimes \pi)(Z)\tilde{\xi}\|^2 &= \Big\|(\delta \otimes \pi)\Big(\sum_k f_k \otimes a_k\Big)\,(\delta_y \otimes \xi)\Big\|^2_{\ell^2(X)\bar{\otimes}\mathcal{H}} \\
&= \Big\|(\delta \otimes \pi)\Big(\sum_k f_k \otimes a_k\Big)\,(\delta_y \otimes \xi)\Big\|^2_{\ell^2(X)\bar{\otimes}\mathcal{H}} \\
&= \sum_{k_1,k_2} f_{k_1}(y)\overline{f_{k_2}(y)}\,\langle \pi(a_{k_1})\xi, \pi(a_{k_2})\xi\rangle = \Big\|\sum_k f_k(y)\pi(a_k)\,\xi\Big\|^2 \\
&\geq \Big\|\sum_k f_k(y)\pi(a_k)\Big\|^2 - \varepsilon \\
&\geq \sup\Big\{\Big\|\sum_k f_k(x)\pi(a_k)\Big\| :\ x \in X\Big\}^2 - 2\varepsilon.
\end{aligned}$$

As ε was arbitrary, the equality is established.

(13.4.25) Let $\mathcal{B} \subset \mathcal{A}$ be C*-algebras, with $\mathcal{B}$ hereditary. Show that if $\mathrm{id}_{\mathcal{A}}$ is nuclear, then so is $\mathrm{id}_{\mathcal{B}}$. *(Hint: use an approximate identity for $\mathcal{B}$)*

Answer. By hypothesis there exit nets of contractive completely positive maps $\{\varphi_\ell\}$ and $\{\psi_\ell\}$ with $\varphi_\ell : \mathcal{A} \to M_{n(\ell)}(\mathbb{C})$ and $\psi_\ell : M_{n(\ell)}(\mathbb{C}) \to \mathcal{A}$ with $\psi_\ell(\varphi_\ell(a)) \to a$ for all $a \in \mathcal{A}$. Let $\{e_j\}$ be an approximate unit for $\mathcal{B}$. Because $\mathcal{B}$ is hereditary, this means that $e_j a e_j \in \mathcal{B}$ for all $a \in \mathcal{A}$ and all j. So we can define maps $\psi_{\ell,j}(X) = e_j\psi_\ell(X)e_j$, still contractive and completely positive, now with codomain $\mathcal{B}$. So $\psi_{\ell,j}(\varphi_\ell(b)) \to b$ for all $b \in \mathcal{B}$ if we make a new net by ordering $(\ell_1, j_1) \leq (\ell_2, j_2)$ if $\ell_1 \leq \ell_2$ and $j_1 \leq j_2$.

(13.4.26) Let $\mathcal{A}$ be a C*-algebra. Show that $\mathrm{id}_{\mathcal{A}}$ is nuclear if and only if $\mathrm{id}_{\tilde{\mathcal{A}}}$ is nuclear.

Answer. Suppose first that $\mathrm{id}_{\mathcal{A}}$ is nuclear. So there exit nets of contractive completely positive maps $\{\varphi_\ell\}$ and $\{\psi_\ell\}$ with $\varphi_\ell : \mathcal{A} \to M_{n(\ell)}(\mathbb{C})$ and $\psi_\ell : M_{n(\ell)}(\mathbb{C}) \to \mathcal{A}$ with $\psi_\ell(\varphi_\ell(a)) \to a$ for all $a \in \mathcal{A}$. We can define maps $\tilde{\varphi}_\ell : \mathcal{A} \to M_{n(\ell)+1}(\mathbb{C})$ given by $\tilde{\varphi}_\ell(a, \lambda) = \varphi_\ell(a) \oplus \lambda$, and similarly define $\tilde{\psi}_\ell(X) = (\psi_\ell(P_\ell X P_\ell), X_{n(\ell)+1,n(\ell)+1})$, where P_ℓ is the compression to the $n(\ell) \times n(\ell)$ upper left corner. In both cases the new maps are completely positive because direct sums of completely positive maps are completely positive. They are also contractive. And

$$\tilde{\psi}_\ell(\tilde{\varphi}_\ell(a,\lambda)) = \tilde{\psi}_\ell(\varphi_\ell(a) \oplus \lambda) = (\psi_\ell(\varphi_\ell(a)), \lambda) \to (a, \lambda).$$

So $\mathrm{id}_{\tilde{\mathcal{A}}}$ is nuclear.

Conversely, suppose that $\mathrm{id}_{\tilde{\mathcal{A}}}$ is nuclear. As $\mathcal{A}$ is an ideal in $\tilde{\mathcal{A}}$, we get that $\mathrm{id}_{\mathcal{A}}$ is nuclear by Exercise 13.4.25.

(13.4.27) Let $\mathcal{M}$ be a von Neumann algebra. Show that

$$(\mathcal{M} \otimes_{\min} I_{\mathcal{K}})' = \mathcal{M}' \,\overline{\otimes} \mathcal{B}(\mathcal{K}),$$

where $\mathcal{M}' \,\overline{\otimes} \mathcal{B}(\mathcal{K})$ is the von Neumann algebra generated by $\mathcal{M}' \otimes \mathcal{B}(\mathcal{K})$ in $\mathcal{B}(\mathcal{H} \,\overline{\otimes}\, \mathcal{K})$. Conclude that $\mathcal{Z}(\mathcal{M} \,\overline{\otimes}\, \mathcal{B}(\mathcal{K})) = \mathcal{Z}(\mathcal{M}) \,\overline{\otimes}\, I_{\mathcal{K}}$.

Answer. If $T \in \mathcal{M}$, $S \in \mathcal{M}'$ and $R \in \mathcal{B}(\mathcal{K})$, then

$$(T \otimes I_{\mathcal{K}})(S \otimes R) = TS \otimes R = ST \otimes R = (S \otimes R)(T \otimes I_{\mathcal{K}}).$$

So (after taking linear combinations and sot-closure, which stay in the commutant) $\mathcal{M}' \otimes \mathcal{B}(\mathcal{K}) \subset (\mathcal{M} \otimes_{\min} I_{\mathcal{K}})'$.

Now consider $X \in (\mathcal{M}' \otimes \mathcal{B}(\mathcal{K}))'$. In particular $X(I_{\mathcal{M}} \otimes E_{kj}) = (I_{\mathcal{M}} \otimes E_{kj})X$ for all $\{E_{kj}\}$ matrix units for a fixed orthonormal basis. Then

$$X(\xi \otimes e_k) = X(I_{\mathcal{M}} \otimes E_{kk})(\xi \otimes e_k) = (I_{\mathcal{M}} \otimes E_{kk})X(\xi \otimes e_k) = \alpha(\xi) \otimes e_k$$

for some function $\alpha : \mathcal{H} \to \mathcal{H}$. By linearity and uniqueness of the tensor product when one side is linearly independent (so in particular for elementary tensors), α is linear. We also have

$$\|\alpha(\xi)\| = \|\alpha(\xi) \otimes e_k\| = \|X(\xi \otimes e_k)\| \leq \|X\| \, \|\xi\|,$$

so $\alpha \in \mathcal{B}(\mathcal{H})$. By linearity and continuity we get that $X = S \otimes I_{\mathcal{K}}$, where $S = \alpha \in \mathcal{B}(\mathcal{H})$. But X also commutes with elements of the form $T \otimes R$ with $T \in \mathcal{M}'$. So $ST \otimes R = TS \otimes R$ for all $R \in \mathcal{B}(\mathcal{K})$. Then

$$\langle (TS - ST)\xi, \eta \rangle = \langle ((TS - ST) \otimes I_{\mathcal{K}})(\xi \otimes e_k), \eta \otimes e_k \rangle = 0$$

for all $\xi, \eta \in \mathcal{H}$, so $TS = ST$ which shows that $S \in (\mathcal{M}')' = \mathcal{M}'' = \mathcal{M}$. We have shown that $(\mathcal{M}' \otimes \mathcal{B}(\mathcal{K}))' \subset \mathcal{M} \otimes I_{\mathcal{K}}$. Taking commutants we get $(\mathcal{M} \otimes_{\min} I_{\mathcal{K}})' \subset (\mathcal{M}' \otimes \mathcal{B}(\mathcal{K}))' = \mathcal{M}' \,\overline{\otimes} \mathcal{B}(\mathcal{K})$.

As for the centre, if $\tilde{T} \in \mathcal{Z}(\mathcal{M} \overline{\otimes} \mathcal{B}(\mathcal{K}))$ then $T \in (\mathcal{M} \overline{\otimes} \mathcal{B}(\mathcal{K}))' = \mathcal{M}' \overline{\otimes} I_{\mathcal{K}}$. So $\tilde{T} = T \otimes I_{\mathcal{K}}$ for some $T \in \mathcal{M}'$. Given $S \in \mathcal{M}'$

$$TS \otimes I_{\mathcal{K}} = (T \otimes I_{\mathcal{K}})(S \otimes I_{\mathcal{K}}) = (S \otimes I_{\mathcal{K}})(T \otimes I_{\mathcal{K}}) = ST \otimes I_{\mathcal{K}}.$$

Then, for any $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$ with $\|\eta\| = 1$,

$$\begin{aligned} \langle TS\xi, \xi \rangle &= \langle (TS \otimes I_{\mathcal{K}})(\xi \otimes \eta), \xi \otimes \eta \rangle \\ &= \langle (ST \otimes I_{\mathcal{K}})(\xi \otimes \eta), \xi \otimes \eta \rangle = \langle ST\xi, \xi \rangle. \end{aligned}$$

By polarization $TS = ST$, so $T \in \mathcal{M}'' = \mathcal{M}$. Hence $T \in \mathcal{Z}(\mathcal{M})$ and $\tilde{T} \in \mathcal{Z}(\mathcal{M}) \otimes I_{\mathcal{K}}$.

13.5. Crossed Products

(13.5.1) Let $(\mathcal{A}, G, \alpha)$ be a C*-dynamical system and (π, U) a covariant representation. Show that $\pi \rtimes U$ is indeed a representation.

Answer. For multiplicativity, let

$$A = \sum_{g \in F} a_g \cdot g, \qquad B = \sum_{h \in F} b_g \cdot h$$

(we use the same F since we can take the larger of two and make new coefficients equal to zero). Then

$$\begin{aligned}
(\pi \rtimes U)[AB] &= (\pi \rtimes U)\Big(\sum_{g,h \in F} (a_g \cdot g)(b_h \cdot h)\Big) \\
&= (\pi \rtimes U)\Big(\sum_{g,h \in F} a_g \alpha_g(b_h) \cdot gh\Big) \\
&= \sum_{g,h \in F} \pi(a_g \alpha_g(b_h)) U_{gh} \\
&= \sum_{g,h \in F} \pi(a_g)\, U_g \pi(b_h) U_g^*\, U_g\, U_h \\
&= \sum_{g,h \in F} \pi(a_g)\, U_g\, \pi(b_h)\, U_h \\
&= \Big(\sum_{g \in G} \pi(a_g) U_g\Big)\Big(\sum_{h \in F} \pi(b_h) U_h\Big) \\
&= (\pi \rtimes U)(A)\, (\pi \rtimes U)(B).
\end{aligned}$$

As for adjoint preservation,

$$\begin{aligned}(\pi \rtimes U)\Big[\Big(\sum_{g\in F} a_g \cdot g\Big)^*\Big] &= (\pi \rtimes U)\Big(\sum_{g\in F} \alpha_g^{-1}(a_g^*)\cdot g^{-1}\Big)\\ &= \sum_{g\in F} \pi(\alpha_g^{-1}(a_g^*))\, U_{g^{-1}}\\ &= \sum_{g\in F} U_{g^{-1}}\pi(a_g^*)U_{g^{-1}}^*\, U_{g^{-1}}\\ &= \sum_{g\in F} U_{g^{-1}}\pi(a_g^*) = \Big(\sum_{g\in F}\pi(a_g)\, U_g\Big)^*\\ &= \Big[(\pi \rtimes U)\Big(\sum_{g\in F} a_g\cdot g\Big)\Big]^*.\end{aligned}$$

(13.5.2) Show that when $\mathcal{A}$ is unital every $*$-representation $\beta : \mathcal{A}\cdot G \to \mathcal{B}(\mathcal{H})$ is $\beta = \pi \rtimes U$ for a covariant representation (π, U).

Answer. Let $\beta : \mathcal{A}\cdot G \to \mathcal{B}(\mathcal{H})$ be a $*$-representation. Define $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ by $\pi(a) = \beta(a\cdot e)$. The equalities in (13.25) and (13.26) show that

$$\pi(ab) = \beta(ab\cdot e) = \beta\big((a\cdot e)(b\cdot e)\big) = \beta(a\cdot e)\beta(b\cdot e) = \pi(a)\pi(b)$$

and

$$\pi(a^*) = \beta(a^*\cdot e) = \beta\big((a\cdot e)^*\big) = \beta(a\cdot e)^* = \pi(a)^*.$$

The additivity is trivial, and hence π is a $*$-representation. Similarly, let $U : G \to \mathcal{B}(\mathcal{H})$ be given by $U_g = \beta(I_{\mathcal{A}}\cdot g)$. Then

$$U_g^{-1} = \beta(I_{\mathcal{A}}\cdot g^{-1}) = U_{g^{-1}}, \qquad U_g^* = \beta\big((I_{\mathcal{A}}\cdot g)^*\big) = \beta(I_{\mathcal{A}}\cdot g^{-1}) = U_{g^{-1}},$$

so U_g is a unitary. We also have

$$U_{gh} = \beta(I_{\mathcal{A}}\cdot gh) = \beta\big((I_{\mathcal{A}}\cdot g)(I_{\mathcal{A}}\cdot h)\big) = U_gU_h,$$

so U is a representation. We have

$$\pi(\alpha_g(a)) = \beta(\alpha_g(a)\cdot e) = \beta\big((I_{\mathcal{A}}\cdot g)(a\cdot e)(I_{\mathcal{A}}\cdot g^{-1})\big) = U_g\pi(a)U_g^*,$$

so the representation is covariant. Finally,

$$\begin{aligned}(\pi \rtimes U)(a\cdot g) &= \pi(a)U_g = \beta(a\cdot e)\beta(I_{\mathcal{A}}\cdot g)\\ &= \beta\big((a\cdot e)(I_{\mathcal{A}}\cdot g)\big) = \beta(a\cdot g),\end{aligned}$$

so by linearity $\beta = \pi \rtimes U$.

(13.5.3) Let $\mathcal{A} = \mathbb{C}$ and $\alpha_g(g) = I_\mathcal{A}$ for al g. Show that $\mathbb{C} \rtimes_\alpha G = C^*(G)$ and $\mathbb{C} \rtimes_\alpha^r G = C^*_\lambda(G)$.

Answer. Because $\alpha_g(g) = I_\mathcal{A}$ for all G we have that the product in $\mathcal{A} \cdot G$ agrees with the plan product in $\mathbb{C}G$, and so does the adjoint. The fact that $\pi(\mathcal{A}) = \mathbb{C}\, I_\mathcal{H}$ makes the covariance condition trivial, too. So the covariant representations are precisely the representations where π is trivial (as it acts on the scalars) and U is any unitary representation of G. Then the norm used for the full crossed product and the norm for the full group algebra are the same, which gives $\mathbb{C} \rtimes_\alpha G = C^*(G)$.

For the regular representations the same happens: as π and α are trivial the norm of the reduced crossed product is calculated over the (unique) norm

$$\Big\| \sum_{g \in G} a_g\, \lambda(g) \Big\|$$

which is precisely the norm in $C^*_\lambda(G)$. Therefore $\mathbb{C} \rtimes_\alpha^r G = C^*_\lambda(G)$.

(13.5.4) Show that $\mathcal{A} \rtimes_\alpha^r G$ is a quotient of $\mathcal{A} \rtimes_\alpha G$ via a surjection that extends the identity map on $\mathcal{A} \cdot G$.

Answer. Since the norm of the reduced product is obtained over less representations that the full one, the identity map $\mathcal{A} \cdot G \to \mathcal{A} \cdot G$ is a bounded $*$-homomorphism when considered with the full norm on the domain and the reduced norm on the codomain. Then it extends to a $*$-homomorphism $\mathcal{A} \rtimes_\alpha G \to \mathcal{A} \rtimes_\alpha^r G$ with dense range and hence surjective.

(13.5.5) Show that (13.28) makes $(C_0(X), G, \alpha)$ a C*-dynamical system.

Answer. We need to check that α is a homomorphism and that $\alpha_s \in \operatorname{Aut} C_0(X)$ for all $s \in G$. For the latter,

$$\begin{aligned}
[\alpha_s(fg + \lambda h^*)](x) &= (fg + \lambda h^*)(s^{-1} \cdot x) \\
&= f(s^{-1} \cdot x) g(s^{-1} \cdot x) + \lambda \overline{h(s^{-1} \cdot x)} \\
&= [\alpha_s f](x)[\alpha_s g](x) + \lambda \overline{[\alpha_s h](x)}.
\end{aligned}$$

This shows that α_s is a $*$-homomorphism. If $\alpha_s f = 0$, then $f(s^{-1} \cdot x) = 0$ for all $x \in X$. In particular $f(x) = f(s^{-1} \cdot (s \cdot x)) = 0$ for all x, so $f = 0$ and α_s is injective. Given $f \in C_0(X)$, let $g(x) = f(s \cdot x)$. Because the map

$x \longmapsto s \cdot x$ is continuous, $g \in C_0(X)$; and $\alpha_s g = f$. So α_s is surjective, and thus $\alpha_s \in \operatorname{Aut} C_0(X)$.

It remains to check that α is a homomorphism. We have

$$[\alpha_{st} f](x) = f\big((st)^{-1} \cdot x\big) = f(t^{-1}s^{-1} \cdot x) = f\big(t^{-1} \cdot (s^{-1} \cdot x)\big)$$
$$= [\alpha_t f](s^{-1} \cdot x) = [\alpha_s \alpha_t f](x).$$

Hence $\alpha_{st} = \alpha_s \alpha_t$.

(13.5.6) Let G be a group, $\mathcal{A} = c_0(G)$, and $\alpha_g(f)(h) = f(g^{-1}h)$. Show that $c_0(G) \rtimes_\alpha G \simeq \mathcal{K}(\ell^2(G))$.

Answer. By remark Remark 13.5.1 we may assume that there is a faithful covariant representation (π, U) of $c_0(G) \rtimes_\alpha G \subset \mathcal{B}(\mathcal{H})$ into some $\mathcal{H}$. Let $P_g = \pi(\delta_g) \in \mathcal{B}(\mathcal{H})$. The family $\{P_g\}$ are pairwise orthogonal projections. We also have

$$U_g \pi(\delta_h) U_g^* = \pi(\alpha_g(\delta_h)) = \pi(\delta_{gh}).$$

Let $E_{g,h} = U_{gh^{-1}} \pi(\delta_h) \in \mathcal{B}(\mathcal{H})$. Then

$$\begin{aligned}
E_{a,b} E_{g,h} &= U_{ab^{-1}} \pi(\delta_b) U_{gh^{-1}} \pi(\delta_h) \\
&= U_{ab^{-1}} \pi(\delta_b) \pi(\delta_{gh^{-1}h}) U_{gh^{-1}} \\
&= \delta_{b,g} \, U_{ag^{-1}} \pi(\delta_g) U_{gh^{-1}} \\
&= \delta_{b,g} \, U_{ag^{-1}} \pi(\delta_{gh^{-1}\,h}) U_{gh^{-1}} \\
&= \delta_{b,g} \, U_{ag^{-1}} U_{gh^{-1}} \pi(\delta_h) \\
&= \delta_{b,g} \, U_{ah^{-1}} \pi(\delta_h) \\
&= \delta_{b,g} \, E_{a,h}.
\end{aligned}$$

That is, the family $\mathcal{F} = \{E_{g,h}\}_{g,h \in G}$ is a family of matrix units, and so $C^*(\mathcal{F}) \simeq \mathcal{K}(\ell^2(G))$. Since $\pi(\delta_g) \in C^*(\mathcal{F})$ for all g and $c_0(G) = C^*(\{\delta_g\}_g)$, we have $\pi(c_0(G)) \subset C^*(\mathcal{F})$. We also have by construction that $\pi(\delta_g) U_h = U_{h^{-1}} \pi(\delta_{h^{-1}g}) \in C^*(\mathcal{F})$. Hence $C^*(\mathcal{F}) = (\pi \rtimes U)(c_0(G) \rtimes_\alpha G)$ and therefore $c_0(G) \rtimes_\alpha G \simeq \mathcal{K}(\ell^2(G))$.

(13.5.7) Let (G, X) be a locally compact transformation group. Show that if $\operatorname{Aut} C_0(X)$ is considered with the pointwise-norm topology, then α as in (13.28) is continuous.

Answer. We want to show that $\lim_{t\to s} \|\alpha(t)f - \alpha(s)f\|_\infty \to 0$. Since we can write

$$\|\alpha_t f - \alpha_s f\|_\infty = \|\alpha_s\,(\alpha_{s^{-1}t} f - f\|_\infty$$

and $t \to s$ if and only if $s^{-1}t \to e$, we only need to show that $\|\alpha_s f - f\|_\infty \to 0$ when $s \to e$. For this latter property to fail we would have an $\varepsilon > 0$ and nets $\{s_j\} \subset G$, $\{x_j\} \subset X$ with $s_j \to e$ and

$$\|f(s_j^{-1} \cdot x_j) - f(x_j)| \geq \varepsilon \tag{AB.13.5}$$

for all j. Let $K = \{|f| \geq \varepsilon/2\}$; this is compact because $f \in C_0(X)$ (Exercise 2.6.2). The inequality (AB.13.5) then means that at least one of $s_j^{-1} \cdot x_j \in K$ or $x_j \in K$ holds. Let V be a neighbourhood of e with compact closure. Then $s_j \in V$ for all big enough j. The set $\overline{V} \cdot K$ is compact, being a continuous image of the compact set $\overline{V} \times K$. If the set of x_j in K is infinite, then it admits a convergent subnet. By picking only those we have that $x_j \to x_0$ for some $x_0 \in K_j$; then $s_j^{-1} \cdot x_j \to x_0$, and by the continuity of f this contradicts (AB.13.5). If instead we have infinitely many $s_j^{-1} \cdot x_j$ in K for j big enough $s_j \in V$ and then $x_j = s_j \cdot (s_j^{-1} \cdot x_j) \in \overline{V} \cdot K$. As this latter set is compact, again we get that x_j admits a convergent subnet and we can repeat the previous argument.

CHAPTER 14

von Neumann Algebras

14.1. Subalgebras

(14.1.1) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $P \in \mathcal{M}$ a projection, $\mathcal{K} = \overline{\mathcal{M}P\mathcal{H}}$, and $Q \in \mathcal{B}(\mathcal{H})$ the orthogonal projection onto $\mathcal{K}$. Show that $\mathcal{K}$ is invariant for $\mathcal{M}$ and $\mathcal{M}'$, and conclude that $Q \in \mathcal{Z}(\mathcal{M})$.

Answer. Fix $S \in \mathcal{M}$. For any $T \in \mathcal{M}$ and $\xi \in \mathcal{H}$, $S(TP\xi) = (ST)P\xi \in \mathcal{M}P\mathcal{H}$. As S is continuous, $S\mathcal{K} \subset \mathcal{K}$. Similarly, if $S \in \mathcal{M}'$ and $T \in \mathcal{M}$, $\xi \in \mathcal{H}$, we have $S(TP\xi) = TP(S\xi) \in \mathcal{M}P\mathcal{H}$; again by linearity and continuity, $S\mathcal{K} \subset \mathcal{K}$. Thus $\mathcal{K}$ is invariant for both $\mathcal{M}$ and $\mathcal{M}'$.

Now consider the orthogonal projection Q onto $\mathcal{K}$. Fix $S \in \mathcal{M}^{\rm sa}$. For any $\xi \in \mathcal{H}$, we have by the invariance $SQ\xi \in \mathcal{K}$, so $QSQ\xi = SQ\xi$. This can be done for all $\xi \in \mathcal{H}$, giving us $QSQ = SQ$. Using that $S = S^*$,

$$QS = (SQ)^* = (QSQ)^* = QSQ = SQ.$$

This says that Q commutes with all sefadjoints in $\mathcal{M}$; but the selfadjoint elements span the whole algebra (as any element $T \in \mathcal{M}$ can be written as $T = \operatorname{Re} T + i \operatorname{Im} T$), so $Q \in \mathcal{M}'$. The previous computations would have been the exact same if $S \in \mathcal{M}'$, selfadjoint, so we also get that $Q \in \mathcal{M}'' = \mathcal{M}$. Thus $Q \in \mathcal{M} \cap \mathcal{M}' = \mathcal{Z}(\mathcal{M})$.

(14.1.2) Let $\mathcal{M}$ be a factor, $S \in \mathcal{M}$, $T \in \mathcal{M}'$. Show that $ST = 0$ if and only if $S = 0$ or $T = 0$.

Answer. Suppose that $ST = 0$. If $T = 0$ we are done. Otherwise, by Exercise 12.4.8 and Corollary 12.4.15 the rank projection P of T is in $\mathcal{M}'$. Since P is a wot limit of polynomials $p_j(T)$, with $p_j(0) = 0$ for all j, we have $SP = \lim_{\text{wot}} Sp_j(T) = 0$. With the notation of Corollary 14.1.4 (where we use that $\mathcal{M}$ is a factor to get $c(P) = I_{\mathcal{M}}$), we have $\gamma(S) = 0$ and hence $S = 0$.

The converse is trivial.

(14.1.3) Let $P, Q, Z \in \mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be projections with Z central and $Q = ZP$. Show that $c(Q) = Zc(P)$.

Answer. Since

$$\mathcal{M}Q\mathcal{H} = \mathcal{M}ZP\mathcal{H} = Z\,\mathcal{M}P\mathcal{H},$$

via taking closures we get $c(Q) = Zc(P)$.

(14.1.4) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{R} \subset \mathcal{M}$ a subset that is closed under multiplication and taking adjoints, and such that $W^*(\mathcal{R}) = \mathcal{M}$. Let $P \in \mathcal{M} \cup \mathcal{M}'$ be a projection. Show that $W^*(P\mathcal{R}P) = P\mathcal{M}P$. Show also that the result is not necessarily true if $\mathcal{R}$ is not closed under multiplication.

Answer. From $P\mathcal{R}P \subset P\mathcal{M}P$ we get $W^*(P\mathcal{R}P) \subset P\mathcal{M}P$. The assumptions on $\mathcal{R}$ guarantee that $\mathcal{M} = \overline{\text{span}}\mathcal{R}$. Then

$$P\mathcal{M}P \subset \overline{\text{span}}P\mathcal{R}P \subset W^*(P\mathcal{R}P).$$

For an example when $\mathcal{R}$ is not closed under multiplication, consider $\mathcal{M} = M_2(\mathbb{C})$, and $\mathcal{R} = \{E_{12}\}$. Then $W^*(\mathcal{R}) = \mathcal{M}$ (because $W^*(\mathcal{R})$ contains E_{12}, $E_{12}^* = E_{21}$, $E_{12}E_{12}^* = E_{11}$). Let $P = E_{11}$. Then $P\mathcal{M}P = \mathbb{C}\,E_{11}$, while $P\mathcal{R}P = \{0\}$.

(14.1.5) Let $\mathcal{M}$ be a von Neumann algebra, $\mathcal{H}$ a Hilbert space, $P \in \mathcal{M}$ a projection. Fix also an orthonormal basis for $\mathcal{H}$ and let $\{E_{kj}\}$ be the associated matrix units. Show that

$$PMP \simeq (P \otimes E_{11})(\mathcal{M} \,\overline{\otimes}\, \mathcal{B}(\mathcal{H}))(P \otimes E_{11}).$$

Answer. Most work was done in Exercise 13.4.4, but we will provide a self-contained version here. Let $\gamma : P\mathcal{M}P \to (P \otimes E_{11})(\mathcal{M} \,\overline{\otimes}\, \mathcal{B}(\mathcal{H}))(P \otimes E_{11})$ be given by

$$\gamma(PTP) = PTP \otimes E_{11}.$$

This is well-defined, since $PTP \otimes E_{11} = (P \otimes E_{11})(PTP \otimes I_{\mathcal{H}})(P \otimes E_{11}) \in (P \otimes E_{11})(\mathcal{M} \,\overline{\otimes}\, \mathcal{B}(\mathcal{H}))(P \otimes E_{11})$. It is linear, for

$$\begin{aligned}\gamma(PTP + \lambda PSP) &= \gamma(P(T + \lambda S)P) = P(T + \lambda S)P \otimes E_{11} \\ &= PTP \otimes E_{11} + \lambda PSP \otimes E_{11} = \gamma(T) + \lambda\gamma(S).\end{aligned}$$

Similarly,

$$\begin{aligned}\gamma(PTPSP) &= PTPSP \otimes E_{11} = (PTP \otimes E_{11})(PSP \otimes E_{11}) \\ &= \gamma(T)\gamma(S)\end{aligned}$$

(this is the first place where we use that E_{11} is a projection). And

$$\gamma((PTP)^*) = (PTP)^* \otimes E_{11} = (PTP \otimes E_{11})^* = \gamma(T)^*$$

(here we use that E_{11} is selfadjoint). So γ is a $*$-homomorphism. If $\gamma(PTP) = 0$, this means that $PTP \otimes E_{11} = 0$ and so $PTP = 0$ since $E_{11} \neq 0$ (Proposition 13.1.3). And γ is surjective by Exercise 13.4.4. Here is an ad-hoc argument for the surjectivity. Given $\tilde{T} \in \mathcal{M} \otimes \mathcal{B}(\mathcal{H})$ we can write $T = \sum_{k=1}^n T_k \otimes S_k$, with $T_1, \ldots, T_n \in \mathcal{M}$ and $S_1, \ldots, S_n \in \mathcal{B}(\mathcal{H})$. For each k we have $E_{11}SE_{11} = \lambda_k E_{11}$ since E_{11} is minimal. Then

$$\begin{aligned}(P \otimes E_{11})\tilde{T}(P \otimes E_{11}) &= \sum_{k=1}^n PT_kP \otimes E_{11}S_kE_{11} = \sum_{k=1}^n \lambda_k PT_kP \otimes E_{11} \\ &= \gamma\Big(\sum_{k=1}^n \lambda_k PT_kP\Big).\end{aligned}$$

Hence γ has dense range. By Proposition 11.4.9, γ is surjective.

14.2. Comparison of Projections

(14.2.1) Let $P \in \mathcal{M}$ be a projection with $P \precsim 0$. Show that $P = 0$.

Answer. By hypothesis we have $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* = 0$. Then

$$\|P\| = \|V^*V\| = \|V\|^2 = \|V^*\|^2 = \|VV^*\| = 0,$$

so $P = 0$. Or, even simpler, we note that $VV^* = 0$ implies $V = 0$ and then $P = V^*V = 0$.

(14.2.2) Let $P, Q \in M_n(\mathbb{C})$ be projections. Show that $P \sim Q$ if and only if $\operatorname{Tr}(P) = \operatorname{Tr}(Q)$.

Answer. If $P \sim Q$, then there exists $V \in M_n(\mathbb{C})$ with $V^*V = P$ and $VV^* = Q$. Therefore

$$\operatorname{Tr}(P) = \operatorname{Tr}(V^*V) = \operatorname{Tr}(VV^*) = \operatorname{Tr}(Q).$$

Now suppose that $\operatorname{Tr}(P) = \operatorname{Tr}(Q)$. By fixing an orthonormal basis for $P\mathcal{H}$ and extending it to an orthonormal basis for the whole $\mathbb{C}^n$, we can see P as a diagonal matrix with $\dim P\mathcal{H}$ ones in the diagonal, and the rest zeros. Thus $\operatorname{Tr}(P) = \dim P\mathcal{H}$. The equality $\operatorname{Tr}(P) = \operatorname{Tr}(Q)$ gives us $\dim P\mathcal{H} = \dim Q\mathcal{H}$. Fix orthonormal bases $\{e_1, \ldots, e_r\}$ and $\{f_1, \ldots, f_r\}$ for $P\mathcal{H}$ and $Q\mathcal{H}$ respectively, and define V to be the linear operator that maps $Ve_k = f_k$, and $V = 0$ on $(P\mathcal{H})^\perp$. Then V^* is the operator that maps f_k to e_k, and is zero on $(Q\mathcal{H})^\perp$. So $V^*V = P$ and $VV^* = Q$.

(14.2.3) Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra and $P, Q \in \mathcal{M}$ be projections. Show that $P \sim Q$ if and only if $\operatorname{Tr}(ZP) = \operatorname{Tr}(ZQ)$ for every central projection Z.

Answer. If $P \sim Q$ there exists $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* = Q$. Then

$$\begin{aligned}\operatorname{Tr}(ZP) &= \operatorname{Tr}(PZ) = \operatorname{Tr}(V^*VZ) = \operatorname{Tr}(VZV^*)\\ &= \operatorname{Tr}(ZVV^*) = \operatorname{Tr}(ZQ).\end{aligned}$$

Conversely, suppose that $\mathrm{Tr}(ZP) = \mathrm{Tr}(ZQ)$ for each central projection. We know from Theorem 11.8.10 that $\mathcal{M} = \bigoplus_{j=1}^{k} M_{n_j}(\mathbb{C})$ (properly, up to isomorphism). For each j we can consider the central projection $Z_j = 0 \oplus \cdots \oplus I_{n_j} \oplus 0 \oplus \cdots \oplus 0$. As $\mathrm{Tr}(Z_j P) = \mathrm{Tr}(Z_j Q)$, we can use Exercise 14.2.2 to get a partial isometry $V_j \in M_{n_j}(\mathbb{C})$ with $V_j^* V_j = Z_j P$ and $V_j V_j^* = Z_j Q$. Then $V = V_1 \oplus \cdots \oplus V_k$ is a partial isometry in $\mathcal{M}$ with $V^*V = P$ and $VV^* = Q$.

(14.2.4) Let $P, Q, R \in \mathcal{M}$ be projections, with $P \preceq Q$ and $Q \preceq R$. Show that $P \preceq R$.

Answer. By hypothesis there exist $V, W \in \mathcal{M}$ with $V^*V = P$, $VV^* \leq Q$, $W^*W = Q$, $WW^* \leq R$. Consider $WV \in \mathcal{M}$; this is a partial isometry, though we don't need to know that in advance. We have

$$(WV)^*WV = V^*W^*WV = V^*QV = V^*VV^*V = P^2 = P$$

and

$$WV(WV)^* = WVV^*W^* \leq WQW^* = WW^* \leq R.$$

Hence $P \preceq R$.

(14.2.5) Let $P, Q, Z \in \mathcal{M}$ be projections, with $Z \in \mathcal{Z}(\mathcal{M})$ and such that $P \preceq Q$. Show that $ZP \preceq ZQ$.

Answer. By definition there exists $V \in \mathcal{M}$ with $V^*V = P$, $VV^* \leq Q$. Then

$$(ZV)^*(ZV) = ZV^*V = ZP, \qquad (ZV)(ZV)^* = ZVV^* \leq ZQ,$$

showing that $ZP \preceq ZQ$.

(14.2.6) Let $P, Q \in \mathcal{M}$ with $P \preceq Q$ and $Q \preceq P$. Prove that $P \sim Q$ by structuring the argument the same as the proof of Schröder–Bernstein (Theorem 1.6.13).

Answer. By hypothesis there exist $V, W \in \mathcal{M}$ with $V^*V = P$, $VV^* \leq Q$, $W^*W = Q$, $WW^* \leq P$. It is enough to show that $P \sim WW^*$, since $WW^* \sim Q$. We define projections

$$P_0 = P, \quad R_1 = WW^*, \quad P_{k+1} = WVR_{k+1}V^*W^*, \quad R_{k+1} = WVP_kV^*W^*.$$

By construction $R_1 \leq P_0$ and

$$P_1 = WVR_1V^*W^* \leq WVPV^*W^* = WVV^*W^* \leq WW^* = R_1.$$

Repeated inductively, we get that $R_{k+1} \le P_k$ and $P_{k+1} \le R_{k+1}$ for all k. Hence

$$P_0 \ge R_1 \ge P_1 \ge R_2 \ge P_2 \ge \cdots$$

Assume for the moment that $P_k - R_{k+1} \sim P_{k+1} - R_{k+2}$ for all k. Let

$$P_\infty = \bigwedge_k P_k = \bigwedge_k R_k.$$

Since $P_\infty = \lim_{\text{sot}} P_k = \lim_{\text{sot}} R_k$,

$$P = P_\infty + \sum_{k=0}^{\infty} \big(P_k - R_{k+1}\big) + \sum_{k=1}^{\infty} \big(R_k - P_k\big)$$

and

$$R_1 = P_\infty + \sum_{k=1}^{\infty} \big(P_k - R_{k+1}\big) + \sum_{k=1}^{\infty} \big(R_k - P_k\big).$$

By Proposition 14.2.7, $P \sim R_1$. So it remains to prove that $P_k - R_{k+1} \sim P_{k+1} - R_{k+2}$. We have

$$WVWVP_kV^*W^*V^*W^* = WVR_{k+1}V^*W^* = P_{k+1}$$

and

$$WVWVR_{k+1}V^*W^*V^*W^* = WVP_{k+1}V^*W^* = R_{k+2}.$$

Also, since $V^*W^*WV = V^*QV = V^*V = P$

$$\begin{aligned} V^*W^*V^*W^*P_{k+1}WVWV &= V^*W^*V^*W^*(WVR_{k+1}V^*W^*)WVWV \\ &= V^*W^*P_{k+1}PWV = V^*W^*R_{k+1}WV \\ &= V^*W^*(WVP_kV^*W^*)WV = PP_kP = P_k \end{aligned}$$

and

$$\begin{aligned} V^*W^*V^*W^*R_{k+2}WVWV &= V^*W^*V^*W^*(WVP_{k+1}V^*W^*)WVWV \\ &= V^*W^*PP_{k+1}PWV = V^*W^*P_{k+1}WV \\ &= V^*W^*(WVR_{k+1}V^*W^*)WV \\ &= PR_{k+1}P = R_{k+1}. \end{aligned}$$

Let $U = (P_k - R_{k+1})V^*W^*V^*W^*(P_{k+1} - R_{k+2})$. Then the equalities above show that

$$U^*U = P_{k+1} - R_{k+2}, \qquad UU^* = P_k - R_{k+1}.$$

(14.2.7) Let $\mathcal{M}$ be a von Neumann algebra and $\{P_j\}$ and $\{Q_j\}$ two families of pairwise orthogonal projections in $\mathcal{M}$, such that $P_j \preceq Q_j$ for all j. Show that $\sum_j P_j \preceq \sum_j Q_j$.

Answer. By hypothesis there exist projections $Q_j' \leq Q_j$ such that $P_j \sim Q_j'$ for all j. The projections $\{Q_j'\}$ are pairwise orthogonal, since the $\{Q_j\}$ are. By Proposition 14.2.7,

$$\sum_j P_j \sim \sum_j Q_j' \leq \sum_j Q_j.$$

(14.2.8) Show that if in Proposition 14.2.16 $\mathcal{M}$ is σ-finite and Q is properly infinite, the result holds without requiring the P_n to be finite.

Answer. Let $P = \bigvee_n P_n$. Since $P_n \preceq Q$, then $P_n \leq c(P_n) \leq c(Q)$ by Proposition 14.2.5. Therefore, using Corollary 14.2.18,

$$P \leq c(P) \leq c(Q) \sim Q,$$

so $P \preceq Q$.

(14.2.9) Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space. Let $P, Q \in \mathcal{B}(\mathcal{H})$ be infinite projections. Show that $P \sim Q$. Is $I_{\mathcal{H}} - Q \sim I_{\mathcal{H}} - Q$?

Answer. This follows directly from Corollary 14.2.18, but we offer here a constructive argument in this simple case.

Since equivalence is transitive, it is enough to show that $P \sim I_{\mathcal{H}}$. We have that $\dim P\mathcal{H} = \infty$, for if $\dim P\mathcal{H} < \infty$ then P is finite (because a partial isometry with finite-dimensional initial space preserves dimension). Fix an orthonormal basis $\{\xi_n\}$ for $\mathcal{H}$, and an orthonormal basis $\{\eta_n\}$ for $P\mathcal{H}$. Let V be the bounded linear operator induced by $V\eta_n = \xi_n$, and $V|_{(P\mathcal{H})^\perp} = 0$. Then, given $\xi \in \mathcal{H}$ and writing $\xi = \xi_0 + \sum_n c_n \eta_n$, with $\xi_0 \in (P\mathcal{H})^\perp$,

$$\langle V^*\xi_n, \xi\rangle = \langle \xi_n, V\xi\rangle = \sum_k c_k \langle \xi_n, V\eta_k\rangle = c_n = \langle \eta_n, \xi\rangle.$$

Thus $V^*\xi_n = \eta_n$ for all n. Then $V^*V\eta_n = V^*\xi_n = \eta_n$, showing that $V^*V = P$, while $VV^*\xi_n = V\eta_n = \xi_n$ for all n, and then $VV^* = I_{\mathcal{H}}$.

The equality of the complements can fail, by taking for instance $P = I_{\mathcal{H}}$ and $Q = I_{\mathcal{H}} - E_{11}$ in $\mathcal{B}(\ell^2(\mathbb{N}))$. Then $I_{\mathcal{H}} - P = 0$ and $I_{\mathcal{H}} - Q = E_{11}$ are not equivalent. Equally dramatic, we can have $P \sim Q$ with $I_{\mathcal{H}} - P$ infinite and $I_{\mathcal{H}} - Q$ finite.

(14.2.10) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $T \in \mathcal{M}$. Show that $[T\mathcal{H}] \sim [|T|\mathcal{H}] = [T^*\mathcal{H}]$ in $\mathcal{M}$.

Answer. The projections are the range projections of T, $|T|$, and T^* respectively, so they are in $\mathcal{M}$ by Corollary 12.3.8. We know from Proposition 10.4.11 that if $T = V|T|$ is the polar decomposition of T, then $V^*V = [T^*\mathcal{H}] = [|T|\mathcal{H}]$ and $VV^* = [T\mathcal{H}]$. This proves the equivalence; said equivalence occurs in $\mathcal{M}$ by Corollary 12.3.8.

(14.2.11) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $P \in \mathcal{M}$ a projection, and $\xi \in \mathcal{H}$. Show that $[P\mathcal{M}\xi] = P[\mathcal{M}\xi]$.

Answer. From $[P\mathcal{M}\xi] \leq P$ (since $P\mathcal{M}\xi \subset P\mathcal{H}$) and $[P\mathcal{M}\xi] \leq [\mathcal{M}\xi]$ (since $P\mathcal{M} \subset \mathcal{M}$) we get that

$$[P\mathcal{M}\xi] \leq P \wedge [\mathcal{M}\xi] = P[\mathcal{M}\xi],$$

the latter equality because $P \in \mathcal{M}$ and $[\mathcal{M}\xi] \in \mathcal{M}'$ so they commute (a proper justification comes from Proposition 12.1.17). For the reverse inequality, since $P[\mathcal{M}\xi]\mathcal{H} \subset \overline{P\mathcal{M}\xi}$, we have $[P\mathcal{M}\xi](P[\mathcal{M}\xi]) = P[\mathcal{M}\xi]$. This implies that $P[\mathcal{M}\xi] \leq [P\mathcal{M}\xi]$, and so the equality is established.

(14.2.12) Let $\mathcal{M}$ be a von Neumann algebra, and $P \in \mathcal{M}$ a projection. Show that P is minimal if and only if $Q \leq P$ for a projection $Q \in \mathcal{M}$, implies that either $Q = P$ or $Q = 0$.

Answer. If P is minimal, then $P\mathcal{M}P = \mathbb{C}\,P$. Then the only projections there are 0 and P.

Conversely, suppose that the only subprojections of P are 0 and P. Any projection $Q \in P\mathcal{M}P$ satisfies $Q \leq P$ (because P is the identity of $P\mathcal{M}P$, and $Q \geq 0$ with $\|Q\| = 1$). So by hypothesis the only projections in $P\mathcal{M}P$ are 0 and P. By Corollary 12.4.16,

$$P\mathcal{M}P = \overline{\operatorname{span}}^{\|\cdot\|}\{0, P\} = \mathbb{C}\,P.$$

(14.2.13) Let $\mathcal{M} = M_n(\mathbb{C})$. Find all the minimal projections. Find all the abelian projections. Show that $\mathcal{M}$ is finite.

Answer. The algebra $P\mathcal{M}P$ consists precisely of the matrices such that they and their adjoint leave $P\mathbb{C}^n$ invariant. The subspace $P\mathbb{C}^n$ has, by definition, dimension $\operatorname{rank} P$. If $\dim P\mathbb{C}^n \geq 2$, $P\mathcal{M}P \simeq M_{\operatorname{rank} P}(\mathbb{C})$ is non-commutative. Thus if P is minimal, then $\operatorname{rank} P = 1$. And if $\operatorname{rank} P = 1$, then $\dim P\mathbb{C}^n = 1$, so $P\mathcal{M}P$ is necessarily one-dimensional, which makes it commutative. The argument shows both that the minimal and abelian projections are the rank-one projections.

To show that I_n is finite, suppose that $V^*V = I_n$. This means that V is injective; and by Exercise 1.7.8 it is surjective. Then it is invertible, and $V^{-1} = V^*$. So $VV^* = I_n$, and hence I_n is finite.

(14.2.14) Let $\{P_j\} \subset \mathcal{M}$ be a family of pairwise equivalent projections. Fix a projection $P \in \mathcal{M}$ with $P \sim P_j$ for all j. Show that

$$c\Big(\bigvee_j P_j\Big) = c(P).$$

Answer. Write $Q = \bigvee_j P_j$. From $P_j \leq Q$ we have $c(P) = c(P_j) \leq c(Q)$. But we also have, for any $\xi \in \mathcal{H}$ and any j, $c(P)P_j\xi = P_j\xi$. It follows that $\bigcup_j P_j\mathcal{H}$ is invariant for $c(P)$. That is, $c(Q) \leq c(P)$. Therefore $c(Q) = c(P)$.

(14.2.15) Let $P, Q \in \mathcal{M}$ be equivalent. Show that $P\mathcal{M}P \simeq Q\mathcal{M}Q$.

Answer. By hypothesis there exists $V \in \mathcal{M}$ with $V^*V = P$, $VV^* = Q$. Define $\gamma : P\mathcal{M}P \to Q\mathcal{M}Q$ by $\gamma(T) = VTV^*$. This map is clearly linear and it maps into $Q\mathcal{M}Q$ since $V = QV$ (Exercise 10.4.12). We have $\gamma(T^*) = VT^*V^* = (VTV^*)^* = \gamma(T)^*$. A key observation is that $T \in P\mathcal{M}P$ if and only if $T = PTP$. And, given $S, T \in P\mathcal{M}P$,

$$\gamma(S)\gamma(T) = VSV^*VTV^* = VSPTV^* = VSTV^* = \gamma(ST),$$

so γ is a $*$-homomorphism. If $T \in P\mathcal{M}P$ and $\gamma(T) = 0$, this means that $VTV^* = 0$; multiplying by V^* on the left and by V on the right, we get $0 = V^*VTV^*V = PTP = T$, so $T = 0$; meaning that γ is injective. Finally, given $S \in Q\mathcal{M}Q$, we have $S = QSQ = VV^*SVV^* = \gamma(V^*SV)$. We will have shown that γ is surjective if we show that $V^*SV \in P\mathcal{M}P$. And this follows from $V = VP$ (Exercise 10.4.12).

(14.2.16) Let $P, Q \in \mathcal{M}$ be equivalent projections. Show that if P is any of minimal, abelian, infinite, finite, purely infinite, properly infinite, then so is Q.

Answer.

(i) If P is minimal, then $P\mathcal{M}P = \mathbb{C}\,P$. By Exercise 14.2.15 $Q\mathcal{M}Q$ is one-dimensional, so $Q\mathcal{M}Q = \mathbb{C}\,Q$ since $\mathbb{C}\,Q$ lies inside it.

(ii) If P is abelian, then $Q\mathcal{M}Q \simeq P\mathcal{M}P$ by Exercise 14.2.15 and so $Q\mathcal{M}Q$ is abelian.

(iii) If P is infinite, there exists $P_0 \leq P$ with $P_0 \sim P$ and $P - P_0 \neq 0$. Let $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* = Q$. Put $Q_0 = VP_0V^* \in \mathcal{M}$. We have $Q_0 \leq VV^* = Q$, and $Q - Q_0 = V(P - P_0)V^* \neq 0$, for if it were 0 we could multiply by V^* on the left and by V on the right to get $P - P_0 = 0$. We have $Q_0 \sim P_0$, since $Q_0 = VP_0V^*$ and $P_0V^*VP_0 = P_0PP_0 = P_0$. So $Q_0 \sim P_0 \sim P \sim Q$, and Q is infinite.

(iv) If P is finite then Q is finite, for if Q were infinite then P would be infinite by the previous paragraph.

(v) Suppose that P is purely infinite and $Q_0 \leq Q$. Writing $P = V^*V$ and $Q = VV^*$, we have $P_0 = V^*Q_0V \sim Q_0$ and $P_0 \leq P$. Then P_0 is infinite (as P is purely infinite) and Q_0 is infinite by (iii). So Q is purely infinite.

(vi) If P is properly infinite, this means that ZP is infinite for all nonzero projections $Z \in \mathcal{Z}(\mathcal{M})$. Then $ZQ \sim ZP$ is infinite, and it follows that Q is properly infinite.

(14.2.17) Let $P, Q \in \mathcal{M}$ be projections with $Q \preceq P$. Show that if P is finite, then Q is finite.

Answer. Suppose that Q is infinite. There exists $Q' \leq P$ with $Q' \sim Q$. So Q' is infinite by Exercise 14.2.16. Then there exists $Q_0 \leq Q' \leq P$ with $Q_0 \neq Q'$ and $Q_0 \sim Q'$. By Proposition 14.2.7,

$$P = P - Q' + Q' \sim P - Q' + Q_0,$$

and $P - (P - Q' + Q_0) = Q' - Q_0 \neq 0$, so P is infinite.

(14.2.18) Let $P, Q \in \mathcal{M}$ be projections with $Q \preceq P$. Show that if P is abelian, then Q is abelian.

Answer. By Exercise 14.2.16 we may assume without loss of generality that $Q \leq P$. By hypothesis, $P\mathcal{M}P$ is abelian and $Q = PQP \in P\mathcal{M}P$. Then

$$Q\mathcal{M}Q = Q(P\mathcal{M}P)Q = (P\mathcal{M}P)Q$$

is abelian, so Q is abelian.

(14.2.19) Let $P, Z \in \mathcal{M}$ be projections with Z central. Show, without using Proposition 14.2.15, that P is finite if and only if ZP and $(I_\mathcal{M} - Z)P$ are finite.

Answer. This is part of the proof of Proposition 14.2.15.

If P is finite, then ZP and $(I_\mathcal{M}-)P$ are finite by Exercise 14.2.17. Conversely, suppose that ZP and $(I_\mathcal{M} - Z)P$ are finite. Let $Q \leq P$ with $Q \sim P$. As Z is central, we have $ZQ \sim ZP$ and $(I_\mathcal{M}-Z)Q \sim (I_\mathcal{M}-Z)P$. The two right-hand-side projections are finite and $ZQ \leq ZP$ and $(I_\mathcal{M} - Z)Q \leq (I_\mathcal{M} - Z)P$, so it follows that $ZQ = ZP$ and $(I_\mathcal{M} - Z)Q = (I_\mathcal{M} - Z)P$. Adding, $Q = P$ and hence P is finite.

(14.2.20) Let $\mathcal{M}$ be a finite von Neumann algebra. Show that the only isometries are the unitaries.

Answer. If V is an isometry, we have $V^*V = I_\mathcal{M}$. Then $VV^* \leq I_\mathcal{M}$ and $VV^* \sim V^*V = I_\mathcal{M}$. As $I_\mathcal{M}$ is finite, $VV^* = I_\mathcal{M}$.

(14.2.21) Let Φ as in the proof of Proposition 14.2.21. Show that it is faithful.

Answer. Let

$$X = \sum_{j_1, j_2} X_{j_1,j_2} \otimes E_{j_1,j_2} \in \mathcal{A} \otimes \mathcal{B}(\ell^2(|J|))$$

and suppose that $\Phi(X^*X) = 0$. This means that

$$\begin{aligned} 0 &= \Phi\Big(\sum_{j_1,j_2,j_3,j_4} X^*_{j_3,j_4} X_{j_1,j_2} \otimes E_{j_4,j_3} E_{j_1,j_2}\Big) \\ &= \Phi\Big(\sum_{j_1,j_2,j_4} X^*_{j_1,j_4} X_{j_1,j_2} \otimes E_{j_4,j_2}\Big) \\ &= \sum_{j_1} \Phi\Big(\sum_{j_2,j_4} X^*_{j_1,j_4} X_{j_1,j_2} \otimes E_{j_4,j_2}\Big) \\ &= \sum_{j_1} \frac{1}{n} \sum_{j_2} X^*_{j_1,j_2} X_{j_1,j_2}. \end{aligned}$$

As every term is non-negative, $X_{j_1,j_2} = 0$ for all j_1, j_2, and thus $X = 0$.

(14.2.22) Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space. Fix an orthonormal basis and consider the associated unilateral shift S. Show that S is a partial isometry, and conclude that $\mathcal{B}(\mathcal{H})$ is infinite.

Answer. If the basis is $\{\xi_n\}$, then the unilateral shift is the bounded linear operator V induced by $V\xi_n = \xi_{n+1}$. So

$$V = \sum_n E_{n+1,n}.$$

This series converges sot, for given $\xi = \sum_k c_k \xi_k$ we have

$$\begin{aligned} \Big\|\sum_{n>n_0} E_{n+1,n}\xi\Big\|^2 &= \Big\|\sum_{n>n_0}\sum_k c_k\, E_{n+1,n}\xi_k\Big\|^2 \\ &= \Big\|\sum_{n>n_0} c_n \xi_{n+1}\Big\|^2 = \sum_{n>n_0} |c_n|^2, \end{aligned}$$

which can be made arbitrarily small by taking n_0 sufficiently large. A similar argument shows that the corresponding series converges for V^* and that the adjoint for V is the series of the adjoints. This gives us, via Proposition 12.1.13,

$$V^*V = \sum_{n,m} E_{m,m+1} E_{n+1,n} = \sum_n E_{n,n} = I_{\mathcal{H}}.$$

Meanwhile,

$$VV^* = \sum_{n,m} E_{n+1,n} E_{m,n+1} = \sum_n E_{n+1,n+1} = I_{\mathcal{H}} - E_{11}.$$

This shows that $I_{\mathcal{H}}$ is infinite, and so $\mathcal{B}(\mathcal{H})$ is infinite.

(14.2.23) Show an example of a infinite projection that cannot be halved.

Answer. The example cannot occur in a factor, for there every infinite projection is properly infinite. Let $\mathcal{M} = \mathbb{C} \oplus \mathcal{B}(\mathcal{H})$, with $\dim \mathcal{H} = \infty$. The projection we consider is the identity $I_{\mathcal{M}} = 1 \oplus I_{\mathcal{H}}$. Suppose that $1 \oplus I_{\mathcal{H}} = P + Q$ for projections $P, Q \in \mathcal{M}$. We have $P = \alpha \oplus P_0$, $Q = \beta \oplus Q_0$, with $\alpha, \beta \in \{0, 1\}$ and $P_0, Q_0 \in \mathcal{B}(\mathcal{H})$ projections. Since $\alpha + \beta = 1$, we may assume without loss of generality that $\alpha = 1$, $\beta = 0$. Suppose that $V^*V = P$ and $VV^* = Q$. Since $V = \lambda \oplus V_0$ with $\lambda \in \mathbb{C}$ and $V_0 \in \mathcal{B}(\mathcal{H})$ a partial isometry, we have

$$1 \oplus P_0 = P = V^*V = |\lambda|^2 \oplus V_0^* V_0.$$

Then $|\lambda| = 1$. This forces $VV^* = |\lambda|^2 \oplus V_0 V_0^* = 1 \oplus V_0 V_0^*$, which can never be equal to Q since the scalar component of Q is 0. Thus P and Q cannot be equivalent.

This shows that $I_{\mathcal{M}}$ fails a weaker form halving, where one does not require $P \sim I_{\mathcal{M}}$. This kind of halving is strictly weaker than the halving of properly infinite projections. For instance $I_2 \in M_2(\mathbb{C})$ can be written as $I_2 = E_{11} + E_{22}$, with $E_{11} \sim E_{22}$.

(14.2.24) Let $\mathcal{M}$ be a finite von Neumann algebra, and $P, Q \in \mathcal{M}$ projections with $P \sim Q$. Show that $I_{\mathcal{M}} - P \sim I_{\mathcal{M}} - Q$.

Answer. If $I_{\mathcal{M}} - P$ is not equivalent to $I_{\mathcal{M}} - Q$, by Comparison there exists a projection $Z \in \mathcal{Z}(\mathcal{M})$ with $Z(I_{\mathcal{M}} - P) \prec Z(I_{\mathcal{M}} - Q)$ (if this fails, then the case with the roles of P, Q exchanged works). So $Z(I_{\mathcal{M}} - P) \sim ZR_0 \leq Z(I_{\mathcal{M}} - Q)$, with ZR_0 a proper subprojection of $Z(I_{\mathcal{M}} - Q)$. This gives (via Proposition 14.2.7)

$$Z = ZP + Z(I_{\mathcal{M}} - P) \sim ZQ + ZR_0 \leq ZQ + Z(I_{\mathcal{M}} - Q) = Z,$$

with $ZQ + ZR_0$ a proper subprojection of Z. As $\mathcal{M}$ is finite this is a contradiction, and so $I_{\mathcal{M}} - P \sim I_{\mathcal{M}} - Q$.

(14.2.25) Let $\mathcal{M}$ be a finite von Neumann algebra, and $P, Q \in \mathcal{M}$ projections with $P \preceq Q$. Show that $I_{\mathcal{M}} - Q \preceq I_{\mathcal{M}} - P$.

Answer. By hypothesis there exists $Q_0 \leq Q$ with $Q_0 \sim P$. By Exercise 14.2.24, $I_{\mathcal{M}} - Q_0 \sim I_{\mathcal{M}} - P$. Then

$$I_{\mathcal{M}} - Q \leq I_{\mathcal{M}} - Q_0 \sim I_{\mathcal{M}} - P.$$

(14.2.26) Let $\mathcal{M}$ be a von Neumann algebra, and $P, Q \in \mathcal{M}$ finite projections with $P \sim Q$. Show that $I_\mathcal{M} - P \sim I_\mathcal{M} - Q$.

Answer. We know that $P \vee Q$ is finite by Proposition 14.2.15. Applying Exercise 14.2.24 to P, Q in the finite algebra $(P \vee Q)\mathcal{M}(P \vee Q)$, we get $P \vee Q - P \sim P \vee Q - Q$ (in $(P \vee Q)\mathcal{M}(P \vee Q)$, hence also in $\mathcal{M}$). Then, using Proposition 14.2.7,

$$\begin{aligned} I_\mathcal{M} - P &= \big(I_\mathcal{M} - P \vee Q\big) + \big(P \vee Q - P\big) \\ &\sim \big(I_\mathcal{M} - P \vee Q\big) + \big(P \vee Q - Q\big) = I_\mathcal{M} - Q. \end{aligned}$$

(14.2.27) Let $P \in \mathcal{M}$ be a finite projection, and $Q \in \mathcal{M}$ a projection. Show that there exists a number $s \in \mathbb{N} \cup \{0\}$ such that s is the maximum such that there exist pairwise orthogonal projections $\{P_1, \ldots, P_s\} \subset \mathcal{M}$ with $P_k \sim Q$ for all k, and $\sum_{k=1}^s P_k \preceq P$.

Answer. Suppose that no such s exists. This means that for any $m \in \mathbb{N}$ there exist pairwise orthogonal projections $\{P_{m,1}, \ldots, P_{m,m}\} \subset \mathcal{M}$ with $P_{m,k} \sim Q$ for all k and $\sum_k P_{m,k} \preceq P$. Assume without loss of generality that $P_1 = P_{1,1} \leq P$. As $P_{2,1} + P_{2,2} \preceq P$ there exists a partial isometry $V \in \mathcal{M}$ with $V^*V = P_{2,1} + P_{2,2}$ and $VV^* \leq P$. Then $P'_{2,1} = VP_{2,1}V^*$ and $P'_{2,2} = VP_{2,2}V^*$ are projections in $\mathcal{M}$ with $P'_{2,1} \sim P_{2,1}$ and $P'_{2,2} \sim P_{2,2}$ and $P'_{2,1} + P'_{2,2} \leq P$. Let $P_2 = P'_{2,2}$.

As $P_1 \sim Q \sim P'_{2,1}$, by Exercise 14.2.25 (applied in the finite von Neumann algebra $P\mathcal{M}P$) we have $P - P_1 \sim P - P'_{2,1} \geq P'_{2,2}$. Then $Q \sim P'_{2,2} \preceq P - P_1$. So there exists $P_2 \in \mathcal{M}$ with $P_2 \leq P$, $P_2 \sim Q$, and $P_2P_1 = 0$. Now we can repeat the argument with $P_1 + P_2 \leq P$ and $P_{3,1} + P_{3,2} + P_{3,3} \preceq P$ to obtain $P_3 \sim Q$ and $P_1 + P_2 + P_3 \leq P$. Continuing inductively, we get a sequence $\{P_n\}$, pairwise orthogonal, $P_n \sim Q$ for all n, and $\sum_n P_n \leq P$. But as the P_n are all pairwise equivalent, the projection $\sum_n P_n$ is infinite (it is equivalent to $\sum_{n\geq 2} P_n$, for instance), a contradiction since P is finite (via Exercise 14.2.17). The contradiction shows that there exists a maximum s as desired.

(14.2.28) Let $\mathcal{M}$ be a finite von Neumann algebra and $T \in \mathcal{M}$ with polar decomposition $T = V|T|$. Show that V can be extended to a unitary U with $T = U|T|$.

Answer. Because $V^*V \sim VV^*$ and $\mathcal{M}$ is finite, by Exercise 14.2.24 there exists a partial isometry $W \in \mathcal{M}$ with $W^*W = I_\mathcal{M} - V^*V$ and $WW^* = I_\mathcal{M} - VV^*$. Let $U = V + W$. We have (using Exercise 10.4.12)

$$V^*W = V^*VV^*(I_\mathcal{M} - VV^*)W = 0.$$

Then U is a unitary, for $U^*U = V^*V + W^*W = I_\mathcal{M}$ and $UU^* = VV^* + WW^* = I_\mathcal{M}$. We also have

$$U|T| = (V + W)|T| = (V + W)V^*V|T| = VV^*V|T| = V|T| = T.$$

(14.2.29) Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra. Show that $\mathcal{M}$ is finite, in two ways:

(i) by using the explicit form of a finite-dimensional von Neumann algebra;

(ii) by a direct argument.

Answer.

(i) We know from Theorem 11.8.10 that $\mathcal{M} = \bigoplus_{j=1}^k M_{n_j}(\mathbb{C})$. So $I_\mathcal{M} = \bigoplus_{j=1}^k I_{n_j}$. Let $V \in \mathcal{M}$ with $V^*V = I_\mathcal{M}$. We can write $V = \bigoplus_{j=1}^k V_j$, and then $V_j^*V_j = I_{n_j}$ for all j. Seen as an element of $M_{n_j}(\mathbb{C})$, this equality gives us that V_j is injective, and so it is surjective by Exercise 1.7.8. Therefore V_j is invertible and $V_j^* = V_j^{-1}$. Thus $V_jV_j^* = I_{n_j}$. It follows that $VV^* = I_\mathcal{M}$, and so $I_\mathcal{M}$ is finite.

(ii) Let $P \leq I_\mathcal{M}$ be a projection with $P \sim I_\mathcal{M}$ and $P \neq I_\mathcal{M}$. So there exists $V \in \mathcal{M}$ with $V^*V = I_\mathcal{M}$ and $VV^* = P$. Let $P_1 = VPV^*$. Then $P_1 \leq VV^* = P$, and $P_1 \neq P$; for if $P_1 = P$ this is $VVV^*V^* = VV^*$, and applying V^* on the left and V on the right, we would have $P = I_\mathcal{M}$. Iterating this construction we get a properly decreasing sequence of projections $P_1 \geq P_2 \geq P_3 \cdots$. If $k > j$,

$$\begin{aligned}(P_k - P_{k+1})(P_j - P_{j+1}) &= P_kP_j + P_{k+1}P_{j+1} - P_{k+1}P_j - P_kP_{j+1} \\ &= P_k + P_{k+1} - P_{k+1} - P_k = 0.\end{aligned}$$

That is, the projections $\{P_k - P_{k+1}\}$ are nonzero and pairwise orthogonal. In particular they are linearly independent, so $\dim \mathcal{M} = \infty$. The contradiction shows that $I_\mathcal{M}$ is finite.

(14.2.30) Let $\mathcal{M}$ be a von Neumann algebra. Show that the following statements are equivalent:

(i) $\mathcal{M}$ is finite;

(ii) for any projections $P, Q \in \mathcal{M}$ with $P \sim Q$, there exists $U \in \mathcal{M}$ unitary with $Q = UPU^*$.

Answer. (i) $\implies$ (ii) Suppose first that $\mathcal{M}$ is finite and $P \sim Q$. So there exists a partial isometry $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* = Q$. By Exercise 14.2.24 there exists a partial isometry $W \in \mathcal{M}$ with $W^*W = I_{\mathcal{M}} - P$ and $WW^* = I_{\mathcal{M}} - Q$. Let $U = V + W$. Since $V^*W = V^*Q(I_{\mathcal{M}} - Q)W = 0$ and $VW^* = VP(I_{\mathcal{M}} - P)W^* = 0$,

$$U^*U = V^*V + W^*W = P + I_{\mathcal{M}} - P = I_{\mathcal{M}}$$

and

$$UU^* = VV^* + WW^* = Q + I_{\mathcal{M}} - Q = I_{\mathcal{M}}.$$

So U is a unitary, and

$$UPU^* = VPV^* = VV^*VV^* = Q.$$

(ii) $\implies$ (i) Suppose that $I_{\mathcal{M}} \sim P$. This means that there exists a unitary $U \in \mathcal{M}$ with $P = UI_{\mathcal{M}}U^* = I_{\mathcal{M}}$. So $I_{\mathcal{M}}$ is finite, and hence $\mathcal{M}$ is finite.

(14.2.31) Let $\mathcal{M}$ be a factor, and $P, Q \in \mathcal{M}$ projections with Q finite and P infinite. Show that $Q \prec P$.

Answer. Because $\mathcal{M}$ is a factor, by Comparison we either have $Q \preceq P$ or $P \preceq Q$. The latter would imply that P is finite (by Exercise 14.2.17); so $Q \preceq P$. We cannot have $Q \sim P$, because that would make P finite or Q infinite; so $Q \prec P$.

(14.2.32) Let $P, Q \in \mathcal{M}$ be projections with P properly infinite and $Q \preceq P$. Show that $P \vee Q \sim P$.

Answer. Because P is properly infinite, by Halving (Proposition 14.2.14) there exists a projection $R \in \mathcal{M}$ with $R \leq P$ and $R \sim P \sim P - R$. We have (using Kaplansky's Formula (14.1))

$$P \vee Q - P \sim Q - P \wedge Q \leq Q \preceq P \sim P - R.$$

Then, using Proposition 14.2.7,

$$P \vee Q = \big(P \vee Q - P\big) + P \preceq (P - R) + R = P.$$

As we also have $P \leq P \vee Q$, we get $P \vee Q \sim P$.

(14.2.33) Let $P, Q \in \mathcal{M}$ be properly infinite, with $P + Q = I_{\mathcal{M}}$ and $P \sim Q$. Show that $P \sim I_{\mathcal{M}}$.

Answer. Since P is properly infinite, by Halving there exists a projection $R \in \mathcal{M}$ with $P \sim R \sim P - R$. Then $Q \sim P \sim R$, and so using Proposition 14.2.7

$$I_{\mathcal{M}} = P + Q \sim P - R + R = P.$$

(14.2.34) Let $P_1, P_2, Q_1, Q_2 \in \mathcal{M}$ be projections with $P_1 + P_2 = Q_1 + Q_2$, $P_1P_2 = 0 = Q_1Q_2$, and $P_1 \sim P_2$, $Q_1 \sim Q_2$. Show that $P_1 \sim Q_1$.

Answer. By working on $(P_1 + P_2)\mathcal{M}(P_1 + P_2)$ we may assume without loss of generality that $P_1 + P_2 = I_{\mathcal{M}}$.

From Proposition 14.2.13 we have a projection $Z \in \mathcal{Z}(\mathcal{M})$ with ZP_1 finite and $(I_{\mathcal{M}} - Z)P_1$ properly infinite. So it is enough that we show that cases P_1 finite and P_1 properly infinite separately.

Suppose first that P_1 is finite. Then $I_{\mathcal{M}} = P_1 + P_2$ is finite by Proposition 14.2.15. By Comparison there exists a projection $Z \in \mathcal{Z}(\mathcal{M})$ with $ZP_1 \preceq ZQ_1$ and $(I_{\mathcal{M}} - Z)Q_1 \preceq (I_{\mathcal{M}} - Z)P_1$. We immediately have

$$ZP_2 \sim ZP_1 \preceq ZQ_1 \sim ZQ_2.$$

Fix a projection $R_1 \leq Q_1$ with $ZP_1 \sim ZR_1$, and $R_2 \leq Q_2$ with $ZP_2 \sim ZR_2$. Then

$$Z = ZP_1 + ZP_2 \sim ZR_1 + ZR_2 \leq ZQ_1 + ZQ_2 = Z.$$

As Z is finite (because $I_{\mathcal{M}} = P_1 + P_2$ is finite), this means that $ZR_1 + ZR_2 = Z$, Then

$$0 \leq Z(Q_1 - R_1) + Z(Q_2 - R_2) = Z - (ZR_1 + ZR_2) = 0.$$

As $Z(Q_1 - R_1) \geq 0$ and $Z(Q_2 - R_2) \geq 0$, this forces $ZR_1 = ZQ_1$ and $ZR_2 = ZQ_2$. But then $ZP_1 \sim ZR_1 = ZQ_1$ and similarly $ZP_2 \sim ZQ_2$. Recalling that the projection Z also satisfies $(I_{\mathcal{M}} - Z)Q_1 \preceq (I_{\mathcal{M}} - Z)P_1$, we can repeat the argument to get $(I_{\mathcal{M}} - Z)P_1 \sim (I_{\mathcal{M}} - Z)Q_1$. Then $P_1 \sim Q_1$ by Proposition 14.2.7.

Now suppose that P_1 is properly infinite. Then P_2 is properly infinite by Exercise 14.2.16 and then $I_{\mathcal{M}} \sim P_1$ by Exercise 14.2.33, which then implies that $I_{\mathcal{M}}$ is properly infinite. It also follows that Q_1 is properly infinite; indeed, if Q_1 is not properly infinite there exists a projection $Z \in \mathcal{Z}(\mathcal{M})$ with ZQ_1 finite. This makes ZQ_2 finite by Exercise 14.2.16 and therefore $Z = ZQ_1 + ZQ_2$ is finite by Proposition 14.2.15. But this would make ZP_1 finite, a contradiction unless $Z = 0$. By Exercise 14.2.33,

$$P_1 \sim I_{\mathcal{M}} \sim Q_1.$$

(14.2.35) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $Q \in \mathcal{M}$ a projection, and $\mathcal{K}$ a Hilbert space. Show that $c(Q \otimes E_{11}) = c(Q) \otimes I_{\mathcal{K}}$ in $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$.

Answer. The projection $c(Q) \otimes I_{\mathcal{K}}$ is central, and

$$c(Q) \otimes I_{\mathcal{K}} - Q \otimes E_{11} = (c(Q) - Q) \otimes I_{\mathcal{K}} + Q \otimes (I_{\mathcal{H}} - E_{11}) \geq 0$$

since both terms are positive. That is, $Q \otimes E_{11} \leq c(Q) \otimes I_{\mathcal{K}}$. Now let $\tilde{P} \in \mathcal{Z}(\mathcal{M})$ with $Q \otimes E_{11} \leq \tilde{P}$. We know from Exercise 13.4.27 that $\mathcal{Z}(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})) = \mathcal{Z}(\mathcal{M}) \otimes I_{\mathcal{H}}$. So $\tilde{P} = P \otimes I_{\mathcal{K}}$ for some $P \in \mathcal{Z}(\mathcal{M})$. Necessarily P is a projection. Indeed, we have

$$P^* \otimes I_{\mathcal{K}} = (P \otimes I_{\mathcal{K}})^* = \tilde{P}^* = \tilde{P} = P \otimes I_{\mathcal{K}}$$

And, for $\eta \in \mathcal{K}$ with $\|\eta\| = 1$ and $\xi \in \mathcal{H}$,

$$\begin{aligned} \langle P^2\xi, \xi\rangle &= \langle (P \otimes I_{\mathcal{K}})^2(\xi \otimes \eta), \xi \otimes \eta\rangle \\ &= \langle (P \otimes I_{\mathcal{K}})(\xi \otimes \eta), \xi \otimes \eta\rangle = \langle P\xi, \xi\rangle. \end{aligned}$$

Using polarization we conclude that $P^2 = P$. We can similarly obtain that $P^* = P$, so P is a central projection. Therefore we have the inequality

$$0 \leq P \otimes I_{\mathcal{K}} - Q \otimes E_{11} = (P - Q) \otimes (I_{\mathcal{K}} - E_{11}).$$

This can only occur if $P \geq Q$ (we can show this using ξ and η as above). Then $P \geq c(Q)$, and so $c(Q) \otimes I_{\mathcal{K}} \leq P \otimes I_{\mathcal{K}} = \tilde{P}$. It follows that $c(Q) \otimes I_{\mathcal{K}}$ is the least central projection above $Q \otimes E_{11}$, and thus $c(Q \otimes E_{11}) = c(Q) \otimes I_{\mathcal{K}}$.

(14.2.36) Let $P \in \mathcal{M}$ be a properly infinite projection, and $n \in \mathbb{N} \cup \{\infty\}$. Show that there exist pairwise orthogonal projections $\{P_k\}_{k=1}^n \subset \mathcal{M}$ with $P_k \sim P$ for all k, and $P = \sum_k P_k$.

Answer. Suppose first that $n \in \mathbb{N}$. We can write $n = \sum_{k\in F} 2^k$ for some finite subset $F \subset \mathbb{N}$. Write $F = \{k_1, \ldots, k_m\}$ with $k_r < k_{r+1}$ for all r. Use Halving to write $P = Q_0 + Q_1$ with $Q_0 \sim Q_1 \sim P$. Now we apply Halving k_1 times to Q_0 to get $Q_0 = \sum_{s=1}^{2^{k_1}} R_{0,j}$, with $R_{0,s} \sim Q_0$ for all s. Next we do the same, but starting with Q_1 and $F_2 = \{k_2, \ldots, k_m\}$. Repeating this inductively we get $2^{k_1} + \cdots + 2^{k_m} = n$ pairwise equivalent projections that add to P.

When $n = \infty$, we subdivide as in the previous paragraph, but always halving the second projection. This way we end up with countably many $\{Q_k\}$, pairwise orthogonal and $P_k \sim P$ for all k. Let $P_0 = P - \sum_k P_k$. Since $P_0 \leq P \sim P_1$, by Exercise 14.2.32 we have $P_1 + P_0 \sim P_1$. So we replace P_1 with $P_1 + P_0$ and now $P = \sum_k P_k$, with all projections equivalent to P.

14.3. Classification of von Neumann Algebras

(14.3.1) Show that the matrix units defined in (14.5) do satisfy the matrix unit relations $E_{kj}E_{ab} = \delta_{j,a}E_{kb}$ and $E_{kj}^* = E_{jk}$.

Answer. We have

$$E_{kj}E_{ab} = E_{1k}^*E_{1j}Q_jQ_aE_{1a}^*E_{1b} = \delta_{j,a}\,E_{1k}^*E_{1j}Q_jE_{1j}^*E_{1b}$$
$$= \delta_{j,a}\,E_{1k}^*E_{11}E_{1b} = \delta_{j,a}\,E_{1k}^*E_{1b} = \delta_{j,a}\,E_{kb}.$$

And $E_{kj}^* = E_{1j}^*E_{1k} = E_{jk}$.

(14.3.2) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Show that $\mathcal{M}$ is type I if and only if there exists a projection $P \in \mathcal{M}$, abelian, with $c(P) = I_{\mathcal{M}}$.

Answer. Suppose first that $\mathcal{M}$ is type I. Then abelian projections exist in $\mathcal{M}$. Via Zorn's Lemma construct a family $\{P_j\} \subset \mathcal{M}$ of abelian projections with pairwise orthogonal central supports. Let $Q = \sum_j c(P_j) \in \mathcal{M}$. If $I_{\mathcal{M}} - Q \neq 0$, by hypothesis there exists $P_0 \in \mathcal{M}$, abelian, with $P_0 \leq I_{\mathcal{M}} - Q$; as this latter projection is central, $c(P_0) \leq I_{\mathcal{M}} - Q$, contradicting the maximality. Then $Q = I_{\mathcal{M}}$.

Let $P = \sum_j P_j \in \mathcal{M}$ (the projections are pairwise orthogonal since their central supports are). Since

$$P_j\mathcal{M}P_k = P_jc(P_j)\mathcal{M}c(P_k)P_k = P_jc(P_j)c(P_k)\mathcal{M}P_k = 0$$

if $j \neq k$, we can write $P\mathcal{M}P = \sum_j P_j\mathcal{M}P_j$, abelian. So P is an abelian projection. Suppose that $Z \in \mathcal{M}$ is a central projection with $P \leq Z$. As $P_j \leq P \leq Z$, we get $c(P_j) \leq Z$. And then $Z \geq \sum_j c(P_j) = I_\mathcal{M}$. This shows that $c(P) = I_\mathcal{M}$ as desired.

Conversely, suppose that there exists $P \in \mathcal{M}$, abelian, with $c(P) = I_\mathcal{M}$. Given any nonzero $Q \in \mathcal{M}$, since $c(P) = I_\mathcal{M}$ by Proposition 14.2.6 there exist nonzero projections $P_0 \leq P$ and $Q_0 \leq Q$ with $P_0 \sim Q_0$. From $P_0 \leq P$ we have that P_0 is abelian; then Q_0 is abelian by Exercise 14.2.16.

(14.3.3) In the proof of Proposition 14.3.6, show that P_0 and Q_0 are in generic position when acting on $(P_0 \vee Q_0)\mathcal{H}$.

Answer. Suppose that $\xi \in (P_0 \wedge Q_0)\mathcal{H}$. This implies that $\xi = P\xi = Q\xi$ (since $P_0 \leq P$ and $Q_0 \leq Q$), so $\xi = (P \wedge Q)\xi$. But we also have $\xi = P_0\xi$, so $\xi = P_0(P \wedge Q)\xi = 0$. That is, $P_0 \wedge Q_0 = 0$.

If now $\xi \in (P_0 \wedge Q_0^\perp)\mathcal{H}$, then $\xi = P_0\xi = P\xi$. As $Q_0^\perp = Q^\perp + P \wedge Q + P^\perp \wedge Q$, the equality $Q_0^\perp\xi = \xi$ leads us to consider the corresponding three components. For $Q^\perp\xi$, as we also have $P\xi = \xi$ we obtain $P_0Q^\perp\xi = P_0(P\wedge Q^\perp)\xi = 0$. For $(P\wedge Q)\xi$ we have $P_0(P\wedge Q)\xi = 0$. And for $(P^\perp\wedge Q)\xi$ we have $P_0(P^\perp \wedge Q)\xi = 0$ since $P_0 \leq P$. This shows that $\xi = P_0\xi = P_0Q^\perp\xi = 0$, and therefore $P_0 \wedge Q_0^\perp = 0$.

The equality $P_0^\perp \wedge Q_0 = 0$ is proven by exchanging the roles of P_0 and Q_0 in the previous paragraph. So it remains to consider $P_0^\perp \wedge Q_0^\perp$. But we are working in a context where $P_0 \vee Q_0 = 1$, so $P_0^\perp \wedge Q_0^\perp = (P_0 \vee Q_0)^\perp = I_\mathcal{H}^\perp = 0$.

(14.3.4) Let $\mathcal{M}$ be a type I von Neumann algebra and $\mathcal{H}$ a Hilbert space. Show that $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$ is type I, without using Theorem 14.3.2.

Answer. Since $\mathcal{M}$ is type I, there exists an abelian projection $Q \in \mathcal{M}$ with $c(Q) = I_\mathcal{M}$ (Exercise 14.3.2). Then $Q \otimes E_{11}$ is abelian in $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$, with central support $c(Q) \otimes I_\mathcal{H} = I_{\mathcal{K}\overline{\otimes}\mathcal{H}}$ (Exercise 14.2.35). If $\tilde{P} \in \mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$ is any nonzero projection, by Proposition 14.2.6 there exist nonzero projections $Q_0 \leq Q \otimes E_{11}$ and $P_0 \leq P$ with $P_0 \sim Q_0$. As Q_0 is abelian (being a subprojection of an abelian projection), so is P_0 by Exercise 14.2.16. Hence $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$ is type I.

(14.3.5) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Show that $\mathcal{M}$ is type II if and only if there exists a projection $P \in \mathcal{M}$, finite, with $c(P) = I_{\mathcal{M}}$.

Answer. Suppose first that $\mathcal{M}$ is type II. Then finite projections exist in $\mathcal{M}$. Via Zorn's Lemma construct a family $\{P_j\} \subset \mathcal{M}$ of finite projections with pairwise orthogonal central supports. Let $Q = \sum_j c(P_j) \in \mathcal{M}$. If $I_{\mathcal{M}} - Q \neq 0$, by hypothesis there exists $P_0 \in \mathcal{M}$, finite, with $P_0 \leq I_{\mathcal{M}} - Q$; as this latter projection is central, $c(P_0) \leq I_{\mathcal{M}} - Q$, contradicting the maximality. Then $Q = I_{\mathcal{M}}$.

Let $P = \sum_j P_j \in \mathcal{M}$ (the projections are pairwise orthogonal since their central supports are). By Lemma 14.2.12, P is a finite projection. Suppose that $Z \in \mathcal{M}$ is a central projection with $P \leq Z$. As $P_j \leq P \leq Z$, we get $c(P_j) \leq Z$. And then $Z \geq \sum_j c(P_j) = I_{\mathcal{M}}$. This shows that $c(P) = I_{\mathcal{M}}$ as desired.

Conversely, suppose that there exists $P \in \mathcal{M}$, finite, with $c(P) = I_{\mathcal{M}}$. Given any nonzero $Q \in \mathcal{M}$, since $c(P) = I_{\mathcal{M}}$ by Proposition 14.2.6 there exist nonzero projections $P_0 \leq P$ and $Q_0 \leq Q$ with $P_0 \sim Q_0$. From $P_0 \leq P$ we have that P_0 is finite (Exercise 14.2.17); then Q_0 is abelian by Exercise 14.2.16.

(14.3.6) Let $\mathcal{M}$ be a II_1-factor. We will outline here a way to "manually" construct the normalized dimension function.

(i) Use Proposition 14.2.22 to construct a family $\{P_{k,n}\} \subset \mathcal{M}$ of projections, with $n \in \mathbb{N}$, $k \in \{1, \ldots, 2^n\}$, and

$$P_{k,n} \sim P_{j,n}, \qquad P_{2k-1,n+1} + P_{2k,n} = P_{k,n}$$

for all n and all $k, j \leq 2^n$, and $\sum_k P_{k,n} = I_{\mathcal{M}}$ (note: $\{P_{k,n}\}_{k=1}^{2^n}$ are pairwise orthogonal by Proposition 10.5.5).

(ii) (Division Algorithm) Show that given $n \in \mathbb{N}$ and $P \in \mathcal{P}(\mathcal{M})$, there exist $s(n) \in \{0, \ldots, 2^n\}$ and projections

$$Q_1, \ldots, Q_{s(n)}, R \in \mathcal{M},$$

with $R \prec P_{1,n}$, $Q_k \sim P_{1,n}$ for all k, and such that $P = R + \sum_{k \leq s(n)} Q_k$.

(iii) Show that if $Q'_1, \ldots, Q'_{s'(n)}, R'$ is another decomposition for P as above, then $s'(n) = s(n)$ and $R' \sim R$.

(iv) Keep considering the same fixed P. Show that $s(n+1) \geq 2s(n)$ for all n.

(v) For the same projection P, let $\alpha_n = 2^{-n}\, s(n)$. Show that the sequence $\{\alpha_n\}$ converges to some number $\tau(P) \in [0,1]$.

(vi) Show that for projections $P, Q \in \mathcal{M}$, we have $P \preceq Q$ if and only if $\tau(P) \leq \tau(Q)$. Conclude that $P \sim Q$ if and only if $\tau(P) = \tau(Q)$ and that $P \prec Q$ if and only if $\tau(P) < \tau(Q)$.

(vii) Use Proposition 14.2.16 to show that if $\{Q_n\} \subset \mathcal{M}$ is a monotone sequence of projections with $Q_n \xrightarrow{\text{sot}} Q$, then $\tau(Q_n) \to \tau(Q)$.

(viii) Show that τ is σ-additive.

(ix) Show that $\tau(\mathcal{P}(\mathcal{M})) = [0,1]$.

Answer.

(i) We let $P_{1,0} = I_{\mathcal{M}}$. By Proposition 14.2.22 there exist equivalent projections $P_{1,1}, P_{1,2} \in \mathcal{M}$ with $P_{1,1} \sim P_{1,2}$ and $P_{1,0} = P_{1,1} + P_{1,2}$. Now we proceed inductively by using Proposition 14.2.22 repeatedly.

(ii) Since we are in a factor, all projections are comparable. We have $P \preceq I_{\mathcal{M}}$. Let $s(n)$ be the largest index such that $\sum_{k=1}^{s(n)} P_{k,n} \preceq P$ (it is possible that $s(n) = 0$ if $P \prec P_{1,n}$); this number exists by Exercise 14.2.27. Let $V_n \in \mathcal{M}$ be a partial isometry with $V_n^* V_n = \sum_{k=1}^{s(n)} P_{k,n}$ and $V_n V_n^* \leq P$. Define $Q_k = V_n P_{k,n} V_n^*$. Then $Q_k \sim P_{k,n}$, and $\sum_{k=1}^{s(n)} Q_k \leq P$. Let $R = P - \sum_{k=1}^{s(n)} Q_k$. If $P_{1,n} \preceq R$ then by Proposition 14.2.7 we have $\sum_{k=1}^{s(n)+1} P_{k,n} \preceq R + \sum_{k=1}^{s(n)} Q_k = P$, contradicting the definition of $s(n)$. Hence $R \prec P_{1,n}$.

(iii) By hypothesis $Q'_k \sim P_{1,n} \sim Q_k$ for all k. If $s'(n) > s(n)$, we would have

$$\sum_{k=1}^{s(n)+1} P_{k,n} \leq \sum_{k=1}^{s'(n)} Q_{k,n} \preceq P,$$

a contradiction. As the roles are equivalent, $s'(n) = s(n)$. Now $R' \sim R$ by Exercise 14.2.24 applied in the II_1-factor $P\mathcal{M}P$, since $\sum_k Q_k \sim \sum_k Q'_k$ by Proposition 14.2.7.

(iv) We have

$$\sum_{k=1}^{2s(n)} P_{k,n+1} = \sum_{k=1}^{s(n)} P_{2k-1,n} + P_{2k,n} = \sum_{k=1}^{s(n)} P_{k,n} \preceq P.$$

Then $s(n+1) \geq 2s(n)$ by definition of $s(n+1)$.

(v) We have, using the estimate $s(n) \leq s(n+1)/2$,

$$\alpha_n = \frac{s(n)}{2^n} \leq \frac{s(n+1)}{2^{n+1}} = \alpha_{n+1} \leq 1.$$

So the sequence $\{\alpha_n\}$ is monotone non-decreasing, and bounded above by 1. Hence $\tau(P) = \lim_n \alpha_n$ exists.

(vi) Suppose that $P \preceq Q$. For each $n \in \mathbb{N}$, by (ii),

$$\sum_{k=1}^{s_P(n)} P_{k,n} \preceq P \preceq Q.$$

Hence $s_P(n) \leq s_Q(n)$. It follows that

$$\alpha_n^P = \frac{s_P(n)}{2^n} \leq \frac{s_Q(n)}{2^n} = \alpha_n^Q.$$

Taking limit, $\tau(P) \leq \tau(Q)$. Exchanging roles we get that $P \sim Q$ implies $\tau(P) = \tau(Q)$.

Now assume that $\tau(P) = \tau(Q)$. We have, writing $\alpha_0^P = 0$,

$$\tau(P) = \sum_{n=1}^{\infty}(\alpha_n^P - \alpha_{n-1}^P) = \sum_{n=1}^{\infty} \frac{s_P(n) - 2s_P(n-1)}{2^n} \tag{AB.14.1}$$

With this idea in mind we have, using (ii) repeatedly, first on P for $n = 1$, then on R_1 for $n = 2$, then on R_2 for $n = 3$, and so on,

$$P = \sum_{k=1}^{s_P(1)} Q_{k,1} + R_1 = \sum_{k=1}^{s_P(1)} Q_{k,1} + \sum_{k=2s_P(1)+1}^{s_P(2)} Q_{k,2} + R_2$$

$$= \cdots = \sum_{n=1}^{m} \sum_{k=2s_P(n-1)+1}^{s_P(n)} Q_{k,n} + R_m,$$

where $Q_{k,n} \sim P_{1,n}$ for all k, n, $R_{m+1} \leq R_m$, and $R_m \prec P_{1,n}$. All the sums on k have either one term (when $2s_P(n-1) + 1 - s_P(n)$) or zero terms (when $2s_P(n-1) = s_P(n)$). Since $\{R_m\}_m$ is a decreasing sequence, $R = \lim_{\text{sot}} R_m$ exists in $\mathcal{M}$ (by Proposition 12.1.10) and it is a projection by Proposition 12.1.13. We have $R \leq R_m \prec P_{1,n}$ for all n. This means that we can put arbitrarily many copies of R below $I_{\mathcal{M}}$, contradicting that $I_{\mathcal{M}}$ is finite (concretely, via Exercise 14.2.27). Hence $R = 0$. Then

$$P = \sum_{n=1}^{\infty} \sum_{k=2s_P(n-1)+1}^{s_P(n)} P_{k,n}. \tag{AB.14.2}$$

We note by definition that $s_P(n) - 2s_P(n-1) \in \{0, 1\}$ for all n. So (AB.14.1) expresses $\tau(P)$ in binary. The representation is unique, for the only ambiguity that can appear is to have an infinite string of 1. But here this is impossible, for if we had $s_P(n_0) = 2s_P(n_0 - 1)$ and

$2s_P(n-1)+1 = s_P(n)$ for all $n \geq n_0+1$, then there is exactly room at the n_0 level to put one more projection of trace $1/2^{n_0}$ and this would force $2s_P(n_0) = s_P(n_0+1)$, a contradiction. It follows, since $s_P(0) = 0 = s_Q(0)$, that $s_P(n) = s_Q(n)$ for all n. Then (AB.14.2) also applies to Q, and we get $P \sim Q$. With this we have shown that $P \sim Q$ if and only if $\tau(P) = \tau(Q)$.

We already know that $P \preceq Q$ implies $\tau(P) \leq \tau(Q)$. Together with the above, we obtain that $P \prec Q$ implies $\tau(P) < \tau(Q)$. And if $\tau(P) < \tau(Q)$ we cannot have $Q \preceq P$ (because we know it implies $\tau(Q) \leq \tau(P)$), so $P \prec Q$.

(vii) We may assume without loss of generality that $Q_n \searrow 0$ (by replacing each Q_n with $Q - Q_n$ or $Q_n - Q$ depending on whether the sequence is increasing or decreasing; the new sequence is still monotone). Let us recall again that, since we are in a factor, all projections are comparable. Fix $k \in \mathbb{N}$. Suppose that $P_{1,k} \preceq Q_n$ for all n. Proposition 14.2.16 gives us, since $I_{\mathcal{M}} - Q_n \preceq I_{\mathcal{M}} - P_{1,k}$ by Exercise 14.2.25,

$$\bigvee_n (I_{\mathcal{M}} - Q_n) \preceq I_{\mathcal{M}} - P_{1,k}.$$

And now using again Exercise 14.2.25, plus Proposition 10.5.9 and Exercise 12.1.24,

$$P_{1,k} \preceq I_{\mathcal{M}} - \bigvee_n (I_{\mathcal{M}} - Q_n) = \bigwedge_n Q_n = 0.$$

As $P_{1,k} \neq 0$, this is a contradiction (Exercise 14.2.1). So there exists n_0 such that $Q_{n_0} \preceq P_{1,k}$. As the sequence is monotone, $Q_n \preceq P_{1,k}$ for all $n \geq n_0$. This says that $\tau(Q_n) \leq 2^{-k}$ for all $n \geq n_0$. As this can be done for all k, we have shown that $\lim_n \tau(Q_n) = 0$.

(viii) Let $\{Q_s\}_{s\in\mathbb{N}} \subset \mathcal{M}$ be a sequence of pairwise orthogonal projections. Write $Q = \bigvee_s Q_s \in \mathcal{M}$. We have by definition that $\tau(I_{\mathcal{M}}) = 1$ and $\tau(P_{k,n}) = 2^{-n}$ for all k, n. Suppose first that $Q_1 + Q_2 = I_{\mathcal{M}}$. Let us write $s_1(n)$ and $s_2(n)$ for the integer counters used to define $\tau(Q_1)$ and $\tau(Q_2)$, and $\{\alpha_n^1\}$ and $\{\alpha_n^2\}$ the corresponding sequences approximating $\tau(Q_1)$ and $\tau(Q_2)$ respectively. For each n we have

$$\sum_{k=1}^{s_1(n)} P_{k,n} \preceq Q_1 \prec \sum_{k=1}^{s_1(n)+1} P_{k,n}. \tag{AB.14.3}$$

So there exists $Q_1' \leq Q_1$ with $Q_1' \sim \sum_{k=1}^{s_1(n)} P_{k,n}$. By Exercise 14.2.24, $I_{\mathcal{M}} - Q_1' \sim \sum_{k=s_1(n)+1}^{2^n} P_{k,n}$. Hence

$$Q_2 = I_{\mathcal{M}} - Q_1 \leq I_{\mathcal{M}} - Q_1' \sim \sum_{k=s_1(n)+1}^{2^n} P_{k,n}.$$

This gives us $\tau(Q_2) \le \frac{2^n - s_1(n)}{2^n} = 1 - \alpha_n^1$. Taking limit, $\tau(Q_2) \le 1 - \tau(Q_1)$. We can also do, with the same idea but using the other side of (AB.14.3),

$$\sum_{k=s_1(n)+2}^{2^n} P_{k,n} \preceq I_{\mathcal{M}} - Q_1 = Q_2,$$

and this gives $\tau(Q_2) \ge 1 - \alpha_n^1 - \frac{1}{2^n}$. Again taking limit, $\tau(Q_2) \ge 1 - \tau(Q_1)$. Therefore $\tau(Q_2) = 1 - \tau(Q_1)$, showing that $\tau(Q_1 + Q_2) = \tau(Q_1) + \tau(Q_2)$.

If we write $\tilde{Q}_n = \sum_{k=n+1}^{\infty} Q_k$, we have

$$\tau(Q) = \tau(Q_1 + \cdots + Q_n + \tilde{Q}_n) = \tau(\tilde{Q}_n) + \sum_{k=1}^{n} \tau(Q_k)$$

for all n. As $\tilde{Q}_n \searrow 0$ (by Exercise 12.1.22, since it is the tail of the convergent series) and using (vii) ,

$$\tau(Q) = \lim_n \tau(\tilde{Q}_n) + \sum_{k=1}^{n} \tau(Q_k) = \sum_{k=1}^{\infty} \tau(Q_k).$$

(ix) We already have that $k/2^n \in \tau(\mathcal{M})$ for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, 2^n\}$. The continuity (AB.14.3), applied to properly chosen subsequences of $\{P_{k,n}\}$, shows that $\tau(\mathcal{M}) = [0, 1]$.

(14.3.7) Let $\mathcal{N}$ be a von Neumann algebra and $\mathcal{K}$ an infinite-dimensional Hilbert space. Show that $\mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}) \simeq M_2(\mathcal{N}) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$.

Answer. Using Exercises 13.1.5, 13.4.9 and 13.4.10,

$$\begin{aligned} \mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}) &\simeq \mathcal{N} \mathbin{\overline{\otimes}} \big(M_2(\mathbb{C}) \otimes \mathcal{B}(\mathcal{K})\big) \\ &\simeq \big(\mathcal{N} \otimes M_2(\mathbb{C})\big) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}) \\ &\simeq M_2(\mathcal{N}) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}). \end{aligned}$$

14.4. Tensor Products of von Neumann Algebras

(14.4.1) Let $S, T \in \mathcal{B}(\mathcal{H})$ be selfadjoint. Show that $ST = TS$ if and only if $S\mathcal{H} \perp_{\mathbb{R}} iT\mathcal{H}$.

Answer. If $ST = TS$ then ST is selfadjoint. For $\xi \in \mathcal{H}$,

$$\operatorname{Re}\langle S\xi, iT\xi\rangle = -\operatorname{Re} i\langle TS\xi, \xi\rangle = 0.$$

By polarization we get that $\operatorname{Re}\langle S\xi, iT\eta\rangle = 0$ for all $\xi, \eta \in \mathcal{H}$; thus $S\mathcal{H} \perp_{\mathbb{R}} iT\mathcal{H}$.

Conversely, if $\operatorname{Re}\langle S\xi, iT\xi\rangle = 0$ then

$$\operatorname{Im}\langle S\xi, T\xi\rangle = -\operatorname{Re} i\langle S\xi, T\xi\rangle = \operatorname{Re}\langle S\xi, iT\xi\rangle = 0.$$

So $\langle S\xi, T\xi\rangle \in \mathbb{R}$. Then

$$\begin{aligned}\langle (ST - TS)\xi, \xi\rangle &= \langle T\xi, S\xi\rangle - \langle S\xi, T\xi\rangle = \overline{\langle T\xi, S\xi\rangle} - \langle S\xi, T\xi\rangle \\ &= \langle S\xi, T\xi\rangle - \langle S\xi, T\xi\rangle = 0.\end{aligned}$$

As this works for all $\xi \in \mathcal{H}$, using Polarization we obtain $ST = TS$.

14.5. The Trace

(14.5.1) Let $\mathcal{M} = M_n(L^\infty[0,1])$ and $\psi \in L^\infty[0,1]_* = L^1[0,1]$. Show that the functional

$$\Psi(T) = \sum_{k=1}^n \psi(T_{kk})$$

is tracial.

Answer. We just compute:

$$\begin{aligned}\Psi(TS) &= \sum_{k=1}^{n} \psi\big((TS)_{kk}\big) = \sum_{k=1}^{n}\sum_{h=1}^{n} \psi(T_{kh}S_{hk}) \\ &= \sum_{k=1}^{n}\sum_{h=1}^{n} \psi(S_{hk}T_{kh}) = \sum_{h=1}^{n} \psi\big((ST)_{hh}\big) \\ &= \Psi(ST).\end{aligned}$$

(14.5.2) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Show that a projection $P \in \mathcal{B}(\mathcal{H})$ is monic in the von Neumann algebra $\mathcal{B}(\mathcal{H})$ if and only if $\dim P\mathcal{H} = \dim(P\mathcal{H})^\perp$.

Answer. Suppose that $\dim P\mathcal{H} = \dim(P\mathcal{H})^\perp$. Considering orthonormal bases for each of these two subspaces, the equal cardinality allows us to construct a partial isometry $V : P\mathcal{H} \to (I_\mathcal{H} - P)\mathcal{H}$. Then $V^*V = P$, $VV^* = I_\mathcal{H} - P$, so $P \simeq I_\mathcal{H} - P$ and $P + (I_\mathcal{H} - P) = I_\mathcal{H} \in \mathcal{Z}(\mathcal{B}(\mathcal{H}))$. So P is monic.

Conversely, suppose that P is monic. Since $\mathcal{Z}(\mathcal{B}(\mathcal{H})) = \mathbb{C}\, I_\mathcal{H}$, the only central projections are 0 and $I_\mathcal{H}$. So there exist projections $P_1, \ldots, P_n \in \mathcal{B}(\mathcal{H})$ with $P_k \sim P$ for all k and $P_1 + \cdots + P_n = I_\mathcal{H}$. If P were finite, then we would get $I_\mathcal{H}$ finite by Proposition 14.2.15, contradicting that $\dim \mathcal{H} = \infty$. So P is infinite, and therefore so are all the P_k. We then have, applying Proposition 1.6.33 to the cardinality of the respective orthonormal bases,

$$\dim(P\mathcal{H})^\perp = \sum_{k=2}^{n} \dim P\mathcal{H} = \dim P\mathcal{H}.$$

(14.5.3) Show that a map $\Psi : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ is a conditional expectation if and only if it is a linear and positive projection.

Answer. We know that a conditional expectation is a positive linear projection by Tomiyama's Theorem (Proposition 13.2.68).

Conversely, suppose that $\Psi : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ is a linear and positive projection. Because $\mathcal{Z}(\mathcal{M})$ is abelian we have that Ψ is completely positive by Proposition 13.2.22. Then Proposition 13.2.24 gives us

$$\|\Psi\| = \|\Psi(I_\mathcal{M})\| = \|I_\mathcal{M}\| = 1.$$

And then Ψ is a conditional expectation by Tomiyama's Theorem (Proposition 13.2.68).

(14.5.4) Let $\mathcal{M}$ be a finite von Neumann algebra and $\mathcal{T}$ its centre-valued trace. Let $P, Q \in \mathcal{M}$ be projections. Show that $P \preceq Q$ if and only if $\mathcal{T}(P) \leq \mathcal{T}(Q)$.

Answer. If $P \preceq Q$, there exists $Q_0 \leq Q$ with $Q_0 \sim P$. Then

$$\mathcal{T}(P) = \mathcal{T}(Q_0) \leq \mathcal{T}(Q).$$

Conversely, suppose that $\mathcal{T}(P) \leq \mathcal{T}(Q)$. By Comparison there exists a nonzero central projection Z with $ZQ \preceq ZP$ and $(I_{\mathcal{M}} - Z)P \preceq (I_{\mathcal{M}} - Z)Q$. From $\mathcal{T}(P) \leq \mathcal{T}(Q)$ we have

$$0 \leq Z(\mathcal{T}(Q) - \mathcal{T}(P)) = \mathcal{T}(ZQ) - \mathcal{T}(ZP).$$

And from $ZQ \preceq ZP$ we have $\mathcal{T}(ZQ) - \mathcal{T}(ZP) \leq 0$. Thus $\mathcal{T}(ZQ) = \mathcal{T}(ZP)$. As $ZQ \sim P_0 \leq ZP$, this implies that $ZQ \sim ZP$ by the faithfulness of $\mathcal{T}$; indeed, $\mathcal{T}(ZP - P_0) = \mathcal{T}(ZP) - \mathcal{T}(ZQ) = 0$, so $P_0 = ZP$. And $ZP \sim ZQ$ and $(I_{\mathcal{M}} - Z)P \preceq (I_{\mathcal{M}} - Z)Q$ together imply $P \preceq Q$ by Proposition 14.2.7.

(14.5.5) Let $\mathcal{M}$ be a finite von Neumann algebra and $\psi, \varphi \in \mathcal{M}_*$ two tracial normal states such that $\psi|_{\mathcal{Z}(\mathcal{M})} = \varphi|_{\mathcal{Z}(\mathcal{M})}$. Show that $\psi = \varphi$.

Answer. Let $P \in \mathcal{M}$ be a monic projection. Then there exist projections $P_1, \ldots, P_r \in \mathcal{M}$ with $P_k \sim P$ for all k and $Z = \sum_k P_k \in \mathcal{Z}(\mathcal{M})$. We have

$$\begin{aligned}\varphi(P) &= r^{-1}\sum_{k=1}^{r} \varphi(P_k) = r^{-1}\varphi(Z) \\ &= r^{-1}\psi(Z) = r^{-1}\sum_{k=1}^{r}\psi(P_k) \\ &= \psi(P).\end{aligned}$$

If now P is an arbitrary projection in $\mathcal{M}$, by Lemma 14.5.4 there exist pairwise orthogonal monic projections $\{P_j\} \subset \mathcal{M}$ with $P = \sum_j P_j$. As both states are normal,

$$\varphi(P) = \sum_j \varphi(P_j) = \sum_j \psi(P_j) = \psi(P).$$

Given $T \in \mathcal{M}^{\mathrm{sa}}$, by the Spectral Theorem it is a norm limit of linear combinations of projections, so $\psi(T) = \varphi(T)$. Finally, any $T \in \mathcal{M}$ is a linear combination of two selfadjoints, so $\psi(T) = \varphi(T)$.

(14.5.6) Let $\mathcal{M}$ be a von Neumann algebra, $T_1, \ldots, T_r \in \mathcal{M}$, and $\varepsilon > 0$. Show that there exists $\gamma \in \mathcal{D}_{\mathcal{M}}$ and $Z_1, \ldots, Z_r \in \mathcal{Z}(\mathcal{M})$ such that

$$\|\gamma(T_k) - Z_k\| < \varepsilon, \qquad k = 1, \ldots, r.$$

Answer. We need to use the idea at the end of the proof of Theorem 14.5.15. We may assume without loss of generality that $T_1, \ldots, T_r$ are selfadjoint, for we may replace the list with a list of their real and imaginary parts. We argue by induction. From the proof of Theorem 14.5.15 up to (14.20) we have the case $r = 1$. So we assume as inductive hypothesis that we have $\beta \in \mathcal{D}_{\mathcal{M}}$ and $Z_1, \ldots, Z_{r-1}$ such that $\|\beta(T_k) - Z_k\| < \varepsilon$ for $k = 1, \ldots, r-1$. Applying again the argument that leads to (14.20) to $\beta(T_r)$, we obtain $\alpha \in \mathcal{D}_{\mathcal{M}}$ and $Z_r \in \mathcal{Z}(\mathcal{M})$ such that $\|\alpha(\beta(T_r)) - Z_r\| < \varepsilon$. Now we put $\gamma = \alpha \circ \beta$. Then for all k,

$$\begin{aligned}\|\gamma(T_k) - Z_k\| &= \|\alpha(\beta(T_k)) - Z_k\| = \|\alpha(\beta(T_k) - Z_k)\| \\ &\le \|\beta(T_k) - Z_k\| < \varepsilon.\end{aligned}$$

This completes the induction.

(14.5.7) Let $\mathcal{M}$ be a von Neumann algebra and $T_1, \ldots, T_r \in \mathcal{M}$. Show that there exist $Z_1, \ldots, Z_r \in \mathcal{Z}(\mathcal{M})$ and $\{\gamma_n\} \subset \mathcal{D}_{\mathcal{M}}$ such that $\lim_n \gamma_n(T_k) = Z_k$, $k = 1, \ldots, r$.

Answer. Given $n \in \mathbb{N}$, we apply Exercise 14.5.6 inductively to obtain $\beta_n \in \mathcal{D}_{\mathcal{M}}$ and $Z_{1,n}, \ldots, Z_{r,n} \in \mathcal{Z}(\mathcal{M})$ with

$$\|\beta_n \circ \cdots \circ \beta_1(T_k) - Z_k\| < 2^{-n}, \qquad k = 1, \ldots, r.$$

Let

$$\gamma_n = \beta_n \circ \beta_{n-1} \circ \cdots \circ \beta_1 \in \mathcal{D}_{\mathcal{M}}, \qquad n \in \mathbb{N}.$$

Then

$$\begin{aligned}\|\gamma_{n+1}(T_k) - Z_{k,n}\| &= \|\beta_{n+1}(\gamma_n(T_k) - Z_{k,n})\| \\ &\le \|\gamma_n(T_k) - Z_{k,n}\| < 2^{-n}.\end{aligned}$$

So

$$\|\gamma_{n+1}(T_k) - \gamma_n(T_k)\| \leq \|\gamma_{n+1}(T_k) - Z_{k,n}\| + \|\gamma_n(T_k) - Z_{k,n}\| < 2^{-n+1}.$$

It follows that

$$\|\gamma_{n+s}(T_k) - \gamma_n(T_k)\| \leq \sum_{j=n}^{n+s-1} \|\gamma_{j+1}(T_k) - \gamma_j(T_k)\| \leq \sum_{j=n}^{n+s-1} 2^{-j+1} < 2^{-n+2}.$$

So $\{\gamma_n(T_k)\}$ is Cauchy for each $k = 1, \ldots, r$. This forces $\{Z_{k,n}\}_n$ to be also Cauchy. Then there exists $Z_k = \lim_n Z_{k,n} \in \mathcal{Z}(\mathcal{M})$, and $Z_k = \lim_n \gamma_n(T_k)$.

(14.5.8) Let $\mathcal{M}$ be a von Neumann algebra and $S, T \in \mathcal{M}$. Show that

$$\mathcal{D}_{\mathcal{M}}(S+T) \cap \mathcal{Z}(\mathcal{M}) \subset \overline{\mathcal{D}_{\mathcal{M}}(S) \cap \mathcal{Z}(\mathcal{M}) + \mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})}.$$

Answer. Let $Z \in \mathcal{D}_{\mathcal{M}}(S+T) \cap \mathcal{Z}(\mathcal{M})$ and fix $\varepsilon > 0$. This means that there exists $\beta \in \mathcal{D}_{\mathcal{M}}$ such that $\|\beta(S+T) - Z\| < \varepsilon$. By Exercise 14.5.7 there exist $\gamma \in \mathcal{D}_{\mathcal{M}}$ and $Z_1 \in \mathcal{D}_{\mathcal{M}}(\beta(S)) \cap \mathcal{Z}(\mathcal{M})$, $Z_2 \in \mathcal{D}_{\mathcal{M}}(\beta(T)) \cap \mathcal{Z}(\mathcal{M})$ such that

$$\|\gamma(\beta(S)) - Z_1\| < \varepsilon, \qquad \|\gamma(\beta(T)) - Z_2\| < \varepsilon.$$

As

$$\|\gamma(\beta(S+T)) - Z\| = \|\gamma(\beta(S+T) - Z)\| \leq \|\beta(S+T) - Z\| < \varepsilon,$$

we obtain

$$\|Z - (Z_1 + Z_2)\| \leq \|Z - \gamma(\beta(S+T))\| + \|\gamma(\beta(S)) - Z_1\| + \|\gamma(\beta(T)) - Z_2\| \leq 3\varepsilon.$$

As this can be done for any ε and

$$\mathcal{D}_{\mathcal{M}}(\beta(S)) \cap \mathcal{Z}(\mathcal{M}) \subset \mathcal{D}_{\mathcal{M}}(S) \cap \mathcal{Z}(\mathcal{M})$$

and

$$\mathcal{D}_{\mathcal{M}}(\beta(T)) \cap \mathcal{Z}(\mathcal{M}) \subset \mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M}),$$

we have shown that

$$Z \in \overline{\mathcal{D}_{\mathcal{M}}(S) \cap \mathcal{Z}(\mathcal{M}) + \mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})}.$$

(14.5.9) Let $\mathcal{M}$ be a von Neumann algebra, $T \in \mathcal{M}$, $Z \in \mathcal{Z}(\mathcal{M})$. Show that
$$\mathcal{D}_{\mathcal{M}}(TZ) \cap \mathcal{Z}(\mathcal{M}) \subset \overline{Z\big(\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})\big)}.$$

Answer. Fix $Y \in \mathcal{D}_{\mathcal{M}}(TZ) \cap \mathcal{Z}(\mathcal{M})$ and $\varepsilon > 0$. Then there exists $\beta \in \mathcal{D}_{\mathcal{M}}$ such that $\|\beta(TZ) - Y\| < \varepsilon$. By Theorem 14.5.15 there exists $\gamma \in \mathcal{D}_{\mathcal{M}}$ and $Y_1 \in \mathcal{D}_{\mathcal{M}}(\beta(T)) \cap \mathcal{Z}(\mathcal{M})$ with $\|\gamma(\beta(T)) - Y_1\| < \varepsilon$. Then
$$\|Z(\gamma(\beta(T)) - Y\| = \|(\gamma(\beta(TZ)) - Y\| \leq \|\beta(TZ) - Y\| < \varepsilon.$$
Also,
$$\|Z(\gamma(\beta(T)) - ZY_1\| \leq \|\gamma(\beta(T)) - Y_1\| \, \|Z\| < \varepsilon \, \|Z\|.$$
Then
$$\|Y - ZY_1\| \leq (1 + \|Z\|)\varepsilon.$$
As Z is fixed and this can be done for all $\varepsilon > 0$, we have shown that $Y \in \overline{Z\big(\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})\big)}$.

(14.5.10) Let $\mathcal{M} = \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$ and $T \in \mathcal{K}(\mathcal{H})^+$. Show that $\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M}) = \{0\}$.

Answer. Fix $\varepsilon > 0$. Using the Spectral Theorem (Theorem 10.6.12) we have $T = \sum_{k=1}^\infty \lambda_k P_k$ where $P_1, P_2, \ldots,$ are rank-one projections that add to the identity. Choosing n big enough we can take $T_0 = \sum_{k=1}^n \lambda_k P_k$ with $\|T - T_0\| < \varepsilon$. Let $\{E_{kj}\}$ be matrix units in $\mathcal{B}(\mathcal{H})$ with $E_{kk} = P_k$ for all k. Fix $m > 1/\varepsilon$. For $j = 1, \ldots, m$ let let U_j be a unitary with
$$U_j E_{kk} U_j^* = P_{(j-1)n+k}, \qquad k = 1, \ldots, n.$$
This can always be done, as what we are doing is just a permutation of projections. If $\{\xi_k\}$ is the orthonormal basis corresponding to the $\{E_{kj}\}$, we have $U_1 = I_{\mathcal{H}}$, U_2 exchanges $\xi_1, \ldots, \xi_n$ with $\xi_{n+1}, \ldots, \xi_{2n}$ and leaves the rest alone, U_3 exchanges $\xi_1, \ldots, \xi_n$ with $\xi_{2n+1}, \ldots, \xi_{3n}$ and leaves the rest alone, etc. We define
$$\gamma(T) = \frac{1}{m} \sum_{j=1}^m U_j T U_j^*.$$
Then
$$\begin{aligned}\gamma(T_0) &= \frac{1}{m} \sum_{j=1}^m U_j T_0 U_j^* = \frac{1}{m} \sum_{j=1}^m \sum_{k=1}^n \lambda_k U_j P_k U_j^* \\ &= \frac{1}{m} \sum_{j=1}^m \sum_{k=1}^n \lambda_k U_j P_{(j-1)n+k} U_j^* = \sum_{j=1}^m \sum_{k=1}^n \frac{\lambda_k}{m} U_j P_{(j-1)n+k} U_j^*.\end{aligned}$$

As the projections $P_{(j-1)n+k}$ are pairwise orthogonal for $k = 1, \ldots, n$ and $j = 1, \ldots, m$, we get that

$$\|\gamma(T_0)\| = \frac{1}{m} \max\{|\lambda_k| : \ k\} \leq \frac{\|T\|}{m} < \varepsilon \, \|T\|.$$

Then

$$\|\gamma(T)\| \leq \|\gamma(T - T_0)\| + \|\gamma(T_0)\| < \varepsilon + \varepsilon \, \|T\|.$$

This can be done for all $\varepsilon > 0$, so $0 \in \mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})$ since the set is closed.

Conversely, for any $\gamma \in \mathcal{D}_{\mathcal{B}(\mathcal{H})}$ we have $\gamma(T) \in \mathcal{K}(\mathcal{H})$ (linear combinations of unitary conjugates of compact are compact). As $\mathcal{K}(\mathcal{H})$ is norm-closed, $\mathcal{D}_{\mathcal{B}(\mathcal{H})}(T) \subset \mathcal{K}(\mathcal{H})$. This means that

$$\mathcal{D}_{\mathcal{B}(\mathcal{H})}(T) \cap \mathcal{Z}(\mathcal{B}(\mathcal{H})) \subset \mathcal{K}(\mathcal{H}) \cap \mathcal{Z}(\mathcal{B}(\mathcal{H})) = \{0\}.$$

Therefore $\mathcal{D}_{\mathcal{B}(\mathcal{H})}(T) \cap \mathcal{Z}(\mathcal{B}(\mathcal{H})) = \{0\}$.

(14.5.11) Let $\mathcal{M} = \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$ and $P \in \mathcal{B}(\mathcal{H})$ an infinite projection with $I_{\mathcal{H}} - P$ infinite. Show that $\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M}) = [0, 1] \, I_{\mathcal{H}}$.

Answer. By working on an orthonormal basis made out of orthonormal bases for $P\mathcal{H}$ and $\ker P$ we may assume without loss of generality that P is diagonal. So we may think of P as an element of $\{0, 1\}^{\mathbb{N}}$ with infinitely many entries equal to 1 and infinitely many entries equal to 0. Via unitary conjugation we can implement any permutation. Fix $m, n \in \mathbb{N}$ with $m < n$. Choose a unitary U such that UPU^* corresponds to

$$\overbrace{0, \ldots, 0}^{n-m \text{ times}}, \overbrace{1, \ldots, 1}^{m \text{ times}}, \ldots$$

where the pattern repeats afterwards. Let $r = \binom{n}{m}$ and $V_1, \ldots, V_r$ unitaries that implement all distinct r permutations (this is the total number of permutations of the first n entries if we ignore permutations that produce the same arrangement of 0 and 1). Let

$$\gamma(T) = \frac{1}{r} \sum_{j=1}^{r} V_j UPU^* V_j^*.$$

In each coordinate the amount of 1 is equal to the amount of configurations of the remaining $m-1$ entries that are equal to 1 distributed over the remaining $n-1$ positions. So there is a total of $\binom{m-1}{n-1}$ entries equal to 1. This shows that

$$\gamma(T) = \frac{\binom{m-1}{n-1}}{\binom{m}{n}} \, I_{\mathcal{H}} = \frac{m}{n} \, I_{\mathcal{H}}.$$

So $\frac{m}{n} \, I_{\mathcal{H}} \in \mathcal{D}_{\mathcal{B}(\mathcal{H})}(P) \cap \mathcal{Z}(\mathcal{B}(\mathcal{H}))$ for all $m, n \in \mathbb{N}$ with $m < n$. As $\mathcal{D}_{\mathcal{B}(\mathcal{H})}(P)$ is closed, $t \, I_{\mathcal{H}} \in \mathcal{D}_{\mathcal{B}(\mathcal{H})}(P) \cap \mathcal{Z}(\mathcal{B}(\mathcal{H}))$ for all $t \in [0, 1]$.

The converse is trivial, for $0 \leq P \leq I_{\mathcal{H}}$ implies $0 \leq \gamma(P) \leq I_{\mathcal{H}}$ for all $\gamma \in \mathcal{D}_{\mathcal{B}(\mathcal{H})}$, so any $Z \in \mathcal{D}_{\mathcal{B}(\mathcal{H})}(P) \cap \mathcal{Z}(\mathcal{B}(\mathcal{H}))$ satisfies $0 \leq Z \leq I_{\mathcal{H}}$. As Z is necessarily a scalar, $Z = t\, I_{\mathcal{H}}$ with $t \in [0,1]$.

(14.5.12) Given a weight $\psi : \mathcal{M}^+ \to [0,\infty]$, show that $\mathcal{F}_\psi$ is a face in $\mathcal{M}^+$, N_ψ is a left ideal in $\mathcal{M}$, $\mathcal{M}_\psi$ is a $*$-subalgebra, and $\mathcal{M}_\psi = \operatorname{span} \mathcal{F}_\psi$.

Answer. Given $S, T \in \mathcal{F}_\psi$ and $t \in [0,1]$, $\psi(tS + (1-t)T) = t\psi(S) + (1-t)\psi(T) < \infty$, so $\mathcal{F}_\psi$ is convex. If $T = tS_1 + (1-t)S_2 \in \mathcal{F}_\psi$, then

$$t\psi(S_1) \leq t\psi(S_1) + (1-t)\psi(S_2) = \psi(T) < \infty,$$

so $S_1 \in \mathcal{F}_\psi$ and similarly $S_2 \in \mathcal{F}_\psi$. Hence $\mathcal{F}_\psi$ is a face.

Given $T \in \mathcal{N}_\psi$ and $S \in \mathcal{N}_\psi$, since $0 \leq (S-T)^*(S-T) = S^*S + T^*T - 2\mathrm{Re}\, S^*T$ we have the inequality

$$\begin{aligned}(S+T)^*(S+T) &= S^*S + T^*T + 2\mathrm{Re}\, S^*T \\ &\leq 2S^*S + 2T^*T.\end{aligned} \tag{AB.14.4}$$

Then

$$\psi((S+T)^*(S+T)) \leq 2\psi(S^*S) + 2\psi(T^*T) < \infty$$

so $\mathcal{N}_\psi$ is a subspace. When $S \in \mathcal{M}$,

$$\psi((ST)^*ST) = \psi(T^*S^*ST) \leq \|S\|^2 \psi(T^*T) < \infty,$$

showing that $\mathcal{N}_\psi$ is a left ideal. Since the adjoint reverses products it follows that $\mathcal{N}_\psi^*$ is a right ideal, and then $\mathcal{M}_\psi = \mathcal{N}_\psi^* \cap \mathcal{N}_\psi$ is an ideal since it is the intersection of a left and a right ideal.

Given $T \in \mathcal{F}_\psi$ we have $T^{1/2} \in \mathcal{N}_\psi \cap \mathcal{N}_\psi^*$, so $T \in \mathcal{M}_\psi$. As $\mathcal{M}_\psi$ is a subspace we get that $\operatorname{span} \mathcal{F}_\psi \subset \mathcal{M}_\psi$. Conversely, an element of $\mathcal{M}_\psi$ is a linear combination of elements of the form S^*T with $S, T \in \mathcal{N}_\psi$. So it is enough to show that $S^*T \in \operatorname{span} \mathcal{F}_\psi$. By (AB.14.4) with both $\pm T$, we get that $|S \pm T|^2 \in \mathcal{F}_\psi$. Then

$$4\mathrm{Re}\, S^*T = (S+T)^*(S+T) - (S-T)^*(S-T) \in \operatorname{span} \mathcal{F}_\psi.$$

Replacing S with iS we get $\mathrm{Im}\, S^*S \in \operatorname{span} \mathcal{F}_\psi$, and so $S^*T \in \operatorname{span} \mathcal{F}_\psi$.

(14.5.13) Show that the converse to Proposition 14.5.22 does not hold in general.

Answer. Let τ be the unique semifinite trace Tr on $\mathcal{B}(\mathcal{H})$. Let $P \in \mathcal{B}(\mathcal{H})$ be an infinite projection such that $I_{\mathcal{H}} - P$ is infinite; for instance, fix a countable orthonormal set and let P be the projection onto the subspace spanned by basis elements with even index. Let $\mathcal{A} = \text{span}\{P, I_{\mathcal{H}} - P\}$. Then $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ is an abelian von Neumann algebra, and $\text{Tr}(Q) = \infty$ for every nonzero projection in $\mathcal{A}$ (the only projections in $\mathcal{A}$ are $I - \mathcal{H}$, P, and $I_{\mathcal{H}} - P$). So $\tau|_{\mathcal{A}}$ is not semifinite.

14.6. Examples of Factors

(14.6.1) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, $\pi : \mathcal{M} \to \mathcal{B}(\mathcal{H})$ the identity representation and G a group that acts on $\mathcal{M}$ via α. Show that if $S \in \mathcal{M}'$, then $S \otimes I_{\ell^2(G)} \in \tilde{\pi}(\mathcal{M})'$.

Answer. For $T \in \mathcal{M}$, $\xi \in \mathcal{H}$, and $g \in G$,

$$\begin{aligned}(S \otimes I_{\ell^2(G)})\tilde{\pi}(T)(\xi \otimes \delta_g) &= (S \otimes I_{\ell^2(G)})(\alpha_g^{-1}(T)\xi \otimes \delta_g) \\ &= S\alpha_g^{-1}(T)\xi \otimes \delta_g = \alpha_g^{-1}(T)S\xi \otimes \delta_g \\ &= \tilde{\pi}(T)(S\xi \otimes \delta_g) = \tilde{\pi}(T)(S \otimes I_{\ell^2(G)})(\xi \otimes \delta_g).\end{aligned}$$

By linearity and taking limits, $(S \otimes I_{\ell^2(G)})\tilde{\pi}(T) = \tilde{\pi}(T)(S \otimes I_{\ell^2(G)})$. This happens for all $T \in \mathcal{M}$, so $(S \otimes I_{\ell^2(G)}) \in \tilde{\pi}(\mathcal{M})'$.

(14.6.2) Let $\tilde{T} \in \mathcal{M} \rtimes_\alpha G$ and $g, h, r, s \in G$. Show that $[U_r\tilde{T}U_s]_{g,h} = \tilde{T}_{r^{-1}g,sh}$.

Answer. For any $\xi, \eta \in \mathcal{H}$,

$$\begin{aligned}\langle \left[U_r\tilde{T}U_s\right]_{g,h}\xi, \eta\rangle &= \langle U_r\tilde{T}U_s(\xi \otimes \delta_h), \eta \otimes \delta_g\rangle \\ &= \langle \tilde{T}(\xi \otimes \delta_{sh}), \eta \otimes \delta_{r^{-1}g}\rangle \\ &= \langle \tilde{T}_{r^{-1}g,sh}\xi, \eta\rangle.\end{aligned}$$

Hence $[U_s\tilde{T}U_s]_{g,h} = \tilde{T}_{r^{-1}g,sh}$.

(14.6.3) Let $\tilde{T} \in \mathcal{M} \rtimes_\alpha G$ and $g, h \in G$. Show that $\tilde{T}_{g,h} = \mathcal{E}(U_g^* \tilde{T} U_h)$.

Answer. Using Exercise 14.6.2,

$$\mathcal{E}(U_g^* \tilde{T} U_h) = \left[U_{g^{-1}} \tilde{T} U_h\right]_{e,e} = \tilde{T}_{g,h}.$$

(14.6.4) Prove (14.22).

Answer. We have

$$\begin{aligned}
\langle \tilde{T}_{g,h}\xi, \eta\rangle &= \langle \tilde{T}(\xi \otimes \delta_h), \eta \otimes \delta_g\rangle \\
&= \langle U_g^* \tilde{T}(\xi \otimes \delta_h), \eta \otimes \delta_e\rangle \\
&= \langle U_g^* \tilde{T} U_g\, U_g^*(\xi \otimes \delta_h), \eta \otimes \delta_e\rangle \\
&= \langle U_g^* \tilde{T} U_g\, (\xi \otimes \delta_{g^{-1}h}), \eta \otimes \delta_e\rangle \\
&= \langle (U_g^* \tilde{T} U_g)_{e,g^{-1}h}\xi, \eta\rangle.
\end{aligned}$$

This holds for all $\xi, \eta \in \mathcal{H}$, so

$$\tilde{T}_{g,h} = \left((U_g^* \tilde{T} U_g\right)_{e,g^{-1}h}.$$

(14.6.5) Use (14.23) to show that the conditional expectation $\mathcal{E} : \mathcal{M} \rtimes_\alpha G \to \mathcal{M}$ is faithful.

Answer. Suppose that $\mathcal{E}(\tilde{T}^*\tilde{T}) = 0$. This means that $[\tilde{T}^*\tilde{T}]_{e,e} = 0$. So, for $\xi \in \mathcal{H}$,

$$0 = \langle [\tilde{T}^*\tilde{T}]_{e,e}\xi, \xi\rangle = \langle \tilde{T}^*\tilde{T}(\xi \otimes \delta_e), \xi \otimes \delta_e\rangle = \|\tilde{T}(\xi \otimes \delta_e)\|^2.$$

Thus $T(\xi \otimes \delta_e) = 0$ for all $\xi \in \mathcal{H}$. This gives us, for any $\xi, \eta \in \mathcal{H}$,

$$\langle \tilde{T}_{g,e}\xi, \eta\rangle = \langle \tilde{T}(\xi \otimes \delta_e), \eta \otimes \delta_g\rangle = 0.$$

It follows that $\tilde{T}_{g,e} = 0$ for all g. By (14.23),

$$\tilde{T}_{g,h} = \alpha_h^{-1}(\tilde{T}_{gh^{-1},e}) = 0$$

for all $g, h \in G$. We can now write

$$\langle \tilde{T}(\xi \otimes \delta_h), \eta \otimes \delta_g\rangle = \langle \tilde{T}_{g,h}\xi, \eta\rangle = 0,$$

and therefore $\tilde{T} = 0$ after using linear combinations and continuity.

(14.6.6) Let θ be an irrational number and α the translation action as in Example 14.6.9. Show that α is free and ergodic.

Answer. Fix $n \in \mathbb{Z}$ and suppose that $fg = \alpha_n(g)f$ for all $g \in L^\infty(\mathbb{T})$. Fix $z \in \mathbb{T}$. We can always construct g (as a polynomial, even) with $g(z) \neq \alpha_n(g)(z)$. Then $f(z) = 0$. This can be done for all z, so $f = 0$ and the action is free.

Now suppose that $\alpha_n(f) = f$ for all n. Fix $\varepsilon > 0$; as f is uniformly continuous by the compactness of $\mathbb{T}$, there exists δ with $|f(z) - f(w)| < \varepsilon$ whenever $|z - w| < \delta$. Since θ is irrational, given $z \in \mathbb{T}$ we can find n such that $|z - e^{2\pi i\theta n}| < \delta$. Then

$$\begin{aligned}|f(z) - f(1)| &\leq |f(z) - f(e^{2\pi i\theta n})| + |f(e^{2\pi i\theta n}) - f(1)|\\ &< \varepsilon + |\alpha_n(f)(1) - f(1)| = \varepsilon.\end{aligned}$$

As this can be done for any $\varepsilon > 0$ we have shown that $f(z) = f(1)$ and thus f is constant. Therefore the action is ergodic.

(14.6.7) Show that the action of $\mathbb{Q}$ on $L^\infty(\mathbb{R})$ by translation is free and ergodic.

Answer. If $fg = \alpha_q(g)f$ for all $g \in L^\infty(\mathbb{R})$, this means that $f(t)g(t) = f(t)g(t+q)$ for all t. We can construct g with $g(t) = 0$ and $g(t+q) = 0$, which shows that $f(t) = 0$. As this can be done for all t, $f = 0$. Properly, the equality $fg = \alpha_q(g)f$ is almost everywhere, so one needs a bit more care, but basically the argument is the above up to a nullset. Hence the action is free.

If $\alpha_q(f) = f$ for all $q \in \mathbb{Q}$, this is $f(t) = f(t+q)$ almost everywhere. Given $r \in \mathbb{R}$, choose $\{q_n\} \subset \mathbb{Q}$ with $q_n \to r$. Using the notation from Lemma 2.8.19, and working on a fixed interval $[-m, m]$ so that f is integrable there,

$$\|f_r - f\|_1 \leq \|f_r - f_{q_n}\|_1 + \|f_{q_n} - f\|_1 = \|f_r - f_{q_n}\|_1 \to 0$$

by Lemma 2.8.19. Hence $f_r = f$ a.e. for each $r \in \mathbb{R}$. For each Lebesgue point r of f, by Theorem 2.11.9

$$\begin{aligned}f(r) &= \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_{r-\varepsilon}^{r+\varepsilon} f(t)\,dt = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} f(t-r)\,dt\\ &= \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} f(t)\,dt = f(0).\end{aligned}$$

Properly 0 might not be a Lebesgue point for f, but we can translate the integrals to any Lebesgue point, and this is all of $\mathbb{R}$ up to a nullset. Finally, this can be done for every interval $[-m, m]$ so f is constant in these overlapping intervals; hence f is constant, and the action is ergodic.

14.7. II_1-Factors

(14.7.1) Where in the proof of Proposition 14.7.2 is the norm-closedness of $\mathcal{J}$ used?

Answer. It is used in the fact that $\mathcal{J}$ is hereditary. The argument after Definition 11.5.19 requires $\mathcal{J}$ to be closed. This was discussed in Exercise 11.5.11.

(14.7.2) Give an alternative proof of Proposition 14.7.2 by using Dixmier's Property.

Answer. Let $\mathcal{M}$ be a II_1-factor and $\mathcal{J} \subset \mathcal{M}$ a nonzero norm-closed ideal. Let $T \in \mathcal{J}$ be positive (recall that $\mathcal{J}$ is a C*-algebra, so it is spanned by its positive elements). As the trace τ is faithful, $\tau(T) > 0$. By Corollary 14.5.16 we have that $\tau(T)\, I_{\mathcal{M}}$ is the unique element of $\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})$. So there exists a sequence $\{\gamma_n\} \subset \mathcal{D}_{\mathcal{M}}$ with $\gamma_n(T) \to \tau(T)\, I_{\mathcal{M}}$. As $T \in \mathcal{J}$ and $\mathcal{J}$ is an ideal, $UTU^* \in \mathcal{J}$ for any unitary U, and henceforth $\gamma_n(T) \in \mathcal{J}$ for all n. Then $\tau(T)\, I_{\mathcal{M}} \in \overline{\mathcal{J}} = \mathcal{J}$, and therefore $I_{\mathcal{M}} \in \mathcal{J}$, showing that $\mathcal{J} = \mathcal{M}$.

(14.7.3) Show that θ in (14.24) is a $*$-homomorphism.

Answer. When we write $\theta(T)\theta(S)$, the 1, 1 entry is

$$\begin{aligned}\gamma^{-1}(PTPSP + PT(I_{\mathcal{M}} - P)SP) &= \gamma^{-1}(PT\big(P + I_{\mathcal{M}} - P\big)SP)\\ &= \gamma^{-1}(PTSP).\end{aligned}$$

The same phenomenon occurs on the other three entries, like the 2, 1 entry is
$\gamma^{-1}(PTPS(I_{\mathcal{M}} - P)V + PT(I_{\mathcal{M}} - P)S(I_{\mathcal{M}} - P)V) = \gamma^{-1}(PTS(I_{\mathcal{M}} - P)V)$.

As for the adjoint,

$$\begin{aligned}\theta(T)^* &= \begin{bmatrix}\gamma^{-1}(PTP) & \gamma^{-1}(PT(I_{\mathcal{M}}-P)V)\\ \gamma^{-1}(V^*(I_{\mathcal{M}}-P)TP) & \gamma^{-1}(V^*(I_{\mathcal{M}}-P)T(I_{\mathcal{M}}-P)V)\end{bmatrix}^*\\
&= \begin{bmatrix}\gamma^{-1}((PTP)^*) & \gamma^{-1}([V^*(I_{\mathcal{M}}-P)TP]^*)\\ \gamma^{-1}([PT(I_{\mathcal{M}}-P)V]^*) & \gamma^{-1}(V^*[(I_{\mathcal{M}}-P)T(I_{\mathcal{M}}-P)]^*V)\end{bmatrix}\\
&= \begin{bmatrix}\gamma^{-1}(PT^*P) & \gamma^{-1}(PT^*(I_{\mathcal{M}}-P)V)\\ \gamma^{-1}(V^*(I_{\mathcal{M}}-P)T^*P) & \gamma^{-1}(V^*(I_{\mathcal{M}}-P)T^*(I_{\mathcal{M}}-P)V)\end{bmatrix}\\
&= \theta(T^*).\end{aligned}$$

And for the injectivity, if $\theta(T)=0$ we immediately get from the injectivity of γ^{-1} that

$$PTP=0,\quad PT^*(I_{\mathcal{M}}-P)V=0,\quad V^*(I_{\mathcal{M}}-P)TP=0,$$

and

$$V^*(I_{\mathcal{M}}-P)T(I_{\mathcal{M}}-P)V=0.$$

Multiplying the second equality by V^* on the right, the third one by V on the left, and the fourth one by V on the left and V^* on the right, we obtain

$$PTP=0,\quad PT^*(I_{\mathcal{M}}-P)=0,\quad (I_{\mathcal{M}}-P)TP=0,\quad (I_{\mathcal{M}}-P)T(I_{\mathcal{M}}-P)=0.$$

And now adding the four equalities yields $T=0$.

(14.7.4) Let $\mathcal{M}$ be a II_1-factor, $n\in\mathbb{N}$,

$$P\in\mathcal{P}(M_n(\mathcal{M})),\qquad\text{and}\qquad Q\in M_n(PM_n(\mathcal{M})P).$$

Show that $QM_{n^2}(\mathcal{M})Q=QM_n(PM_n(\mathcal{M})P)Q$.

Answer. Thinking about the natural identification $M_{n^2}(\mathcal{M})\simeq M_n(M_n(\mathcal{M}))$ we clearly have $QM_{n^2}(\mathcal{M})Q\supset QM_n(PM_n(\mathcal{M})P)Q$, so we only need to prove the inclusion $QM_{n^2}(\mathcal{M})Q\subset QM_n(PM_n(\mathcal{M})P)Q$. Let $X\in M_{n^2}(\mathcal{M})$, seen as $M_n(M_n(\mathcal{M}))$. The hypothesis is that Q is an $n\times n$ block matrix $[Q_{kj}]$ with $Q_{kj}\in PM_n(\mathcal{M})P$ for all k,j. Then

$$\begin{aligned}(QXQ)_{k,j} &= \sum_{r,s}Q_{k,r}X_{r,s}Q_{s,j}=\sum_{r,s}PQ_{k,r}X_{r,s}Q_{s,j}P\\
&= P\Big(\sum_{r,s}Q_{k,r}X_{r,s}Q_{s,j}\Big)P\in PM_n(\mathcal{M})P.\end{aligned}$$

APPENDIX

A

The Determinant

A.1. Preliminaries on Permutations

A.2. Preliminaries on Multilinear Maps

A.3. The Determinant

(A.3.1) Given $j \in \{1, \ldots, n\}$ and $\nu \in \mathbb{S}_{n-1}$, let $\sigma_{j,\nu}$ as in (A.3). Show that $\sigma_{j,\nu} \in \mathbb{S}_n$.

Answer. Let $\alpha : \{2, \ldots, n\} \to \{1, \ldots, n-1\}$ be given by $\alpha(k) = k - 1$. Let $\beta_j : \{1, \ldots, n-1\} \to \{1, \ldots, n\} \setminus \{j\}$ be

$$\beta_j(k) = \begin{cases} k, & k < j \\ k+1, & k \geq j \end{cases}$$

Then $\sigma_{j,\nu}$ is given by $\sigma_{j,\nu}(1) = j$ and for $k \geq 2$,

$$\sigma_{j,\nu}(k) = \beta_j \circ \nu \circ \alpha(k).$$

Being a composition of bijections, it is a bijection.

(A.3.2) Given $j \in \{1, \ldots, n\}$ and $\nu \in \mathbb{S}_{n-1}$, let $\sigma_{j,\nu}$ as in (A.3). Show that $\operatorname{sgn} \sigma_{j,\nu} = (-1)^{j-1} \operatorname{sgn} \nu$.

Answer. The number $\operatorname{sgn} \sigma_{j,\nu}$ is the parity of the cardinality of the set

$$P_{\sigma_{j,\nu}} = \{(r,s) : \ r < s, \ \sigma_{j,\nu}(r) > \sigma_{j,\nu}(s)\}.$$

The formula for $\sigma_{j,\nu}$ will usually apply a "+1‘" to both $\nu(r-1)$ and $\nu(s-1)$; and it if it only applies it to one of them, it will be to the largest. Then

$$\begin{aligned} P_{\sigma_{j,\nu}} &= \{(r,s) : \ 2 \leq r < s, \ \sigma_{j,\nu}(r) > \sigma(s)\} \\ &\qquad \cup \{(1,s) : \ 1 < s, \ \sigma_{j,\nu}(1) > \sigma_{j,\nu}(s)\} \\ &= \{(r,s) : \ 2 \leq r < s, \ \nu(r-1) > \nu(s-1)\} \\ &\qquad \cup \{(1,s) : \ 1 < s, \ j > \sigma_{j,\nu}(s)\} \\ &= \{(r,s) : \ 2 \leq r < s, \ \nu(r-1) > \nu(s-1)\} \cup \{1, \ldots, j-1\} \end{aligned}$$

(the set $\{1, \ldots, j-1\}$ pops up because there are precisely since ν is a bijection there are precisely $j-1$ points that end up below j). Thus $\operatorname{sgn} \sigma_{j,\nu} = (-1)^{j-1} \operatorname{sgn} \nu$.

(A.3.3) With the notation of (A.3), show that $\mathbb{S}_n = \{\sigma_{j,\nu} : \ j \in \{1, \ldots, n\}, \ \nu \in \mathbb{S}_{n-1}\}$.

Answer. Given $\sigma \in \mathbb{S}_n$, let $j = \sigma(1)$, and put

$$\nu(k) = \begin{cases} \sigma(k+1), & \sigma(k+1) < j \\ \sigma(k+1) - 1, & \sigma(k+1) > j \end{cases}$$

Then $\sigma_{j,\nu} = \sigma$.

APPENDIX

B

Getting to Know Majorization

B.1. Preliminaries on Majorization

(B.1.1) Let $x \in \mathbb{R}^n$ with $x_j \geq 0$ for all j. Show that
$$\frac{\operatorname{Tr}(x)}{n}\, e \prec x \prec (\operatorname{Tr}(x), 0, \ldots, 0)$$

Answer. For $x \prec (\operatorname{Tr}(x), 0, \ldots, 0)$, both vectors have the same trace, and since the entries of x are non-negative,
$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{k} x_j = \operatorname{Tr}(x), \qquad k = 1, \ldots, n.$$

And for $\frac{\operatorname{Tr}(x)}{n} e \prec x$, again they have the same trace. If we had, for some k,
$$\frac{k}{n} \operatorname{Tr}(x) = \sum_{j=1}^{k} \frac{\operatorname{Tr}(x)}{n} > \sum_{j=1}^{k} x_j^{\downarrow},$$

then

$$\sum_{j=1}^{n} x_j^{\downarrow} < \frac{k}{n}\operatorname{Tr}(x) + \sum_{j=k+1}^{n} x_j^{\downarrow} \leq \frac{k}{n}\operatorname{Tr}(x) + (n-k)x_k^{\downarrow}$$

$$\leq \frac{k}{n}\operatorname{Tr}(x) + (n-k)\,\frac{1}{n}\sum_{j=1}^{n} x_j^{\downarrow}$$

$$\leq \frac{k}{n}\operatorname{Tr}(x) + \Big(1-\frac{k}{n}\Big)\operatorname{Tr}(x) = \operatorname{Tr}(x),$$

a contradiction. Thus $\frac{\operatorname{Tr}(x)}{n}\, e \prec_w x$ and, as they have the same trace, we get majorization.

(B.1.2) Show that $t_1, \ldots, t_n$ are convex coefficients (that, is $t_j \geq 0$ for all j and $\sum_j t_j = 1$) if and only if $t \prec e_1$, where $t = (t_1, \ldots, t_n)$.

Answer. Suppose that $t_j \geq 0$ for all j and $\sum_j t_j = 1$. Then $\sum_{j=1}^{k} t_j \leq 1 = \sum_{j=1}^{k} (e_1)_j$ for all k, with equality for $k = n$. Thus $t \prec e_1$.

Conversely, if $t \prec e_1$, assuming without loss of generality that $t = t^{\downarrow}$, we have $t_n \geq (e_1)_n = 0$, so $t_j \geq 0$ for all j. And $\sum_j t_j = \operatorname{Tr}(t) = \operatorname{Tr}(e_1) = 1$.

(B.1.3) Prove (B.1), that is $x \prec_w y$ and $y \prec_w x$ if and only if there exists a permutation σ with $y = P_\sigma x$.

Answer. For each k we have

$$\sum_{j=1}^{k} x_j^{\downarrow} = \sum_{j=1}^{k} y_j^{\downarrow},$$

for all $k = 1, \ldots, n$, since the double submajorization gives us the inequality both ways. When $k = 1$ we get $x_1^{\downarrow} = y_1^{\downarrow}$. This equality together with the equality for $k = 2$ give $x_2^{\downarrow} = y_2^{\downarrow}$. Continuing this way, we obtain $x_j^{\downarrow} = y_j^{\downarrow}$ for all j. That is, x and y have the exact same entries, possibly in different order. Hence there exists $\sigma \in \mathbb{S}_n$ with $x = P_\sigma y$.

(B.1.4) Let $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. Show that if $x \prec y$ then $\alpha x \prec \alpha y$.

Answer. Suppose that $x \prec y$. When $\alpha \geq 0$, the equality and inequalities defining majorization are preserved, so $\alpha x \prec \alpha y$. To deal with the case where $\alpha < 0$, it is enough to show that $-x \prec -y$. The condition $\mathrm{Tr}(-x) = \mathrm{Tr}(-y)$ is satisfied trivially by linearity. Using that $(-x)^{\downarrow}_j = -x^{\uparrow}_{n-j+1}$,

$$\sum_{j=1}^{k} (-x)^{\downarrow}_j = -\sum_{j=1}^{k} x^{\uparrow}_{n-j+1} = -\sum_{j=n-k+1}^{n} x^{\uparrow}_j$$
$$\leq -\sum_{j=n-k+1}^{n} y^{\uparrow}_j = \sum_{j=1}^{k} (-y)^{\downarrow}_j.$$

Hence $(-x) \prec (-y)$.

(B.1.5) Show that $A \in \mathrm{DS}(n)$ if and only if A has non-negative entries, $Ae = e$, and $A^\top e = e$.

Answer. The equality $Ae = e$ means that $(Ae)_k = e_k = 1$ for all k, and this is

$$\sum_{j=1}^{n} A_{kj} = 1.$$

In other words, $Ae = e$ describes exactly row stochasticity. Similarly, $A^\top e = e$ is

$$\sum_{k=1}^{n} A_{kj} = 1,$$

which is column stochasticity.

(B.1.6) Show that $\mathrm{DS}(n)$ is convex.

Answer. If $A, B \in M_n(\mathbb{R})$ are doubly stochastic and $t \in [0,1]$, then $[tA + (1-t)B]_{kj} = tA_{kj} + (1-t)B_{kj} \geq 0$, and

$$\sum_{j=1}^{n} [tA + (1-t)B]_{kj} = t\sum_{j=1}^{n} A_{kj} + (1-t)\sum_{j=1}^{n} B_{kj} = t + 1 - t = 1.$$

Similarly,

$$\sum_{k=1}^{n} [tA + (1-t)B]_{kj} = t\sum_{k=1}^{n} A_{kj} + (1-t)\sum_{k=1}^{n} B_{kj} = t + 1 - t = 1.$$

(B.1.7) Show that a T-transform is doubly stochastic.

Answer. Suppose that $T = tI + (1-t)P_\sigma$, with $\sigma = (r\ s)$. When $h \notin \{r, s\}$ we have $Te_h = e_h$. Also,

$$Te_s = te_s + (1-t)e_r, \qquad Te_r = te_r + (1-t)e_s.$$

It follows that $T_{kj} \in \{0, 1, t, 1-t\}$. And, if we reorder the canonical basis so that e_s and e_r are the first two elements,

$$T = \begin{bmatrix} t & 1-t & \\ 1-t & t & \\ & & I_{n-2} \end{bmatrix}.$$

So T is doubly stochastic.

(B.1.8) Show that if $A, B \in \mathrm{DS}(n)$ then $AB \in \mathrm{DS}(n)$.

Answer. When all entries of A and B are non-negative, the formula for the product of matrices guarantees that all entries of AB are non-negative. The rest follows directly from Exercise B.1.5. For $(AB)e = A(Be) = Ae = e$, and $(AB)^\top e = B^\top A^\top e = B^\top e = e$.

(B.1.9) Let $A \in M_n(\mathbb{R})$ be doubly stochastic.

(i) Show that A has at least n nonzero entries.

(ii) Show that A has precisely n nonzero entries if and only if A is a permutation.

Answer. Because the entries of each row of A add to 1, this implies that each row has at least one nonzero entry. So A has at least n nonzero entries.

Now suppose that A has precisely n nonzero entries. Because each row has at least one nonzero entry, this means that each row (and column, by analogy) has precisely one nonzero entries. This also shows that all nonzero entries of A are 1. Now we proceed as follows. Let $\sigma(j)$ be the row in which column j has its entry equal to 1. The numbers $\sigma(1), \ldots, \sigma(n)$ have to be all different, for otherwise there would be a column with two nonzero entries. Thus $\sigma \in \mathbb{S}_n$, and then $A = P_\sigma$ is a permutation matrix.

(B.1.10) Show that (ii) and (i) in Proposition B.1.4 are equivalent.

Answer. (i) $\Longrightarrow$ (ii) Suppose that $x \prec y$. Since $x_j^\uparrow = x_{n-j+1}^\downarrow$,

$$\sum_{j=1}^k x_j^\uparrow = \sum_{j=1}^k x_{n-j+1}^\downarrow = \sum_{j=n-k+1}^n x_j^\downarrow = \operatorname{Tr}(x) - \sum_{j=1}^{n-k} x_j^\downarrow.$$

Thus

$$\sum_{j=1}^k x_j^\uparrow = \operatorname{Tr}(x) - \sum_{j=1}^{n-k} x_j^\downarrow \geq \operatorname{Tr}(x) - \sum_{j=1}^{n-k} y_j^\downarrow = \operatorname{Tr}(y) - \sum_{j=1}^{n-k} y_j^\downarrow = \sum_{j=1}^k y_j^\uparrow$$

(ii) $\Longrightarrow$ (i) The case $k = n$ gives us $\operatorname{Tr}(x) = \operatorname{Tr}(y)$. And $x \prec_w y$ is a given, so $x \prec y$.

(B.1.11) Show directly that (i) $\Longrightarrow$ (iv) in Proposition B.1.4

Answer. This also follows directly from Proposition B.3.4, but it is not hard to write an ad-hoc proof. Since both sums involve all terms of x and y, we may assume without loss of generality that $x = x^\downarrow$ and $y = y^\downarrow$. Fix $t \in \mathbb{R}$. We consider three cases:

- $t \leq x_n$. Then
$$\sum_{j=1}^n |x_j - t| = \sum_{j=1}^n x_j - t = \sum_{j=1}^n y_j - t \leq \sum_{j=1}^n |y_j - t|.$$
- $t \geq x_1$. In this case,
$$\sum_{j=1}^n |x_j - t| = \sum_{j=1}^n t - x_j = \sum_{j=1}^n t - y_j \leq \sum_{j=1}^n |y_j - t|.$$

- $x_k \geq t \geq x_{k+1}$. We have

$$\begin{aligned}\sum_{j=1}^{n}|x_j - t| &= \sum_{j=1}^{k}(x_j - t) + \sum_{j=k+1}^{n}(t - x_j) = (n-2k)t + \sum_{j=1}^{k}x_j - \sum_{j=k+1}^{n}x_j \\ &= (n-2k)t + \sum_{j=1}^{k}x_j - \Big(\operatorname{Tr}(x) - \sum_{j=1}^{k}x_j\Big) \\ &= (n-2k)t - \operatorname{Tr}(x) + 2\sum_{j=1}^{k}x_j \\ &\leq (n-2k)t - \operatorname{Tr}(y) + 2\sum_{j=1}^{k}y_j \\ &= \sum_{j=1}^{k}(y_j - t) + \sum_{j=k+1}^{n}(t - y_j) \\ &\leq \sum_{j=1}^{n}|y_j - t|.\end{aligned}$$

(B.1.12) Let $x, y \in \mathbb{R}^n$. Show that the following statements are equivalent:

(i) $x \prec_w y$;

(ii) $\operatorname{Tr} f(x) \leq \operatorname{Tr} f(y)$ for all f convex and non-decreasing;

(iii)

$$\sum_{j=1}^{n}(x_j - t)^+ \leq \sum_{j=1}^{n}(y_j - t)^+, \qquad t \in \mathbb{R}. \tag{B.3}$$

Answer. (i) $\implies$ (ii) Suppose that $x \prec_w y$. By Proposition B.1.7 there exists v with $x \prec v \leq y$. Then

$$\operatorname{Tr} f(x) \leq \operatorname{Tr} f(v) \leq \operatorname{Tr} f(y),$$

the first inequality by the convexity and Proposition B.1.4, and the second inequality by the monotonicity.

(ii) $\implies$ (iii) We have that $s \longmapsto (s-t)^+$ is convex and non-decreasing.

(iii) $\implies$ (i) If we take $t = \min\{x_j, y_j : j\}$ then (B.3) gives

$$\begin{aligned}\operatorname{Tr}(x) &= nt + \sum_{j=1}^n (x_j - t) = nt + \sum_{j=1}^n (x_j - t)^+ \\ &\leq nt + \sum_{j=1}^n (y_j - t)^+ = nt + \sum_{j=1}^n (y_j - t) = \operatorname{Tr}(y).\end{aligned}$$

If instead we take $t = \max\{x_j, y_j : j\}$, the same idea gives us $\operatorname{Tr}(x) \geq \operatorname{Tr}(y)$. Thus $\operatorname{Tr}(x) = \operatorname{Tr}(y)$. If we take $t = y_k$ then

$$\begin{aligned}\sum_{j=1}^k x_j - kt = \sum_{j=1}^k (x_j - t) &\leq \sum_{j=1}^k (x_j - t)^+ \\ &\leq \sum_{j=1}^n (x_j - t)^+ \leq \sum_{j=1}^n (y_j - t)^+ \\ &= \sum_{j=1}^k (y_j - t) = \sum_{j=1}^k y_j - kt.\end{aligned}$$

So $x \prec_w y$.

B.2. Some Combinatorics

(B.2.1) Prove Corollary B.2.3.

Answer. We will use Theorem B.2.1. We take $D = H = \{1, \ldots, n\}$, and

$$R = \{(k, j) : A_{kj} \neq 0\}.$$

Each diagonal without zero entries is a matching. So (i) in Corollary B.2.3 says that there is no matching, and then Theorem B.2.1 gives us indices $k_1, \ldots, k_r$ such that $|H_{k_1,\ldots,k_r}| < r$. Write $H_{k_1,\ldots,k_r} = \{j_1, \ldots, j_s\}$, with $s < r$. This means that for $j \in \{1, \ldots, n\} \setminus H_{k_1,\ldots,k_r}$ we have $A_{k_t,j} = 0$. This is a zero submatrix with r rows and $n-s$ columns. We have $r+(n-s) > r+(n-r) = n$ and (ii) in Corollary B.2.3 holds.

This process can done in reverse. If A admits a zero submatrix with r rows $k_1, \ldots, k_r$ and $n - s$ columns such that $s < r$, taking the remaining columns $j_1, \ldots, j_s$ (which are the only possible nonzero for the given rows) we get $|H_{k_1,\dots,k_r}| < r$, and by Theorem B.2.1 no matching is possible, which means that every diagonal has a zero entry.

B.3. Birkhoff's Theorem and Convex Functions

(B.3.1) A matrix $A \in M_n(\mathbb{C})$ is **doubly substochastic** if $A_{kj} \geq 0$ for all k, j and

$$\sum_{j=1}^{n} A_{kj} \leq 1, \qquad k = 1, \dots, n$$

and

$$\sum_{k=1}^{n} A_{kj} \leq 1, \qquad j = 1, \dots, n.$$

Let $A \in \text{DSS}(n)$ such that there exist k, j with $A_{kj} \in (0, 1)$. Show that A is not extreme.

Answer. The result is trivial for $n = 1$ so we assume $n \geq 2$. We we will write A as a nontrivial convex combination of two matrices by modifying a 2×2 submatrix inside A. If after modifying the submatrix the sum its the rows and columns stay the same, or they decrease while staying non-negative, the modified full matrix will still be doubly substochastic. We consider several cases:

- There exists a submatrix
$$A_0 = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$
with $B_{rs} > 0$ for $r, s = 1, 2$. Since $A_{kj} > 0$, from the doubly stochasticity we also have $B_{rs} < 1$ for $r, s = 1, 2$. We can choose $\delta = \frac{1}{2} \min\{B_{rs}, 1 - B_{rs} : \ r, s = 1, 2\}$ and form
$$A_0^+ = \begin{bmatrix} B_{11} + \delta & B_{12} - \delta \\ B_{21} - \delta & B_{22} + \delta \end{bmatrix}, \qquad A_0^- = \begin{bmatrix} B_{11} - \delta & B_{12} + \delta \\ B_{21} + \delta & B_{22} - \delta \end{bmatrix}.$$
The choice of δ guarantees that $A_0^\pm \in \text{DSS}(n)$, and $A_0 = \frac{1}{2}(A_0^+ + A_0^-)$. Replacing the submatrix A_0 in A by $A^\pm$ respectively we get matrices $A^\pm \in \text{DSS}(n)$ with $A = \frac{1}{2}(A^+ + A^-)$ and therefore A is not extreme.

- The previous case does not hold and there exists a submatrix
$$A_0 = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}$$
with $B_{11}, B_{12}, B_{22} \in (0,1)$. The failure of the previous case means that the column of A corresponding to B_{11} is zero with the only exception of B_{11}; this guarantees that if we add a small enough positive number of B_{11}, the sum of the column will still be less than 1. Then we can form, for small enough $\delta > 0$,
$$A_0^+ = \begin{bmatrix} B_{11}+\delta & B_{12}-\delta \\ 0 & B_{22}+\delta \end{bmatrix}, \qquad A_0^- = \begin{bmatrix} B_{11}-\delta & B_{12}+\delta \\ 0 & B_{22}-\delta \end{bmatrix}.$$
As before, this allows us to write $A = \frac{1}{2}(A^+ + A^-)$ for $A^\pm \in \mathrm{DSS}(n)$, showing that A is not extreme.

- Every column of A has at most one nonzero element, and there exists a row with at least two nonzero elements. So there exists a submatrix
$$A_0 = \begin{bmatrix} B_{11} & B_{12} \\ 0 & 0 \end{bmatrix}$$
with $B_{11}, B_{12} \in (0,1)$ (we know that $B_{12} < 1$ from $B_{11} > 0$). Now we take
$$A_0^+ = \begin{bmatrix} B_{11}+\delta & B_{12}-\delta \\ 0 & 0 \end{bmatrix}, \qquad A_0^- = \begin{bmatrix} B_{11}-\delta & B_{12}+\delta \\ 0 & 0 \end{bmatrix}.$$
Again this allows us to write $A = \frac{1}{2}(A^+ + A^-)$ for $A^\pm \in \mathrm{DSS}(n)$, showing that A is not extreme.

- Every 2×2 submatrix of A that contains A_{kj} is of the form
$$A_0 = \begin{bmatrix} B_{11} & 0 \\ 0 & 0 \end{bmatrix}$$
with $B_{11} \in (0,1)$. Now we take
$$A_0^+ = \begin{bmatrix} B_{11}+\delta & 0 \\ 0 & 0 \end{bmatrix}, \qquad A_0^- = \begin{bmatrix} B_{11}-\delta & 0 \\ 0 & 0 \end{bmatrix}.$$
Again this allows us to write $A = \frac{1}{2}(A^+ + A^-)$ for $A^\pm \in \mathrm{DSS}(n)$, showing that A is not extreme.

The cases above cover all possible situations where A has a nonzero entry, so A is not extreme regardless.

(B.3.2) Show that the set $\mathrm{DSS}(n)$ of all doubly substochastic $n \times n$ matrices is convex, and its extreme points are those $A \in M_n(\mathbb{C})$ with at most an entry 1 in each row, and all other entries equal to zero (in particular, the zero matrix is an extreme point of $\mathrm{DSS}(n)$).

Answer. If $t \in [0,1]$ and $A, B \in \mathrm{DSS}(n)$, then

$$\sum_{j=1}^n tA_{kj} + (1-t)B_{kj} = t\sum_{j=1}^n A_{kj} + (1-t)\sum_{j=1}^n B_{kj} \leq t + 1 - t = 1.$$

Similarly,

$$\sum_{k=1}^n tA_{kj} + (1-t)B_{kj} = t\sum_{k=1}^n A_{kj} + (1-t)\sum_{k=1}^n B_{kj} \leq t + 1 - t = 1,$$

so $\mathrm{DSS}(n)$ is convex. When A has some entry in $(0,1)$, then A is not extreme in $\mathrm{DSS}(n)$ by Exercise B.3.1. When A has at most an entry equal to 1 per row (and hence per column) and zeroes elsewhere, if $A = tB + (1-t)C$ with $t \in [0,1]$ and $B, C \in \mathrm{DSS}(n)$, we do the following. If $A_{rs} = 1$, then $1 = tB_{rs} + (1-t)C_{rs}$ forces $B_{rs} = C_{rs} = 1$. This immediately forces the rest of the r^{th} row and the s^{th} column of B, C to be zero. When a full row k of A is zero, we have $0 = tB_{kj} + (1-t)C_{kj}$ and this forces $B_{kj} = C_{kj} = 0$ for all j. Hence $B = C = A$.

(B.3.3) Show that B is doubly substochastic if and only if there exists $A \in \mathrm{DS}(n)$ with $0 \leq B_{kj} \leq A_{kj}$ for all k, j.

Answer.

If $B \in \mathrm{DS}(n)$ then we may take $A = B$. So we assume without loss of generality that $B \in \mathrm{DSS}(n) \setminus \mathrm{DS}(n)$. If B every entry of B is either 0 or 1, then there are at most $n-1$ entries equal to 1. We could list all rows and columns that have all zeros and that would form a square submatrix. By making the said submatrix equal to I we get a matrix with exactly n entries equal to 1, and such matrix is a permutation P that satisfies $B_{kj} \leq P_{kj}$ for all k, j. For arbitrary B we could try to do the same, but trying to come up with the algorithm doesn't look appealing; for every time we modify an entry both the sum of the row and the column changes, and we need to account for that. Instead, here is fairly sleek idea.

Combining Carathéodory (Theorem 7.5.12) with Exercise B.3.1 we get that $B = \sum_k t_k R_k$, where $t \prec e_1$ are convex coefficients and $R_1, \ldots, R_m$ are doubly substochastic matrices with all their entries equal to either 0 or

1. Using the previous paragraph we know that for each R_k there exists a permutation P_k with $R_k \leq P_k$ entrywise. Then we can take

$$A = \sum_{k=1}^{m} t_k P_k.$$

This is doubly stochastic and $B \leq A$ entrywise.

(B.3.4) Let $x, y \in \mathbb{C}^n$ with non-negative coordinates. Show that $x \prec_w y$ if and only if $x = By$ for some $B \in \mathrm{DSS}(n)$.

Answer. Assume first that $x \prec_w y$. By Proposition B.1.7 there exists z with $x \leq z \prec y$. Then Proposition B.1.4 gives us $A \in \mathrm{DS}(n)$ with $z = Ay$. Since $x \leq z$, for each k there exists $\alpha_k \in [0,1]$ with $x_k = \alpha_k z_k$. Let $B \in M_n(\mathbb{C})$ be given by $B_{kj} = \alpha_k A_{kj}$. Then $B \in \mathrm{DSS}(n)$ and $(By)_k = \alpha_k (Ay)_k = \alpha_k z_k = x_k$, so $x = By$.

Conversely, if $x = By$ with $B \in \mathrm{DSS}(n)$, then for any r

$$\sum_{k=1}^{r} x_k = \sum_{k=1}^{r} \sum_{j=1}^{r} B_{kj} y_j = \sum_{j=1}^{r} \Big(\sum_{k=1}^{r} B_{kj} \Big) y_j \leq \sum_{j=1}^{r} y_j,$$

so $x \prec_w y$.

APPENDIX

C

Lidskii's Theorem

C.1. Antisymmetric Tensor Products and the Determinant

(C.1.1) Let $\mathcal{H}$ be a Hilbert space and $\{\xi_k\}$ an orthonormal basis. Show that

$$\{\xi_{j_1} \wedge \cdots \wedge \xi_{j_n} : j_1 < \cdots < j_n\}$$

is an orthonormal basis for $\bigwedge^n \mathcal{H}$.

Answer. We have

$$\langle \xi_{j_1} \wedge \cdots \wedge \xi_{j_n}, \xi_{h_1} \wedge \cdots \wedge \xi_{h_n} \rangle = \frac{1}{n!} \sum_{\sigma,\sigma' \in \mathbb{S}_n} \operatorname{sgn}\sigma \operatorname{sgn}\sigma' \prod_{t=1}^{n} \langle \xi_{j_{\sigma(t)}}, \xi_{h_{\sigma'(t)}} \rangle$$

$$= \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}\sigma \prod_{t=1}^{n} \langle \xi_{j_t}, \xi_{h_{\sigma(t)}} \rangle$$

The only way any of the products can be nonzero is that $j_t = h_{\sigma(t)}$ for $t = 1, \ldots, n$. Because the j_t and the h_t are ordered, this forces $h_t = j_t$ for all

$t = 1, \ldots, n$. So the product can be nonzero only when $\sigma = \mathrm{id}$. Thus

$$\{\xi_{j_1} \wedge \cdots \wedge \xi_{j_n} : \ j_1 < \cdots < j_n\}$$

is orthonormal. And it has dense span in $\bigwedge^n(\mathcal{H})$, so it is an orthonormal basis.

(C.1.2) Let $a \in \ell^1(\mathbb{N})$. Show that

$$\sum_{j_1 < \cdots < j_k} a_{j_1} \cdots a_{j_n} \leq \frac{1}{k!}\Big(\sum_{j=1}^{\infty} a_j\Big)^k.$$

Answer. We have, as all the coefficients are non-negative,

$$\Big(\sum_{j=1}^{\infty} a_j\Big)^k = \sum_{j_1,\ldots,j_k} a_{j_1} \cdots a_{j_k} \geq k! \sum_{j_1 < \cdots < j_k} a_{j_1} \cdots a_{j_k}$$

(since each product $a_{j_1} \cdots a_{j_k}$ appears $k!$ times in the full product).

(C.1.3) Let $\mathcal{H} = \mathbb{C}^n$ and $T \in \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C})$. Show that $\dim \bigwedge^n \mathcal{H} = 1$, and that $\bigwedge^n T$ is the operator of multiplication by $\det T$.

Answer. With $e_1, \ldots, e_n$ the canonical basis, we know from Exercise C.1.1 that

$$\{e_{j_1} \wedge \cdots \wedge e_{j_n} : \ j_1 < \cdots < j_n\}$$

is an orthonormal basis for $\bigwedge^n \mathcal{H}$. The only possible choice $j_1 < \cdots < j_n$ for indices in $\{1, \ldots, n\}$ is $j_k = k$ for all k. So the orthonormal basis is $\{e_1 \wedge \cdots \wedge e_n\}$ and $\bigwedge^n \mathcal{H}$ is one-dimensional.

We have, expanding each Te_k in terms of the entries of T with respect to the canonical basis and using that the exterior products are zero when there is any repetition so we need $k_1, \ldots, k_n$ distinct,

$$\begin{aligned}
(\textstyle\bigwedge^n T)(e_1 \wedge \cdots \wedge e_n) &= Te_1 \wedge \cdots \wedge Te_n \\
&= \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} T_{k_1,1} \cdots T_{k_n,n}\, e_{k_1} \wedge \cdots \wedge e_{k_n} \\
&= \sum_{\sigma \in \mathbb{S}_n} T_{\sigma(1),1} \cdots T_{\sigma(n),n}\, e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} \\
&= \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}\sigma\, T_{\sigma(1),1} \cdots T_{\sigma(n),n}\, e_1 \wedge \cdots \wedge e_n \\
&= (\det T)\, e_1 \wedge \cdots \wedge e_n.
\end{aligned}$$

(C.1.4) Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Show that if rank $T = n$, then $\bigwedge^{n+k}(T) = 0$ for all $k \in \mathbb{N}$.

Answer. Given $\xi_1, \dots, \xi_{n+k} \in \mathcal{H}$, the set $\{T\xi_1, \dots, T\xi_{n+k}\}$ is linearly dependent by hypothesis. We can take $\{\eta_1, \dots, \eta_n\}$ to be an orthonormal basis of $T\mathcal{H}$, and so there exists coefficients $c_{j,s}$ such that

$$T\xi_j = \sum_{s=1}^{n} c_{j,s}\eta_s.$$

Then

$$\begin{aligned}\textstyle\bigwedge^{n+k} T(\xi_1 \wedge \cdots \wedge \xi_{n+k}) &= T\xi_1 \wedge \cdots \wedge T\xi_{n+k} \\ &= \sum_{s_1,\dots,s_{n+k}} \prod_{j=1}^{n+k} c_{j,s_j} \eta_{s_1} \wedge \cdots \wedge \eta_{s_{n+k}} = 0,\end{aligned}$$

for each product $\eta_{s_1} \wedge \cdots \wedge \eta_{s_{n+k}}$ contains at least a repetition.

C.2. Lidskii's Formula

APPENDIX D

The Banach–Tarski Paradox

D.1. The Construction

D.2. The Axiomatic Issue

APPENDIX E

Ultrafilters

E.1. First abstract approach: Gelfand–Naimark

E.2. Second abstract approach: the Stone Čech compactification

E.3. A more intuitive approach: Ultrafilters

(E.3.1) Let $A_0 \subset \mathbb{N}$. Show that $\mathcal{U} = \{A \subset \mathbb{N} : A \supset A_0\}$ is an ultrafilter. Show also that $\mathcal{U}$ is free if and only if A_0 is infinite.

Answer. We have

(i) $\mathbb{N} \in \mathcal{U}$ because $\mathbb{N} \supset A_0$.

(ii) Suppose that $A \in \mathcal{U}$. Then $A_0 \subset A$, so $\mathbb{N} \setminus A \subset \mathbb{N} \setminus A_0$; then $A_0 \not\subset \mathbb{N} \setminus A$ and thus $\mathbb{N} \setminus A \notin \mathcal{U}$.

(iii) Suppose that $A, B \in \mathcal{U}$. Then $A_0 \subset A$ and $A_0 \subset B$, which implies that $A_0 \subset A \cap B$; therefore $A \cap B \in \mathcal{U}$.

(iv) If $A \in \mathcal{U}$ and $A \subset B$, then $A_0 \subset A \subset B$, so $B \in \mathcal{U}$.

This shows that $\mathcal{U}$ is an ultrafilter. If A_0 is finite, let $n = 1 + \max A_0$; then $A_0 \subset \mathbb{N} \setminus \{k : k \geq n\}$, showing that $\{k : k \geq n\} \notin \mathcal{U}$ and so $\mathcal{U}$ is not free. Conversely, if A_0 is infinite then $A_0 \not\subset \{1, \dots, n-1\}$ and thus $\{k : k \geq n\} \in \mathcal{U}$ by Lemma E.3.1, and hence $\mathcal{U}$ is free.

(E.3.2) Show that $\varphi_{\mathcal{U}}$, as defined in (E.1) and extended by linearity to $\operatorname{span}\{e_k : k\}$, is well-defined.

Answer. We need to show that if

$$\sum_{k=1}^{n} \beta_k 1_{A_n} = \sum_{j=1}^{m} \gamma_j 1_{B_j}, \tag{AB.5.1}$$

then $\varphi_{\mathcal{U}}$ agrees on both. By extending with zero coefficients if needed we may assume that $\sum_k 1_{A_k} = \sum_j 1_{B_j} = 1$. The key property is that $A_k \cap B_j \in \mathcal{U}$ if and only if $A_k \in \mathcal{U}$ and $B_j \in \mathcal{U}$, which happens by definition of ultrafilter. This implies that $\varphi_{\mathcal{U}}(1_{A_k \cap B_j}) = \varphi_{\mathcal{U}}(1_{A_k})\varphi_{\mathcal{U}}(1_{B_j})$ for all k, j. From Lemma E.3.1 we know that $\varphi(1_{B_j}) = 1$ for precisely one j, so

$$\varphi(1_{A_k}) = \sum_{k=1}^{n} \varphi(1_{A_k})\varphi(1_{B_j}).$$

This, together with $\varphi_{\mathcal{U}}(1) = 1$ (because $\mathbb{N} \in \mathcal{U}$) and the fact from (AB.5.1) that $\beta_k = \gamma_j$ if $A_k \cap B_j \neq \varnothing$ gives

$$\varphi\Big(\sum_{k=1}^n \beta_k 1_{A_n}\Big) = \sum_{k=1}^n \beta_k \varphi_{\mathcal{U}}(1_{A_n}) = \sum_{k=1}^n \beta_k \sum_{j=1}^n \varphi_{\mathcal{U}}(1_{A_n})\varphi(1_{B_j})$$
$$= \sum_{j=1}^n \gamma_j \sum_{k=1}^n \varphi_{\mathcal{U}}(1_{A_n})\varphi(1_{B_j}) = \sum_{j=1}^n \gamma_j \varphi(1_{B_j})$$
$$= \varphi\Big(\sum_{j=1}^m \gamma_j 1_{B_j}\Big)$$

(E.3.3) Let $x \in \ell^\infty(\mathbb{N})$. Show that $\lim_{n\to\omega} x_n = \varphi_\omega(x)$.

Answer. Let $\alpha = \lim_{n\to\omega} x_n$. Fix $\varepsilon > 0$ and let $A_\varepsilon = \{n : \ x_n \in B_\varepsilon(\alpha)\}$. By hypothesis, $A_\varepsilon \in \omega$. Then, using that $|\varphi(y)|^2 \leq \varphi(y^*y)$ by Cauchy–Schwarz since φ is positive and unital,

$$\Big|\varphi_\omega\Big(\sum_{n\notin A_\varepsilon} x_n\, 1_{\{n\}}\Big)\Big|^2 \leq \varphi_\omega\Big(\sum_{n\notin A_\varepsilon} |x_n|^2\, 1_{\{n\}}\Big) \leq \|x\|^2\, \varphi_\omega(1_{\{n\notin A_\varepsilon\}}) = 0.$$

So, as $\varphi_\omega(1_{A_\varepsilon}) = 1$,

$$|\alpha - \varphi_\omega(x)| = \Big|\alpha - \varphi_\omega\Big(\sum_{n=1}^\infty x_n\, 1_{\{n\}}\Big)\Big| = \Big|\alpha - \varphi_\omega\Big(\sum_{n\in A_\varepsilon} x_n\, 1_{\{n\}}\Big)\Big|$$
$$= \Big|\sum_{n\in A_\varepsilon} (\alpha - x_n)\, 1_{\{n\}}\Big)\Big| \leq \varepsilon\, \varphi_\omega(1_{A_\varepsilon}) = \varepsilon.$$

As this can be done for all $\varepsilon > 0$, we have shown that $\alpha = \varphi_\omega(x)$.

(E.3.4) Let $n_0 \in \mathbb{N}$ and $\omega = \{A \subset \mathbb{N} : \ n_0 \in A\}$ the associated principal ultrafilter. Show that $\varphi_\omega(x) = x_{n_0}$.

Answer. We have $\varphi_\omega(1_{\{n_0\}}) = 1$. Then

$$\varphi_\omega(x) = \varphi_\omega(x)\, \varphi_\omega(1_{\{n_0\}}) = \varphi_\omega(x\, 1_{n_0}) = x_{n_0}\, \varphi(1_{n_0}) = x_{n_0}.$$

(E.3.5) Let $x \in \ell^\infty(\mathbb{N})$ be given by $x(n) = (-1)^n$. Show that there exist free ultrafilters ω_1 and ω_2 such that $\varphi_{\omega_1}(x) = 1$ and $\varphi_{\omega_2}(x) = -1$.

Answer. Let

$$\omega_1 = \{A \subset \mathbb{N} : \ 2\mathbb{N} \subset A\}, \qquad \omega_2 = \{A \subset \mathbb{N} : \ 2\mathbb{N} + 1 \subset A\}.$$

These are free ultrafilters by Exercise E.3.1. Let $x \in \ell^\infty(\mathbb{N})$ be given by $x(n) = (-1)^n$. We consider the states φ_{ω_1} and φ_{ω_2}. We have

$$\varphi_{\omega_1}(x) = \varphi_{\omega_1}(x)\, \varphi_{\omega_1}(1_{2\mathbb{N}}) = \varphi_{\omega_1}(x\, 1_{2\mathbb{N}}) = \varphi_{\omega_1}(1_{2\mathbb{N}}) = 1,$$

while

$$\varphi_{\omega_2}(x) = \varphi_{\omega_2}(x)\, \varphi_{\omega_2}(1_{2\mathbb{N}+1}) = \varphi_{\omega_2}(x\, 1_{2\mathbb{N}+1}) = \varphi_{\omega_2}(1_{2\mathbb{N}+1}) = -1.$$

APPENDIX

F

Unbounded Operators

(F.0.1) Prove Proposition F.0.3.

Answer. Suppose that T is closed and let $\{x_n\} \subset \mathcal{D}(T)$ be Cauchy for $\|\cdot\|_G$. Then $\{(x_n, Tx_n)\}$ is Cauchy in $\mathcal{G}(T)$; so there exists $(x, Tx) = \lim_n (x_n, Tx_n) \in \mathcal{G}(T)$. Then $\|x - x_n\|_G \to 0$ and $\mathcal{D}(T)$ is complete.

Conversely, suppose that $\mathcal{D}(T)$ is complete for $\|\cdot\|_G$ and let

$$\{(x_n, Tx_n)\} \subset \mathcal{G}(T)$$

be Cauchy. Then $\|x_n - x_m\|_G = \|x_n - x_m\| + \|Tx_n - Tx_m\|$ is Cauchy in $\mathcal{D}(T)$. By the completeness, there exists $x \in \mathcal{D}(T)$ with $\|x - x_n\|_G \to 0$. In particular, $x_n \to x$ and $Tx_n \to Tx$, so $(x_n, Tx_n) \to (x, Tx)$ in $\mathcal{G}(T)$.

(F.0.2) Show that V, defined in Proposition F.0.5 is an isometry, and $V\mathcal{H}_0^\perp = (V\mathcal{H}_0)^\perp$ for any subspace $\mathcal{H}_0 \subset \mathcal{H} \times \mathcal{K}$.

Answer.

$$\|V(\eta, \xi)\|^2 = \|(\xi, -\eta)\|^2 = \|\xi\|^2 + \|\eta\|^2 = \|(\eta, \xi)\|^2.$$

So V is an isometry. Also, if $(\rho, \gamma) \in \mathcal{H}_0^\perp$ and $(\xi, \eta) \in \mathcal{H}_0$, then

$$\begin{aligned}\langle V(\rho, \gamma), V(\xi, \eta)\rangle &= \langle(\gamma, -\rho), (\eta, -\xi)\rangle \\ &= \langle\gamma, \eta\rangle + \langle\rho, \xi\rangle = \langle(\rho, \gamma), (\xi, \eta)\rangle = 0.\end{aligned}$$

Thus $V\mathcal{H}_0^\perp \subset (V\mathcal{H}_0)^\perp$. Conversely, if $(\rho,\gamma) \in (V\mathcal{H}_0)^\perp$, this means that for all $(\xi,\eta) \in \mathcal{H}_0$,

$$\langle V(\rho,\gamma),(\xi,\eta)\rangle = \langle(\rho,\gamma),V(\xi,\eta)\rangle = 0.$$

Thus $V(\rho,\gamma) \in \mathcal{H}_0^\perp$, and therefore $(\rho,\gamma) = -V^2(\rho,\gamma) \in V\mathcal{H}_0^\perp$. Then $(V\mathcal{H}_0)^\perp \subset V\mathcal{H}_0^\perp$, which proves the equality.

(F.0.3) Show that the map T in Example F.0.1, that is $T : \mathcal{X} \to \mathcal{Y}$ given by $Tf = f$, is unbounded.

Answer. Let $g_n(t) = t^n$, $n \in \mathbb{N}$. As g_n is monotone, $\|g_n\|_{\mathcal{X}} = g_n(1) = 1$, while $\|g_n\|_{\mathcal{Y}} = g(3) = 3^n$. Hence,

$$\frac{\|Tg_n\|_{\mathcal{Y}}}{\|g_n\|_x} = 3^n,$$

showing that T is unbounded.

(F.0.4) Let $\mathcal{X} = \mathcal{Y} = C[0,1]$, with the infinity norm. Let $\mathcal{D} = C^1[0,1]$ and $T : \mathcal{D} \to \mathcal{Y}$ the operator $Tf = f'$. Show that T is unbounded and closed. If instead we consider the operator $Sf = f'$ but now $\mathcal{D}(S) = C^\infty[0,1]$, show that this operator is closable with closure T.

Answer. We know that T is unbounded by considering the usual example of $\|x^n\|_\infty = 1$ while $\|Tx^n\|_\infty = n$. Suppose that (g_n, g_n') is a Cauchy sequence in $\mathcal{G}(T)$. This means that both $\{g_n\}$ and $\{g_n'\}$ are uniformly Cauchy; this guarantees that $\lim_n g_n' = (\lim_n g_n)'$. So there exists $g \in C[0,1]$ with $g = \lim_n g_n$, and g_n is differentiable with $g' = \lim g_n'$. Then $(g,g') \in \mathcal{G}(T)$, showing that it is closed.

In the case of S, a Cauchy sequence in its graph will now be (h_n, h_n') with some $h \in C[0,1]$ such that $h = \lim_n h_n$, $h' = \lim_n h_n'$. This implies that $h \in C^1[0,1]$. The closure of $\{(h,h') : \ h \in C^\infty[0,1]\}$ in $C[0,1] \times C[0,1]$ is $\{f,f') : f \in C^1[0,1]\}$. Indeed, we have

$$\{(h,h') : \ h \in C^\infty[0,1]\} \subset \{f,f') : f \in C^1[0,1]\}.$$

And if $f \in C^1[0,1]$, let $\{h_n\} \subset C^\infty[0,1]$ with $h_n \to f'$ uniformly. Then

$$f(x) = f(0) + \int_0^x f' = \lim_n f(0) + \int_0^x h_n.$$

Thus $(f,f') = \lim_n (f(0) + \int_0^x h_n, h_n) \in \overline{\{(h,h') : \ h \in C^\infty[0,1]\}}$.

(F.0.5) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, $\mathcal{D} \subset \mathcal{X}$ a subspace, and $T : \mathcal{D} \to \mathcal{Y}$ linear. Show that the following statements are equivalent:

(i) T is closable;

(ii) $(0, y) \in \overline{\mathcal{G}(T)}$ implies $y = 0$.

Answer. Suppose that T is closable. Then $\overline{\mathcal{G}(T)}$ is the graph of a linear operator $\overline{T}$. If $(0, y) \in \mathcal{G}(\overline{T})$, then $y = \overline{T}(0) = 0$.

Conversely, suppose that $(0, 0)$ is the only element on $\overline{\mathcal{G}(T)}$ with first coordinate zero. If $(x, y), (x, z) \in \overline{\mathcal{G}(T)}$, as this is a vector space we have $(0, y - z) \in \overline{\mathcal{G}(T)}$. Then $y = z$. Hence $\overline{\mathcal{G}(T)}$ is the graph of a function $\overline{T}$. When $x \in \mathcal{D}$ we have $(x, Tx) \in \mathcal{G}(T) \subset \overline{\mathcal{G}(T)} = \mathcal{G}(\overline{T})$, so $\overline{T}x = Tx$. Given $x_1, x_1 \in \mathcal{X}$ such that there exist $y_1, y_2 \in \mathcal{Y}$ with $(x_1, y_1), (x_2, y_2) \in \mathcal{G}(\overline{T}) = \overline{\mathcal{G}(T)}$, there exist sequences $\{x_n'\}, \{x_n''\} \subset \mathcal{D}$ such that

$$x_n' \to x_1, \qquad x_n'' \to x_2, \qquad Tx_n' \to y_1, \quad Tx_n'' \to y_2.$$

Then for any $\alpha \in \mathbb{C}$

$$(\alpha x_1 + x_2, y_1 + y_2) = \lim_n (\alpha x_n' + x_n'', Tx_n' + Tx_n'') \in \overline{\mathcal{G}(T)} = \mathcal{G}(\overline{T}),$$

showing that $\overline{T}(\alpha x_1 + x_2) = \alpha \overline{T} x_1 + \overline{T} x_2$. That is $\overline{T}$ is linear and so T is closable.

(F.0.6) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, $\mathcal{D} \subset \mathcal{X}$ a subspace, and $T : \mathcal{D} \to \mathcal{Y}$ linear. Show that T is closable if and only if $\overline{\mathcal{G}(T)}$ is the graph of an operator.

Answer. Suppose that T is closable. Then $\mathcal{G}(\overline{T})$ is closed. A Cauchy sequence $\{(x_n, Tx_n)\}$ in $\mathcal{G}(T)$ is also Cauchy in $\mathcal{G}(\overline{T})$ which is closed, so $\overline{\mathcal{G}(T)} \subset \mathcal{G}(\overline{T})$. By Exercise F.0.5 we have that $\overline{\mathcal{G}(T)}$ is the graph of an operator.

Conversely, if $\overline{\mathcal{G}(T)}$ is the graph of an operator T', we can take $\overline{T} = T'$. As $\{(x, Tx) : x \in \mathcal{D}(T)\} \subset \mathcal{G}(\overline{T})$, we have that $\mathcal{D}(T) \subset \mathcal{D}(\overline{T})$ and that $\overline{T}|_{\mathcal{D}(T)} = T$.

(F.0.7) Let $\mathcal{D}(T) \subset \mathcal{H}$ be dense, and $T : \mathcal{D}(T) \to \mathcal{K}$ linear. Show that T^* is closed.

Answer. Let $\{\eta_n\} \subset \mathcal{D}(T^*)$ such that $(\eta_n, T^*\eta_n) \subset \mathcal{G}(T^*)$ is Cauchy. This means that there exist $\eta \in \mathcal{K}$ and $\nu \in \mathcal{H}$ such that $\eta_n \to \eta$ and $T^*\eta_n \to \nu$.

We have, for any $\xi \in \mathcal{H}$ and since $\eta_n \in \mathcal{D}(T^*)$,

$$\langle \xi, \nu \rangle = \lim_n \langle \xi, T^* \eta_n \rangle = \lim_n T \langle \xi, \eta_n \rangle = \langle T\xi, \eta \rangle.$$

This shows that γ_η is bounded, for $\gamma_\eta(\xi) = \langle \xi, \nu \rangle$, and that $\nu = T^*\eta$. So T^* is closed.

(F.0.8) Let $\mathcal{X} = L^2[0,1]$ with its dense subspace $\mathcal{D} = C[0,1]$. Define $T : \mathcal{D} \to \mathcal{D}$ by $Tf = f(0)$. Show that this operator is unbounded. Also, consider the functions $f_n = (1 - nt)\, 1_{[0,\frac{1}{n}]}$ and show that $f_n \in C[0,1]$ for all n, $\|f_n\|_2 \to 0$, and $Tf_n = 1$ for all n, so $\mathcal{G}(T)$ is not closed; conclude that T is not closed and that it is not even closable.

Answer. If we show that T is not closed, it cannot be bounded.

If $f_n = (1-nt)\, 1_{[0,\frac{1}{n}]}(t)$, we have $f_n(1/n) = 0$ so f_n is continuous. Also

$$\|f_n\|_2^2 = \int_0^{1/n} (1-nt)^2 \, dt = \frac{1}{3n},$$

showing that $f_n \to 0$. Meanwhile, $Tf_n = f_n(0) = 1$ for all n. Hence $(f_n, Tf_n) \to (0,1)$; this point cannot be in the graph of any linear operator, so T is not closable and in particular it is not closed.

(F.0.9) Give an example of a Banach space $\mathcal{X}$, subspaces $\mathcal{D}, M \subset \mathcal{X}$, and an idempotent $E : \mathcal{D} \to M$ which is unbounded.

Answer. Let $\mathcal{X} = c_0$, $\mathcal{D} = c_{00}$, and $M = \mathbb{C}\, e_1$. Let

$$Ex = \Big(\sum_{k=1}^\infty x_k, 0, 0, \dots \Big).$$

Then $E\mathcal{D} = M$, and if $x = \sum_{k=1}^n e_k$ then $\|x\| = 1$ and $\|Ex\| = n$; as this can be done for all n, E is unbounded.

(F.0.10) Let $p \in [1, \infty)$, $\mathcal{X} = \mathcal{Y} = \ell^p(\mathbb{N})$, $\mathcal{D} = c_{00}$, and $T : \mathcal{D} \to \mathcal{X}$ given by

$$T\Big(\sum_{k=1}^n c_k e_k \Big) = \Big(\sum_{k=1}^n k\, c_k \Big)\, e_1 + \sum_{k=2}^n c_k e_k.$$

Decide if T is closable. Find T^*.

Answer. We have $Te_k = ke_1$ for all k, so T is unbounded. We have $(0, e_1) = \lim_k(k^{-1}e_k, T(k^{-1}e_k))$, so T is not closable.

As for the adjoint, if $y \in \ell^q(\mathbb{N})$ and $\gamma_y(x) = \langle Tx, y\rangle$ is bounded, there exists $c > 0$ such that

$$\begin{aligned} c = c\|e_n\|_p \geq |\langle Te_n, y\rangle| &= \left|\sum_k (Te_n)_k y_k\right| \\ &\geq \left|y_1\, n\right| - \left|\sum_{k\geq 2}(e_n)_k\, y_k\right| = n|y_1| - |y_n| \\ &\geq n|y_1| - \|y\|_q. \end{aligned}$$

Thus $y \in \mathcal{D}(T^*)$ if and only if $y_1 = 0$. We have, for such y,

$$\langle Tx, y\rangle = \sum_{k\geq 2} x_k y_k = \langle x, y\rangle,$$

so $T^*y = y$.

(F.0.11) Let $T : \mathcal{D} \to C[0,1]$ be the map $Tf = f(0)$, where $\mathcal{D} = C[0,1] \subset L^2[0,1]$. Show that $\mathcal{D}_* \neq \{0\}$, but $T^* = 0$.

Answer.

$$\gamma_g(f) = \langle Tf, g\rangle = \langle f(0), g\rangle = f(0)\int_{[0,1]} \bar{g}.$$

The only way this can be bounded is if $\int_0^1 g = 0$. Hence

$$\mathcal{D}_* = \left\{g \in L^2[0,1] : \int_{[0,1]} g = 0\right\}$$

and, for any $g \in \mathcal{D}_*$,

$$\langle T^*g, h\rangle = \langle g, Th\rangle = \overline{h(0)}\int_{[0,1]} g = 0.$$

Hence $T^* = 0$.

(F.0.12) Let T be as in Example F.0.1. Show that $\mathcal{D}_* = \{0\}$.

Answer. If $g \in \mathcal{D}_*$ and $f \in \mathbb{C}[x]$, let $h_n \in C[0,3]$ with $h = 1$ on $[0,1]$ and $h = ng$ on $[2,3]$ (this is achieved by simply joining the points $(1,1)$ and $(2, ng(2))$ with a straight segment). By Stone–Weierstrass (Theorem 7.4.20) there exists $f_n \in \mathbb{C}[x]$ with $\|f - h_n\|_\infty < 1/n$ on $[0,3]$. As an element of

$C[0,1]$ we have $\|f_n\| \leq \|h_n\| + 1/n = 1 + 1/n < 2$, while

$$\gamma_g(f_n) = \langle Tf_n, g\rangle = \int_{[2,3]} n|g|^2 = n\|g\|_2^2.$$

So γ is unbounded unless $g = 0$. Thus $\mathcal{D}_* = 0$.

(F.0.13) Let $\mathcal{H} = \mathcal{K} = \ell^2(\mathbb{N})$, $\mathcal{D} = c_{00}$, and $T : \mathcal{D} \to \mathcal{K}$ the linear operator induced by $Te_{p^n} = e_p$ for each $p \in \mathbb{N}$ prime and $n \in \mathbb{N}$. Find $\mathcal{D}(T^*)$.

Answer. Let $y \in \mathcal{D}(T^*)$. We have, for all n

$$(T^*y)_{p^n} = \langle e_{p^n}, T^*y\rangle = \langle Te_{p^n}, y\rangle = y_p.$$

This means that if $y_p \neq 0$, then T^*y has to have infinitely many entries equal to y_p; this prevents it from being in $\ell^2(\mathbb{N})$. Thus $\mathcal{D}(T^*) = \{y : \ y_p = 0 \ \text{ if } p \text{ prime}\}$.

(F.0.14) Let $\mathcal{D} \subset \mathcal{H}$ be dense and $T : \mathcal{D} \to \mathcal{K}$ linear. Show that the following statements are equivalent:

(i) $\mathcal{D}(T^*) = \{0\}$;

(ii) $\overline{\mathcal{G}(T)} = \mathcal{H} \times \mathcal{K}$.

Answer. We get a direct proof from (F.3). Since $\mathcal{G}(T^*) = V\mathcal{G}(T)^\perp$, if $\mathcal{D}(T^*) = \{0\}$ then $\mathcal{G}(T^*) = \{(0,0)\}$ and it follows after applying V that $\mathcal{G}(T)^\perp = \{(0,0)\}$; hence $\overline{\mathcal{G}(T)} = \mathcal{H} \times \mathcal{K}$. Conversely, if $\mathcal{G}(T)$ is dense then $\mathcal{G}(T)^\perp = \{(0,0)\}$, and therefore $\mathcal{G}(T^*) = \{(0,0)\}$; as $(\eta, T^*\eta) \in \mathcal{G}(T^*)$ for all $\eta \in \mathcal{D}(T^*)$, we conclude that $\mathcal{D}(T^*) = \{0\}$.

Below we show another proof without using (F.3), though the ideas are not really different.

(i) $\implies$ (ii) If $(\nu, \eta) \in \overline{\mathcal{G}(T)}^\perp$ and nonzero, this means that for all $\xi \in \mathcal{D}(T)$,

$$0 = \langle \xi, \nu\rangle + \langle T\xi, \eta\rangle.$$

This gives us

$$\gamma_\eta(\xi) = \langle T\xi, \eta\rangle = -\langle \xi, \nu\rangle,$$

so γ_η is bounded since $\mathcal{D}(T)$ is dense. Note that we cannot have $\eta = 0$, for in that case we get $\langle \xi, \nu\rangle = 0$ for ξ in a dense set, and so $\nu = 0$ contradicting that (ν, η) was nonzero. Hence $\mathcal{D}(T^*) \neq \{0\}$.

(ii) $\implies$ (i) Suppose that $\mathcal{D}(T^*) \neq \{0\}$. Let $\eta \in \mathcal{D}(T^*)$ be nonzero. Then

$$\langle T\xi, \eta \rangle = \langle \xi, T^*\eta \rangle, \qquad \xi \in \mathcal{D}(T).$$

We can read the above equality as saying that $(-T^*\eta, \eta) \in \mathcal{G}(T)^\perp$. So $\overline{\mathcal{G}(T)} \neq \mathcal{H} \times \mathcal{K}$.

(F.0.15) Let $g \in L^\infty(\mathbb{R})$, and such that $\int_\mathbb{R} |g|^2 = \infty$. Fix $h_0 \in L^2(\mathbb{R})$. Let $T : \mathcal{D}(T) \to L^2(\mathbb{R})$ be given by

$$Tf = \langle f, g \rangle h_0,$$

where

$$\mathcal{D}(T) = \left\{ f \in L^2(\mathbb{R}) : \int_\mathbb{R} |fg| < \infty \right\}.$$

Show that T is densely defined, and find T^*.

Answer. Since g is bounded, for any measurable $E \subset \mathbb{R}$ with finite measure,

$$\int_\mathbb{R} |g\,1_E| \leq \|g\|_\infty\, m(E) < \infty.$$

So $\mathcal{D}(T)$ contains all integrable simple functions and hence it is dense in $L^2(\mathbb{R})$ (this can be seen by combining Proposition 2.8.14 and Theorem 2.4.13).

To find $\mathcal{D}(T^*)$, if $h \in \mathcal{D}(T^*)$ and $f \in \mathcal{D}(T)$, we have

$$\langle f, T^*h \rangle = \langle Tf, h \rangle = \langle f, g \rangle \, \langle h_0, h \rangle = \langle f, \langle h, h_0 \rangle \, g \rangle.$$

Thus

$$T^*h = \langle h, h_0 \rangle \, g.$$

Even though the way that adjoint was defined guarantees it, it might not be obvious at first sight that T^* maps into $L^2(\mathbb{R})$. But from $h \in \mathcal{D}(T^*)$ we know that the linear functional $\gamma_h : f \longmapsto \langle f, \langle h, h_0 \rangle \, g \rangle$ is bounded. This implies, via Proposition 5.6.8, that $\langle h, h_0 \rangle \, g \in L^2(\mathbb{R})$. Hence

$$\mathcal{D}(T^*) = \left\{ h \in L^2(\mathbb{R}) : \int_\mathbb{R} |hh_0| < \infty \right\}.$$

(F.0.16) For an unbounded operator $T : \mathcal{D}(T) \to \mathcal{Y}$, where $\mathcal{D}(T) \subset \mathcal{X}$ is dense and $\mathcal{X}, \mathcal{Y}$ are normed spaces, write a definition for T^*, and explore how much of the results in the text can be made to work. Reflexivity might be needed in some cases.

Answer. We can mimic exactly the criterion in (F.1). Indeed, let

$$\mathcal{D}(T^*) = \{\varphi \in \mathcal{Y}^* : \ \gamma_\varphi \ \text{ is bounded}\},$$

where $\gamma_\varphi(x) = \varphi(Tx)$. When $\mathcal{D}(T)$ is dense γ_φ is defined everywhere for $\varphi \in \mathcal{D}(T^*)$, so we can define $T^*\varphi = \gamma_\varphi \in \mathcal{X}^*$. That is, for $\varphi \in \mathcal{D}(T^*)$ and $x \in \mathcal{X}$, we have $(T^*\varphi)x = \varphi(Tx)$.

In analogy with (F.3) we have, when T is densely defined,

$$\mathcal{G}(T^*) = V\mathcal{G}(T)^o,$$

where $V : \mathcal{X}^* \times \mathcal{Y}^* \to \mathcal{Y}^* \times \mathcal{X}^*$ is the isometry $V(\varphi, \psi) = (\psi, -\varphi)$. Indeed, if $(\varphi, T^*\varphi) \in \mathcal{G}(T^*)$, then

$$V^{-1}(\varphi, T^*\varphi)(x, Tx) = (-T^*\varphi, \varphi)(x, Tx) = -(T^*\varphi)x + \varphi(Tx) = 0$$

since $\varphi \in \mathcal{D}(T^*)$ (the dual of the direct sum was considered in Proposition 5.6.5). This shows that $V^{-1}\mathcal{G}(T^*) \subset \mathcal{G}(T)^o$. Conversely, if $(\psi, \varphi) \in \mathcal{G}(T)^o$, we have $\psi(x) + \varphi(Tx) = 0$ for all $x \in \mathcal{X}$; then $\varphi \in \mathcal{D}(T^*)$ and $\psi = -T^*\varphi$. Thus $\mathcal{G}(T)^o \subset V^{-1}\mathcal{G}(T^*)$ and therefore $\mathcal{G}(T^*) = V\mathcal{G}(T)^o$.

Next we show that when $\mathcal{X}, \mathcal{Y}$ are reflexive T is closable if and only if T^* is densely defined. Suppose that T is not closable. Then there exists $z \in \mathcal{Y}$ such that $(0, z) \in \overline{\mathcal{G}(T)}$. This means that there exists a sequence $\{x_n\} \subset \mathcal{D}(T)$ with $x_n \to 0$ and $Tx_n \to z$. Then, for any $\varphi \in \mathcal{D}(T^*)$,

$$\varphi(z) = \lim_n \varphi(Tx_n) = \lim_n (T^*\varphi)x_n = 0,$$

since $T^*\varphi \in \mathcal{X}^*$. We cannot have $\mathcal{D}(T^*)$ dense in $\mathcal{Y}^*$, for in such case we would get $\varphi(z) = 0$ for all $\varphi \in \mathcal{Y}^*$, contradicting that $z \neq 0$ (via Corollary 5.7.7). Conversely, if $\mathcal{D}(T^*)$ is not dense, there exists nonzero $\Phi \in \mathcal{D}(T^*)^o \subset \mathcal{Y}^{**}$. So $(\Phi, 0) \in \mathcal{G}(T^*)^o$; then $(0, \Phi) \in V^{-1}\mathcal{G}(T^*)^o$, meaning that $V^{-1}\mathcal{G}(T^*)^o$ is not the graph of an operator. But

$$V^{-1}\mathcal{G}(T^*)^o = V^{-1}(V\mathcal{G}(T)^o)^o = \mathcal{G}(T)^{oo},$$

since V preserves polars (this is straightforward to check). Here is where we use the reflexivity of $\mathcal{X}$ and $\mathcal{Y}$. We have that $\mathcal{G}(T)^{oo} = \overline{J\mathcal{G}(T)}^{w^*}$, using that $\mathcal{G}(T)$ is a subspace and Exercise 7.3.5. Here $J = J_\mathcal{X} \times J_\mathcal{Y}$ is the canonical embedding $\mathcal{X} \times \mathcal{Y} \hookrightarrow \mathcal{X}^{**} \times \mathcal{Y}^{**}$. By the reflexivity, the weak*-closure agrees with the weak closure, and being a subspace we end up with the norm closure. Thus $\mathcal{G}(T)^{oo} = \overline{\mathcal{G}(T)}$, implying that T is not closable.

(F.0.17) Continuing from Exercise F.0.16, show that if T is closable and $\mathcal{X}, \mathcal{Y}$ are reflexive, then

$$\overline{T} = T^{**}.$$

This requires showing first that because T is closable then T^* is densely defined, which hopefully was done in Exercise F.0.16.

Answer. We did show in Exercise F.0.16 that $\mathcal{D}(T^*)$ is dense when T is closable. For $x \in \mathcal{D}(T)$ and $\psi \in \mathcal{D}(T^*)$,

$$(T^{**}J_{\mathcal{X}}x)\psi = (T^{**}\hat{x})\psi = \hat{x}(T^*\psi) = (T^*\psi)x = \psi(Tx) = (J_{\mathcal{Y}}Tx)\psi.$$

Hence, as $\mathcal{D}(T^*)$ is dense, on $\mathcal{D}(T)$ we have $T = J_{\mathcal{Y}}^{-1}T^{**}J_{\mathcal{X}}$. As T^{**} is closed, so is $J_{\mathcal{Y}}^{-1}T^{**}J_{\mathcal{X}}$; this means that $J_{\mathcal{Y}}^{-1}T^{**}J_{\mathcal{X}}$ is a closed operator that agrees with T on $\mathcal{D}(T)$, and so $J_{\mathcal{Y}}^{-1}T^{**}J_{\mathcal{X}} = \overline{T}$.

(F.0.18) Let $\mathcal{X} = \mathcal{Y} = c_0$, and $T : \mathcal{X} \to \mathcal{Y}$ given by

$$Tx = (x_1, 2x_2, 3x_3, \ldots),$$

with $\mathcal{D}(T) = c_{00}$. Decide if T is closed or not, and find T^*.

Answer. We have

$$\mathcal{G}(T) = \{(x, Tx) : \ x \in c_{00}\}.$$

Intuitively, we can extend the domain of T a bit, as there are full nonzero sequences where T makes sense; for instance, $T(1/k^2) = (1/k)$. If P_n is the projection onto the first n coordinates, we have $P_n x \in \mathcal{D}(T)$ for all $x \in c_0$. And $TP_n(1/k^2) = P_n(1/k) \xrightarrow[n]{} (1/k)$. So $((1/k^2), (1/k)) \in \overline{\mathcal{G}(T)}$ and T is not closed. It is still densely defined, though, so T^* exists.

The domain of T^* is

$$\mathcal{D}(T^*) = \{z \in \ell^1(\mathbb{N}) : \ \eta_z \ \text{ is bounded}\},$$

where $\eta_z(x) = \langle Tx, z\rangle$. Since

$$\langle Tx, z\rangle = \sum_n nx(n)z(n)$$

needs to be bounded for all $x \in c_0$, we need $Tz \in \ell^1(\mathbb{N})$. Thus

$$\mathcal{D}(T^*) = \{z \in \ell^1(\mathbb{N}) : \ (nz(n))_n \in \ell^1(\mathbb{N})\}.$$

And it is dense, as it contains all finitely supported sequences. As a formula, T^* is the same as T, now with domain $\mathcal{D}(T^*)$.

(F.0.19) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{D}(T) \to \mathcal{Y}$ linear and densely defined. Show that $\mathcal{D}(T^*) = \mathcal{Y}^*$ if and only if T is bounded.

Answer. Assume first that $\mathcal{D}(T^*) = \mathcal{Y}^*$. We have that for each $\varphi \in \mathcal{Y}^*$ there exists $c_\varphi > 0$ such that

$$|\varphi(Tx)| \leq c_\varphi \, \|x\|, \qquad \varphi \in \mathcal{Y}^*, \; x \in \mathcal{D}(T).$$

We can read this inequality as saying that if $\|x\| \leq 1$.

$$\sup_{x \in \mathcal{D}(T) \cap \overline{B_1^{\mathcal{X}}(0)}} |\widehat{Tx}\varphi| \leq c_\varphi.$$

By the Uniform Boundedness Principle (Theorem 6.3.16), applied on the Banach space $\mathcal{X}^{**}$, there exists $c > 0$ such that $\|Tx\| = \|\widehat{Tx}\| \leq c$ for all x with $\|x\| \leq 1$. Thus T is bounded.

For the converse, if T is bounded then for all $\varphi \in \mathcal{Y}^*$ we have $|\varphi(Tx)| \leq \|\varphi\| \, \|T\| \, \|x\|$, so γ_φ is bounded and $\varphi \in \mathcal{D}(T^*)$.