Solutions to Axler's Measure, Integration & Real Analysis

Aayush Bajaj

2026-08-29T12:30:00+10:00

Contents

Solutions to every exercise in Sheldon Axler's *Measure, Integration & Real Analysis* (Springer GTM 282, 2020) — 587 exercises across sections 1A–12. The book is open access at measure.axler.net and is filed at Measure, Integration & Real Analysis (Axler).

1 Riemann Integration

Contents

1.1 Exercises 1A

Problem 1A.1. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function such that

$$L(f, P, [a, b]) = U(f, P, [a, b])$$

for some partition P of $[a, b]$. Prove that f is a constant function on $[a, b]$.

Solution. Write P as $a = x_0 < x_1 < \cdots < x_n = b$. By 1.3,

$$U(f, P, [a, b]) - L(f, P, [a, b]) = \sum_{j=1}^{n} (x_j - x_{j-1}) \Big(\sup_{[x_{j-1}, x_j]} f - \inf_{[x_{j-1}, x_j]} f \Big),$$

and the hypothesis says this equals 0. Every summand is nonnegative, so every summand vanishes; since $x_j - x_{j-1} > 0$, this forces

$$\sup_{[x_{j-1}, x_j]} f = \inf_{[x_{j-1}, x_j]} f \quad \text{for } j = 1, \ldots, n.$$

Hence f is constant on each $[x_{j-1}, x_j]$, say with value c_j. Consecutive subintervals share the point x_j, so $c_j = f(x_j) = c_{j+1}$ for $1 \le j \le n - 1$, and all the c_j equal a common value c. The subintervals cover $[a, b]$, so f is constant on $[a, b]$.

Problem 1A.2. Suppose $a \le s < t \le b$. Define $f : [a, b] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & \text{if } s < x < t, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that f is Riemann integrable on $[a, b]$ and that $\int_a^b f = t - s$.

Solution. Squeeze with the partition P_ε of $[a, b]$ formed by the distinct elements of $\{a,\ s,\ s+\varepsilon,\ t-\varepsilon,\ t,\ b\}$ in increasing order, where $0 < \varepsilon < \frac{t-s}{2}$. Its subintervals are $[a, s]$ (present only if $a < s$), $[s, s + \varepsilon]$, $[s + \varepsilon, t - \varepsilon]$, $[t - \varepsilon, t]$, and $[t, b]$ (present only if $t < b$). Now $f \equiv 1$ on $[s + \varepsilon, t - \varepsilon]$, while $f \equiv 0$ on $[a, s]$ and on $[t, b]$, and $0 \le f \le 1$ throughout; hence

$$L(f, P_\varepsilon, [a, b]) \ge (t - \varepsilon) - (s + \varepsilon) = (t - s) - 2\varepsilon,$$
$$U(f, P_\varepsilon, [a, b]) \le \varepsilon + \big((t - \varepsilon) - (s + \varepsilon)\big) + \varepsilon = t - s.$$

Letting $\varepsilon \to 0$ in the first estimate and using 1.7 and 1.8,

$$t - s \le L(f, [a, b]) \le U(f, [a, b]) \le t - s.$$

By 1.9, f is Riemann integrable on $[a, b]$ with $\int_a^b f = t - s$.

Problem 1A.3. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. Prove that f is Riemann integrable if and only if for each $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that

$$U(f, P, [a, b]) - L(f, P, [a, b]) < \varepsilon.$$

Solution. Both directions run off 1.7, which gives $L(f, P, [a, b]) \le L(f, [a, b])$ and $U(f, [a, b]) \le U(f, P, [a, b])$ for every partition P of $[a, b]$.
(i) The ε-condition implies integrability. Given $\varepsilon > 0$, choose P as in the statement; then with 1.8,

$$0 \le U(f, [a, b]) - L(f, [a, b]) \le U(f, P, [a, b]) - L(f, P, [a, b]) < \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $L(f,[a,b]) = U(f,[a,b])$, so f is Riemann integrable by 1.9.
(ii) Integrability implies the ε-condition. Let $\varepsilon > 0$. Since $L(f,[a,b])$ is a supremum and $U(f,[a,b])$ an infimum over partitions (1.7), there are partitions P_1, P_2 of $[a,b]$ with

$$L(f,P_1,[a,b]) > L(f,[a,b]) - \tfrac{\varepsilon}{2}, \qquad U(f,P_2,[a,b]) < U(f,[a,b]) + \tfrac{\varepsilon}{2}.$$

Let P merge the lists defining P_1 and P_2; each of those lists is a sublist of P, so 1.5 gives $L(f,P,[a,b]) \geq L(f,P_1,[a,b])$ and $U(f,P,[a,b]) \leq U(f,P_2,[a,b])$. Hence, using $L(f,[a,b]) = U(f,[a,b])$,

$$\begin{aligned} U(f,P,[a,b]) - L(f,P,[a,b]) &\leq U(f,P_2,[a,b]) - L(f,P_1,[a,b]) \\ &< \Big(U(f,[a,b]) + \tfrac{\varepsilon}{2}\Big) - \Big(L(f,[a,b]) - \tfrac{\varepsilon}{2}\Big) = \varepsilon. \end{aligned}$$

Problem 1A.4. Suppose $f, g : [a,b] \to \mathbb{R}$ are Riemann integrable. Prove that $f+g$ is Riemann integrable on $[a,b]$ and

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

Solution. Write $L(h,P)$ and $U(h,P)$ for $L(h,P,[a,b])$ and $U(h,P,[a,b])$. For every nonempty $A \subseteq [a,b]$,

$$\inf_A f + \inf_A g \leq \inf_A (f+g) \leq \sup_A (f+g) \leq \sup_A f + \sup_A g,$$

since $\inf_A f + \inf_A g$ is a lower bound and $\sup_A f + \sup_A g$ an upper bound for the values of $f+g$ on A. Applying this with $A = [x_{j-1}, x_j]$ for each subinterval of a partition P, multiplying by $x_j - x_{j-1} > 0$ and summing gives, by 1.3,

$$L(f,P) + L(g,P) \leq L(f+g,P) \leq U(f+g,P) \leq U(f,P) + U(g,P). \tag{$*$}$$

Let $\varepsilon > 0$. By Exercise 3 in this section there are partitions P_1, P_2 of $[a,b]$ with $U(f,P_1) - L(f,P_1) < \frac{\varepsilon}{2}$ and $U(g,P_2) - L(g,P_2) < \frac{\varepsilon}{2}$; let P merge their lists, so by 1.5 both estimates persist at P. Then $(*)$ gives

$$U(f+g,P) - L(f+g,P) \leq \big(U(f,P) - L(f,P)\big) + \big(U(g,P) - L(g,P)\big) < \varepsilon,$$

so $f+g$ is Riemann integrable by Exercise 3 again. For its value, keep that P and use $(*)$ together with $L(f,P) \leq L(f,[a,b]) = \int_a^b f$ and $U(f,P) \geq U(f,[a,b]) = \int_a^b f$ from 1.7 and 1.9 (and the same for g):

$$\begin{aligned} \int_a^b (f+g) \leq U(f+g,P) &\leq U(f,P) + U(g,P) \\ &< \Big(L(f,P) + \tfrac{\varepsilon}{2}\Big) + \Big(L(g,P) + \tfrac{\varepsilon}{2}\Big) \\ &\leq \int_a^b f + \int_a^b g + \varepsilon, \end{aligned}$$

and symmetrically $\int_a^b (f+g) \geq L(f+g,P) \geq L(f,P) + L(g,P) > \int_a^b f + \int_a^b g - \varepsilon$. As $\varepsilon > 0$ was arbitrary, $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

Problem 1A.5. Suppose $f : [a,b] \to \mathbb{R}$ is Riemann integrable. Prove that the function $-f$ is Riemann integrable on $[a,b]$ and

$$\int_a^b (-f) = -\int_a^b f.$$

Solution. Negation swaps infima and suprema: for every nonempty $A \subseteq [a,b]$,

$$\inf_A (-f) = -\sup_A f, \qquad \sup_A (-f) = -\inf_A f,$$

because α is an upper bound for $\{f(x) : x \in A\}$ exactly when $-\alpha$ is a lower bound for $\{-f(x) : x \in A\}$, and conversely. Applying this with $A = [x_{j-1}, x_j]$ on each subinterval of a partition P of $[a, b]$ gives, by 1.3,

$$L(-f, P, [a, b]) = -U(f, P, [a, b]), \qquad U(-f, P, [a, b]) = -L(f, P, [a, b]).$$

The same swap, applied now to the set of upper Riemann sums of f and to the set of lower ones, turns 1.7 into

$$L(-f, [a, b]) = -U(f, [a, b]), \qquad U(-f, [a, b]) = -L(f, [a, b]).$$

Since f is Riemann integrable, $L(f, [a, b]) = U(f, [a, b]) = \int_a^b f$, so both left sides equal $-\int_a^b f$. By 1.9, $-f$ is Riemann integrable on $[a, b]$ with $\int_a^b(-f) = -\int_a^b f$.

Problem 1A.6. Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable. Suppose $g : [a, b] \to \mathbb{R}$ is a function such that $g(x) = f(x)$ for all except finitely many $x \in [a, b]$. Prove that g is Riemann integrable on $[a, b]$ and

$$\int_a^b g = \int_a^b f.$$

Solution. By Exercise 4 in this section it suffices to prove that $h = g - f$, which vanishes off a finite set, is Riemann integrable with $\int_a^b h = 0$; then $g = f + h$ is Riemann integrable with $\int_a^b g = \int_a^b f$.
Let $c_1 < \cdots < c_m$ be the points where $h \neq 0$ (if there are none, every Riemann sum of h is 0 and we are done), and put $M = \max\{|h(c_1)|, \ldots, |h(c_m)|\} > 0$, so $|h| \le M$ on $[a, b]$. For $\delta > 0$ let P_δ be the partition of $[a, b]$ whose points are the distinct elements of

$$\{a, b\} \cup \big(\{c_i - \delta,\ c_i + \delta : i = 1, \ldots, m\} \cap (a, b)\big)$$

in increasing order. If a subinterval $[u, w]$ of P_δ contains some c_i, then $[u, w] \subseteq [c_i - \delta, c_i + \delta]$: otherwise $c_i - \delta$ or $c_i + \delta$ would lie in $(u, w) \subseteq (a, b)$ and hence be a point of P_δ strictly between the consecutive points u and w.
On a subinterval containing no c_i we have $h \equiv 0$, so it contributes 0 to both $L(h, P_\delta, [a, b])$ and $U(h, P_\delta, [a, b])$. For each i at most two subintervals contain c_i, and they have disjoint interiors and lie in $[c_i - \delta, c_i + \delta]$, so their lengths sum to at most 2δ; hence the subintervals meeting $\{c_1, \ldots, c_m\}$ have total length at most $2m\delta$, and there $|\sup h| \le M$, $|\inf h| \le M$. Therefore

$$|L(h, P_\delta, [a, b])| \le 2m\delta M \quad \text{and} \quad |U(h, P_\delta, [a, b])| \le 2m\delta M.$$

Given $\varepsilon > 0$, take $\delta = \frac{\varepsilon}{2mM}$; then 1.7 gives $U(h, [a, b]) \le \varepsilon$ and $L(h, [a, b]) \ge -\varepsilon$. As $\varepsilon > 0$ was arbitrary and $L(h, [a, b]) \le U(h, [a, b])$ by 1.8, both are 0. By 1.9, h is Riemann integrable with $\int_a^b h = 0$.

Problem 1A.7. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. For $n \in \mathbb{Z}^+$, let P_n denote the partition that divides $[a, b]$ into 2^n intervals of equal size. Prove that

$$L(f, [a, b]) = \lim_{n\to\infty} L(f, P_n, [a, b]) \quad \text{and} \quad U(f, [a, b]) = \lim_{n\to\infty} U(f, P_n, [a, b]).$$

Solution. Put $M = \sup_{[a,b]} |f| < \infty$, and let P_n have the points $x_j^{(n)} = a + \frac{j(b-a)}{2^n}$ for $j = 0, \ldots, 2^n$, so each of its subintervals has length $\frac{b-a}{2^n}$.
Refinement estimate: if every subinterval of a partition Q has length at most ℓ and Q' is obtained by adjoining p points to the list defining Q, then

$$L(f, Q', [a, b]) \le L(f, Q, [a, b]) + 2pM\ell.$$

For $p = 1$ the adjoined point v either lies in Q already, when the two sums agree, or lies interior to one subinterval $[u,w]$ of Q, when all other terms agree and, using $|\inf_A f| \le M$,

$$(v-u)\inf_{[u,v]} f + (w-v)\inf_{[v,w]} f - (w-u)\inf_{[u,w]} f \le 2M(w-u) \le 2M\ell.$$

For general p, adjoin the points one at a time; each intermediate partition refines Q, so its subintervals still have length at most ℓ.
Since $x_j^{(n)} = x_{2j}^{(n+1)}$, the list defining P_n is a sublist of that defining P_{n+1}, so 1.5 makes $\big(L(f,P_n,[a,b])\big)_n$ increasing while 1.7 bounds it above by $L(f,[a,b])$; hence it converges, with limit at most $L(f,[a,b])$.
For the reverse inequality let $\varepsilon > 0$ and use 1.7 to choose a partition P, say $a = y_0 < \cdots < y_k = b$, with $L(f,P,[a,b]) > L(f,[a,b]) - \varepsilon$. Let Q_n merge the lists of P_n and P; this adjoins at most $k-1$ points to P_n, whose subintervals have length $\frac{b-a}{2^n}$, so the refinement estimate together with $L(f,P,[a,b]) \le L(f,Q_n,[a,b])$ from 1.5 gives

$$L(f,P,[a,b]) \le L(f,P_n,[a,b]) + \frac{2(k-1)M(b-a)}{2^n}.$$

Letting $n \to \infty$ gives $\lim_{n\to\infty} L(f,P_n,[a,b]) \ge L(f,P,[a,b]) > L(f,[a,b]) - \varepsilon$; as $\varepsilon > 0$ was arbitrary, $\lim_{n\to\infty} L(f,P_n,[a,b]) = L(f,[a,b])$.
For the upper integral apply this to the bounded function $-f$, then use the identities $L(-f,P,[a,b]) = -U(f,P,[a,b])$ and $L(-f,[a,b]) = -U(f,[a,b])$ from Exercise 5 in this section, whose proofs use only boundedness:

$$U(f,[a,b]) = -L(-f,[a,b]) = -\lim_{n\to\infty} L(-f,P_n,[a,b]) = \lim_{n\to\infty} U(f,P_n,[a,b]).$$

Problem 1A.8. Suppose $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. Prove that

$$\int_a^b f = \lim_{n\to\infty} \frac{b-a}{n} \sum_{j=1}^{n} f\Big(a + \tfrac{j(b-a)}{n}\Big).$$

Solution. Squeeze the sum between the Riemann sums of the equally spaced partition. Let P_n have points $x_j = a + \frac{j(b-a)}{n}$ for $j = 0,\dots,n$, and write $R_n = \frac{b-a}{n}\sum_{j=1}^{n} f(x_j)$. Since $x_j \in [x_{j-1},x_j]$, multiplying $\inf_{[x_{j-1},x_j]} f \le f(x_j) \le \sup_{[x_{j-1},x_j]} f$ by $x_j - x_{j-1} = \frac{b-a}{n} > 0$ and summing gives

$$L(f,P_n,[a,b]) \le R_n \le U(f,P_n,[a,b]).$$

Both outer terms converge to $\int_a^b f$. Put $M = \sup_{[a,b]} |f| < \infty$ (finite since f is Riemann integrable, hence bounded), let $\varepsilon > 0$, and use 1.7 to choose a partition P, say $a = y_0 < \cdots < y_k = b$, with $L(f,P,[a,b]) > L(f,[a,b]) - \varepsilon$. Merging P into P_n adjoins at most $k-1$ points to a partition of mesh $\frac{b-a}{n}$, so the refinement estimate proved in the solution to Exercise 7 in this section, together with 1.5, gives

$$L(f,[a,b]) - \varepsilon < L(f,P,[a,b]) \le L(f,P_n,[a,b]) + \frac{2(k-1)M(b-a)}{n},$$

while $L(f,P_n,[a,b]) \le L(f,[a,b])$ by 1.7. Hence $L(f,P_n,[a,b])$ lies within 2ε of $L(f,[a,b])$ for all large n, and $\varepsilon > 0$ was arbitrary, so $L(f,P_n,[a,b]) \to L(f,[a,b])$. Applying this to $-f$ and using $L(-f,P,[a,b]) = -U(f,P,[a,b])$ and $L(-f,[a,b]) = -U(f,[a,b])$ from Exercise 5 in this section gives $U(f,P_n,[a,b]) \to U(f,[a,b])$.
Since f is Riemann integrable, $L(f,[a,b]) = U(f,[a,b]) = \int_a^b f$ by 1.9, so the squeeze yields $R_n \to \int_a^b f$.

Problem 1A.9. Suppose $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. Prove that if $c,d \in \mathbb{R}$ and $a \le c < d \le b$, then f is Riemann integrable on $[c,d]$.
[To say that f is Riemann integrable on $[c,d]$ means that f with its domain restricted to $[c,d]$ is Riemann integrable.]

Solution. Cut a good partition of $[a, b]$ at c and d and keep the middle piece; Exercise 3 in this section does the rest.
Let $\varepsilon > 0$. By Exercise 3 there is a partition P of $[a, b]$ with $U(f, P, [a, b]) - L(f, P, [a, b]) < \varepsilon$. Let P', say $a = x_0 < \cdots < x_n = b$, be obtained by adjoining c and d to the list defining P; since that list is a sublist of the new one, 1.5 gives

$$U(f, P', [a, b]) - L(f, P', [a, b]) \leq U(f, P, [a, b]) - L(f, P, [a, b]) < \varepsilon.$$

Choose $p < q$ with $x_p = c$ and $x_q = d$, and let Q be the partition $x_p, x_{p+1}, \ldots, x_q$ of $[c, d]$. Each subinterval $[x_{j-1}, x_j]$ with $p < j \leq q$ lies inside $[c, d]$, so restricting f to $[c, d]$ changes none of the suprema and infima below, and every omitted term is nonnegative:

$$\begin{aligned} U(f, Q, [c, d]) - L(f, Q, [c, d]) &= \sum_{j=p+1}^{q} (x_j - x_{j-1}) \Big(\sup_{[x_{j-1}, x_j]} f - \inf_{[x_{j-1}, x_j]} f \Big) \\ &\leq \sum_{j=1}^{n} (x_j - x_{j-1}) \Big(\sup_{[x_{j-1}, x_j]} f - \inf_{[x_{j-1}, x_j]} f \Big) \\ &= U(f, P', [a, b]) - L(f, P', [a, b]) < \varepsilon. \end{aligned}$$

By Exercise 3 again, f is Riemann integrable on $[c, d]$.

Problem 1A.10. Suppose $f\colon [a, b] \to \mathbb{R}$ is a bounded function and $c \in (a, b)$. Prove that f is Riemann integrable on $[a, b]$ if and only if f is Riemann integrable on $[a, c]$ and f is Riemann integrable on $[c, b]$. Furthermore, prove that if these conditions hold, then

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Solution. Everything comes from the splitting identity. If P, say $x_0 < \cdots < x_n$, is a partition of $[a, b]$ with $c = x_p$ for some $0 < p < n$, let $P_1 : x_0, \ldots, x_p$ and $P_2 : x_p, \ldots, x_n$ be the induced partitions of $[a, c]$ and $[c, b]$. Each subinterval of P lies wholly in $[a, c]$ or wholly in $[c, b]$, so restricting f changes none of the infima and suprema, and splitting the sums of 1.3 at $j = p$ gives

$$\begin{aligned} L(f, P, [a, b]) &= L(f, P_1, [a, c]) + L(f, P_2, [c, b]), \\ U(f, P, [a, b]) &= U(f, P_1, [a, c]) + U(f, P_2, [c, b]). \end{aligned} \tag{$*$}$$

Conversely, any partition P_1 of $[a, c]$ and P_2 of $[c, b]$ concatenate (at the shared endpoint c) to such a P, and $(*)$ again holds.
Lower integrals add. For $\geq$: given P_1, P_2 with concatenation P, $(*)$ and 1.7 give $L(f, P_1, [a, c]) + L(f, P_2, [c, b]) \leq L(f, [a, b])$; take suprema over P_1 and over P_2, which vary independently. For $\leq$: given any partition P of $[a, b]$, adjoin c to get P' and apply 1.5, then $(*)$, then 1.7,

$$L(f, P, [a, b]) \leq L(f, P', [a, b]) = L(f, P_1', [a, c]) + L(f, P_2', [c, b]) \leq L(f, [a, c]) + L(f, [c, b]),$$

and take the supremum over P. The mirror-image argument gives the same for upper integrals, so

$$L(f, [a, b]) = L(f, [a, c]) + L(f, [c, b]), \qquad U(f, [a, b]) = U(f, [a, c]) + U(f, [c, b]).$$

Subtracting,

$$U(f, [a, b]) - L(f, [a, b]) = \Big(U(f, [a, c]) - L(f, [a, c]) \Big) + \Big(U(f, [c, b]) - L(f, [c, b]) \Big),$$

where by 1.8 all three differences are nonnegative; so the left side vanishes exactly when both terms on the right do, which by 1.9 is precisely the asserted equivalence. Granting it, 1.9 and the additivity of the lower integral give

$$\int_a^b f = L(f, [a, b]) = L(f, [a, c]) + L(f, [c, b]) = \int_a^c f + \int_c^b f.$$

Problem 1A.11. Suppose $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. Define $F\colon [a,b] \to \mathbb{R}$ by

$$F(t) = \begin{cases} 0 & \text{if } t = a, \\ \displaystyle\int_a^t f & \text{if } t \in (a,b]. \end{cases}$$

Prove that F is continuous on $[a,b]$.

Solution. F is Lipschitz with constant $M = \sup_{[a,b]} |f| < \infty$, hence uniformly continuous on $[a,b]$. Let $a \le s < t \le b$. By Exercise 9 in this section, f is Riemann integrable on $[a,t]$ and on $[s,t]$, so F is well defined and

$$F(t) - F(s) = \int_s^t f :$$

for $s = a$ this is $F(t) = \int_a^t f$ with $F(a) = 0$, and for $s > a$ it is Exercise 10 in this section applied on $[a,t]$ at the interior point s. Since $-M \le \inf_{[s,t]} f$ and $\sup_{[s,t]} f \le M$, applying 1.13 on $[s,t]$ gives

$$|F(t) - F(s)| = \left| \int_s^t f \right| \le M(t-s).$$

Thus $|F(t) - F(s)| \le M|t-s|$ for all $s,t \in [a,b]$. Given $\varepsilon > 0$, any s,t with $|s-t| < \frac{\varepsilon}{M+1}$ then satisfy $|F(t) - F(s)| \le \frac{M}{M+1}\varepsilon < \varepsilon$, so F is continuous on $[a,b]$.

Problem 1A.12. Suppose $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. Prove that $|f|$ is Riemann integrable and that

$$\left| \int_a^b f \right| \le \int_a^b |f|.$$

Solution. Taking absolute values cannot increase oscillation: for nonempty $A \subseteq [a,b]$ and $x, y \in A$,

$$|f(x)| - |f(y)| \le |f(x) - f(y)| = \max\{f(x) - f(y),\ f(y) - f(x)\} \le \sup_A f - \inf_A f,$$

and taking the supremum over x and over y gives $\sup_A |f| - \inf_A |f| \le \sup_A f - \inf_A f$. Applying this with $A = [x_{j-1}, x_j]$ on each subinterval of a partition P of $[a,b]$, multiplying by $x_j - x_{j-1} > 0$ and summing gives, by 1.3,

$$U(|f|, P, [a,b]) - L(|f|, P, [a,b]) \le U(f, P, [a,b]) - L(f, P, [a,b]).$$

Given $\varepsilon > 0$, Exercise 3 in this section supplies P making the right side less than ε, hence the left side too; by Exercise 3 again, $|f|$ is Riemann integrable.
For the inequality, note first that Riemann integration is monotone: if $g \le h$ are Riemann integrable on $[a,b]$, then $\inf_{[x_{j-1},x_j]} g \le \inf_{[x_{j-1},x_j]} h$ on every subinterval, so $L(g, P, [a,b]) \le L(h, P, [a,b])$ for every P, and taking suprema, by 1.7 and 1.9,

$$\int_a^b g = L(g, [a,b]) \le L(h, [a,b]) = \int_a^b h.$$

By Exercise 5 in this section $-|f|$ is Riemann integrable with $\int_a^b (-|f|) = -\int_a^b |f|$, so applying monotonicity to $-|f| \le f$ and to $f \le |f|$,

$$-\int_a^b |f| \le \int_a^b f \le \int_a^b |f|, \quad \text{that is,} \quad \left| \int_a^b f \right| \le \int_a^b |f|.$$

Problem 1A.13. Suppose $f\colon [a,b] \to \mathbb{R}$ is an increasing function, meaning that $c, d \in [a,b]$ with $c < d$ implies $f(c) \le f(d)$. Prove that f is Riemann integrable on $[a,b]$.

Solution. On an equally spaced partition the oscillation sum telescopes. Monotonicity gives $f(a) \le f \le f(b)$, so f is bounded. Fix $n \in \mathbb{Z}^+$ and let P_n be the partition with points $x_j = a + \frac{j(b-a)}{n}$, $j = 0, \dots, n$. On $[x_{j-1}, x_j]$ monotonicity gives $f(x_{j-1}) \le f \le f(x_j)$, and both bounds are attained, so

$$\inf_{[x_{j-1},x_j]} f = f(x_{j-1}) \qquad \text{and} \qquad \sup_{[x_{j-1},x_j]} f = f(x_j).$$

Hence, by 1.3 and telescoping,

$$U(f, P_n, [a,b]) - L(f, P_n, [a,b]) = \frac{b-a}{n} \sum_{j=1}^{n} \bigl(f(x_j) - f(x_{j-1})\bigr) = \frac{(b-a)\bigl(f(b) - f(a)\bigr)}{n}.$$

Given $\varepsilon > 0$, the Archimedean property supplies n making the right side less than ε, so f is Riemann integrable on $[a,b]$ by Exercise 3 in this section.

Problem 1A.14. Suppose $f_1, f_2, \dots$ is a sequence of Riemann integrable functions on $[a,b]$ such that $f_1, f_2, \dots$ converges uniformly on $[a,b]$ to a function $f\colon [a,b] \to \mathbb{R}$. Prove that f is Riemann integrable and

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

Solution. Put $\varepsilon_n = \sup_{[a,b]} |f_n - f|$; uniform convergence says exactly that $\varepsilon_n \to 0$. Choosing N with $\varepsilon_N \le 1$ and using that f_N is bounded (being Riemann integrable) gives $|f| \le 1 + \sup_{[a,b]} |f_N| < \infty$, so f is bounded and every ε_n is finite.
On a nonempty $A \subseteq [a,b]$ we have $f_n - \varepsilon_n \le f \le f_n + \varepsilon_n$, so $\sup_A f \le \sup_A f_n + \varepsilon_n$ and $\inf_A f \ge \inf_A f_n - \varepsilon_n$; subtracting and applying the result with $A = [x_{j-1}, x_j]$ on the subintervals of a partition P of $[a,b]$, whose lengths total $b-a$, gives by 1.3

$$U(f, P, [a,b]) - L(f, P, [a,b]) \le \Bigl(U(f_n, P, [a,b]) - L(f_n, P, [a,b])\Bigr) + 2(b-a)\varepsilon_n. \tag{$\star$}$$

Let $\eta > 0$. Choose n with $\varepsilon_n < \frac{\eta}{4(b-a)}$, then use Exercise 3 in this section to choose a partition P with $U(f_n, P, [a,b]) - L(f_n, P, [a,b]) < \frac{\eta}{2}$; now ($\star$) gives $U(f, P, [a,b]) - L(f, P, [a,b]) < \eta$, so f is Riemann integrable by Exercise 3 again.
Hence Exercises 4, 5 and 12 in this section make $f - f_n$ and $|f - f_n|$ Riemann integrable, and

$$\left| \int_a^b f - \int_a^b f_n \right| = \left| \int_a^b (f - f_n) \right| \le \int_a^b |f - f_n| \le (b-a)\varepsilon_n,$$

the last step by 1.13 applied to $|f - f_n|$. Since $\varepsilon_n \to 0$, $\int_a^b f = \lim_{n\to\infty} \int_a^b f_n$.

1.2 Exercises 1B

Problem 1B.1. Define $f\colon [0,1] \to \mathbb{R}$ as follows:

$$f(a) = \begin{cases} 0 & \text{if } a \text{ is irrational,} \\ \dfrac{1}{n} & \text{if } a \text{ is rational and } n \text{ is the smallest positive} \\ & \text{integer such that } a = \frac{m}{n} \text{ for some integer } m. \end{cases}$$

Show that f is Riemann integrable and compute $\int_0^1 f$.

Solution. f is Riemann integrable with $\int_0^1 f = 0$.
Every subinterval of positive length contains an irrational, where f vanishes, and $f \ge 0$; so every lower Riemann sum of f is 0 and $L(f,[0,1]) = 0$ by 1.7.
For the upper integral, note $0 \le f \le 1$ and that for each $N \in \mathbb{Z}^+$ the set $A_N = \{a \in [0,1] : f(a) \ge \frac{1}{N}\}$ is finite: such an a is rational, its least denominator n satisfies $\frac{1}{n} = f(a) \ge \frac{1}{N}$, and $a = \frac{m}{n} \in [0,1]$ forces $0 \le m \le n$, so

$$A_N \subseteq \left\{ \tfrac{m}{n} : n \in \{1,\dots,N\},\ m \in \{0,1,\dots,n\} \right\}.$$

Write K for the number of its elements ($K \ge 1$, as $0 \in A_N$); off A_N we have $f < \frac{1}{N}$.
Let $\varepsilon > 0$. Choose N with $\frac{1}{N} < \frac{\varepsilon}{2}$, then $n > \frac{4K}{\varepsilon}$, and let P_n split $[0,1]$ into the n equal subintervals I_j of length $\frac{1}{n}$. Each point of A_N lies in at most two of the I_j, so at most $2K$ indices j have $I_j \cap A_N \ne \emptyset$; for those $\sup_{I_j} f \le 1$, and for the others $\sup_{I_j} f \le \frac{1}{N}$. Since the lengths of the others total at most 1,

$$U(f, P_n, [0,1]) \le \frac{2K}{n} + \frac{1}{N} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

So $U(f,[0,1]) < \varepsilon$ by 1.7 for every $\varepsilon > 0$, giving $U(f,[0,1]) \le 0$; with $L(f,[0,1]) = 0$ and 1.8 this forces $L(f,[0,1]) = U(f,[0,1]) = 0$, so 1.9 gives the claim.

Problem 1B.2. Suppose $f\colon [a,b] \to \mathbb{R}$ is a bounded function. Prove that f is Riemann integrable if and only if

$$L(-f,[a,b]) = -L(f,[a,b]).$$

Solution. Everything follows from the identity

$$U(f,[a,b]) = -L(-f,[a,b]). \tag{$*$}$$

For a nonempty $E \subseteq [a,b]$ we have $\sup_E f = -\inf_E(-f)$, since α is an upper bound for the values of f on E exactly when $-\alpha$ is a lower bound for the values of $-f$ on E. Applying this on each subinterval of a partition P of $[a,b]$ gives, by 1.3,

$$U(f,P,[a,b]) = -L(-f,P,[a,b]),$$

and the same swap, applied now to the (nonempty, bounded) set of these numbers, turns 1.7 into

$$U(f,[a,b]) = \inf_P U(f,P,[a,b]) = -\sup_P L(-f,P,[a,b]) = -L(-f,[a,b]),$$

which is $(*)$. By 1.9, f is Riemann integrable if and only if $L(f,[a,b]) = U(f,[a,b])$, which by $(*)$ reads $L(f,[a,b]) = -L(-f,[a,b])$, that is, $L(-f,[a,b]) = -L(f,[a,b])$.

Problem 1B.3. Suppose $f, g\colon [a,b] \to \mathbb{R}$ are bounded functions. Prove that

$$L(f,[a,b]) + L(g,[a,b]) \le L(f+g,[a,b])$$

and

$$U(f+g,[a,b]) \le U(f,[a,b]) + U(g,[a,b]).$$

Solution. Both inequalities come from the single-partition versions plus a common refinement. For nonempty $E \subseteq [a,b]$, $\inf_E f + \inf_E g$ is a lower bound and $\sup_E f + \sup_E g$ an upper bound for the values of $f+g$ on E, so

$$\inf_E f + \inf_E g \le \inf_E (f+g), \qquad \sup_E (f+g) \le \sup_E f + \sup_E g.$$

Applying these with $E = [x_{j-1}, x_j]$ on the subintervals of a partition P of $[a,b]$, multiplying by $x_j - x_{j-1} > 0$ and summing gives, by 1.3,

$$L(f,P,[a,b]) + L(g,P,[a,b]) \le L(f+g,P,[a,b]), \tag{1}$$

$$U(f+g,P,[a,b]) \le U(f,P,[a,b]) + U(g,P,[a,b]). \tag{2}$$

Let $\varepsilon > 0$. By 1.7 there are partitions P_1, P_2 of $[a,b]$ with $L(f,P_1,[a,b]) > L(f,[a,b]) - \frac{\varepsilon}{2}$ and $L(g,P_2,[a,b]) > L(g,[a,b]) - \frac{\varepsilon}{2}$; let P merge their lists, so that 1.5 gives $L(f,P,[a,b]) \ge L(f,P_1,[a,b])$ and $L(g,P,[a,b]) \ge L(g,P_2,[a,b])$. (The refinement is what lets the two suprema be approached along one partition.) Then, using 1.7 and (1),

$$\begin{aligned} L(f+g,[a,b]) &\ge L(f+g,P,[a,b]) \ge L(f,P,[a,b]) + L(g,P,[a,b]) \\ &\ge L(f,P_1,[a,b]) + L(g,P_2,[a,b]) > L(f,[a,b]) + L(g,[a,b]) - \varepsilon, \end{aligned}$$

and $\varepsilon > 0$ was arbitrary. The mirror-image argument, with partitions Q_1, Q_2 chosen by 1.7 so that $U(f,Q_1,[a,b]) < U(f,[a,b]) + \frac{\varepsilon}{2}$ and $U(g,Q_2,[a,b]) < U(g,[a,b]) + \frac{\varepsilon}{2}$, their merge Q, 1.5 and (2), gives $U(f+g,[a,b]) \le U(f,[a,b]) + U(g,[a,b])$.

Problem 1B.4. Give an example of bounded functions $f, g\colon [0,1] \to \mathbb{R}$ such that

$$L(f,[0,1]) + L(g,[0,1]) < L(f+g,[0,1])$$

and

$$U(f+g,[0,1]) < U(f,[0,1]) + U(g,[0,1]).$$

Solution. Take f to be the function of 1.14 and $g = 1 - f$:

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational,} \end{cases} \qquad g(x) = \begin{cases} 0 & \text{if } x \text{ is rational,} \\ 1 & \text{if } x \text{ is irrational.} \end{cases}$$

Every $[c,d] \subseteq [0,1]$ with $c < d$ contains a rational and an irrational, so $\inf_{[c,d]} f = \inf_{[c,d]} g = 0$ and $\sup_{[c,d]} f = \sup_{[c,d]} g = 1$. Since the subinterval lengths of a partition of $[0,1]$ total 1, every lower sum of f and of g is 0 and every upper sum is 1, so 1.7 gives

$$L(f,[0,1]) = L(g,[0,1]) = 0, \qquad U(f,[0,1]) = U(g,[0,1]) = 1.$$

But $f+g$ is the constant function 1, so all its lower and upper sums equal 1 and $L(f+g,[0,1]) = U(f+g,[0,1]) = 1$. Hence

$$\begin{aligned} L(f,[0,1]) + L(g,[0,1]) = 0 < 1 = L(f+g,[0,1]), \\ U(f+g,[0,1]) = 1 < 2 = U(f,[0,1]) + U(g,[0,1]). \end{aligned}$$

Problem 1B.5. Give an example of a sequence of continuous real-valued functions $f_1, f_2, \dots$ on $[0,1]$ and a continuous real-valued function f on $[0,1]$ such that

$$f(x) = \lim_{k\to\infty} f_k(x)$$

for each $x \in [0,1]$ but

$$\int_0^1 f \neq \lim_{k\to\infty} \int_0^1 f_k.$$

Solution. Take $f = 0$ and let f_k be the triangular spike of height $2k$ and area 1 on $\left[0, \frac{1}{k}\right]$:

$$f_k(x) = \begin{cases} 4k^2 x & \text{if } 0 \le x \le \frac{1}{2k}, \\ 4k^2\left(\frac{1}{k} - x\right) & \text{if } \frac{1}{2k} \le x \le \frac{1}{k}, \\ 0 & \text{if } \frac{1}{k} \le x \le 1. \end{cases}$$

The formulas agree where the cases overlap, both giving $2k$ at $x = \frac{1}{2k}$ and 0 at $x = \frac{1}{k}$, so each f_k is continuous and hence Riemann integrable by 1.11, as is the constant f.
Pointwise convergence: $f_k(0) = 0$ for every k, and for $x > 0$ we have $f_k(x) = 0$ as soon as $\frac{1}{k} < x$, which happens for all large k by the Archimedean property. So $f_k(x) \to 0 = f(x)$ for every $x \in [0,1]$.
The integrals: $\int_0^1 f = 0$ by 1.13. For f_k, split $[0,1]$ at $\frac{1}{k}$ and $\left[0, \frac{1}{k}\right]$ at $\frac{1}{2k}$ using Exercises 9 and 10 in Section 1A (for $k = 1$ only the split at $\frac{1}{2}$ is needed), and integrate the two linear pieces by the fundamental theorem of calculus, with antiderivatives $2k^2x^2$ and $-2k^2\left(\frac{1}{k} - x\right)^2$:

$$\int_0^1 f_k = 2k^2\left(\frac{1}{2k}\right)^2 + \left[-2k^2\left(\frac{1}{k} - x\right)^2\right]_{x=1/(2k)}^{x=1/k} + 0 = \tfrac{1}{2} + \tfrac{1}{2} = 1.$$

Hence $\int_0^1 f = 0 \neq 1 = \lim_{k\to\infty} \int_0^1 f_k$.

2 Measures

Contents

2.1 Exercises 2A

Problem 2A.1. Prove that if A and B are subsets of $\mathbb{R}$ and $|B| = 0$, then $|A \cup B| = |A|$.

Solution. Two inequalities. Since $A \subseteq A \cup B$, order preservation (2.5) gives $|A| \le |A \cup B|$. In the other direction, countable subadditivity (2.8) applied to the sequence $A, B, \emptyset, \emptyset, \ldots$, whose union is $A \cup B$, gives

$$|A \cup B| \le |A| + |B| = |A| + 0 = |A|,$$

using $|\emptyset| = 0$ from 2.3. Hence $|A \cup B| = |A|$.

Problem 2A.2. Suppose $A \subseteq \mathbb{R}$ and $t \in \mathbb{R}$. Let $tA = \{ta : a \in A\}$. Prove that $|tA| = |t|\,|A|$. [Assume that $0 \cdot \infty$ is defined to be 0.]

Solution. (i) $t = 0$. Then tA is $\emptyset$ or $\{0\}$, so $|tA| = 0$ by 2.3, while $|t|\,|A| = 0 \cdot |A| = 0$ by the stated convention.
(ii) $t \neq 0$. Dilation carries open intervals to open intervals with

$$\ell(tI) = |t|\,\ell(I),$$

since $t(\alpha, \beta)$ is $(t\alpha, t\beta)$ or $(t\beta, t\alpha)$ according to the sign of t, of length $|t|(\beta - \alpha)$, while $\emptyset$ maps to $\emptyset$ and an unbounded open interval maps to an unbounded one, both sides then being $\infty = |t| \cdot \infty$. So if $I_1, I_2, \ldots$ are open intervals with $A \subseteq \bigcup_{k=1}^{\infty} I_k$, then $tA \subseteq \bigcup_{k=1}^{\infty} tI_k$ (as $x \mapsto tx$ is a bijection of $\mathbb{R}$), and 2.2 gives

$$|tA| \le \sum_{k=1}^{\infty} \ell(tI_k) = |t| \sum_{k=1}^{\infty} \ell(I_k).$$

Taking the infimum over all such covers yields $|tA| \le |t|\,|A|$. Applying this inequality with $\frac{1}{t}$ in place of t and tA in place of A, and using $\frac{1}{t}(tA) = A$, gives $|A| \le \frac{1}{|t|}|tA|$, that is, $|t|\,|A| \le |tA|$. Hence $|tA| = |t|\,|A|$.

Problem 2A.3. Prove that if $A, B \subseteq \mathbb{R}$ and $|A| < \infty$, then $|B \setminus A| \ge |B| - |A|$.

Solution. Since $B \subseteq (B \setminus A) \cup A$, order preservation (2.5) and subadditivity (2.8, applied to $B \setminus A$, A, $\emptyset$, $\emptyset, \ldots$) give

$$|B| \le \bigl|(B \setminus A) \cup A\bigr| \le |B \setminus A| + |A|.$$

If $|B \setminus A| = \infty$ the desired inequality is trivial. Otherwise both terms on the right are finite, since $|A| < \infty$ by hypothesis, so subtracting $|A|$ is legitimate and gives $|B| - |A| \le |B \setminus A|$.

Problem 2A.4. Suppose F is a subset of $\mathbb{R}$ with the property that every open cover of F has a finite subcover. Prove that F is closed and bounded.

Solution. Two open covers, chosen so that a finite subcover collapses to a single set.
(i) F is bounded. The intervals $(-k, k)$, $k \in \mathbb{Z}^+$, are open and their union is $\mathbb{R} \supseteq F$, so they form an open cover of F in the sense of 2.10. A finite subcover gives $k_1, \ldots, k_n$ with $F \subseteq (-k_1, k_1) \cup \cdots \cup (-k_n, k_n) = (-K, K)$, where $K = \max\{k_1, \ldots, k_n\}$, the union collapsing because these intervals increase with k. (An empty subcover would make $F = \emptyset$, which is bounded.) So $|x| < K$ for all $x \in F$.
(ii) F is closed. Let $b \in \mathbb{R} \setminus F$ and put

$$G_k = \Bigl(-\infty,\, b - \tfrac{1}{k}\Bigr) \cup \Bigl(b + \tfrac{1}{k},\, \infty\Bigr) = \Bigl\{x \in \mathbb{R} : |x - b| > \tfrac{1}{k}\Bigr\},$$

an open set, with $G_1 \subseteq G_2 \subseteq \cdots$ and $\bigcup_{k=1}^{\infty} G_k = \mathbb{R} \setminus \{b\} \supseteq F$. So $\{G_k\}$ is an open cover of F, and a finite subcover collapses, as in (i), to $F \subseteq G_K$ for some K (an empty subcover again forcing $F = \emptyset$, which is closed). Thus

$$\left(b - \tfrac{1}{K},\, b + \tfrac{1}{K}\right) \subseteq \mathbb{R} \setminus F.$$

Hence $\mathbb{R} \setminus F$ is open, so F is closed.

Problem 2A.5. Suppose $\mathcal{A}$ is a set of closed subsets of $\mathbb{R}$ such that $\bigcap_{F \in \mathcal{A}} F = \emptyset$. Prove that if $\mathcal{A}$ contains at least one bounded set, then there exist $n \in \mathbb{Z}^+$ and $F_1, \ldots, F_n \in \mathcal{A}$ such that $F_1 \cap \cdots \cap F_n = \emptyset$.

Solution. Apply Heine–Borel to a bounded $F_0 \in \mathcal{A}$, covered by the complements of the members of $\mathcal{A}$. If $F_0 = \emptyset$, take $n = 1$ and $F_1 = F_0$. So assume $F_0 \neq \emptyset$; being a member of $\mathcal{A}$ it is closed, and it is bounded by choice. Put $\mathcal{C} = \{\mathbb{R} \setminus F : F \in \mathcal{A}\}$, a collection of open sets. It covers F_0: each $x \in F_0$ fails to lie in some $F \in \mathcal{A}$, since $\bigcap_{F \in \mathcal{A}} F = \emptyset$, and then $x \in \mathbb{R} \setminus F \in \mathcal{C}$.
By the Heine–Borel Theorem (2.12) applied to the closed bounded set F_0, there are $F_1, \ldots, F_m \in \mathcal{A}$ with

$$F_0 \subseteq (\mathbb{R} \setminus F_1) \cup \cdots \cup (\mathbb{R} \setminus F_m) = \mathbb{R} \setminus (F_1 \cap \cdots \cap F_m)$$

by De Morgan (the subcover is nonempty because $F_0 \neq \emptyset$). That says exactly $F_0 \cap F_1 \cap \cdots \cap F_m = \emptyset$, so the $m + 1$ sets $F_0, F_1, \ldots, F_m$ of $\mathcal{A}$ are as required.

Problem 2A.6. Prove that if $a, b \in \mathbb{R}$ and $a < b$, then

$$|(a, b)| = |[a, b)| = |(a, b]| = b - a.$$

Solution. Each of the three sets becomes $[a, b]$ after adjoining a finite set, which has outer measure 0 by 2.3, so Exercise 1 in this section and $|[a, b]| = b - a$ from 2.14 give

$$\begin{aligned}
|(a, b)| &= \big|(a, b) \cup \{a, b\}\big| = |[a, b]| = b - a, \\
|[a, b)| &= \big|[a, b) \cup \{b\}\big| = |[a, b]| = b - a, \\
|(a, b]| &= \big|(a, b] \cup \{a\}\big| = |[a, b]| = b - a.
\end{aligned}$$

Method (2): for $0 < \varepsilon < \frac{b-a}{2}$ we have $[a + \varepsilon, b - \varepsilon] \subseteq (a, b) \subseteq [a, b]$, so 2.5 and 2.14 give

$$b - a - 2\varepsilon = \big|[a + \varepsilon, b - \varepsilon]\big| \leq |(a, b)| \leq |[a, b]| = b - a,$$

and letting $\varepsilon \to 0$ gives $|(a, b)| = b - a$; then 2.5 squeezes $|[a, b)|$ and $|(a, b]|$ between $|(a, b)|$ and $|[a, b]|$.

Problem 2A.7. Suppose a, b, c, d are real numbers with $a < b$ and $c < d$. Prove that

$$|(a, b) \cup (c, d)| = (b - a) + (d - c) \quad \text{if and only if} \quad (a, b) \cap (c, d) = \emptyset.$$

Solution. Write $E = (a, b) \cup (c, d)$; the inequality $|E| \leq (b - a) + (d - c)$ always holds, by subadditivity (2.8) and $|(a, b)| = b - a$, $|(c, d)| = d - c$ from Exercise 6 of this section. So only the equality case is at issue. Since x lies in both intervals exactly when $u < x < v$, where

$$u = \max\{a, c\}, \qquad v = \min\{b, d\},$$

we have $(a, b) \cap (c, d) = (u, v)$, nonempty precisely when $u < v$, i.e. (given $a < b$ and $c < d$) precisely when $c < b$ and $a < d$.
(i) *Disjoint case.* Then $b \leq c$ or $d \leq a$; interchanging the two intervals changes neither E nor $(b - a) + (d - c)$, so assume $b \leq c$. Splitting (a, d) at b and at c gives

$$(a, d) \subseteq (a, b) \cup [b, c] \cup (c, d) = E \cup [b, c]$$

(Check!), and $|[b, c]| = c - b$ by 2.14 if $b < c$, by 2.3 if $b = c$. Since $a < d$, Exercise 6 gives $|(a, d)| = d - a$, so 2.5 and 2.8 yield

$$d - a = |(a, d)| \le |E| + (c - b).$$

All terms are finite, so $|E| \ge (b-a) + (d-c)$; with the upper bound, equality.
(ii) *Overlapping case.* Put $p = \min\{a, c\}$ and $q = \max\{b, d\}$. Because the two intervals meet, their union is the single interval $E = (p, q)$ (Check!), so $|E| = q - p$ by Exercise 6. As $\{p, u\} = \{a, c\}$ and $\{v, q\} = \{b, d\}$ as multisets,

$$(b-a) + (d-c) = (v+q) - (p+u) = (q-p) + (v-u) = |E| + (v-u),$$

and $v - u > 0$ with $|E|$ finite, so $|E| < (b-a) + (d-c)$.

Problem 2A.8. Prove that if $A \subseteq \mathbb{R}$ and $t > 0$, then

$$|A| = |A \cap (-t,t)| + \bigl|A \cap \bigl(\mathbb{R} \setminus (-t,t)\bigr)\bigr|.$$

Solution. With $B = A \cap (-t,t)$ and $C = A \setminus (-t,t)$, the inequality $|A| \le |B| + |C|$ is subadditivity (2.8) applied to $B, C, \emptyset, \ldots$, so only $|B| + |C| \le |A|$ needs proof; assume $|A| < \infty$, fix $\varepsilon > 0$, and by 2.2 choose open intervals $I_1, I_2, \ldots$ with

$$A \subseteq \bigcup_{k=1}^{\infty} I_k \qquad \text{and} \qquad \sum_{k=1}^{\infty} \ell(I_k) \le |A| + \varepsilon.$$

Cut each I_k at $-t$ and at t, putting

$$J_k = I_k \cap (-t,t), \qquad K_k = I_k \cap (-\infty, -t), \qquad L_k = I_k \cap (t, \infty),$$

again open intervals. Then

$$\ell(J_k) + \ell(K_k) + \ell(L_k) = \ell(I_k) \qquad \text{for each } k:$$

cutting an open interval I at a single point c gives $\ell\bigl(I \cap (-\infty, c)\bigr) + \ell\bigl(I \cap (c, \infty)\bigr) = \ell(I)$ (Check! the cases $c \le a$, $a < c < b$, $c \ge b$, and I unbounded), and applying this to I_k at $c = -t$ and then to $I_k \cap (-t, \infty)$ at $c = t$ and adding gives the display.
Now $B \subseteq \bigcup_k J_k$, so $|B| \le \sum_k \ell(J_k)$ by 2.2; and a point of A outside $(-t,t)$ lies in some I_k, hence in K_k or L_k unless it is $-t$ or t, so

$$C \subseteq \Bigl(\bigcup_{k=1}^{\infty} K_k\Bigr) \cup \Bigl(\bigcup_{k=1}^{\infty} L_k\Bigr) \cup \{-t, t\},$$

whence $|C| \le \sum_k \bigl(\ell(K_k) + \ell(L_k)\bigr)$ by 2.2, 2.3 and 2.8. Adding,

$$|B| + |C| \le \sum_{k=1}^{\infty} \ell(I_k) \le |A| + \varepsilon,$$

and $\varepsilon > 0$ was arbitrary.

Problem 2A.9. Prove that $|A| = \lim_{t \to \infty} |A \cap (-t,t)|$ for all $A \subseteq \mathbb{R}$.

Solution. By 2.5 the function $t \mapsto |A \cap (-t,t)|$ is nondecreasing on $(0, \infty)$, so the limit exists in $[0, \infty]$ and

$$L := \lim_{t \to \infty} |A \cap (-t,t)| = \sup_{t > 0} |A \cap (-t,t)| = \lim_{n \to \infty} |A \cap (-n,n)|.$$

Since $A \cap (-t,t) \subseteq A$, 2.5 gives $|A \cap (-t,t)| \le |A|$ for every t, hence $L \le |A|$.

For the reverse, cut A into integer annuli: $C_1 = A \cap (-1,1)$ and $C_n = A \cap \{x : n-1 \le |x| < n\}$ for $n \ge 2$. Applying Exercise 8 of this section to the set $A \cap (-n,n)$ with $t = n-1 > 0$, whose two pieces are $A \cap (-(n-1), n-1)$ and C_n (Check!), gives

$$|A \cap (-n,n)| = |A \cap (-(n-1), n-1)| + |C_n| \qquad (n \ge 2),$$

so induction from $|A \cap (-1,1)| = |C_1|$ yields $|A \cap (-n,n)| = \sum_{k=1}^{n} |C_k|$ and therefore $L = \sum_{k=1}^{\infty} |C_k|$. Each $x \in A$ lies in C_m for the least m with $|x| < m$, so $A = \bigcup_{k=1}^{\infty} C_k$ and 2.8 gives

$$|A| \le \sum_{k=1}^{\infty} |C_k| = L.$$

Problem 2A.10. Prove that $|[0,1] \setminus \mathbb{Q}| = 1$.

Solution. Write $E = [0,1] \setminus \mathbb{Q}$. Since $E \subseteq [0,1]$, the inequality $|E| \le |[0,1]| = 1$ holds by 2.5 and 2.14. Conversely $[0,1] \cap \mathbb{Q}$ is countable, so $|[0,1] \cap \mathbb{Q}| = 0$ by 2.4, and applying subadditivity (2.8) to $E,\ [0,1] \cap \mathbb{Q},\ \emptyset, \ldots$, whose union is $[0,1]$, gives

$$1 = |[0,1]| \le |E| + 0 = |E|.$$

Hence $|[0,1] \setminus \mathbb{Q}| = 1$.

Problem 2A.11. Prove that if $I_1, I_2, \ldots$ is a disjoint sequence of open intervals, then

$$\Big|\bigcup_{k=1}^{\infty} I_k\Big| = \sum_{k=1}^{\infty} \ell(I_k).$$

Solution. Write $A = \bigcup_{k=1}^{\infty} I_k$. The bound $|A| \le \sum_{k=1}^{\infty} \ell(I_k)$ is immediate from 2.2 (disjointness unused), so the content is the reverse bound, which comes from splitting outer measure at the endpoints of $I_1, \ldots, I_n$. Recorded once: $|I| = \ell(I)$ for every open interval I, by Exercise 6 of this section when I is bounded and nonempty, and because an unbounded I contains closed intervals of every length, so $|I| = \infty$ by 2.5 and 2.14.
Splitting lemma. For every $S \subseteq \mathbb{R}$ and $c \in \mathbb{R}$,

$$|S| = \big|S \cap (-\infty, c]\big| + \big|S \cap (c, \infty)\big|.$$

Here $\le$ is 2.8, and $\ge$ follows exactly as in Exercise 8 of this section: assuming $|S| < \infty$, take open intervals J_k with $S \subseteq \bigcup_k J_k$ and $\sum_k \ell(J_k) \le |S| + \varepsilon$, cut each at c into $J_k' = J_k \cap (-\infty, c)$ and $J_k'' = J_k \cap (c, \infty)$ with $\ell(J_k') + \ell(J_k'') = \ell(J_k)$ (Check!), and apply 2.2, 2.3 and 2.8 to

$$S \cap (-\infty, c] \subseteq \Big(\bigcup_{k=1}^{\infty} J_k'\Big) \cup \{c\} \qquad \text{and} \qquad S \cap (c, \infty) \subseteq \bigcup_{k=1}^{\infty} J_k''.$$

Iterating the lemma on $S \cap (c_1, \infty)$ and inducting on m gives, for $c_1 < \cdots < c_m$,

$$|S| = \big|S \cap (-\infty, c_1]\big| + \sum_{j=1}^{m-1} \big|S \cap (c_j, c_{j+1}]\big| + \big|S \cap (c_m, \infty)\big|.$$

The bound $|A| \ge \sum_{k=1}^{n} \ell(I_k)$. Fix n. If some I_k with $k \le n$ is unbounded then $|A| \ge |I_k| = \infty$ by 2.5, so assume $I_1, \ldots, I_n$ bounded and discard the empty ones (they contribute 0), leaving $(a_1, b_1), \ldots, (a_p, b_p)$ with $a_j < b_j$; the case $p = 0$ is trivial. Disjoint nonempty open intervals (a,b), (a', b') satisfy $b \le a'$ or $b' \le a$, since otherwise $\max\{a, a'\} < \min\{b, b'\}$ and they meet; so after relabelling

$$a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_p < b_p.$$

Let $c_1 < \cdots < c_m$ list the distinct endpoints. No endpoint lies strictly between a_j and b_j, so a_j, b_j are consecutive in that list and each $(a_j, b_j]$ is one of the intervals $(c_l, c_{l+1}]$ above. Applying the displayed identity with $S = A$ and discarding the remaining terms (all nonnegative, and distinct from the kept ones),

$$|A| \geq \sum_{j=1}^{p} \big|A \cap (a_j, b_j]\big| \geq \sum_{j=1}^{p} (b_j - a_j) = \sum_{k=1}^{n} \ell(I_k),$$

the second inequality because $(a_j, b_j) \subseteq A \cap (a_j, b_j]$ has outer measure $b_j - a_j$. Letting $n \to \infty$ gives $|A| \geq \sum_{k=1}^{\infty} \ell(I_k)$, hence equality.

Problem 2A.12. Suppose $r_1, r_2, \ldots$ is a sequence that contains every rational number. Let

$$F = \mathbb{R} \setminus \bigcup_{k=1}^{\infty} \Big(r_k - \frac{1}{2^k},\ r_k + \frac{1}{2^k}\Big).$$

(a) Show that F is a closed subset of $\mathbb{R}$.
(b) Prove that if I is an interval contained in F, then I contains at most one element.
(c) Prove that $|F| = \infty$.

Solution. Write $G = \bigcup_{k=1}^{\infty} \big(r_k - 2^{-k},\ r_k + 2^{-k}\big)$, so that $F = \mathbb{R} \setminus G$.
(a) G is a union of open sets, hence open, so its complement F is closed.
(b) If an interval $I \subseteq F$ contained $x < y$, then $[x, y] \subseteq I$, and by density of $\mathbb{Q}$ there is a rational q with $x < q < y$, so $q \in F$. But $q = r_j$ for some j, and

$$q = r_j \in \big(r_j - 2^{-j},\ r_j + 2^{-j}\big) \subseteq G,$$

contradicting $q \notin G$.
(c) By 2.2 applied to the defining intervals,

$$|G| \leq \sum_{k=1}^{\infty} \frac{2}{2^k} = 2.$$

Fix $n \in \mathbb{Z}^+$. Since $[-n, n] \subseteq F \cup G$, 2.5 and 2.8 (applied to $F, G, \emptyset, \ldots$) with $|[-n, n]| = 2n$ from 2.14 give

$$2n \leq |F| + |G| \leq |F| + 2,$$

and $|G| \leq 2 < \infty$ may be subtracted, so $|F| \geq 2n - 2$ for every n. Hence $|F| = \infty$.

Problem 2A.13. Suppose $\varepsilon > 0$. Prove that there exists a subset F of $[0, 1]$ such that F is closed, every element of F is an irrational number, and $|F| > 1 - \varepsilon$.

Solution. Take $F = [0, 1] \setminus G$, where $q_1, q_2, \ldots$ lists the countably infinite set $[0, 1] \cap \mathbb{Q}$ and

$$G = \bigcup_{k=1}^{\infty} \Big(q_k - \frac{\varepsilon}{2^{k+2}},\ q_k + \frac{\varepsilon}{2^{k+2}}\Big).$$

Then G is open, so $F = [0, 1] \cap (\mathbb{R} \setminus G)$ is an intersection of two closed sets, hence closed, and $F \subseteq [0, 1]$. Every element of F is irrational, since a rational $x \in [0, 1]$ equals some q_k and therefore lies in G. Finally 2.2 gives

$$|G| \leq \sum_{k=1}^{\infty} \frac{2\varepsilon}{2^{k+2}} = \frac{\varepsilon}{2},$$

and $[0, 1] \subseteq F \cup G$, so 2.5, 2.8 and $|[0, 1]| = 1$ (2.14) give

$$1 \leq |F| + |G| \leq |F| + \frac{\varepsilon}{2}.$$

Since $|G| \le \varepsilon/2 < \infty$ may be subtracted, $|F| \ge 1 - \varepsilon/2 > 1 - \varepsilon$.

Problem 2A.14. Consider the following figure, which is drawn accurately to scale.
[Description of the figure. In the coordinate plane, with the horizontal axis marked at $2, 5, 8, 11, 14, 17, 20$ and the vertical axis marked at $1, 3, 5, 7, 9$, three colored pieces are drawn: a yellow right triangle with vertices $(0, 0)$, $(11, 0)$, $(11, 5)$; a red rectangle with vertices $(11, 0)$, $(20, 0)$, $(20, 5)$, $(11, 5)$; and a blue right triangle with vertices $(11, 5)$, $(20, 5)$, $(20, 9)$. Their upper boundary runs from $(0, 0)$ to $(11, 5)$ to $(20, 9)$ and looks to the eye like a single straight segment, so the three pieces together appear to fill the right triangle with vertices $(0, 0)$, $(20, 0)$, $(20, 9)$.]
(a) Show that the right triangle whose vertices are $(0, 0)$, $(20, 0)$, and $(20, 9)$ has area 90.
[We have not defined area yet, but just use the elementary formulas for the areas of triangles and rectangles that you learned long ago.]
(b) Show that the yellow (lower) right triangle has area 27.5.
(c) Show that the red rectangle has area 45.
(d) Show that the blue (upper) right triangle has area 18.
(e) Add the results of parts (b), (c), and (d), showing that the area of the colored region is 90.5.
(f) Seeing the figure above, most people expect parts (a) and (e) to have the same result. Yet in (a) we found area 90, and in (e) we found area 90.5. Explain why these results differ.
[You may be tempted to think that what we have here is a two-dimensional example similar to the result about the nonadditivity of outer measure (2.18). However, genuine examples of nonadditivity require much more complicated sets than in this example.]

Solution. (a) The triangle T has its right angle at $(20, 0)$, with legs 20 and 9, so $\text{area}(T) = \frac{1}{2} \cdot 20 \cdot 9 = 90$.
(b) Right angle at $(11, 0)$, legs 11 and 5: area $= \frac{1}{2} \cdot 11 \cdot 5 = 27.5$.
(c) Sides $20 - 11 = 9$ and $5 - 0 = 5$: area $= 9 \cdot 5 = 45$.
(d) Right angle at $(20, 5)$, legs $20 - 11 = 9$ and $9 - 5 = 4$: area $= \frac{1}{2} \cdot 9 \cdot 4 = 18$.
(e) The three pieces meet only in the segments $\{11\} \times [0, 5]$ and $[11, 20] \times \{5\}$ and the point $(11, 5)$, which contribute no area, so the colored region has area $27.5 + 45 + 18 = 90.5$.
(f) The colored region is not the triangle T: its upper boundary bends upward at $(11, 5)$. Indeed the line through $(0, 0)$ and $(20, 9)$ is $y = \frac{9}{20}x$, of height $\frac{99}{20} = 4.95 < 5$ at $x = 11$; equivalently the three slopes $\frac{5}{11}$, $\frac{9}{20}$, $\frac{4}{9}$ are distinct. So the union Q of the pieces is the quadrilateral $(0, 0)$, $(20, 0)$, $(20, 9)$, $(11, 5)$, which strictly contains T, and $Q \setminus T$ is the thin triangle with vertices $(0, 0)$, $(11, 5)$, $(20, 9)$: base $\sqrt{20^2 + 9^2} = \sqrt{481}$ and height the distance from $(11, 5)$ to $9x - 20y = 0$, namely

$$\frac{|9 \cdot 11 - 20 \cdot 5|}{\sqrt{481}} = \frac{1}{\sqrt{481}},$$

hence area $\frac{1}{2}$. Thus $\text{area}(Q) = 90 + 0.5 = 90.5$, agreeing with (e). Nonadditivity (2.18) is not at issue: the pieces do not overlap and their areas do add to the area of their union.

2.2 Exercises 2B

Problem 2B.1. Show that $\mathcal{S} = \left\{\bigcup_{n \in K}(n, n + 1] : K \subseteq \mathbb{Z}\right\}$ is a σ-algebra on $\mathbb{R}$.

Solution. Everything rests on the fact that the intervals $(n, n + 1]$, $n \in \mathbb{Z}$, partition $\mathbb{R}$: each x lies in $(n_x, n_x + 1]$ where $n_x = \lceil x \rceil - 1$ (Check! separately for $x \in \mathbb{Z}$ and $x \notin \mathbb{Z}$), and in no other such interval, since $m < n$ in $\mathbb{Z}$ forces $x \le m + 1 \le n$ for $x \in (m, m + 1]$. Hence, writing $E_K = \bigcup_{n \in K}(n, n + 1]$ so that $\mathcal{S} = \{E_K : K \subseteq \mathbb{Z}\}$,

$$x \in E_K \iff n_x \in K.$$

The three conditions of 2.23 follow. First, $\emptyset = E_\emptyset \in \mathcal{S}$. Second,

$$\mathbb{R} \setminus E_K = E_{\mathbb{Z} \setminus K},$$

because $x \notin E_K$ iff $n_x \notin K$ iff $n_x \in \mathbb{Z} \setminus K$. Third, given $K_1, K_2, \ldots \subseteq \mathbb{Z}$ and $K = \bigcup_{j=1}^{\infty} K_j$,

$$x \in \bigcup_{j=1}^{\infty} E_{K_j} \iff n_x \in K_j \text{ for some } j \iff x \in E_K,$$

so $\bigcup_{j=1}^{\infty} E_{K_j} = E_K \in \mathcal{S}$. Thus $\mathcal{S}$ is a σ-algebra on $\mathbb{R}$.

Problem 2B.2. Verify both bullet points in Example 2.28, which assert the following.

- (a) Suppose X is a set and $\mathcal{A}$ is the set of subsets of X that consist of exactly one element, $\mathcal{A} = \big\{\{x\} : x \in X\big\}$. Then the smallest σ-algebra on X containing $\mathcal{A}$ is the set of all subsets E of X such that E is countable or $X \setminus E$ is countable.
- (b) Suppose $\mathcal{A} = \{(0,1), (0,\infty)\}$. Then the smallest σ-algebra on $\mathbb{R}$ containing $\mathcal{A}$ is $\{\emptyset,\ (0,1),\ (0,\infty),\ (-\infty,0] \cup [1,\infty),\ (-\infty,0],\ [1,\infty),\ (-\infty,1),\ \mathbb{R}\}$.

Solution. In each part we exhibit the asserted collection $\mathcal{T}$ and check (i) $\mathcal{T}$ is a σ-algebra, (ii) $\mathcal{A} \subseteq \mathcal{T}$, (iii) every σ-algebra containing $\mathcal{A}$ contains $\mathcal{T}$; by 2.27 these three facts identify $\mathcal{T}$ as the smallest σ-algebra containing $\mathcal{A}$.
Part (a). Let

$$\mathcal{T} = \{E \subseteq X : E \text{ is countable or } X \setminus E \text{ is countable}\}.$$

(i) $\emptyset$ is countable; the defining condition is symmetric in E and $X \setminus E$, so $\mathcal{T}$ is closed under complementation. If $E_1, E_2, \ldots \in \mathcal{T}$ and $E = \bigcup_{k=1}^{\infty} E_k$, then either every E_k is countable, whence E is countable; or some E_m is uncountable, whence $X \setminus E_m$ is countable and $X \setminus E \subseteq X \setminus E_m$ is countable. Either way $E \in \mathcal{T}$.
(ii) Each $\{x\}$ is finite, hence countable.
(iii) Let $\mathcal{S} \supseteq \mathcal{A}$ be a σ-algebra and $E \in \mathcal{T}$. If E is countable then $E = \bigcup_k \{x_k\} \in \mathcal{S}$, a countable union of elements of $\mathcal{A}$ (pad a finite list with copies of $\emptyset$). If $X \setminus E$ is countable, that case gives $X \setminus E \in \mathcal{S}$, so $E \in \mathcal{S}$.
Part (b). Set

$$A_1 = (0,1), \qquad A_2 = [1,\infty), \qquad A_3 = (-\infty, 0].$$

These are nonempty, pairwise disjoint, and partition $\mathbb{R}$. For $K \subseteq \{1,2,3\}$ put $E_K = \bigcup_{i \in K} A_i$; the eight subsets K give

$$\begin{aligned} E_{\emptyset} &= \emptyset, & E_{\{1\}} &= (0,1),\\ E_{\{2\}} &= [1,\infty), & E_{\{3\}} &= (-\infty,0],\\ E_{\{1,2\}} &= (0,\infty), & E_{\{1,3\}} &= (-\infty,1),\\ E_{\{2,3\}} &= (-\infty,0] \cup [1,\infty), & E_{\{1,2,3\}} &= \mathbb{R}, \end{aligned}$$

so $\mathcal{T} = \{E_K : K \subseteq \{1,2,3\}\}$ is exactly the eight-element collection displayed in the exercise (the eight sets are distinct, since $i \in K \iff A_i \subseteq E_K$).
(i) Each x lies in exactly one $A_{i(x)}$, and $x \in E_K \iff i(x) \in K$; exactly as in Exercise 1 of this section this gives $\emptyset = E_{\emptyset} \in \mathcal{T}$ together with

$$\mathbb{R} \setminus E_K = E_{\{1,2,3\} \setminus K} \qquad \text{and} \qquad \bigcup_{j=1}^{\infty} E_{K_j} = E_{K_1 \cup K_2 \cup \cdots}.$$

(ii) $(0,1) = E_{\{1\}}$ and $(0,\infty) = E_{\{1,2\}}$.
(iii) If $\mathcal{S}$ is a σ-algebra containing $(0,1)$ and $(0,\infty)$, then by 2.25(a) and 2.25(b)

$$A_2 = (0,\infty) \setminus (0,1) \in \mathcal{S}, \qquad A_3 = \mathbb{R} \setminus (0,\infty) \in \mathcal{S},$$

and $A_1 = (0,1) \in \mathcal{S}$; closure under finite unions (2.25(b)) puts every E_K in $\mathcal{S}$.

Problem 2B.3. Suppose $\mathcal{S}$ is the smallest σ-algebra on $\mathbb{R}$ containing $\{(r,s] : r,s \in \mathbb{Q}\}$. Prove that $\mathcal{S}$ is the collection of Borel subsets of $\mathbb{R}$.

Solution. Let $\mathcal{B}$ be the collection of Borel subsets of $\mathbb{R}$, the smallest σ-algebra containing the open sets (2.29).
$\mathcal{S} \subseteq \mathcal{B}$. Every generator is Borel: $(r,s] = \emptyset$ if $r \geq s$, while if $r < s$ then

$$(r,s] = \bigcap_{k=1}^{\infty} \left(r,\ s + \tfrac{1}{k}\right)$$

(since $x < s + \frac{1}{k}$ for all k forces $x \leq s$), a countable intersection of open sets, which lies in $\mathcal{B}$ by 2.25(c). So $\mathcal{B}$ is a σ-algebra containing the generating collection and hence contains the smallest one, $\mathcal{S}$.
$\mathcal{B} \subseteq \mathcal{S}$. By 2.29 it suffices that every open $G \subseteq \mathbb{R}$ lie in $\mathcal{S}$. With $P = \{(r,s) \in \mathbb{Q} \times \mathbb{Q} : (r,s] \subseteq G\}$,

$$G = \bigcup_{(r,s) \in P} (r,s] :$$

the inclusion $\supseteq$ is the definition of P, and for $\subseteq$, given $x \in G$ pick $\varepsilon > 0$ with $(x-\varepsilon, x+\varepsilon) \subseteq G$ and then, by density of $\mathbb{Q}$, rationals $r \in (x-\varepsilon, x)$ and $s \in (x, x+\varepsilon)$, so that $x \in (r,s] \subseteq G$. Since $P \subseteq \mathbb{Q} \times \mathbb{Q}$ is countable, this is a countable union of elements of $\mathcal{S}$ (and $G = \emptyset \in \mathcal{S}$ if $P = \emptyset$), so $G \in \mathcal{S}$.

Problem 2B.4. Suppose $\mathcal{S}$ is the smallest σ-algebra on $\mathbb{R}$ containing $\{(r,n] : r \in \mathbb{Q},\ n \in \mathbb{Z}\}$. Prove that $\mathcal{S}$ is the collection of Borel subsets of $\mathbb{R}$.

Solution. Let $\mathcal{B}$ denote the collection of Borel subsets of $\mathbb{R}$.
$\mathcal{S} \subseteq \mathcal{B}$. Every integer is rational, so each generator $(r,n]$ is among the intervals $(r,s]$ with $r,s \in \mathbb{Q}$ shown to be Borel in Exercise 3 of this section. Hence $\mathcal{B}$ is a σ-algebra containing the generating collection, so it contains $\mathcal{S}$.
$\mathcal{B} \subseteq \mathcal{S}$. By Exercise 3, $\mathcal{B}$ is the smallest σ-algebra containing all $(r,s]$ with $r,s \in \mathbb{Q}$, so it suffices to put those intervals in $\mathcal{S}$. If $r \geq s$ then $(r,s] = \emptyset \in \mathcal{S}$. If $r < s$, choose $n \in \mathbb{Z}$ with $n \geq s$ (Archimedean property); then

$$(r,s] = (r,n] \setminus (s,n],$$

since for $r < x \leq n$ the failure of $s < x$ is exactly $x \leq s$. Both sets on the right are generators of $\mathcal{S}$, and σ-algebras are closed under differences by 2.25(b), so $(r,s] \in \mathcal{S}$.

Problem 2B.5. Suppose $\mathcal{S}$ is the smallest σ-algebra on $\mathbb{R}$ containing $\{(r, r+1) : r \in \mathbb{Q}\}$. Prove that $\mathcal{S}$ is the collection of Borel subsets of $\mathbb{R}$.

Solution. Let $\mathcal{B}$ denote the collection of Borel subsets of $\mathbb{R}$.
$\mathcal{S} \subseteq \mathcal{B}$. Each generator $(r, r+1)$ is open, hence Borel, so $\mathcal{B}$ is a σ-algebra containing the generating collection and therefore contains $\mathcal{S}$.
Short rational intervals lie in $\mathcal{S}$. If $p,q \in \mathbb{Q}$ with $p < q$ and $q - p \leq 1$, then $q - 1 \leq p$ and $q \leq p+1$, so

$$(p, p+1) \cap (q-1, q) = \big(\max\{p,\ q-1\},\ \min\{p+1,\ q\}\big) = (p,q),$$

and both factors are generators (p and $q-1$ being rational), so $(p,q) \in \mathcal{S}$ by closure under finite intersections (2.25(b)).
$\mathcal{B} \subseteq \mathcal{S}$. By 2.29 it suffices that every open $G \subseteq \mathbb{R}$ lie in $\mathcal{S}$. Put

$$P = \{(p,q) \in \mathbb{Q} \times \mathbb{Q} : p < q,\ q - p \leq 1,\ (p,q) \subseteq G\},$$

so that $G = \bigcup_{(p,q) \in P} (p,q)$: the inclusion $\supseteq$ is the definition of P, and for $\subseteq$, given $x \in G$ choose $\varepsilon \in \left(0, \frac{1}{2}\right)$ with $(x-\varepsilon, x+\varepsilon) \subseteq G$ and rationals $p \in (x-\varepsilon, x)$, $q \in (x, x+\varepsilon)$, whence $q - p < 2\varepsilon < 1$ and $x \in (p,q) \subseteq G$. As $P \subseteq \mathbb{Q} \times \mathbb{Q}$ is countable, the previous paragraph and closure under countable unions give $G \in \mathcal{S}$ (and $G = \emptyset \in \mathcal{S}$ if $P = \emptyset$).

Problem 2B.6. Suppose $\mathcal{S}$ is the smallest σ-algebra on $\mathbb{R}$ containing $\{[r,\infty) : r \in \mathbb{Q}\}$. Prove that $\mathcal{S}$ is the collection of Borel subsets of $\mathbb{R}$.

Solution. Let $\mathcal{B}$ denote the collection of Borel subsets of $\mathbb{R}$.
$\mathcal{S} \subseteq \mathcal{B}$. Each generator $[r,\infty)$ is closed, its complement $(-\infty,r)$ being open, hence Borel by the first bullet point of 2.30; so $\mathcal{B}$ is a σ-algebra containing the generating collection and therefore contains $\mathcal{S}$.
Bounded rational open intervals lie in $\mathcal{S}$. For $p \in \mathbb{Q}$,

$$(p,\infty) = \bigcup_{k=1}^{\infty} \bigl[p + \tfrac{1}{k},\ \infty\bigr) \in \mathcal{S},$$

since $x > p$ gives $\frac{1}{k} \leq x - p$ for some k by the Archimedean property, each $p + \frac{1}{k}$ is rational, and $\mathcal{S}$ is closed under countable unions. Also $(-\infty,q) = \mathbb{R} \setminus [q,\infty) \in \mathcal{S}$ for $q \in \mathbb{Q}$, so

$$(p,q) = (p,\infty) \cap (-\infty,q) \in \mathcal{S}$$

for all $p,q \in \mathbb{Q}$, by closure under finite intersections (2.25(b)).
$\mathcal{B} \subseteq \mathcal{S}$. Exactly as in Exercise 5 of this section, every open $G \subseteq \mathbb{R}$ is the union of the countably many intervals (p,q) with $p,q \in \mathbb{Q}$, $p < q$ and $(p,q) \subseteq G$; by the previous paragraph and closure under countable unions, $G \in \mathcal{S}$. Now 2.29 gives $\mathcal{B} \subseteq \mathcal{S}$.

Problem 2B.7. Prove that the collection of Borel subsets of $\mathbb{R}$ is translation invariant. More precisely, prove that if $B \subseteq \mathbb{R}$ is a Borel set and $t \in \mathbb{R}$, then $t + B$ is a Borel set.

Solution. Fix $t \in \mathbb{R}$ and let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x - t$, which is continuous and hence Borel measurable by 2.41. Since $x - t \in B$ exactly when $x = t + b$ for some $b \in B$,

$$f^{-1}(B) = \{x \in \mathbb{R} : x - t \in B\} = t + B,$$

so $t + B$ is a Borel set for every Borel set B, by the definition 2.40 of Borel measurability.
Method (2) (avoiding 2.41). Let $\mathcal{B}$ be the collection of Borel subsets of $\mathbb{R}$ and put $\mathcal{T} = \{A \subseteq \mathbb{R} : t + A \in \mathcal{B}\}$. Because $x \in t + A$ iff $x - t \in A$,

$$t + (\mathbb{R} \setminus A) = \mathbb{R} \setminus (t + A), \qquad t + \bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} (t + A_k),$$

and $t + \emptyset = \emptyset$; since $\mathcal{B}$ is a σ-algebra, so is $\mathcal{T}$. If G is open then $t + G$ is open, translation being an isometry ($|y - x| < \varepsilon$ iff $|(y-t) - (x-t)| < \varepsilon$), so $t + G \in \mathcal{B}$ and $G \in \mathcal{T}$. Thus $\mathcal{T}$ is a σ-algebra containing every open set, and 2.29 gives $\mathcal{B} \subseteq \mathcal{T}$.

Problem 2B.8. Prove that the collection of Borel subsets of $\mathbb{R}$ is dilation invariant. More precisely, prove that if $B \subseteq \mathbb{R}$ is a Borel set and $t \in \mathbb{R}$, then tB (which is defined to be $\{tb : b \in B\}$) is a Borel set.

Solution. Suppose $B \subseteq \mathbb{R}$ is a Borel set and $t \in \mathbb{R}$.
(i) $t = 0$. Then tB is $\emptyset$ or $\{0\}$, both Borel ($\{0\}$ is closed, hence Borel by 2.30).
(ii) $t \neq 0$. The function $m(x) = x/t$ is continuous, hence Borel measurable by 2.41, so $m^{-1}(B)$ is a Borel set by the definition 2.40. And $m^{-1}(B) = tB$: if $x/t \in B$ then $x = t(x/t) \in tB$, while if $x = tb$ with $b \in B$ then $x/t = b \in B$, using $t \neq 0$.
Method (2) (avoiding 2.41). Fix $t \neq 0$ and put $\mathcal{S} = \{B \subseteq \mathbb{R} : tB \text{ is Borel}\}$. Multiplication by t is a bijection of $\mathbb{R}$, so

$$t(\mathbb{R} \setminus B) = \mathbb{R} \setminus tB, \qquad t\Bigl(\bigcup_{k=1}^{\infty} B_k\Bigr) = \bigcup_{k=1}^{\infty} tB_k,$$

and $t\emptyset = \emptyset$; hence $\mathcal{S}$ is a σ-algebra. It contains every open G, since multiplication by t is a homeomorphism and so tG is open, hence Borel. By 2.29, $\mathcal{S}$ contains every Borel set.

Problem 2B.9. Give an example of a measurable space $(X, \mathcal{S})$ and a function $f : X \to \mathbb{R}$ such that $|f|$ is $\mathcal{S}$-measurable but f is not $\mathcal{S}$-measurable.

Solution. Take $X = \mathbb{R}$, $\mathcal{S} = \{\emptyset, \mathbb{R}\}$ (a σ-algebra; Check!), and

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Then $|f|$ is constant with value 1, so $|f|^{-1}(B)$ is $\mathbb{R}$ or $\emptyset$ according as $1 \in B$ or not; either way it lies in $\mathcal{S}$, so $|f|$ is $\mathcal{S}$-measurable. But $(0, \infty)$ is open, hence Borel, and

$$f^{-1}\big((0,\infty)\big) = [0, \infty) \notin \mathcal{S},$$

so f fails the definition 2.35 of $\mathcal{S}$-measurability.

Problem 2B.10. Show that the set of real numbers that have a decimal expansion with the digit 5 appearing infinitely often is a Borel set.

Solution. The set in question is $A = A_+ \cup (-A_+)$, where

$$A_+ = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} D_k, \qquad D_k = \bigcup_{m=0}^{\infty} \left[\frac{10m+5}{10^k}, \frac{10m+6}{10^k}\right),$$

and this is a Borel set. Here a decimal expansion of $x \geq 0$ means a representation $x = N + \sum_{j=1}^{\infty} e_j 10^{-j}$ with $N \in \mathbb{Z}$, $N \geq 0$ and $e_j \in \{0, 1, \ldots, 9\}$, written $x = N.e_1e_2e_3\ldots$, while a negative number carries the expansion of its absolute value; the integer part is a finite string, so the digit 5 can recur infinitely often only among $e_1, e_2, \ldots$.
Reduction to $[0, \infty)$. Since x and $-x$ carry the same digit string, $x \in A$ iff $|x| \in A$, so $A = A_+ \cup (-A_+)$ with $A_+ = A \cap [0, \infty)$, and $-A_+$ is Borel whenever A_+ is, by Exercise 8 of this section with $t = -1$.
Canonical digits. For $x \geq 0$ and $k \in \mathbb{Z}^+$ set $d_k(x) = \lfloor 10^k x \rfloor - 10\lfloor 10^{k-1}x \rfloor$. Writing $q = \lfloor 10^{k-1}x \rfloor$ gives $10q \leq 10^k x < 10q + 10$, hence $d_k(x) \in \{0, 1, \ldots, 9\}$, and the sum telescopes to

$$\lfloor x \rfloor + \sum_{k=1}^{n} d_k(x)10^{-k} = \frac{\lfloor 10^n x \rfloor}{10^n} \longrightarrow x$$

because $x - 10^{-n} < \lfloor 10^n x \rfloor / 10^n \leq x$. So $\lfloor x \rfloor.d_1(x)d_2(x)\ldots$ is a decimal expansion of x, the canonical one.
An expansion with infinitely many 5*s is the canonical one.* Let $x = N.e_1e_2e_3\ldots$ with $e_j = 5$ for infinitely many j, and fix $k \geq 0$. Some $i > k$ has $e_i = 5$, so the digits beyond position k are not all 9 and

$$0 \leq \sum_{j>k} e_j 10^{-j} \leq 10^{-k} - 4 \cdot 10^{-i} < 10^{-k}.$$

With $M_k = 10^k N + \sum_{j=1}^{k} e_j 10^{k-j} \in \mathbb{Z}$ this reads $M_k \leq 10^k x < M_k + 1$, so $\lfloor 10^k x \rfloor = M_k$ for every $k \geq 0$, whence $\lfloor x \rfloor = M_0 = N$ and $d_k(x) = M_k - 10M_{k-1} = e_k$. The converse direction is immediate, the canonical string being itself an expansion, so

$$A_+ = \{x \geq 0 : d_k(x) = 5 \text{ for infinitely many } k \in \mathbb{Z}^+\}.$$

The sets D_k. Fix k. Then $d_k(x) = 5$ iff $\lfloor 10^k x \rfloor \in \{10m + 5 : m \in \mathbb{Z}\}$: forwards because $\lfloor 10^k x \rfloor = 10\lfloor 10^{k-1}x \rfloor + 5$; backwards because $\lfloor 10^k x \rfloor = 10m + 5$ gives $m + \frac{1}{2} \leq 10^{k-1}x < m + \frac{3}{5}$, so $\lfloor 10^{k-1}x \rfloor = m$

and $d_k(x) = 5$. Since $x \geq 0$ forces $\lfloor 10^k x \rfloor \geq 0$, only $m \geq 0$ occurs, so $\{x \geq 0 : d_k(x) = 5\}$ is exactly the set D_k displayed at the start, a countable union of half-open intervals and hence Borel by 2.30.
Assembling. The digit occurs infinitely often exactly when for every n some $k \geq n$ has $d_k(x) = 5$, which is the displayed formula for A_+; it is Borel by 2.23 (countable unions) and 2.25(c) (countable intersections), and hence so is $A = A_+ \cup (-A_+)$.

Problem 2B.11. Suppose $\mathcal{T}$ is a σ-algebra on a set Y and $X \in \mathcal{T}$. Let $\mathcal{S} = \{E \in \mathcal{T} : E \subseteq X\}$.
(a) Show that $\mathcal{S} = \{F \cap X : F \in \mathcal{T}\}$.
(b) Show that $\mathcal{S}$ is a σ-algebra on X.

Solution. (a) If $E \in \mathcal{S}$ then $E \in \mathcal{T}$ and $E = E \cap X$, so E has the required form. Conversely, if $F \in \mathcal{T}$ then $F \cap X \in \mathcal{T}$ by 2.25(b) (using $X \in \mathcal{T}$) and $F \cap X \subseteq X$, so $F \cap X \in \mathcal{S}$.
(b) Each element of $\mathcal{S}$ is a subset of X, and complementation is taken in X. The three conditions of 2.23 hold: $\emptyset \in \mathcal{T}$ and $\emptyset \subseteq X$; for $E \in \mathcal{S}$, both E and X lie in $\mathcal{T}$, so $X \setminus E \in \mathcal{T}$ by 2.25(b) and $X \setminus E \subseteq X$; and for $E_1, E_2, \ldots \in \mathcal{S}$, the union $\bigcup_{k=1}^{\infty} E_k$ lies in $\mathcal{T}$ and is contained in X. Hence $\mathcal{S}$ is a σ-algebra on X.

Problem 2B.12. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a function.
(a) For $k \in \mathbb{Z}^+$, let

$$G_k = \Big\{a \in \mathbb{R} : \text{there exists } \delta > 0 \text{ such that } |f(b) - f(c)| < \tfrac{1}{k} \text{ for all } b, c \in (a - \delta, a + \delta)\Big\}.$$

Prove that G_k is an open subset of $\mathbb{R}$ for each $k \in \mathbb{Z}^+$.
(b) Prove that the set of points at which f is continuous equals $\bigcap_{k=1}^{\infty} G_k$.
(c) Conclude that the set of points at which f is continuous is a Borel set.

Solution. (a) Let $a \in G_k$, with $\delta > 0$ as in the definition. Then $\big(a - \frac{\delta}{2}, a + \frac{\delta}{2}\big) \subseteq G_k$: for a' in that interval put $\delta' = \frac{\delta}{2} - |a' - a| > 0$, so that

$$|y - a| \leq |y - a'| + |a' - a| < \delta' + |a' - a| = \tfrac{\delta}{2} < \delta$$

for $y \in (a' - \delta', a' + \delta')$; hence that interval sits inside $(a - \delta, a + \delta)$, where the oscillation bound $\frac{1}{k}$ already holds, and $a' \in G_k$. So G_k contains an open interval about each of its points and is open.
(b) Write C for the set of continuity points. If $a \in C$ and $k \in \mathbb{Z}^+$, continuity at a with $\varepsilon = \frac{1}{2k}$ gives $\delta > 0$ with $|f(y) - f(a)| < \frac{1}{2k}$ on $(a - \delta, a + \delta)$, so

$$|f(b) - f(c)| \leq |f(b) - f(a)| + |f(a) - f(c)| < \tfrac{1}{k}$$

there, and $a \in G_k$. Conversely, if $a \in \bigcap_{k=1}^{\infty} G_k$ and $\varepsilon > 0$, choose k with $\frac{1}{k} < \varepsilon$ and take $c = a$ in the defining property of G_k: then $|f(b) - f(a)| < \frac{1}{k} < \varepsilon$ for all $b \in (a - \delta, a + \delta)$, so $a \in C$. Hence $C = \bigcap_{k=1}^{\infty} G_k$.
(c) Each G_k is open by (a), hence Borel by 2.29, and C is a countable intersection of Borel sets, hence Borel by 2.25(c).

Problem 2B.13. Suppose $(X, \mathcal{S})$ is a measurable space, $E_1, \ldots, E_n$ are disjoint subsets of X, and $c_1, \ldots, c_n$ are distinct nonzero real numbers. Prove that $c_1\chi_{E_1} + \cdots + c_n\chi_{E_n}$ is an $\mathcal{S}$-measurable function if and only if $E_1, \ldots, E_n \in \mathcal{S}$.

Solution. Write $f = c_1\chi_{E_1} + \cdots + c_n\chi_{E_n}$.
($\Leftarrow$) Suppose $E_1, \ldots, E_n \in \mathcal{S}$. By 2.38 each $\chi_{E_j}^{-1}(B)$ is one of E_j, $X \setminus E_j$, X, $\emptyset$, all in $\mathcal{S}$ by hypothesis together with 2.23 and 2.25(a), so each χ_{E_j} is $\mathcal{S}$-measurable. Constant functions are $\mathcal{S}$-measurable (inverse images are X or $\emptyset$), so 2.46(a) makes each $c_j\chi_{E_j}$ measurable and then, inducting on the

number of summands, makes f measurable.
($\Rightarrow$) Suppose f is $\mathcal{S}$-measurable. Disjointness gives $f(x) = c_i$ for $x \in E_i$ and $f(x) = 0$ off $E_1 \cup \cdots \cup E_n$, so for each j,

$$f^{-1}(\{c_j\}) = E_j :$$

$E_j \subseteq f^{-1}(\{c_j\})$ is clear, and if $f(x) = c_j$ then x lies in some E_i (else $f(x) = 0 \neq c_j$, the cs being nonzero) with $c_i = c_j$, forcing $i = j$ since the cs are distinct. As $\{c_j\}$ is closed, hence Borel by 2.30, the definition 2.35 gives $E_j = f^{-1}(\{c_j\}) \in \mathcal{S}$.

Problem 2B.14. (a) Suppose $f_1, f_2, \ldots$ is a sequence of functions from a set X to $\mathbb{R}$. Explain why

$$\{x \in X : \text{the sequence } f_1(x), f_2(x), \ldots \text{ has a limit in } \mathbb{R}\} = \bigcap_{n=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} (f_j - f_k)^{-1}\Big(\big(-\tfrac{1}{n}, \tfrac{1}{n}\big)\Big).$$

(b) Suppose $(X, \mathcal{S})$ is a measurable space and $f_1, f_2, \ldots$ is a sequence of $\mathcal{S}$-measurable functions from X to $\mathbb{R}$. Prove that

$$\{x \in X : \text{the sequence } f_1(x), f_2(x), \ldots \text{ has a limit in } \mathbb{R}\}$$

is an $\mathcal{S}$-measurable subset of X.

Solution. (a) The identity is the Cauchy criterion, written with a single anchor index. Indeed, by the definition 2.31 of inverse image, x lies in the right side R exactly when

$$\text{for every } n \text{ there is } j \text{ with } |f_j(x) - f_k(x)| < \tfrac{1}{n} \text{ for all } k \geq j.$$

If $f_1(x), f_2(x), \ldots$ converges it is Cauchy, and taking $p = j$ in the Cauchy condition for $\varepsilon = \frac{1}{n}$ puts $x \in R$. Conversely, let $x \in R$ and $\varepsilon > 0$; choose n with $\frac{2}{n} < \varepsilon$ and the corresponding j, so that for all $p, q \geq j$,

$$|f_p(x) - f_q(x)| \leq |f_p(x) - f_j(x)| + |f_j(x) - f_q(x)| < \tfrac{2}{n} < \varepsilon.$$

Thus the sequence is Cauchy, hence convergent by completeness of $\mathbb{R}$.
(b) Each $f_j - f_k$ is $\mathcal{S}$-measurable by 2.46(a), and $\big(-\frac{1}{n}, \frac{1}{n}\big)$ is open, hence Borel by 2.29, so the definition 2.35 gives

$$(f_j - f_k)^{-1}\Big(\big(-\tfrac{1}{n}, \tfrac{1}{n}\big)\Big) \in \mathcal{S}$$

for all n, j, k. The set of (a) is assembled from these by a countable intersection over k, a countable union over j, and a countable intersection over n; $\mathcal{S}$ is closed under the first and third by 2.25(c) and under the second by 2.23, so that set lies in $\mathcal{S}$.

Problem 2B.15. Suppose X is a set and $E_1, E_2, \ldots$ is a disjoint sequence of subsets of X such that $\bigcup_{k=1}^{\infty} E_k = X$. Let $\mathcal{S} = \big\{\bigcup_{k \in K} E_k : K \subseteq \mathbb{Z}^+\big\}$.
(a) Show that $\mathcal{S}$ is a σ-algebra on X.
(b) Prove that a function from X to $\mathbb{R}$ is $\mathcal{S}$-measurable if and only if the function is constant on E_k for every $k \in \mathbb{Z}^+$.

Solution. Since the E_k are disjoint with union X, every $x \in X$ lies in $E_{k(x)}$ for exactly one index $k(x)$, and $x \in \bigcup_{k \in K} E_k \iff k(x) \in K$.
(a) Taking $K = \emptyset$ gives $\emptyset \in \mathcal{S}$. For $E = \bigcup_{k \in K} E_k$, the equivalence above gives

$$X \setminus E = \bigcup_{k \in \mathbb{Z}^+ \setminus K} E_k \in \mathcal{S},$$

and for $A_n = \bigcup_{k\in K_n} E_k$ with $K = \bigcup_{n=1}^{\infty} K_n$,

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{k\in K} E_k \in \mathcal{S}.$$

Hence $\mathcal{S}$ is a σ-algebra on X.
(b) ($\Leftarrow$) Let f be constant on each E_k, let B be Borel, and put $K = \{k : E_k \subseteq f^{-1}(B)\}$. Then $f^{-1}(B) = \bigcup_{k\in K} E_k \in \mathcal{S}$: the inclusion $\supseteq$ is the definition of K, and if $x \in f^{-1}(B)$ then constancy on $E_{k(x)}$ gives $E_{k(x)} \subseteq f^{-1}(B)$, so $k(x) \in K$.
($\Rightarrow$) Let f be $\mathcal{S}$-measurable and $x, y \in E_k$. The singleton $\{f(x)\}$ is closed, hence Borel, so $f^{-1}(\{f(x)\}) = \bigcup_{j\in K} E_j$ for some $K \subseteq \mathbb{Z}^+$; since x lies in it, $k = k(x) \in K$, so $E_k \subseteq f^{-1}(\{f(x)\})$ and $f(y) = f(x)$.

Problem 2B.16. Suppose $\mathcal{S}$ is a σ-algebra on a set X and $A \subseteq X$. Let

$$\mathcal{S}_A = \{E \in \mathcal{S} : A \subseteq E \text{ or } A \cap E = \emptyset\}.$$

(a) Prove that $\mathcal{S}_A$ is a σ-algebra on X.
(b) Suppose $f : X \to \mathbb{R}$ is a function. Prove that f is measurable with respect to $\mathcal{S}_A$ if and only if f is measurable with respect to $\mathcal{S}$ and f is constant on A.

Solution. (a) Since $\emptyset \in \mathcal{S}$ and $A \cap \emptyset = \emptyset$, we get $\emptyset \in \mathcal{S}_A$. For $E \in \mathcal{S}_A$ we have $X \setminus E \in \mathcal{S}$, and $A \subseteq E$ gives $A \cap (X \setminus E) = \emptyset$ while $A \cap E = \emptyset$ gives $A \subseteq X \setminus E$; either way $X \setminus E \in \mathcal{S}_A$. For $E_1, E_2, \ldots \in \mathcal{S}_A$ and $E = \bigcup_{k=1}^{\infty} E_k \in \mathcal{S}$: either $A \subseteq E_k$ for some k, whence $A \subseteq E$, or $A \cap E_k = \emptyset$ for every k, whence

$$A \cap E = \bigcup_{k=1}^{\infty} (A \cap E_k) = \emptyset.$$

Hence $\mathcal{S}_A$ is a σ-algebra on X.
(b) ($\Rightarrow$) If f is $\mathcal{S}_A$-measurable, then $\mathcal{S}_A \subseteq \mathcal{S}$ makes f $\mathcal{S}$-measurable. Assuming $A \neq \emptyset$, pick $x \in A$; the singleton $\{f(x)\}$ is closed, hence Borel, so $f^{-1}(\{f(x)\}) \in \mathcal{S}_A$, and it meets A (at x), so the definition of $\mathcal{S}_A$ forces $A \subseteq f^{-1}(\{f(x)\})$; that is, f is constant on A.
($\Leftarrow$) Let f be $\mathcal{S}$-measurable and constant on A. If $A = \emptyset$ then $\mathcal{S}_A = \mathcal{S}$; otherwise let c be the value of f on A and B Borel. Then $f^{-1}(B) \in \mathcal{S}$, with $A \subseteq f^{-1}(B)$ if $c \in B$ and $A \cap f^{-1}(B) = \emptyset$ if $c \notin B$; either way $f^{-1}(B) \in \mathcal{S}_A$.

Problem 2B.17. Suppose X is a Borel subset of $\mathbb{R}$ and $f : X \to \mathbb{R}$ is a function such that $\{x \in X : f \text{ is not continuous at } x\}$ is a countable set. Prove f is a Borel measurable function.

Solution. By 2.39, applied with $\mathcal{S}$ the σ-algebra of Borel sets contained in X, it suffices to show $f^{-1}\big((a, \infty)\big)$ is Borel for every $a \in \mathbb{R}$. Write D for the countable set of discontinuity points; every countable set is Borel, being a countable union of closed singletons. Fix a and split

$$f^{-1}\big((a,\infty)\big) = U \cup V, \qquad U = \{x \in X \setminus D : f(x) > a\}, \qquad V = \{x \in D : f(x) > a\},$$

so $V \subseteq D$ is countable, hence Borel. For $x \in U$, continuity of f at x relative to X, applied with $\varepsilon = f(x) - a > 0$, yields $\delta_x > 0$ with

$$f(y) > f(x) - \big(f(x) - a\big) = a \quad \text{for all } y \in (x - \delta_x, x + \delta_x) \cap X,$$

points of D in that intersection included. Then $G = \bigcup_{x\in U} (x - \delta_x, x + \delta_x)$ is open, hence Borel, so $G \cap X$ is Borel, and

$$f^{-1}\big((a,\infty)\big) = (G \cap X) \cup V.$$

Indeed $\subseteq$ holds because a point of the left side lies in V if it is in D and otherwise lies in $U \subseteq G \cap X$; and $\supseteq$ holds because V is part of the left side while any $y \in G \cap X$ lies in $(x-\delta_x, x+\delta_x) \cap X$ for some $x \in U$, so $f(y) > a$.

Problem 2B.18. Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable at every element of $\mathbb{R}$. Prove that f' is a Borel measurable function from $\mathbb{R}$ to $\mathbb{R}$.

Solution. f' is the pointwise limit of the difference quotients

$$g_k(x) = \frac{f\left(x+\frac{1}{k}\right) - f(x)}{1/k}, \qquad k \in \mathbb{Z}^+.$$

Each g_k is continuous on $\mathbb{R}$, since f is continuous (being differentiable everywhere) and $x \mapsto x + \frac{1}{k}$ is continuous; hence each g_k is Borel measurable by 2.41. Because f is differentiable at each x, the limit

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = f'(x)$$

exists in $\mathbb{R}$, so evaluating it along $h = \frac{1}{k} \to 0$ gives $g_k(x) \to f'(x)$ for every $x \in \mathbb{R}$. Thus f' is Borel measurable by 2.48, applied with $\mathcal{S}$ the Borel σ-algebra on $\mathbb{R}$.

Problem 2B.19. Suppose X is a nonempty set and $\mathcal{S}$ is the σ-algebra on X consisting of all subsets of X that are either countable or have a countable complement in X. Give a characterization of the $\mathcal{S}$-measurable real-valued functions on X.

Solution. A function $f : X \to \mathbb{R}$ is $\mathcal{S}$-measurable if and only if f is constant outside some countable subset of X, that is, if and only if there is $c \in \mathbb{R}$ with $\{x \in X : f(x) \neq c\}$ countable.
Suppose $C = \{x \in X : f(x) \neq c\}$ is countable and $B \subseteq \mathbb{R}$ is Borel. If $c \in B$, then $X \setminus f^{-1}(B) \subseteq C$ is countable; if $c \notin B$, then $f^{-1}(B) \subseteq C$ is countable. Either way $f^{-1}(B) \in \mathcal{S}$.
Conversely, suppose f is $\mathcal{S}$-measurable. If X is countable, take $c = 0$; so assume X is uncountable, in which case no subset of X is both countable and cocountable. Thus each $E_a := f^{-1}\big((a,\infty)\big) \in \mathcal{S}$ is countable or cocountable but not both, and

$$T = \{a \in \mathbb{R} : E_a \text{ is countable}\}$$

is closed upward, since $b > a$ gives $E_b \subseteq E_a$.

- (i) $T \neq \emptyset$: otherwise every E_n with $n \in \mathbb{Z}^+$ is cocountable, so $\bigcap_{n=1}^{\infty} E_n$ has countable complement $\bigcup_{n=1}^{\infty}(X \setminus E_n)$ and hence is nonempty as X is uncountable; but that intersection is empty because f is real-valued.
- (ii) T is bounded below: $X = \bigcup_{n=1}^{\infty} E_{-n}$ is uncountable, so E_{-m} is uncountable for some $m \in \mathbb{Z}^+$, i.e. $-m \notin T$; upward closure then forces $a > -m$ for every $a \in T$.

Let $c = \inf T \in \mathbb{R}$. For each $n \in \mathbb{Z}^+$ there is $a \in T$ with $a < c + \frac{1}{n}$, so $c + \frac{1}{n} \in T$ by upward closure, whence

$$\{x \in X : f(x) > c\} = \bigcup_{n=1}^{\infty} E_{c+\frac{1}{n}}$$

is countable. Also $c - \frac{1}{n} \notin T$, so $E_{c-1/n}$ is cocountable and $X \setminus E_{c-1/n} = \{f \leq c - \frac{1}{n}\}$ is countable, whence

$$\{x \in X : f(x) < c\} = \bigcup_{n=1}^{\infty} \left\{x \in X : f(x) \leq c - \tfrac{1}{n}\right\}$$

is countable. Hence $\{x \in X : f(x) \neq c\} = \{f > c\} \cup \{f < c\}$ is countable.

Problem 2B.20. Suppose $(X, \mathcal{S})$ is a measurable space and $f, g : X \to \mathbb{R}$ are $\mathcal{S}$-measurable functions. Prove that if $f(x) > 0$ for all $x \in X$, then f^g (which is the function whose value at $x \in X$ equals $f(x)^{g(x)}$) is an $\mathcal{S}$-measurable function.

Solution. Because $t^s = e^{s \ln t}$ for $t > 0$ and f is positive everywhere,

$$f^g(x) = e^{\,g(x) \ln f(x)} \qquad \text{for all } x \in X,$$

so $f^g = \exp \circ h$ with $h = g \cdot (\ln \circ f)$.
Now ln is continuous on the Borel set $(0, \infty)$, hence Borel measurable by 2.41, and the range of f lies in $(0, \infty)$, so $\ln \circ f$ is $\mathcal{S}$-measurable by 2.44. Hence h is $\mathcal{S}$-measurable by 2.46(a). Finally exp is continuous on $\mathbb{R}$, hence Borel measurable by 2.41, and its domain $\mathbb{R}$ contains the range of h, so $f^g = \exp \circ h$ is $\mathcal{S}$-measurable by 2.44.

Problem 2B.21. Prove 2.52. [2.52, condition for measurable function: Suppose $(X, \mathcal{S})$ is a measurable space and $f : X \to [-\infty, \infty]$ is a function such that $f^{-1}\big((a, \infty]\big) \in \mathcal{S}$ for all $a \in \mathbb{R}$. Then f is an $\mathcal{S}$-measurable function.]

Solution. Let

$$\mathcal{T} = \{A \subseteq [-\infty, \infty] : f^{-1}(A) \in \mathcal{S}\};$$

it suffices to show $\mathcal{T}$ contains every Borel subset of $[-\infty, \infty]$.
Because inverse images commute with complementation and countable unions, $\mathcal{T}$ is a σ-algebra on $[-\infty, \infty]$ (the computation in the proof of 2.39, verbatim), hence closed also under countable intersections and set differences by 2.25. By hypothesis $(a, \infty] \in \mathcal{T}$ for all $a \in \mathbb{R}$, so

$$\{\infty\} = \bigcap_{n=1}^{\infty} (n, \infty] \in \mathcal{T},$$
$$(-\infty, \infty] = \bigcup_{n=1}^{\infty} (-n, \infty] \in \mathcal{T},$$

the second giving $\{-\infty\} \in \mathcal{T}$ by complementation and hence

$$\mathbb{R} = [-\infty, \infty] \setminus \big(\{\infty\} \cup \{-\infty\}\big) \in \mathcal{T}.$$

Consequently $\mathcal{T}_{\mathbb{R}} = \{A \in \mathcal{T} : A \subseteq \mathbb{R}\}$ is a σ-algebra on $\mathbb{R}$: it is closed under countable unions because $\mathcal{T}$ is, and under complementation within $\mathbb{R}$ because $\mathbb{R} \setminus A = \mathbb{R} \cap ([-\infty, \infty] \setminus A) \in \mathcal{T}$, using $\mathbb{R} \in \mathcal{T}$. It contains $(a, \infty) = (a, \infty] \setminus \{\infty\}$ for every $a \in \mathbb{R}$, so by the second half of the proof of 2.39 it contains every Borel subset of $\mathbb{R}$.
Now let $C \subseteq [-\infty, \infty]$ be Borel. By 2.50 the set $B = C \cap \mathbb{R}$ is a Borel subset of $\mathbb{R}$, and C equals one of

$$B, \quad B \cup \{\infty\}, \quad B \cup \{-\infty\}, \quad B \cup \{\infty, -\infty\}.$$

Each of these lies in $\mathcal{T}$, since $B \in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}$, both $\{\infty\}$ and $\{-\infty\}$ lie in $\mathcal{T}$, and $\mathcal{T}$ is closed under finite unions. Hence $f^{-1}(C) \in \mathcal{S}$ for every Borel C, so f is $\mathcal{S}$-measurable.

Problem 2B.22. Suppose $B \subseteq \mathbb{R}$ and $f : B \to \mathbb{R}$ is an increasing function. Prove that f is continuous at every element of B except for a countable subset of B.

Solution. The set D of points of B at which f is discontinuous carries a pairwise disjoint family of nonempty open intervals, hence injects into $\mathbb{Q}$ and is countable.
For $x \in B$ put $\alpha(x) = \sup\{f(t) : t \in B,\ t < x\}$ if $B \cap (-\infty, x) \neq \emptyset$ and $\alpha(x) = f(x)$ otherwise, and dually $\beta(x) = \inf\{f(t) : t \in B,\ t > x\}$ if $B \cap (x, \infty) \neq \emptyset$ and $\beta(x) = f(x)$ otherwise; monotonicity

bounds the first set above by $f(x)$ and the second below by $f(x)$, so $\alpha(x), \beta(x) \in \mathbb{R}$ with $\alpha(x) \leq f(x) \leq \beta(x)$. Call x a left (right) limit point of B when $B \cap (x-\delta, x) \neq \emptyset$ (respectively $B \cap (x, x+\delta) \neq \emptyset$) for every $\delta > 0$. Continuity is relative to the domain B throughout.
Then $x \in D$ if and only if at least one of

- (i) x is a left limit point of B and $\alpha(x) < f(x)$;
- (ii) x is a right limit point of B and $f(x) < \beta(x)$

holds. Indeed, suppose both fail and let $\varepsilon > 0$. If x is not a left limit point, pick $\delta_1 > 0$ with $B \cap (x-\delta_1, x) = \emptyset$, so x is the only point of $B \cap (x-\delta_1, x]$; otherwise $\alpha(x) = f(x)$, and choosing $t_0 \in B$ with $t_0 < x$ and $f(t_0) > f(x) - \varepsilon$ and setting $\delta_1 = x - t_0$ gives $f(x) - \varepsilon < f(t_0) \leq f(t) \leq f(x)$ for all $t \in B \cap (x-\delta_1, x]$. The failure of (ii) yields δ_2 symmetrically, and $\delta = \min\{\delta_1, \delta_2\}$ witnesses continuity at x. Conversely, if (i) holds, take $\varepsilon = f(x) - \alpha(x) > 0$; every $\delta > 0$ admits $t \in B \cap (x-\delta, x)$, and $f(t) \leq \alpha(x) = f(x) - \varepsilon$, so f is discontinuous at x. The case (ii) is symmetric.
For $x \in D$ set $I_x = (\alpha(x), f(x))$ if (i) holds and $I_x = (f(x), \beta(x))$ otherwise (so (ii) holds); either way I_x is a nonempty bounded open interval. For $x < y$ in D we have $\sup I_x \leq \inf I_y$ in all four cases:

- $I_x = (\alpha(x), f(x))$ and $I_y = (\alpha(y), f(y))$: $\alpha(y) \geq f(x)$, as $x \in B$ and $x < y$.
- $I_x = (\alpha(x), f(x))$ and $I_y = (f(y), \beta(y))$: $f(y) \geq f(x)$, as f is increasing.
- $I_x = (f(x), \beta(x))$ and $I_y = (f(y), \beta(y))$: $\beta(x) \leq f(y)$, as $y \in B$ and $y > x$.
- $I_x = (f(x), \beta(x))$ and $I_y = (\alpha(y), f(y))$: here (ii) holds for x, so x is a right limit point and some $s \in B$ has $x < s < y$; then $\beta(x) \leq f(s) \leq \alpha(y)$.

So the I_x are pairwise disjoint, and choosing $q_x \in I_x \cap \mathbb{Q}$ for each $x \in D$ (density of $\mathbb{Q}$) gives an injection $D \to \mathbb{Q}$. Hence D is countable.

Problem 2B.23. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function. Prove that the inverse function $f^{-1} : f(\mathbb{R}) \to \mathbb{R}$ is a continuous function.
[Note that this exercise does not have as a hypothesis that f is continuous.]

Solution. Write $g = f^{-1}$, which is defined on $f(\mathbb{R})$ because a strictly increasing function is injective, and note that $f(x) < f(u)$ forces $x < u$ (otherwise $x \geq u$ would give $f(x) \geq f(u)$).
Fix $y_0 \in f(\mathbb{R})$, put $x_0 = g(y_0)$, and let $\varepsilon > 0$. Strict increase gives

$$f(x_0 - \varepsilon) < f(x_0) = y_0 < f(x_0 + \varepsilon),$$

the two outer values being defined since the domain of f is all of $\mathbb{R}$, so

$$\delta = \min\{\, y_0 - f(x_0 - \varepsilon),\ f(x_0 + \varepsilon) - y_0 \,\} > 0.$$

If $y \in f(\mathbb{R})$ with $|y - y_0| < \delta$, then writing $x = g(y)$,

$$f(x_0 - \varepsilon) \leq y_0 - \delta < f(x) < y_0 + \delta \leq f(x_0 + \varepsilon),$$

so $x_0 - \varepsilon < x < x_0 + \varepsilon$, that is, $|g(y) - g(y_0)| < \varepsilon$. Hence g is continuous at every point of $f(\mathbb{R})$.

Problem 2B.24. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function and $B \subseteq \mathbb{R}$ is a Borel set. Prove that $f(B)$ is a Borel set.

Solution. Let $g = f^{-1} : f(\mathbb{R}) \to \mathbb{R}$, which is defined because f is injective and is continuous by Exercise 23 of this section, and observe that

$$g^{-1}(B) = \{\, y \in f(\mathbb{R}) : g(y) \in B \,\} = \{\, f(x) : x \in B \,\} = f(B),$$

since every $y \in f(\mathbb{R})$ equals $f(x)$ for a unique x, and then $g(y) = x$. So by 2.41 it suffices to show that the domain $f(\mathbb{R})$ of g is a Borel set: then g is Borel measurable and $f(B) = g^{-1}(B)$ is Borel.
Put $f(x^-) = \sup\{f(t) : t < x\}$ and $f(x^+) = \inf\{f(t) : t > x\}$, which are real with $f(x^-) \leq f(x) \leq f(x^+)$, and

$$L = \inf\{f(t) : t \in \mathbb{R}\}, \qquad M = \sup\{f(t) : t \in \mathbb{R}\},$$

with the conventions $(-\infty, L] = \emptyset$ if $L = -\infty$ and $[M, \infty) = \emptyset$ if $M = \infty$. Since every real number is both a left and a right limit point of $\mathbb{R}$, the characterization in Exercise 22 of this section (with $B = \mathbb{R}$, $\alpha = f(\cdot^-)$, $\beta = f(\cdot^+)$) identifies

$$D = \{\, x \in \mathbb{R} : f(x^-) < f(x) \text{ or } f(x) < f(x^+) \,\}$$

as the discontinuity set of f, which that exercise shows is countable. We claim

$$\mathbb{R} \setminus f(\mathbb{R}) = (-\infty, L] \cup [M, \infty) \cup \bigcup_{x \in D} \Big(\big[f(x^-), f(x)\big) \cup \big(f(x), f(x^+)\big] \Big).$$

No set on the right meets $f(\mathbb{R})$. The infimum L is not attained ($f(u) = L$ would give $f(u-1) < L$) and $f(t) \geq L$ for all t, so $(-\infty, L] \cap f(\mathbb{R}) = \emptyset$; symmetrically for $[M, \infty)$. If $f(x) < f(x^+)$ and $f(u) \in (f(x), f(x^+)]$, then $u > x$ (as $u \leq x$ gives $f(u) \leq f(x)$), so $f(u) \geq f(x^+)$ and hence $f(u) = f(x^+)$ — impossible, since $f(u) = f(x^+)$ with $u > x$ would give $f(t) < f(x^+) = \inf\{f(s) : s > x\}$ for $x < t < u$. So $(f(x), f(x^+)] \cap f(\mathbb{R}) = \emptyset$, and symmetrically $[f(x^-), f(x)) \cap f(\mathbb{R}) = \emptyset$.
Conversely, let $y \notin f(\mathbb{R})$ with $L < y < M$ (otherwise y lies in one of the first two sets). Then $T = \{t \in \mathbb{R} : f(t) < y\}$ is nonempty because $y > L$, and bounded above by any s with $f(s) > y$ (such s exists as $y < M$, and $t \geq s$ would give $f(t) \geq f(s) > y$), so $x = \sup T \in \mathbb{R}$. For $t < x$ some $t' \in T$ has $t < t' \leq x$, so $f(t) \leq f(t') < y$ and hence $f(x^-) \leq y$; for $t > x$ we have $t \notin T$, so $f(t) \geq y$, indeed $f(t) > y$ as $y \notin f(\mathbb{R})$, giving $f(x^+) \geq y$. Also $f(x) \neq y$.
- (i) $f(x) > y$: then $y \in [f(x^-), f(x))$ and $f(x^-) < f(x)$, so $x \in D$.
- (ii) $f(x) < y$: then $y \in (f(x), f(x^+)]$ and $f(x) < f(x^+)$, so $x \in D$.

The right side is a countable union of intervals, hence Borel, so $f(\mathbb{R})$ is Borel.

Problem 2B.25. Suppose $B \subseteq \mathbb{R}$ and $f : B \to \mathbb{R}$ is an increasing function. Prove that there exists a sequence $f_1, f_2, \ldots$ of strictly increasing functions from B to $\mathbb{R}$ such that

$$f(x) = \lim_{k \to \infty} f_k(x)$$

for every $x \in B$.

Solution. Take $f_k : B \to \mathbb{R}$ defined by

$$f_k(x) = f(x) + \frac{x}{k} \qquad \text{for } k \in \mathbb{Z}^+.$$

If $x, y \in B$ with $x < y$, then $f(x) \leq f(y)$ because f is increasing and $x/k < y/k$ because $k > 0$, so

$$f_k(x) = f(x) + \frac{x}{k} < f(y) + \frac{y}{k} = f_k(y);$$

thus each f_k is strictly increasing. For fixed $x \in B$ we have $x/k \to 0$, so $f_k(x) \to f(x)$.

Problem 2B.26. Suppose $B \subseteq \mathbb{R}$ and $f : B \to \mathbb{R}$ is a bounded increasing function. Prove that there exists an increasing function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = f(x)$ for all $x \in B$.

Solution. Take $g = 0$ if $B = \emptyset$, and otherwise

$$g(x) = \begin{cases} m & \text{if } B \cap (-\infty, x] = \emptyset, \\ \sup\{\, f(t) : t \in B,\ t \leq x \,\} & \text{if } B \cap (-\infty, x] \neq \emptyset, \end{cases}$$

where $m = \inf f(B)$ and $M = \sup f(B)$ are real because f is bounded and $B \neq \emptyset$. In the second case the set is a nonempty subset of $f(B)$, so its supremum lies in $[m, M] \subseteq \mathbb{R}$.
g is increasing: let $x < y$. If $B \cap (-\infty, x] = \emptyset$, then $g(x) = m \leq g(y)$, since $g(y)$ is either m or a supremum of a subset of $f(B)$, every element of which is at least m. Otherwise $B \cap (-\infty, y] \neq \emptyset$ as

well and

$$\{\, f(t) : t \in B,\ t \le x \,\} \subseteq \{\, f(t) : t \in B,\ t \le y \,\},$$

so $g(x) \le g(y)$.
g extends f: for $x \in B$ the second case applies, and $f(x)$ both belongs to $\{f(t) : t \in B,\ t \le x\}$ and bounds it above (as f is increasing), so $g(x) = f(x)$.

Problem 2B.27. Prove or give a counterexample: If $(X, \mathcal{S})$ is a measurable space and

$$f : X \to [-\infty, \infty]$$

is a function such that $f^{-1}\big((a, \infty)\big) \in \mathcal{S}$ for every $a \in \mathbb{R}$, then f is an $\mathcal{S}$-measurable function.

Solution. False. Take $X = \{1, 2\}$, $\mathcal{S} = \{\emptyset, X\}$, and $f : X \to [-\infty, \infty]$ given by

$$f(1) = \infty, \qquad f(2) = -\infty.$$

Neither value lies in $\mathbb{R}$, so $f^{-1}\big((a, \infty)\big) = \emptyset \in \mathcal{S}$ for every $a \in \mathbb{R}$. But $\{\infty\}$ is a Borel subset of $[-\infty, \infty]$ by 2.50, since $\{\infty\} \cap \mathbb{R} = \emptyset$ is Borel in $\mathbb{R}$, while

$$f^{-1}\big(\{\infty\}\big) = \{1\} \notin \mathcal{S}.$$

So f is not $\mathcal{S}$-measurable in the sense of 2.51.

Problem 2B.28. Suppose $f : B \to \mathbb{R}$ is a Borel measurable function. Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \begin{cases} f(x) & \text{if } x \in B, \\ 0 & \text{if } x \in \mathbb{R} \setminus B. \end{cases}$$

Prove that g is a Borel measurable function.

Solution. Because $g = f$ on B and $g = 0$ off B, every Borel set $A \subseteq \mathbb{R}$ satisfies

$$g^{-1}(A) = \begin{cases} f^{-1}(A) \cup (\mathbb{R} \setminus B) & \text{if } 0 \in A, \\ f^{-1}(A) & \text{if } 0 \notin A, \end{cases}$$

noting $f^{-1}(A) \subseteq B$ since B is the domain of f. Here $f^{-1}(A)$ is Borel because f is Borel measurable, and $\mathbb{R} \setminus B$ is Borel because $B = f^{-1}(\mathbb{R})$ is Borel (as observed after 2.40) and the Borel sets form a σ-algebra. So $g^{-1}(A)$ is Borel in both cases, and hence g is Borel measurable.

Problem 2B.29. Give an example of a measurable space $(X, \mathcal{S})$ and a family $\{f_t\}_{t \in \mathbb{R}}$ such that each f_t is an $\mathcal{S}$-measurable function from X to $[0, 1]$, but the function $f : X \to [0, 1]$ defined by

$$f(x) = \sup\{f_t(x) : t \in \mathbb{R}\}$$

is not $\mathcal{S}$-measurable.
[Compare this exercise to 2.53, where the index set is $\mathbb{Z}^+$ rather than $\mathbb{R}$.]

Solution. Take $X = \mathbb{R}$ with

$$\mathcal{S} = \{E \subseteq \mathbb{R} : E \text{ or } \mathbb{R} \setminus E \text{ is countable}\},$$

a σ-algebra by the first bullet point of 2.28, and for $t \in \mathbb{R}$ take

$$f_t = \chi_{\{t\}} \text{ if } t \in [0, 1], \qquad f_t = 0 \text{ if } t \notin [0, 1],$$

each mapping $\mathbb{R}$ into $\{0,1\} \subseteq [0,1]$. Each f_t is $\mathcal{S}$-measurable, since $f_t^{-1}(B)$ is one of $\emptyset$, $\{t\}$, $\mathbb{R} \setminus \{t\}$, $\mathbb{R}$, all of which are countable or cocountable.
The pointwise supremum is $f = \chi_{[0,1]}$: for $x \in [0,1]$ the index $t = x$ gives $f_x(x) = 1$, the largest value available, while for $x \notin [0,1]$ every $f_t(x) = 0$ (either $f_t \equiv 0$ or $t \in [0,1]$, so $t \neq x$). But

$$f^{-1}\big((\tfrac{1}{2}, \infty)\big) = [0,1] \notin \mathcal{S},$$

since $[0,1]$ is uncountable and so is its complement, which contains $(1,2)$. Hence f is not $\mathcal{S}$-measurable.

Problem 2B.30. Show that

$$\lim_{j\to\infty} \left(\lim_{k\to\infty} \big(\cos(j!\,\pi x)\big)^{2k} \right) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

for every $x \in \mathbb{R}$.
[This example is due to Henri Lebesgue.]

Solution. Fix $x \in \mathbb{R}$ and $j \in \mathbb{Z}^+$, and set $c = \cos(j!\,\pi x) \in [-1,1]$, so that $c^{2k} = (c^2)^k$ with $c^2 \in [0,1]$. Since $1^k = 1$ for all k while $0 \le r < 1$ forces $r^k \to 0$, the inner limit

$$g_j(x) = \lim_{k\to\infty} \big(\cos(j!\,\pi x)\big)^{2k} = \begin{cases} 1 & \text{if } c^2 = 1, \\ 0 & \text{if } c^2 < 1 \end{cases}$$

exists for every j and x. Because $|\cos\theta| = 1$ exactly when θ is an integer multiple of π,

$$|\cos(j!\,\pi x)| = 1 \iff j!\,\pi x \in \{m\pi : m \in \mathbb{Z}\} \iff j!\,x \in \mathbb{Z},$$

so $g_j = \chi_{A_j}$ with $A_j = \{m/j! : m \in \mathbb{Z}\}$. Now take $j \to \infty$.

- (i) x irrational: $j!\,x \in \mathbb{Z}$ would make $x = (j!\,x)/j!$ rational, so $g_j(x) = 0$ for every j and the outer limit is 0.
- (ii) $x = p/q$ with $p \in \mathbb{Z}$, $q \in \mathbb{Z}^+$: for $j \ge q$ the integer q is a factor of $j!$, so $j!\,x = p \cdot (j!/q) \in \mathbb{Z}$ and $g_j(x) = 1$; the outer limit is 1.

2.3 Exercises 2C

Problem 2C.1. Explain why there does not exist a measure space $(X, \mathcal{S}, \mu)$ with the property that $\{\mu(E) : E \in \mathcal{S}\} = [0,1)$.

Solution. The set $\{\mu(E) : E \in \mathcal{S}\}$ always contains its own upper bound $\mu(X)$, whereas $[0,1)$ has no largest element.
Explicitly, if $\{\mu(E) : E \in \mathcal{S}\} = [0,1)$, then $\mu(X) < 1$ since $X \in \mathcal{S}$, so there is t with $\mu(X) < t < 1$, and $t = \mu(E)$ for some $E \in \mathcal{S}$. But $E \subseteq X$, so 2.57(a) gives

$$t = \mu(E) \le \mu(X) < t,$$

a contradiction.

Problem 2C.2. Let $2^{\mathbb{Z}^+}$ denote the σ-algebra on $\mathbb{Z}^+$ consisting of all subsets of $\mathbb{Z}^+$.
Suppose μ is a measure on $(\mathbb{Z}^+, 2^{\mathbb{Z}^+})$. Prove that there is a sequence $w_1, w_2, \ldots$ in $[0,\infty]$ such that

$$\mu(E) = \sum_{k\in E} w_k$$

for every set $E \subseteq \mathbb{Z}^+$.

Solution. Take
$$w_k = \mu(\{k\}) \in [0,\infty] \qquad \text{for } k \in \mathbf{Z}^+.$$
As in 2.55, $\sum_{k\in E} w_k$ denotes the supremum of the subsums $\sum_{k\in D} w_k$ over finite $D \subseteq E$. Fix $E \subseteq \mathbf{Z}^+$.

- (i) E finite, say $E = \{k_1, \dots, k_n\}$: finite additivity (from countable additivity by appending $\emptyset, \emptyset, \dots$, as noted after 2.54) applied to the disjoint singletons gives $\mu(E) = \sum_{j=1}^n w_{k_j}$, and the supremum over finite $D \subseteq E$ is attained at $D = E$ since the w_k are nonnegative. (For $E = \emptyset$ both sides are 0.)
- (ii) E infinite, listed increasingly as $k_1 < k_2 < \cdots$: countable additivity applied to $E = \bigcup_{j=1}^\infty \{k_j\}$ gives $\mu(E) = \sum_{j=1}^\infty w_{k_j} = \sup_n \sum_{j=1}^n w_{k_j}$, the partial sums being increasing in $[0,\infty]$. That supremum equals the supremum over all finite $D \subseteq E$: each $\{k_1, \dots, k_n\}$ is one such D, and conversely each finite $D \subseteq E$ lies inside some $\{k_1, \dots, k_n\}$, whose subsum is then at least that of D.

Hence $\mu(E) = \sum_{k\in E} w_k$ for every $E \subseteq \mathbf{Z}^+$.

Problem 2C.3. Give an example of a measure μ on $(\mathbf{Z}^+, 2^{\mathbf{Z}^+})$ such that
$$\{\mu(E) : E \subseteq \mathbf{Z}^+\} = [0,1].$$

Solution. Take
$$\mu(E) = \sum_{k\in E} 2^{-k} \qquad \text{for } E \subseteq \mathbf{Z}^+,$$
a measure by the third bullet point of 2.55.
One inclusion: $0 \le \mu(E) \le \mu(\mathbf{Z}^+) = \sum_{k=1}^\infty 2^{-k} = 1$, using 2.57(a).
For the other, let $t \in [0,1]$; take $E = \mathbf{Z}^+$ if $t = 1$, so assume $0 \le t < 1$. Define $b_k \in \{0,1\}$ recursively by $b_k = 1$ if $s_{k-1} + 2^{-k} \le t$ and $b_k = 0$ otherwise, where $s_k = \sum_{j=1}^k b_j 2^{-j}$ and $s_0 = 0$. Then
$$s_k \le t < s_k + 2^{-k} \qquad \text{for every } k \ge 0,$$
by induction: at $k = 0$ this reads $0 \le t < 1$; granting it at $k-1$, if $b_k = 1$ then $s_k = s_{k-1} + 2^{-k} \le t$ and $t < s_{k-1} + 2^{-(k-1)} = s_k + 2^{-k}$, while if $b_k = 0$ then $s_k = s_{k-1} \le t < s_{k-1} + 2^{-k} = s_k + 2^{-k}$. Hence $0 \le t - s_k < 2^{-k}$, so $s_k \to t$.
Now put $E = \{k \in \mathbf{Z}^+ : b_k = 1\}$, so s_k is the subsum of $\sum_{k\in E} 2^{-k}$ over the finite set $E \cap \{1, \dots, k\}$; thus $s_k \le \mu(E)$, while every finite $D \subseteq E$ satisfies $D \subseteq E \cap \{1, \dots, k\}$ for large k and hence $\sum_{j\in D} 2^{-j} \le s_k$. Therefore
$$\mu(E) = \sup_k s_k = \lim_{k\to\infty} s_k = t,$$
the suprema and limits agreeing because $s_1 \le s_2 \le \cdots$.

Problem 2C.4. Give an example of a measure space $(X, \mathcal{S}, \mu)$ such that
$$\{\mu(E) : E \in \mathcal{S}\} = \{\infty\} \cup \bigcup_{k=0}^\infty [3k, 3k+1].$$

Solution. Take $X = \mathbf{Z}^+ \times \{1,2\}$ with $\mathcal{S}$ all subsets of X and $\mu(E) = \sum_{x\in E} w(x)$, where
$$w(k,1) = 2^{-k}, \qquad w(k,2) = 3 \qquad \text{for } k \in \mathbf{Z}^+;$$
this is a measure by the third bullet point of 2.55 (the sums being suprema of finite subsums).
Writing $E_1 = \{k : (k,1) \in E\}$ and $E_2 = \{k : (k,2) \in E\}$, finite additivity applied to the disjoint pieces $E_1 \times \{1\}$ and $E_2 \times \{2\}$ gives
$$\mu(E) = s + b, \qquad s = \sum_{k\in E_1} 2^{-k}, \quad b = \sum_{k\in E_2} 3,$$
where $b = 3n$ if E_2 has exactly n elements and $b = \infty$ if E_2 is infinite, the finite subsums $3, 6, 9, \dots$ being unbounded. By Exercise 3 of this section, s ranges over exactly $[0,1]$ as E_1 ranges over the subsets of

$\mathbb{Z}^+$; so b ranges over exactly $\{0, 3, 6, \ldots\} \cup \{\infty\}$, and the two choices are independent because every pair (E_1, E_2) arises from $E = (E_1 \times \{1\}) \cup (E_2 \times \{2\})$. Since $3k + s$ sweeps $[3k, 3k+1]$ as s sweeps $[0, 1]$, while $b = \infty$ forces $\mu(E) = \infty$,

$$\{\mu(E) : E \in \mathcal{S}\} = \{\infty\} \cup \bigcup_{k=0}^{\infty} [3k, 3k+1].$$

Problem 2C.5. Suppose $(X, \mathcal{S}, \mu)$ is a measure space such that $\mu(X) < \infty$. Prove that if $\mathcal{A}$ is a set of disjoint sets in $\mathcal{S}$ such that $\mu(A) > 0$ for every $A \in \mathcal{A}$, then $\mathcal{A}$ is a countable set.

Solution. Write $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$, where $\mathcal{A}_n = \{A \in \mathcal{A} : \mu(A) > \frac{1}{n}\}$; this is an equality because every $A \in \mathcal{A}$ has $\mu(A) > 0$ and hence $\mu(A) > \frac{1}{n}$ for some n by the Archimedean property.
Each $\mathcal{A}_n$ is finite: if $A_1, \ldots, A_m$ are distinct elements of $\mathcal{A}_n$, they are pairwise disjoint sets in $\mathcal{S}$, so finite additivity (from countable additivity, as noted after 2.54) together with $A_1 \cup \cdots \cup A_m \subseteq X$ and 2.57(a) gives

$$\frac{m}{n} < \sum_{j=1}^{m} \mu(A_j) = \mu(A_1 \cup \cdots \cup A_m) \le \mu(X) < \infty,$$

so $m < n\,\mu(X)$; were $\mathcal{A}_n$ infinite it would supply m distinct elements with $m \ge n\,\mu(X)$. Hence $\mathcal{A}$ is a countable union of finite sets, so countable.

Problem 2C.6. Find all $c \in [3, \infty)$ such that there exists a measure space $(X, \mathcal{S}, \mu)$ with

$$\{\mu(E) : E \in \mathcal{S}\} = [0, 1] \cup [3, c].$$

Solution. Only $c = 4$.
Necessity. Suppose $R := \{\mu(E) : E \in \mathcal{S}\} = [0,1] \cup [3, c]$. Since $c \in R$, say $c = \mu(E_0)$, and $E_0 \subseteq X$ gives $c \le \mu(X)$ by 2.57(a), while $\mu(X) \in R$ gives $\mu(X) \le c$, we get $\mu(X) = c < \infty$. So for any $t = \mu(E) \in R$ we have $\mu(E) \le c < \infty$, and 2.57(b) applied to $E \subseteq X$ yields

$$\mu(X \setminus E) = \mu(X) - \mu(E) = c - t,$$

giving the reflection property $t \in R \implies c - t \in R$.

- (i) $t = 1$: $c - 1 \in R$, and $c - 1 \ge 2 > 1$ puts it in $[3, c]$, so $c \ge 4$.
- (ii) $c \ge 5$ is impossible: then $c - 2 \in [3, c] \subseteq R$, so reflection gives $2 \in R$, false.
- (iii) $t = 3$: $c - 3 \in R$, and $4 \le c < 5$ gives $1 \le c - 3 < 2$, whose only element in R is 1. So $c = 4$.

Sufficiency. Take $X = \{0, 1, 2, \ldots\}$ with $\mathcal{S}$ all subsets of X and $\mu(E) = \sum_{x \in E} w(x)$, where $w(0) = 3$ and $w(k) = 2^{-k}$ for $k \in \mathbb{Z}^+$; this is a measure by the third bullet point of 2.55. Splitting E into $E \cap \{0\}$ and $E' = E \cap \mathbb{Z}^+$, finite additivity gives $\mu(E) = \sum_{k \in E'} 2^{-k}$ if $0 \notin E$ and $\mu(E) = 3 + \sum_{k \in E'} 2^{-k}$ if $0 \in E$. By Exercise 3 of this section that sum ranges over exactly $[0, 1]$, independently of whether $0 \in E$, so

$$\{\mu(E) : E \in \mathcal{S}\} = [0, 1] \cup \bigl(3 + [0, 1]\bigr) = [0, 1] \cup [3, 4].$$

Problem 2C.7. Give an example of a measure space $(X, \mathcal{S}, \mu)$ such that

$$\{\mu(E) : E \in \mathcal{S}\} = [0, 1] \cup [3, \infty].$$

Solution. Take $X = \mathbb{Z}^+ \times \{1, 2\}$ with $\mathcal{S}$ all subsets of X and $\mu(E) = \sum_{x \in E} w(x)$, where

$$w(k, 1) = 2^{-k}, \qquad w(k, 2) = k + 2 \qquad \text{for } k \in \mathbb{Z}^+;$$

this is a measure by the third bullet point of 2.55.
With $E_1 = \{k : (k, 1) \in E\}$ and $E_2 = \{k : (k, 2) \in E\}$, finite additivity applied to the disjoint pieces

$E_1 \times \{1\}$ and $E_2 \times \{2\}$ gives $\mu(E) = s + b$, where $s = \sum_{k \in E_1} 2^{-k}$ and $b = \sum_{k \in E_2}(k+2)$; the two choices are independent, since every pair (E_1, E_2) arises from $E = (E_1 \times \{1\}) \cup (E_2 \times \{2\})$. By Exercise 3 of this section s ranges over exactly $[0, 1]$, and the set B of possible values of b satisfies:

- (i) $B \subseteq \{0\} \cup [3, \infty]$, since $E_2 = \emptyset$ gives $b = 0$ while any $k \in E_2$ gives $b \geq k + 2 \geq 3$;
- (ii) every integer $n \geq 3$ lies in B, via $E_2 = \{n - 2\}$;
- (iii) $\infty \in B$, via $E_2 = \mathbb{Z}^+$, whose finite subsums $\sum_{k=1}^{m}(k + 2)$ are unbounded.

So by (i) each $\mu(E)$ is either $s \in [0, 1]$ or $s + b \geq 3$, putting the range inside $[0, 1] \cup [3, \infty]$. Conversely, $t \in [0, 1]$ arises from $E_2 = \emptyset$ with E_1 chosen by Exercise 3 so that $s = t$; $t = \infty$ arises from $E_2 = \mathbb{Z}^+$ by (iii); and $t \in [3, \infty)$ arises from $E_2 = \{n - 2\}$ with $n = \lfloor t \rfloor \geq 3$, giving $b = n$ by (ii), together with E_1 chosen so that $s = t - n \in [0, 1)$. Hence

$$\{\mu(E) : E \in \mathcal{S}\} = [0, 1] \cup [3, \infty].$$

Problem 2C.8. Give an example of a set X, a σ-algebra $\mathcal{S}$ of subsets of X, a set $\mathcal{A}$ of subsets of X such that the smallest σ-algebra on X containing $\mathcal{A}$ is $\mathcal{S}$, and two measures μ and ν on $(X, \mathcal{S})$ such that $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$ and $\mu(X) = \nu(X) < \infty$, but $\mu \neq \nu$.

Solution. Take

$$X = \{1, 2, 3, 4\}, \qquad \mathcal{S} = \{E : E \subseteq X\}, \qquad \mathcal{A} = \big\{\{1, 2\}, \{2, 3\}\big\},$$

and let μ and ν be the measures on $(X, \mathcal{S})$ defined by

$$\mu = \delta_2 + \delta_4, \qquad \nu = \delta_1 + \delta_3,$$

where δ_c is the Dirac measure at c (2.55), so that μ and ν are measures by Exercise 9 of this section. Explicitly, for $E \subseteq X$,

$$\mu(E) = \#\big(E \cap \{2, 4\}\big), \qquad \nu(E) = \#\big(E \cap \{1, 3\}\big).$$

Here $\mathcal{S}$ is the smallest σ-algebra on X containing $\mathcal{A}$: any σ-algebra $\mathcal{T} \supseteq \mathcal{A}$ is closed under finite intersections and set differences, so

$$\begin{aligned} &\{2\} = \{1, 2\} \cap \{2, 3\}, \qquad \{1\} = \{1, 2\} \setminus \{2\}, \\ &\{3\} = \{2, 3\} \setminus \{2\} \end{aligned}$$

lie in $\mathcal{T}$, hence so does $\{4\} = X \setminus (\{1\} \cup \{2\} \cup \{3\})$; as X is finite, every subset is a finite union of singletons and $\mathcal{T} = \mathcal{S}$.
Finally μ and ν agree on $\mathcal{A}$ and on X but differ at $\{1\}$:

$$\mu(\{1, 2\}) = \nu(\{1, 2\}) = \mu(\{2, 3\}) = \nu(\{2, 3\}) = 1, \qquad \mu(X) = \nu(X) = 2,$$

while $\mu(\{1\}) = 0 \neq 1 = \nu(\{1\})$.

Problem 2C.9. Suppose μ and ν are measures on a measurable space $(X, \mathcal{S})$. Prove that $\mu + \nu$ is a measure on $(X, \mathcal{S})$. [Here $\mu + \nu$ is the usual sum of two functions: if $E \in \mathcal{S}$, then $(\mu + \nu)(E) = \mu(E) + \nu(E)$.]

Solution. The map $\mu + \nu$ sends $\mathcal{S}$ into $[0, \infty]$, which is closed under addition with the convention $a + \infty = \infty$, and $(\mu + \nu)(\varnothing) = 0 + 0 = 0$; only countable additivity needs work.
Series in $[0, \infty]$ add termwise: for sequences (a_k), (b_k) in $[0, \infty]$,

$$\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k,$$

each series being the supremum of its increasing partial sums. Indeed, writing A_n, B_n, C_n for the n-th partial sums and A, B, C for their suprema, rearranging a finite sum gives $C_n = A_n + B_n$. If $A = \infty$, then $C_n \geq A_n$ is unbounded, so $C = \infty = A + B$, and symmetrically if $B = \infty$. If $A, B < \infty$, then $C_n \leq A + B$ gives $C \leq A + B$, while for $\varepsilon > 0$ a single n with $A_n > A - \varepsilon/2$ and $B_n > B - \varepsilon/2$ (monotonicity) gives $C_n > A + B - \varepsilon$; hence $C = A + B$.
So for a disjoint sequence $E_1, E_2, \ldots$ in $\mathcal{S}$, countable additivity of μ and of ν gives

$$\begin{aligned}(\mu+\nu)\Big(\bigcup_{k=1}^{\infty} E_k\Big) &= \sum_{k=1}^{\infty} \mu(E_k) + \sum_{k=1}^{\infty} \nu(E_k) \\ &= \sum_{k=1}^{\infty} \big(\mu(E_k) + \nu(E_k)\big) = \sum_{k=1}^{\infty} (\mu+\nu)(E_k).\end{aligned}$$

Hence $\mu + \nu$ is a measure on $(X, \mathcal{S})$.

Problem 2C.10. Give an example of a measure space $(X, \mathcal{S}, \mu)$ and a decreasing sequence $E_1 \supseteq E_2 \supseteq \cdots$ of sets in $\mathcal{S}$ such that

$$\mu\Big(\bigcap_{k=1}^{\infty} E_k\Big) \neq \lim_{k\to\infty} \mu(E_k).$$

Solution. Take counting measure μ on $X = \mathbf{Z}^+$ with $\mathcal{S}$ the σ-algebra of all subsets of X (a measure by 2.55), and

$$E_k = \{k, k+1, k+2, \ldots\} \quad \text{for } k \in \mathbf{Z}^+,$$

a decreasing sequence in $\mathcal{S}$. Each E_k is infinite, so $\mu(E_k) = \infty$ for every k, while $\bigcap_{k=1}^{\infty} E_k = \varnothing$ because $n \notin E_{n+1}$ for every $n \in \mathbf{Z}^+$. Hence

$$\mu\Big(\bigcap_{k=1}^{\infty} E_k\Big) = 0 \neq \infty = \lim_{k\to\infty} \mu(E_k).$$

Problem 2C.11. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $C, D, E \in \mathcal{S}$ are such that

$$\mu(C \cap D) < \infty, \qquad \mu(C \cap E) < \infty, \qquad \mu(D \cap E) < \infty.$$

Find and prove a formula for $\mu(C \cup D \cup E)$ in terms of $\mu(C)$, $\mu(D)$, $\mu(E)$, $\mu(C \cap D)$, $\mu(C \cap E)$, $\mu(D \cap E)$, and $\mu(C \cap D \cap E)$.

Solution. The formula is inclusion-exclusion for three sets:

$$\begin{aligned}\mu(C \cup D \cup E) &= \mu(C) + \mu(D) + \mu(E) \\ &\qquad - \mu(C \cap D) - \mu(C \cap E) - \mu(D \cap E) + \mu(C \cap D \cap E).\end{aligned}$$

Write $p = \mu(C \cap D)$, $q = \mu(C \cap E)$, $r = \mu(D \cap E)$, $s = \mu(C \cap D \cap E)$. The first three are finite by hypothesis and $s \leq p < \infty$ by 2.57(a), so only finite quantities are subtracted above (convention $\infty - t = \infty$ for real t) and the formula is equivalent to the subtraction-free identity

$$\mu(C \cup D \cup E) + p + q + r = \mu(C) + \mu(D) + \mu(E) + s.$$

Since $p < \infty$, 2.61 applied to C and D gives $\mu(C \cup D) + p = \mu(C) + \mu(D)$. Since $(C \cap E) \cap (D \cap E) = C \cap D \cap E$ has measure $s < \infty$, 2.61 applied to $C \cap E$ and $D \cap E$, whose union is $(C \cup D) \cap E$, gives

$$\mu\big((C \cup D) \cap E\big) + s = q + r;$$

in particular $\mu\big((C \cup D) \cap E\big) \leq q + r < \infty$, so 2.61 applied to $C \cup D$ and E gives

$$\mu(C \cup D \cup E) + \mu\big((C \cup D) \cap E\big) = \mu(C \cup D) + \mu(E).$$

Adding s and then p to this last identity and substituting the previous two yields

$$\mu(C \cup D \cup E) + p + q + r = \mu(C) + \mu(D) + \mu(E) + s,$$

every rearrangement being legitimate because addition on $[0, \infty]$ is commutative and associative. Subtracting the finite quantity $p + q + r$ gives the formula.

Problem 2C.12. Suppose X is a set and $\mathcal{S}$ is the σ-algebra of all subsets E of X such that E is countable or $X \setminus E$ is countable. Give a complete description of the set of all measures on $(X, \mathcal{S})$.

Solution. The measures on $(X, \mathcal{S})$ are exactly the functions

$$\mu_{w,c}(E) = \sum_{x \in E} w(x) \; + \; c \cdot [E \text{ uncountable}],$$

with $w : X \to [0, \infty]$ and $c \in [0, \infty]$ arbitrary: a weighted counting measure with weights $w(x) = \mu(\{x\})$, plus a lump of mass c carried by every uncountable set in $\mathcal{S}$. Here $\sum_{x \in A} w(x)$ is the supremum of the finite subsums, as in the third bullet of 2.55; $[E \text{ uncountable}]$ is 1 if E is uncountable and 0 otherwise; and $0 \cdot \infty = 0$, so a countable E carries no lump even when $c = \infty$.

Three facts carry the proof.

(L1) Unordered sums are countably additive: if $A = \bigcup_{k=1}^{\infty} A_k$ with the A_k disjoint, then $\sum_{x \in A} w(x) = \sum_{k=1}^{\infty} \sum_{x \in A_k} w(x)$. Indeed a finite $D \subseteq A$ meets only finitely many A_k, say those with $k \leq n$, and is the disjoint union of the $D \cap A_k$, so $\sum_{x \in D} w(x) \leq \sum_{k=1}^{n} \sum_{x \in A_k} w(x)$; taking the supremum over D gives $\leq$. Conversely finite sets $D_k \subseteq A_k$ for $k \leq n$ have union a finite subset of A, so $\sum_{k=1}^{n} \sum_{x \in D_k} w(x) \leq \sum_{x \in A} w(x)$; taking the n suprema one at a time (legitimate since addition on $[0, \infty]$ is order preserving and commutes with suprema of increasing families in one variable) and letting $n \to \infty$ gives $\geq$.

(L2) If N is uncountable and $w > 0$ on N, then $\sum_{x \in N} w(x) = \infty$: from $N = \bigcup_{n=1}^{\infty} \{x \in N : w(x) > 1/n\}$ and the countability of a countable union of countable sets, some $\{x \in N : w(x) > 1/n\}$ is uncountable, hence contains finite sets D of arbitrarily many elements, with $\sum_{x \in D} w(x) > \#D/n$.

(L3) If $E_1, E_2, \dots \in \mathcal{S}$ are disjoint with union E, then E is uncountable if and only if exactly one E_k is. At most one can be, since an uncountable set of $\mathcal{S}$ has countable complement while $E_i \cap E_j = \varnothing$ forces $E_i \subseteq X \setminus E_j$; and E is countable when all E_k are, uncountable when some E_k is.

Each $\mu_{w,c}$ is a measure: it maps $\mathcal{S}$ into $[0, \infty]$ with $\mu_{w,c}(\varnothing) = 0$, and for disjoint $E_1, E_2, \dots$ with union E, (L3) gives $[E \text{ uncountable}] = \sum_{k=1}^{\infty} [E_k \text{ uncountable}]$ with at most one nonzero term, so this may be multiplied by c termwise unambiguously even when $c = \infty$; combining that with (L1) and with termwise addition of series in $[0, \infty]$ (Exercise 9 of this section),

$$\mu_{w,c}(E) = \sum_{k=1}^{\infty} \Big(\sum_{x \in E_k} w(x) + c \, [E_k \text{ uncountable}] \Big) = \sum_{k=1}^{\infty} \mu_{w,c}(E_k).$$

Conversely, let μ be a measure, put $w(x) = \mu(\{x\})$ (singletons are countable, hence in $\mathcal{S}$) and $N = \{x \in X : w(x) > 0\}$. For countable $E \in \mathcal{S}$, countable additivity along the disjoint singletons $\{x_1\}, \{x_2\}, \dots$ exhausting E gives $\mu(E) = \sum_k w(x_k)$, which equals $\sum_{x \in E} w(x)$ because the partial sums are the subsums over $\{x_1, \dots, x_n\}$ and these are cofinal among finite subsets of E. Also $\mu(E) \geq \mu(D) = \sum_{x \in D} w(x)$ for finite $D \subseteq E$ by 2.57(a) and finite additivity, so $\mu(E) \geq \sum_{x \in E} w(x)$ for every $E \in \mathcal{S}$.

- (i) N uncountable. If $G \in \mathcal{S}$ is uncountable then $X \setminus G$ is countable, so $G \cap N = N \setminus (X \setminus G)$ is uncountable and (L2) gives $\sum_{x \in G} w(x) = \infty$, whence $\mu(G) = \infty$ by the inequality above. So $\mu = \mu_{w,0}$.
- (ii) N countable. Put $c = \mu(X \setminus N)$. For uncountable $G \in \mathcal{S}$, the sets $G \cap N$ and $G \setminus N$ lie in $\mathcal{S}$ (the first countable, the second with countable complement $(X \setminus G) \cup N$) and partition G, so $\mu(G) = \mu(G \cap N) + \mu(G \setminus N)$. Here $\mu(G \cap N) = \sum_{x \in G \cap N} w(x) = \sum_{x \in G} w(x)$, as w vanishes off N; and splitting $X \setminus N$ into $G \setminus N$ and $(X \setminus N) \cap (X \setminus G)$, the latter countable with $w \equiv 0$ on it and hence of measure 0, gives $\mu(G \setminus N) = c$. So $\mu = \mu_{w,c}$.

2.4 Exercises 2D

Problem 2D.1. (a) Show that the set consisting of those numbers in $(0,1)$ that have a decimal expansion containing one hundred consecutive 4s is a Borel subset of $\mathbb{R}$.
(b) What is the Lebesgue measure of the set in (a)?

Solution. The set E of (a) is a countable union of closed intervals intersected with $(0,1)$, and $|E| = 1$. For a digit string $s = (d_1, \ldots, d_m)$ put $a_s = \sum_{j=1}^{m} d_j 10^{-j}$ and $I_s = [a_s, a_s + 10^{-m}]$. Then $x \in [0,1]$ has a decimal expansion beginning with s if and only if $x \in I_s$, since the tail $\sum_{j>m} d_j 10^{-j}$ ranges over exactly $[0, 10^{-m}]$. (Check!)
(a). Let F be the string of one hundred 4s and let pF denote concatenation. A hundred consecutive 4s must begin in some position n, preceded by some $p \in \{0, \ldots, 9\}^{n-1}$, so

$$E = (0,1) \cap \bigcup_{n=1}^{\infty} \bigcup_{p \in \{0,\ldots,9\}^{n-1}} I_{pF}.$$

Only 10^{n-1} strings p occur at each n, so this is a countable union of closed intervals; hence E is Borel.
(b). Fix N and call $s \in \{0, \ldots, 9\}^{100N}$ *good* if none of its N blocks of length 100 is a hundred 4s; there are $(10^{100} - 1)^N$ good strings, one choice per block. If $x \in (0,1) \setminus E$, the first $100N$ digits of any expansion of x form a good s (else $x \in E$), and then $x \in I_s$; so $(0,1) \setminus E$ is covered by the good I_s, whence by 2.5, 2.8, and $|I_s| = 10^{-100N}$ (2.14),

$$|(0,1) \setminus E| \le (10^{100} - 1)^N \cdot 10^{-100N} = \left(1 - 10^{-100}\right)^N.$$

Letting $N \to \infty$ gives $|(0,1) \setminus E| = 0$. Now $E \subseteq [0,1]$ gives $|E| \le 1$ by 2.5 and 2.14, while $[0,1] = E \cup ((0,1) \setminus E) \cup \{0,1\}$ with $|\{0,1\}| = 0$ (2.4) gives $1 \le |E|$ by 2.8. Hence $|E| = 1$.

Problem 2D.2. Prove that there exists a bounded set $A \subseteq \mathbb{R}$ such that $|F| \le |A| - 1$ for every closed set $F \subseteq A$.

Solution. Take $A = \frac{1}{c}V$, where $V \subseteq [-1,1]$ is the set constructed in the proof of 2.18 and $c = |V|$.
Two facts from that proof: $|V| > 0$, and V contains exactly one element of each $\widetilde{a} = \{t \in [-1,1] : a - t \in \mathbb{Q}\}$, so that $(r + V) \cap (r' + V) = \varnothing$ for distinct $r, r' \in \mathbb{Q}$ (if $r + v_1 = r' + v_2$ then $v_1 - v_2 \in \mathbb{Q}$, placing v_1, v_2 in one $\widetilde{a}$, so $v_1 = v_2$ and $r = r'$). Note $0 < c \le 2$ by 2.5 and 2.14.
Every closed $F \subseteq V$ has $|F| = 0$. Let $r_1, r_2, \ldots$ be distinct rationals in $[0,1]$; the sets $r_k + F$ are then disjoint closed, hence Borel, subsets of $[-1,2]$. Since outer measure is a measure on the Borel sets (2.68) and is translation invariant (2.7), while 2.5 and 2.14 bound the union,

$$3 = |[-1,2]| \ \ge \ \left|\bigcup_{k=1}^{\infty} (r_k + F)\right| = \sum_{k=1}^{\infty} |F|,$$

which forces $|F| = 0$.
Now $A = \frac{1}{c}V \subseteq [-\frac{1}{c}, \frac{1}{c}]$ is bounded, and the dilation formula $|tB| = |t|\,|B|$ of Exercise 2 in Section 2A gives $|A| = \frac{1}{c}|V| = 1$. If $F \subseteq A$ is closed, then cF is a closed subset of $cA = V$, so $|cF| = 0$ by the previous paragraph and hence $|F| = \frac{1}{c}|cF| = 0$. Thus

$$|F| = 0 = |A| - 1$$

for every closed $F \subseteq A$.

Problem 2D.3. Prove that there exists a set $A \subseteq \mathbb{R}$ such that $|G \setminus A| = \infty$ for every open set G that contains A.

Solution. Take $A = \bigcup_{n=1}^{\infty}(n+V)$, where $V \subseteq (0,1)$ contains exactly one element of each set $\widetilde{a} = \{t \in (0,1) : a - t \in \mathbb{Q}\}$, $a \in (0,1)$. Throughout, $|(a,b)| = b - a$ is Exercise 6 in Section 2A.
Exactly as in the proof of 2.18, $(r+V) \cap (r'+V) = \varnothing$ for distinct $r, r' \in \mathbb{Q}$. Each $a \in (0,1)$ satisfies $a - v \in \mathbb{Q} \cap (-1,1)$ for its representative $v \in V$, so $(0,1) \subseteq \bigcup_k (q_k + V)$ with $q_1, q_2, \dots$ enumerating $\mathbb{Q} \cap (-1,1)$; then 2.8 and 2.7 give $1 \le \sum_k |V|$, forcing $|V| > 0$.
V is not Lebesgue measurable. First, every closed $F \subseteq V$ has $|F| = 0$: for distinct rationals $r_1, r_2, \ldots \in [0,1]$ the sets $r_k + F$ are disjoint closed, hence Borel, subsets of $(0,2)$, so by 2.68, 2.7, 2.5, and 2.14,

$$2 \ge \Big|\bigcup_{k=1}^{\infty}(r_k + F)\Big| = \sum_{k=1}^{\infty}|F|.$$

Were V measurable, then since $0 < |V| \le 1$, condition (b) of 2.71 with $\varepsilon = |V|/2$ would give a closed $F \subseteq V$ with $|V \setminus F| < |V|/2$, whence $|V| \le |F| + |V \setminus F| < |V|/2$ by 2.8, a contradiction.
The constant. Put $c = |V| + |(0,1) \setminus V| - 1$, a real number since both terms lie in $[0,1]$ by 2.5. Subadditivity (2.8) gives $c \ge 0$, and Exercise 12 of this section gives $c \ne 0$ because $V \subseteq (0,1)$ is not measurable. So $c > 0$.
The estimate on one unit interval. Since $V \subseteq (0,1)$, we have $A \cap (n, n+1) = n + V$. Let $G \supseteq A$ be open, fix $n \in \mathbb{Z}^+$, and put $W = G \cap (n, n+1)$ and $E = (n, n+1) \setminus W$, a Borel (hence measurable) set disjoint from $n + V \subseteq W$. Exercise 10 of this section, applied with E as the measurable set, and 2.5 give

$$|E| + |n+V| = \big|E \cup (n+V)\big| \le |(n,n+1)| = 1,$$

so $|E| \le 1 - |V|$ by 2.7. Since $(n,n+1) \setminus (n+V) \subseteq E \cup \big(W \setminus (n+V)\big)$, subadditivity (2.8) and 2.7 give

$$|(0,1) \setminus V| \le |E| + \big|W \setminus (n+V)\big| \le 1 - |V| + \big|W \setminus (n+V)\big|,$$

that is, $\big|(G \setminus A) \cap (n, n+1)\big| = \big|W \setminus (n+V)\big| \ge c$.
Summing. Fix $N \in \mathbb{Z}^+$ and put $S = (G \setminus A) \cap (1, N+1)$. By the Lemma in Exercise 10 of this section there is a Borel set $B \supseteq S$ with $|B| = |S|$; the sets $B \cap (n, n+1)$ for $1 \le n \le N$ are disjoint Borel sets, the n-th containing $(G \setminus A) \cap (n, n+1)$, so by 2.5 and 2.68,

$$|G \setminus A| \ge |S| = |B| \ge \sum_{n=1}^{N}\big|B \cap (n, n+1)\big| \ge Nc.$$

Letting $N \to \infty$ gives $|G \setminus A| = \infty$.

Problem 2D.4. The phrase *nontrivial interval* is used to denote an interval of $\mathbb{R}$ that contains more than one element. Recall that an interval might be open, closed, or neither.
(a) Prove that the union of each collection of nontrivial intervals of $\mathbb{R}$ is the union of a countable subset of that collection.
(b) Prove that the union of each collection of nontrivial intervals of $\mathbb{R}$ is a Borel set.
(c) Prove that there exists a collection of closed intervals of $\mathbb{R}$ whose union is not a Borel set.

Solution. (a). Let $\mathcal{C}$ be such a collection, $U = \bigcup_{I \in \mathcal{C}} I$, and $W = \bigcup_{I \in \mathcal{C}} I^{\circ}$; since each $I \in \mathcal{C}$ has at least two points, I° is a nonempty open interval and $I \setminus I^{\circ}$ consists of at most the two endpoints of I.
Let D be the countable set of rational pairs (p,q), $p < q$, with $(p,q) \subseteq I^{\circ}$ for at least one $I \in \mathcal{C}$, and fix one such $I_{p,q}$ for each. If $x \in W$ then $x \in I^{\circ}$ for some I, and I° is open, so $p < x < q$ with $(p,q) \subseteq I^{\circ}$ for some rationals p, q; hence

$$W \subseteq \bigcup_{(p,q) \in D} I_{p,q} \subseteq U.$$

Next, $U \setminus W$ is countable. If $x \in U \setminus W$, say $x \in I \in \mathcal{C}$, then $x \notin I^{\circ}$, so x is an endpoint of I and I° is a nonempty open interval with endpoint x; thus $(x, x+\delta_x) \subseteq W$ or $(x - \delta_x, x) \subseteq W$ for some $\delta_x > 0$. Write $U \setminus W = U_+ \cup U_-$ accordingly. For $x \in U_+$ pick a rational $r_x \in (x, x+\delta_x)$; the map $x \mapsto r_x$ is injective on U_+, since $x < y$ and $r_x = r_y$ would give $y \in (x, x+\delta_x) \subseteq W$. So U_+ is countable, and

symmetrically so is U_-.
Choosing $J_x \in \mathcal{C}$ with $x \in J_x$ for each $x \in U \setminus W$, the collection $\mathcal{D} = \{I_{p,q}\} \cup \{J_x\}$ is a countable subcollection of $\mathcal{C}$ with

$$U = W \cup (U \setminus W) \subseteq \bigcup_{J \in \mathcal{D}} J \subseteq U.$$

(b). By (a), U is a countable union of intervals, and every interval $J \neq \varnothing$ is Borel: with $a = \inf J$ and $b = \sup J$, the set $\{x \in \mathbb{R} : a < x < b\}$ is open and J differs from it by at most the points a, b, each Borel because $\{a\} = \bigcap_{m=1}^{\infty}(a - \frac{1}{m}, a + \frac{1}{m})$. As the Borel sets form a σ-algebra, U is Borel.
(c). Part (c) drops *nontrivial*, so singletons are allowed. By 2.67 there is a non-Borel set $S \subseteq \mathbb{R}$; the collection $\mathcal{C} = \{[x, x] : x \in S\}$ consists of closed intervals and has union S.

Problem 2D.5. Prove that if $A \subseteq \mathbb{R}$ is Lebesgue measurable, then there exists an increasing sequence $F_1 \subseteq F_2 \subseteq \cdots$ of closed sets contained in A such that

$$\left| A \setminus \bigcup_{k=1}^{\infty} F_k \right| = 0.$$

Solution. Take $F_k = E_1 \cup \cdots \cup E_k$, where (as A is Lebesgue measurable, so that (a) implies (b) in 2.71 applies) $E_k \subseteq A$ is closed with $|A \setminus E_k| < \frac{1}{k}$.
Each F_k is a finite union of closed subsets of A, hence is closed and contained in A; clearly $F_1 \subseteq F_2 \subseteq \cdots$ and $\bigcup_k F_k = \bigcup_k E_k$. Fixing n and using $E_n \subseteq \bigcup_k F_k$ with 2.5,

$$\left| A \setminus \bigcup_{k=1}^{\infty} F_k \right| \leq |A \setminus E_n| < \tfrac{1}{n}.$$

Letting $n \to \infty$ gives $\left| A \setminus \bigcup_k F_k \right| = 0$.

Problem 2D.6. Suppose $A \subseteq \mathbb{R}$ and $|A| < \infty$. Prove that A is Lebesgue measurable if and only if for every $\varepsilon > 0$ there exists a set G that is the union of finitely many disjoint bounded open intervals such that $|A \setminus G| + |G \setminus A| < \varepsilon$.

Solution. Suppose A is Lebesgue measurable, and let $\varepsilon > 0$. Because (a) implies (e) in 2.71, there is an open $U \supseteq A$ with $|U \setminus A| < \frac{\varepsilon}{2}$, and then $|U| \leq |A| + |U \setminus A| < \infty$ by 2.8 and the hypothesis $|A| < \infty$. If $U = \varnothing$ then $A = \varnothing$ and $G = (0, \varepsilon/3)$ serves; so assume $U \neq \varnothing$ and write $U = \bigcup_n I_n$ with $I_1, I_2, \ldots$ disjoint nonempty open intervals (the connected components, as in the proof of 2.63), padding a finite list with copies of $\varnothing$. Each I_n is bounded, since an unbounded one contains closed intervals of arbitrary length and would force $|U| = \infty$ by 2.5 and 2.14. The I_n are disjoint Borel sets, so $\sum_n |I_n| = |U| < \infty$ by 2.68; choose N with $\sum_{n>N} |I_n| < \frac{\varepsilon}{2}$ and put $G = I_1 \cup \cdots \cup I_N$. Since $G \setminus A \subseteq U \setminus A$ and $A \setminus G \subseteq U \setminus G = \bigcup_{n>N} I_n$, 2.5 and 2.8 give

$$|A \setminus G| + |G \setminus A| \leq \sum_{n>N} |I_n| + |U \setminus A| < \varepsilon.$$

Conversely, let $\varepsilon > 0$ and choose such a G with $|A \setminus G| + |G \setminus A| < \frac{\varepsilon}{2}$. As $|A \setminus G| < \frac{\varepsilon}{2}$, the definition of outer measure (2.1) supplies open intervals J_k with $A \setminus G \subseteq H = \bigcup_k J_k$ and $\sum_k \ell(J_k) < \frac{\varepsilon}{2}$, so $|H| < \frac{\varepsilon}{2}$ by 2.8 and 2.14. Then $V = G \cup H$ is open, $A = (A \cap G) \cup (A \setminus G) \subseteq V$, and $V \setminus A \subseteq (G \setminus A) \cup H$, so by 2.5 and 2.8,

$$|V \setminus A| \leq |G \setminus A| + |H| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Thus A satisfies condition (e) of 2.71, and (e) implies (a) there, so A is Lebesgue measurable.

Problem 2D.7. Prove that if $A \subseteq \mathbb{R}$ is Lebesgue measurable, then there exists a decreasing sequence $G_1 \supseteq G_2 \supseteq \cdots$ of open sets containing A such that

$$\Big|\Big(\bigcap_{k=1}^{\infty} G_k\Big) \setminus A\Big| = 0.$$

Solution. Take $G_k = U_1 \cap \cdots \cap U_k$, where (as A is Lebesgue measurable, so that (a) implies (e) in 2.71 applies) $U_k \supseteq A$ is open with $|U_k \setminus A| < \frac{1}{k}$.
Each G_k is a finite intersection of open sets containing A, hence is open and contains A; clearly $G_1 \supseteq G_2 \supseteq \cdots$ and $\bigcap_k G_k = \bigcap_k U_k$. Fixing n and using $\bigcap_k G_k \subseteq U_n$ with 2.5,

$$\Big|\Big(\bigcap_{k=1}^{\infty} G_k\Big) \setminus A\Big| \le |U_n \setminus A| < \tfrac{1}{n}.$$

Letting $n \to \infty$ gives the stated equality.

Problem 2D.8. Prove that the collection of Lebesgue measurable subsets of $\mathbb{R}$ is translation invariant. More precisely, prove that if $A \subseteq \mathbb{R}$ is Lebesgue measurable and $t \in \mathbb{R}$, then $t + A$ is Lebesgue measurable.

Solution. Translate the Borel witness of 2.70. Take a Borel set $B \subseteq A$ with $|A \setminus B| = 0$ (2.70). Then $t + B$ is Borel by the paragraph below and $t + B \subseteq t + A$; moreover $(t + A) \setminus (t + B) = t + (A \setminus B)$, since x belongs to the left side exactly when $x - t \in A \setminus B$. Hence by translation invariance of outer measure (2.7),

$$\big|(t + A) \setminus (t + B)\big| = \big|t + (A \setminus B)\big| = |A \setminus B| = 0,$$

so $t + A$ is Lebesgue measurable by 2.70.
The Borel sets are translation invariant. Fix $t \in \mathbb{R}$ and put $\mathcal{S} = \{E \subseteq \mathbb{R} : t + E \text{ is Borel}\}$. Since $x \mapsto t+x$ is a bijection of $\mathbb{R}$, we have $t+\varnothing = \varnothing$, $t+(\mathbb{R}\setminus E) = \mathbb{R}\setminus(t+E)$, and $t+\bigcup_k E_k = \bigcup_k(t+E_k)$, so $\mathcal{S}$ is a σ-algebra on $\mathbb{R}$. It contains every open G, because $t + G$ is open: if $x \in t + G$ then $(x - t - \delta, x - t + \delta) \subseteq G$ for some $\delta > 0$, whence $(x - \delta, x + \delta) \subseteq t + G$. As the Borel sets form the smallest σ-algebra containing the open sets (2.29), every Borel set lies in $\mathcal{S}$.

Problem 2D.9. Prove that the collection of Lebesgue measurable subsets of $\mathbb{R}$ is dilation invariant. More precisely, prove that if $A \subseteq \mathbb{R}$ is Lebesgue measurable and $t \in \mathbb{R}$, then tA (which is defined to be $\{ta : a \in A\}$) is Lebesgue measurable.

Solution. Dilate the Borel witness of 2.70, exactly as in Exercise 8 of this section, using the dilation formula $|tA| = |t|\,|A|$ of Exercise 2 in Section 2A in place of 2.7.
(i) $t = 0$: then $0A$ is $\varnothing$ or $\{0\}$, in either case closed, hence Borel, hence Lebesgue measurable.
(ii) $t \neq 0$: take a Borel set $B \subseteq A$ with $|A \setminus B| = 0$ (2.70). Then tB is Borel by the paragraph below and $tB \subseteq tA$; also $tA \setminus tB = t(A \setminus B)$, since x belongs to the left side exactly when $\frac{x}{t} \in A \setminus B$. So by Exercise 2 in Section 2A,

$$|tA \setminus tB| = \big|t(A \setminus B)\big| = |t|\,|A \setminus B| = 0,$$

and tA is Lebesgue measurable by 2.70.
The Borel sets are dilation invariant. Fix $t \neq 0$ and put $\mathcal{S} = \{E \subseteq \mathbb{R} : tE \text{ is Borel}\}$. Since $x \mapsto tx$ is a bijection of $\mathbb{R}$ with inverse $x \mapsto \frac{x}{t}$, we have $t\varnothing = \varnothing$, $t(\mathbb{R} \setminus E) = \mathbb{R} \setminus tE$, and $t\bigcup_k E_k = \bigcup_k tE_k$, so $\mathcal{S}$ is a σ-algebra on $\mathbb{R}$. It contains every open G, because tG is open: if $x \in tG$ then $\big(\frac{x}{t} - \delta, \frac{x}{t} + \delta\big) \subseteq G$ for some $\delta > 0$, so the interval of radius $|t|\delta$ centred at x lies in tG. As the Borel sets form the smallest σ-algebra containing the open sets (2.29), every Borel set lies in $\mathcal{S}$.

Problem 2D.10. Prove that if A and B are disjoint subsets of $\mathbb{R}$ and B is Lebesgue measurable, then $|A \cup B| = |A| + |B|$.

Solution. Subadditivity (2.8) gives $|A \cup B| \le |A| + |B|$, so only the reverse inequality is at issue.
Lemma (Borel hull; used again in Exercises 3 and 12 of this section). Every $S \subseteq \mathbb{R}$ lies in a Borel set E with $|E| = |S|$. Indeed, if $|S| = \infty$ take $E = \mathbb{R}$. Otherwise, for each $n \in \mathbb{Z}^+$ the infimum defining outer measure (2.2) supplies open intervals $I_1, I_2, \dots$ with $S \subseteq G_n = \bigcup_k I_k$ and $\sum_k \ell(I_k) < |S| + \frac{1}{n}$, whence $|G_n| < |S| + \frac{1}{n}$ by 2.8 and $|I_k| \le \ell(I_k)$. Then $E = \bigcap_n G_n$ is Borel, contains S, and by 2.5 satisfies $|S| \le |E| \le |G_n| < |S| + \frac{1}{n}$ for every n.
Now suppose $A, B \subseteq \mathbb{R}$ are disjoint with B Lebesgue measurable, and take a Borel set $D \subseteq B$ with $|B \setminus D| = 0$ (2.70); then

$$|B| \le |D| + |B \setminus D| = |D| \le |B|$$

by 2.8 and 2.5, so $|D| = |B|$. By the Lemma take a Borel set $E \supseteq A \cup B$ with $|E| = |A \cup B|$. Since $D \subseteq B \subseteq E$, the sets D and $E \setminus D$ are disjoint Borel sets with union E, so $|E| = |D| + |E \setminus D|$ because outer measure is a measure on the Borel sets (2.68). Finally $A \subseteq E \setminus D$, since $A \subseteq E$ and $A \cap D \subseteq A \cap B = \emptyset$; hence $|A| \le |E \setminus D|$ by 2.5 and

$$|A \cup B| = |E| = |D| + |E \setminus D| \ge |B| + |A|.$$

Problem 2D.11. Prove that if $A \subseteq \mathbb{R}$ and $|A| > 0$, then there exists a subset of A that is not Lebesgue measurable.

Solution. Suppose toward a contradiction that every subset of A is Lebesgue measurable.
The Vitali set. On $\mathbb{R}$, the relation $x \sim y$ when $x - y \in \mathbb{Q}$ is an equivalence relation, and each class $\tilde{x} = x + \mathbb{Q}$ meets $[0,1]$ (take a rational $q \in [-x, -x+1]$). So the Axiom of Choice, as in 2.18, provides $V \subseteq [0,1]$ containing exactly one element of each class. Two consequences: (i) $\mathbb{R} = \bigcup_{r \in \mathbb{Q}} (r + V)$, since $x = (x - v) + v$ for the representative v of $\tilde{x}$ and $x - v \in \mathbb{Q}$; (ii) the sets $r + V$, $r \in \mathbb{Q}$, are pairwise disjoint, since $r + v_1 = s + v_2$ gives $v_1 - v_2 \in \mathbb{Q}$, hence $v_1 = v_2$ and $r = s$.
Put $A_r = A \cap (r + V)$ for $r \in \mathbb{Q}$. Each $A_r \subseteq A$ is Lebesgue measurable by assumption, and by (i) and (ii) the countable union $A = \bigcup_{r \in \mathbb{Q}} A_r$ is disjoint, so countable additivity on the Lebesgue measurable sets (2.72) gives

$$0 < |A| = \sum_{r \in \mathbb{Q}} |A_r|.$$

Fix r with $|A_r| > 0$ and put $W = (-r) + A_r$. Then $W \subseteq V \subseteq [0,1]$, W is Lebesgue measurable by Exercise 8 in this section, and $|W| = |A_r| > 0$ by 2.7.
Let $q_1, q_2, \dots$ enumerate $\mathbb{Q} \cap [0,1]$. The sets $q_k + W$ are Lebesgue measurable (Exercise 8) with $|q_k + W| = |W|$ (2.7), are pairwise disjoint by (ii) since $q_k + W \subseteq q_k + V$, and lie in $[0,2]$. Hence by 2.72, 2.5, and 2.14,

$$\sum_{k=1}^{\infty} |W| = \Big| \bigcup_{k=1}^{\infty} (q_k + W) \Big| \le |[0,2]| = 2,$$

while $|W| > 0$ makes the left side ∞ – a contradiction. So some subset of A is not Lebesgue measurable.

Problem 2D.12. Suppose $b < c$ and $A \subseteq (b,c)$. Prove that A is Lebesgue measurable if and only if $|A| + |(b,c) \setminus A| = c - b$.

Solution. Throughout, $|(b,c)| = c - b$ (Exercise 6 in Section 2A), and both $|A|$ and $|(b,c) \setminus A|$ are at most $c - b$ by 2.5, hence finite, so the arithmetic below is legitimate.
Suppose A is Lebesgue measurable. Exercise 10 of this section, applied to the disjoint sets $(b,c) \setminus A$ and A with A in the role of the measurable set, gives

$$|A| + |(b,c) \setminus A| = \big|A \cup \big((b,c) \setminus A\big)\big| = |(b,c)| = c - b,$$

the middle union being (b,c) because $A \subseteq (b,c)$.
Conversely, suppose $|A| + |(b,c) \setminus A| = c - b$. By the Lemma in Exercise 10 of this section there is a Borel set $E_0 \supseteq (b,c) \setminus A$ with $|E_0| = |(b,c) \setminus A|$; replacing E_0 by $E = E_0 \cap (b,c)$ keeps both the inclusion and, by 2.5, the outer measure. Put $B = (b,c) \setminus E$, a Borel set contained in A (if $x \in B$ then $x \in (b,c)$ and $x \notin (b,c) \setminus A$). The disjoint Borel sets B and E have union (b,c), so $|B| + |E| = c - b$ by 2.68; subtracting the finite $|E| = |(b,c) \setminus A|$ and invoking the hypothesis gives $|B| = |A|$. Now Exercise 10 again, applied to the disjoint sets $A \setminus B$ and the Borel set B, gives

$$|A| = \big|(A \setminus B) \cup B\big| = |A \setminus B| + |B| = |A \setminus B| + |A|,$$

so cancelling the finite $|A|$ yields $|A \setminus B| = 0$. By 2.70, A is Lebesgue measurable.

Problem 2D.13. Suppose $A \subseteq \mathbb{R}$. Prove that A is Lebesgue measurable if and only if

$$|(-n,n) \cap A| + |(-n,n) \setminus A| = 2n$$

for every $n \in \mathbb{Z}^+$.

Solution. Put $A_n = (-n,n) \cap A \subseteq (-n,n)$. Since $(-n,n) \setminus A_n = (-n,n) \setminus A$, the displayed condition says exactly that

$$|A_n| + |(-n,n) \setminus A_n| = n - (-n),$$

which by Exercise 12 of this section (with $b = -n$, $c = n$) holds if and only if A_n is Lebesgue measurable. So the exercise reduces to: A is Lebesgue measurable if and only if every A_n is.
(i) If A is Lebesgue measurable, then each $(-n,n)$ is open, hence Borel, hence Lebesgue measurable, and the Lebesgue measurable sets form a σ-algebra (2.72), which is closed under intersection (2.25); so each A_n is Lebesgue measurable.
(ii) Conversely, every real lies in $(-n,n)$ for all large n, so $A = \bigcup_{n=1}^{\infty} A_n$; that σ-algebra is closed under countable unions, so A is Lebesgue measurable.

Problem 2D.14. Show that $\frac{1}{4}$ and $\frac{9}{13}$ are both in the Cantor set.

Solution. Both numbers have a base 3 representation using only 0s and 2s, namely $\frac{1}{4} = 0.\overline{02}_3$ and $\frac{9}{13} = 0.\overline{200}_3$, so both lie in the Cantor set C by 2.75. (The repeating blocks come from the greedy digit algorithm, whose remainders return to $\frac{1}{4}$ after two steps and to $\frac{9}{13}$ after three.)
Summing the geometric series confirms the two representations:

$$\sum_{k=1}^{\infty} \frac{2}{3^{2k}} = 2 \cdot \frac{1/9}{1 - 1/9} = \frac{1}{4},$$

$$\sum_{k=0}^{\infty} \frac{2}{3^{3k+1}} = \frac{2}{3} \cdot \frac{1}{1 - 1/27} = \frac{18}{26} = \frac{9}{13}.$$

Problem 2D.15. Show that $\frac{13}{17}$ is not in the Cantor set.

Solution. The unique base 3 representation of $\frac{13}{17}$ is $0.2021\ldots_3$, whose fourth digit is 1, so $\frac{13}{17} \notin C$ by 2.75.
Step 1: $\frac{13}{17}$ has only one base 3 representation. Suppose $x \in [0,1]$ has two different base 3 representations, say

$$x = \sum_{k=1}^{\infty} \frac{a_k}{3^k} = \sum_{k=1}^{\infty} \frac{b_k}{3^k},$$

with all $a_k, b_k \in \{0,1,2\}$, and let n be the smallest index with $a_n \neq b_n$; say $a_n < b_n$. Then

$$0 = \sum_{k=n}^{\infty} \frac{b_k - a_k}{3^k} \geq \frac{b_n - a_n}{3^n} - \sum_{k=n+1}^{\infty} \frac{2}{3^k} = \frac{b_n - a_n}{3^n} - \frac{1}{3^n} \geq 0,$$

because $b_n - a_n \geq 1$. Hence equality holds throughout, which forces $b_n - a_n = 1$ and $b_k - a_k = -2$ for all $k > n$, i.e. $b_k = 0$ and $a_k = 2$ for all $k > n$. In particular

$$x = \sum_{k=1}^{n} \frac{b_k}{3^k} = \frac{m}{3^n} \quad \text{where } m = \sum_{k=1}^{n} b_k 3^{n-k} \in \mathbb{Z}.$$

Thus a number with two base 3 representations must be of the form $m/3^n$. But $\frac{13}{17} = \frac{m}{3^n}$ would give $17 \mid 13 \cdot 3^n$, impossible since the prime 17 divides neither 13 nor 3.

Step 2: that representation contains a 1. The digit algorithm $x_0 = \frac{13}{17}$, $a_k = \lfloor 3x_{k-1} \rfloor$, $x_k = 3x_{k-1} - a_k$ gives

$$\begin{aligned}
3 \cdot \tfrac{13}{17} &= \tfrac{39}{17} = 2 + \tfrac{5}{17}, \quad a_1 = 2,\\
3 \cdot \tfrac{5}{17} &= \tfrac{15}{17} = 0 + \tfrac{15}{17}, \quad a_2 = 0,\\
3 \cdot \tfrac{15}{17} &= \tfrac{45}{17} = 2 + \tfrac{11}{17}, \quad a_3 = 2,\\
3 \cdot \tfrac{11}{17} &= \tfrac{33}{17} = 1 + \tfrac{16}{17}, \quad a_4 = 1.
\end{aligned}$$

Every x_k lies in $[0,1)$, so these digits do represent $\frac{13}{17}$; explicitly

$$\frac{13}{17} = \frac{61}{81} + \frac{1}{3^4} \cdot \frac{16}{17}. \qquad \text{(Check!)}$$

By Step 1 this representation is the only one, and its fourth digit is 1; so no base 3 representation of $\frac{13}{17}$ consists only of 0s and 2s.

Problem 2D.16. List the eight open intervals whose union is G_4 in the definition of the Cantor set (2.74).

Solution. The eight intervals are those $\left(\frac{k}{81}, \frac{k+1}{81}\right)$ with $k \in \{1, 7, 19, 25, 55, 61, 73, 79\}$:

$$\begin{aligned}
G_4 = &\left(\tfrac{1}{81}, \tfrac{2}{81}\right) \cup \left(\tfrac{7}{81}, \tfrac{8}{81}\right) \cup \left(\tfrac{19}{81}, \tfrac{20}{81}\right) \cup \left(\tfrac{25}{81}, \tfrac{26}{81}\right)\\
&\cup \left(\tfrac{55}{81}, \tfrac{56}{81}\right) \cup \left(\tfrac{61}{81}, \tfrac{62}{81}\right) \cup \left(\tfrac{73}{81}, \tfrac{74}{81}\right) \cup \left(\tfrac{79}{81}, \tfrac{80}{81}\right).
\end{aligned}$$

Indeed, $[0,1] \setminus (G_1 \cup G_2) = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$ by 2.74, and deleting from each of these its middle third (the corresponding interval of G_3) leaves the eight closed intervals $\left[\frac{3j}{81}, \frac{3j+3}{81}\right]$ with $j \in \{0, 2, 6, 8, 18, 20, 24, 26\}$. By 2.74, G_4 is the union of their middle thirds $\left(\frac{3j+1}{81}, \frac{3j+2}{81}\right)$, and $3j+1$ runs over the eight values of k listed above.

Problem 2D.17. Let C denote the Cantor set. Prove that $\left\{\frac{1}{2}x + \frac{1}{2}y : x, y \in C\right\} = [0,1]$.

Solution. Split the base 3 digits of t between two $\{0,2\}$-strings. Write $S = \left\{\frac{1}{2}x + \frac{1}{2}y : x, y \in C\right\}$. Since $C \subseteq [0,1]$ (2.74) and $[0,1]$ is convex, $S \subseteq [0,1]$. Conversely let $t \in [0,1]$. If $t = 1$, take $x = y = 1 = 0.222\ldots_3 \in C$ (2.75). Otherwise fix a base 3 representation $t = \sum_k t_k 3^{-k}$ and put

$$(a_k, b_k) = \begin{cases} (0,0) & \text{if } t_k = 0,\\ (0,2) & \text{if } t_k = 1,\\ (2,2) & \text{if } t_k = 2, \end{cases}$$

so $a_k, b_k \in \{0,2\}$ and $a_k + b_k = 2t_k$ in every case. The series $x = \sum_k a_k 3^{-k}$ and $y = \sum_k b_k 3^{-k}$ converge, being dominated by $\sum_k 2\cdot 3^{-k}=1$, and exhibit base 3 representations of x and y using only 0s and 2s, so $x,y\in C$ by 2.75. Adding the two absolutely convergent series termwise,

$$\tfrac{1}{2}x+\tfrac{1}{2}y=\sum_{k=1}^{\infty}\frac{a_k+b_k}{2\cdot 3^k}=\sum_{k=1}^{\infty}\frac{t_k}{3^k}=t,$$

so $t\in S$. Hence $S=[0,1]$.

Problem 2D.18. Prove that every open interval of $\mathbb{R}$ contains either infinitely many or no elements in the Cantor set.

Solution. It suffices to show that the Cantor set C has no isolated points: if an open interval I contains some $x\in C$, then $(x-\varepsilon,x+\varepsilon)\subseteq I$ for some $\varepsilon>0$, so $I\cap C$ is infinite.
Fix $x\in C$ and $\varepsilon>0$, and by 2.75 write $x=\sum_k a_k3^{-k}$ with every $a_k\in\{0,2\}$. Flip the n-th digit:

$$x^{(n)}=\sum_{k\neq n}\frac{a_k}{3^k}+\frac{2-a_n}{3^n}.$$

Since $2-a_n\in\{0,2\}$, this display is a base 3 representation of $x^{(n)}$ using only 0s and 2s, so $x^{(n)}\in C$ by 2.75. The two strings differ only in position n, so

$$\left|x-x^{(n)}\right|=\frac{|a_n-(2-a_n)|}{3^n}=\frac{2}{3^n},$$

whence $x^{(1)},x^{(2)},\ldots$ are pairwise distinct and all differ from x. Choosing N with $\frac{2}{3^N}<\varepsilon$, the points $x^{(N)},x^{(N+1)},\ldots$ are infinitely many elements of $C\cap(x-\varepsilon,x+\varepsilon)$.

Problem 2D.19. Evaluate $\int_0^1\Lambda$, where Λ is the Cantor function.

Solution. $\int_0^1\Lambda=\frac{1}{2}$, forced by the symmetry $\Lambda(1-x)=1-\Lambda(x)$ on $[0,1]$.
By 2.79 the Cantor function Λ is continuous on $[0,1]$, hence Riemann integrable there; being bounded and continuous it is Lebesgue measurable with the same integral, so the value below settles either reading.
(i) $x\in C$. By 2.75 write $x=\sum_k a_k3^{-k}$ with every $a_k\in\{0,2\}$; this $\{0,2\}$-representation is unique, since by Step 1 of Exercise 15 two representations of one number differ by 1 at their first disagreement, whereas two digits from $\{0,2\}$ differ by 0 or 2. As $\sum_k 2\cdot 3^{-k}=1$,

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

again a $\{0,2\}$-representation, so $1-x\in C$; with $e_k=a_k/2\in\{0,1\}$, 2.77 gives

$$\Lambda(1-x)=\sum_{k=1}^{\infty}\frac{1-e_k}{2^k}=1-\sum_{k=1}^{\infty}\frac{e_k}{2^k}=1-\Lambda(x).$$

(ii) $x\in[0,1]\setminus C$. Then $x\in G_n$ for some n, since $[0,1]\setminus C=\bigcup_n G_n$ (2.74).
Subclaim. For every $n\geq 1$, the set $[0,1]\setminus(G_1\cup\cdots\cup G_n)$ is the union of the 2^n closed intervals $\left[c,c+\frac{1}{3^n}\right]$ with $c=\sum_{i\leq n}d_i3^{-i}$ and $d_i\in\{0,2\}$, while G_n is the union of the 2^{n-1} open intervals $\left(a,a+\frac{1}{3^n}\right)$ with $a=\sum_{i<n}d_i3^{-i}+3^{-n}$ and $d_i\in\{0,2\}$. This is induction on n from 2.74: the middle third of $\left[c,c+3^{-(n-1)}\right]$ is $(c+3^{-n},c+2\cdot 3^{-n})$, which is the interval attached to $(d_1,\ldots,d_{n-1})$, and removing it leaves the two intervals attached to $(d_1,\ldots,d_{n-1},0)$ and $(d_1,\ldots,d_{n-1},2)$. (Check!)
Let (a,b) be the interval of G_n containing x, with a as in the subclaim, $b=a+3^{-n}$, and $e_i=d_i/2$. Using $3^{-n}=\sum_{k>n}2\cdot 3^{-k}$,

$$a=0.d_1\ldots d_{n-1}0\,2\,2\,2\ldots_3\,,\qquad b=0.d_1\ldots d_{n-1}2\,0\,0\,0\ldots_3\,,$$

so $a, b \in C$ and, by 2.77,

$$\Lambda(a) = \sum_{i=1}^{n-1} \frac{e_i}{2^i} + \sum_{k>n} \frac{1}{2^k} = \sum_{i=1}^{n-1} \frac{e_i}{2^i} + \frac{1}{2^n} = \Lambda(b).$$

Moreover every base 3 representation of x begins $0.d_1 \dots d_{n-1}1$: the numbers admitting a representation with a given prefix $c_1 \dots c_n$ form exactly the interval $\left[\frac{k}{3^n}, \frac{k+1}{3^n}\right]$, where $k = \sum_{i \le n} c_i 3^{n-i}$, because the tails $\sum_{k>n} c_k 3^{-k}$ fill $[0, 3^{-n}]$; distinct prefixes give distinct k, the prefix $d_1 \dots d_{n-1}1$ gives exactly $[a, b]$, and x lies in the interior of $[a, b]$, hence in no other prefix interval. So 2.77 gives $\Lambda(x) = \sum_{i<n} e_i 2^{-i} + 2^{-n}$, which is $\Lambda(a) = \Lambda(b)$.
Now $1 - x \in (1 - b, 1 - a)$ with $1 - a, 1 - b \in C$, so by (i) and monotonicity of Λ (2.79),

$$1 - \Lambda(a) = \Lambda(1 - b) \le \Lambda(1 - x) \le \Lambda(1 - a) = 1 - \Lambda(a),$$

whence $\Lambda(1 - x) = 1 - \Lambda(a) = 1 - \Lambda(x)$.
Substituting $u = 1 - x$ in the integral now gives

$$\int_0^1 \Lambda(x)\, dx = \int_0^1 \bigl(1 - \Lambda(u)\bigr)\, du = 1 - \int_0^1 \Lambda(u)\, du,$$

so $\int_0^1 \Lambda = \frac{1}{2}$.

Problem 2D.20. Evaluate each of the following:
(a) $\Lambda\left(\frac{9}{13}\right)$;
(b) $\Lambda(0.93)$.

Solution. (a) $\Lambda\left(\frac{9}{13}\right) = \frac{4}{7}$. (b) $\Lambda(0.93) = \frac{7}{8}$.
(a). By Exercise 14, $\frac{9}{13} = 0.\overline{200}_3 \in C$, so 2.77 replaces each 2 by 1 and reads the result in base 2:

$$\Lambda\left(\frac{9}{13}\right) = 0.\overline{100}_2 = \sum_{j=0}^{\infty} \frac{1}{2^{3j+1}} = \frac{1/2}{1 - 1/8} = \frac{4}{7}.$$

(b). Here $0.93 = \frac{93}{100}$, whose base 3 digits start

$$\begin{aligned} 3 \cdot \tfrac{93}{100} &= \tfrac{279}{100} = 2 + \tfrac{79}{100}, && \text{digit } 2,\\ 3 \cdot \tfrac{79}{100} &= \tfrac{237}{100} = 2 + \tfrac{37}{100}, && \text{digit } 2,\\ 3 \cdot \tfrac{37}{100} &= \tfrac{111}{100} = 1 + \tfrac{11}{100}, && \text{digit } 1, \end{aligned}$$

so $\frac{93}{100} = 0.221\dots_3$. This representation is unique by Step 1 of Exercise 15, since $\frac{93}{100} = m/3^k$ would force $100 \mid 93 \cdot 3^k$, impossible as $\gcd(93 \cdot 3^k, 100) = 1$; it contains a 1, so $\frac{93}{100} \notin C$ by 2.75. By 2.77, truncate after that first 1 and replace each earlier 2 by 1:

$$\Lambda(0.93) = 0.111_2 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}.$$

Problem 2D.21. Find each of the following sets:
(a) $\Lambda^{-1}\left(\{\frac{1}{3}\}\right)$;
(b) $\Lambda^{-1}\left(\{\frac{5}{16}\}\right)$.

Solution. $\Lambda^{-1}\left(\{\frac{1}{3}\}\right) = \left\{\frac{1}{4}\right\}$ and $\Lambda^{-1}\left(\{\frac{5}{16}\}\right) = \left[\frac{19}{81}, \frac{20}{81}\right]$.
Fact 1 (points of C). By 2.75 and the uniqueness of the $\{0, 2\}$-representation noted in Exercise 19, the map $x \mapsto (e_k)$, $e_k = a_k/2$, is a bijection from C onto the set of binary digit strings, and 2.77 says $\Lambda(x) = \sum_k e_k 2^{-k}$. So for $y \in [0, 1]$ the preimages of y lying in C correspond exactly to the base 2

representations of y.
Fact 2 (points off C). By the subclaim in case (ii) of Exercise 19, together with $[0,1]\setminus C=\bigcup_n G_n$ (2.74) and the disjointness of $G_1, G_2, \dots$ noted in the proof of 2.76, the set $[0,1]\setminus C$ is the disjoint union, over $n\ge 1$ and $(d_1,\dots,d_{n-1})\in\{0,2\}^{n-1}$, of the intervals $(a, a+3^{-n})$ with $a=\sum_{i<n} d_i 3^{-i}+3^{-n}$; and on such an interval Λ is constantly

$$\sum_{i=1}^{n-1}\frac{e_i}{2^i}+\frac{1}{2^n}=\frac{2m+1}{2^n},\qquad e_i=\tfrac{d_i}{2},\quad m=\sum_{i<n} e_i 2^{n-1-i}.$$

As $(d_1,\dots,d_{n-1})$ ranges over $\{0,2\}^{n-1}$, the integer m ranges bijectively over $0,1,\dots,2^{n-1}-1$. So the values of Λ off C are exactly the dyadic rationals in $(0,1)$, and such a value determines n, then m, then the interval.
Fact 3 (base 2 representations). Step 1 of Exercise 15, run with 3 replaced by 2 and the digit bound 2 by 1, shows that a number with two base 2 representations is dyadic. Conversely, if $y=m/2^n\in(0,1)$ with m odd then $y=0.b_1\dots b_n000\dots_2$ with $b_n=1$, and also $y=0.b_1\dots b_{n-1}0111\dots_2$ since $\sum_{k>n}2^{-k}=2^{-n}$. So a dyadic rational in $(0,1)$ has exactly two base 2 representations, every other number in $[0,1]$ exactly one.
(a). $\frac13$ is not dyadic, since $\frac13=\frac{m}{2^n}$ would give $3\mid 2^n$; so by Fact 2 no point off C maps to it, and by Fact 3 it has the single base 2 representation $0.\overline{01}_2$ (indeed $\sum_{j\ge1}4^{-j}=\frac13$). By Fact 1 its only preimage is the point of C with $\{0,2\}$-representation $0.\overline{02}_3$, namely

$$\sum_{j=1}^{\infty}\frac{2}{9^j}=\frac{2/9}{1-1/9}=\frac14.$$

(b). $\frac{5}{16}$ is dyadic, so both cases contribute.
Off C: $\frac{5}{16}=\frac{2m+1}{2^n}$ forces $n=4$ and $m=2$, so $e_1e_2e_3=010$ and $(d_1,d_2,d_3)=(0,2,0)$; by Fact 2 the interval is $(a,a+\frac{1}{81})$ with $a=\frac29+\frac{1}{81}=\frac{19}{81}$, one of the eight intervals of G_4 found in Exercise 16.
On C: by Fact 3 the two base 2 representations of $\frac{5}{16}$ are $0.0101\,000\dots_2$ and $0.0100\,111\dots_2$, so by Fact 1 the two Cantor set preimages are

$$0.0202\,000\dots_3=\frac29+\frac{2}{81}=\frac{20}{81},$$

$$0.0200\,222\dots_3=\frac{18}{81}+\frac{1}{81}=\frac{19}{81}.$$

Combining the two cases,

$$\Lambda^{-1}\left(\{\tfrac{5}{16}\}\right)=\{\tfrac{19}{81}\}\cup(\tfrac{19}{81},\tfrac{20}{81})\cup\{\tfrac{20}{81}\}=\left[\frac{19}{81},\frac{20}{81}\right].$$

Problem 2D.22. (a) Suppose x is a rational number in $[0,1]$. Explain why $\Lambda(x)$ is rational.
(b) Suppose $x\in C$ is such that $\Lambda(x)$ is rational. Explain why x is rational.

Solution. Both parts reduce to one fact: a base b representation of $t\in[0,1]$ – digits $d_j\in\{0,\dots,b-1\}$ with $t=\sum_{j\ge1}d_jb^{-j}$ – has eventually periodic digit string exactly when t is rational.
Lemma 1 (near-uniqueness). The argument of Step 1 in Exercise 15, run in base b, gives: if $(c_j)\ne(c'_j)$ are base b representations of the same number, n is the first index where they differ, and $c_n>c'_n$, then $c_n=c'_n+1$ and $c_j=0$, $c'_j=b-1$ for every $j>n$. In particular each string is eventually constant.
Lemma 2. Let (d_j) be a base b representation of $t\in[0,1]$. Then t is rational if and only if (d_j) is eventually periodic.
Suppose $d_{j+k}=d_j$ for all $j>N$, and put $z=\sum_{i\ge1}d_{N+i}b^{-i}$, so that $t=\sum_{j\le N}d_jb^{-j}+b^{-N}z$.

Periodicity of $(d_{N+i})_{i\geq 1}$ gives

$$b^k z = \sum_{i=1}^{\infty} d_{N+i} b^{k-i} = m + z, \qquad m = \sum_{i=1}^{k} d_{N+i} b^{k-i} \in \mathbb{Z},$$

so $z = m/(b^k - 1)$ and hence t is rational. Conversely let $t = p/q$ with p, q integers, $q \geq 1$, and put

$$s_n = \sum_{i=1}^{\infty} \frac{d_{n+i}}{b^i} = b^n t - \sum_{j=1}^{n} d_j b^{n-j}.$$

Then $s_n \in [0, 1]$ and qs_n is an integer, hence $qs_n \in \{0, 1, \dots, q\}$; by the pigeonhole principle $s_n = s_m$ for some $n < m$. Now $(d_{n+i})_{i\geq 1}$ and $(d_{m+i})_{i\geq 1}$ are base b representations of that one number: if they are equal then $d_{j+(m-n)} = d_j$ for every $j > n$, and if not then Lemma 1 makes each eventually constant, so (d_j) is eventually constant. Either way (d_j) is eventually periodic.
(a). Let $x \in [0, 1]$ be rational; the two bullet points of 2.77 give two cases.
(i) $x \notin C$: by 2.75 every base 3 representation of x contains a digit 1, and 2.77 truncates after the first such 1, so $\Lambda(x) = \sum_{j=1}^{n} e_j 2^{-j}$ is a finite sum of rationals.
(ii) $x \in C$: let (d_j) be the base 3 representation of x using only 0s and 2s, which is the one used in 2.77. Lemma 2 with $b = 3$ makes (d_j) eventually periodic, hence so is $(d_j/2)$, which by 2.77 is a base 2 representation of $\Lambda(x)$; Lemma 2 with $b = 2$ then makes $\Lambda(x)$ rational.
(b). Let $x \in C$ with $\Lambda(x)$ rational, and let (d_j) be its base 3 representation using only 0s and 2s; this one is unique, since by Lemma 1 two representations of one number differ by 1 at their first disagreement whereas digits in $\{0, 2\}$ differ by 0 or 2. By 2.77 the string $(d_j/2)$, with terms in $\{0, 1\}$, is a base 2 representation of $\Lambda(x)$, so Lemma 2 with $b = 2$ makes it eventually periodic; then (d_j) is eventually periodic and Lemma 2 with $b = 3$ makes $x = \sum_j d_j 3^{-j}$ rational.

Problem 2D.23. Show that there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that the image under f of every nonempty open interval is $\mathbb{R}$.

Solution. Take $f = h \circ \varphi$, where φ reads off the limit superior of the decimal digit averages of the fractional part and h maps $[0, 9]$ onto $\mathbb{R}$:

$$\varphi(x) = \limsup_{n\to\infty} \frac{d_1(x) + d_2(x) + \cdots + d_n(x)}{n} \in [0, 9],$$

$$h(t) = \tan\left(\pi\left(\frac{t}{9} - \frac{1}{2}\right)\right) \text{ for } t \in (0, 9), \qquad h(0) = h(9) = 0.$$

Here $d_j(x)$ is the j-th digit of the decimal representation of $\langle x\rangle = x - \lfloor x\rfloor \in [0, 1)$ that is not eventually 9; it exists by the greedy algorithm $d_j = \lfloor 10^j\langle x\rangle\rfloor - 10\lfloor 10^{j-1}\langle x\rangle\rfloor$ (Check!), and it is unique because by Lemma 1 of Exercise 22 with $b = 10$, of two representations of one number one is eventually 0 and the other eventually 9. As t runs over $(0, 9)$, $\pi(t/9 - 1/2)$ runs over $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, on which tan is a bijection onto $\mathbb{R}$; so $h([0, 9]) = \mathbb{R}$, and it suffices to prove $\varphi(I) = [0, 9]$ for every nonempty open interval I.
Step 1: I contains a full decimal cylinder. Shrink I to an interval strictly between consecutive integers: choose $m \in \mathbb{Z}$ and $u < v$ with $m < u < v < m+1$ and $(u, v) \subseteq I$, so that $\langle x\rangle = x - m$ on (u, v). Choose $k \in \mathbb{Z}^+$ with $10^{-k} < (v - u)/2$ and put $M = \lfloor(u - m)10^k\rfloor + 1$, so that $(u - m)10^k < M \leq (u - m)10^k + 1$ and hence

$$u - m < \frac{M}{10^k} \quad\text{and}\quad \frac{M+1}{10^k} \leq (u - m) + \frac{2}{10^k} < v - m.$$

The second inequality gives $(M + 1)10^{-k} < 1$, so $1 \leq M < 10^k$; let $c_1 \dots c_k$ be the k-digit form of M, leading zeros allowed. For any digit string (e_i) that is not eventually 9, put

$$x = m + \frac{M}{10^k} + \frac{1}{10^k}\sum_{i=1}^{\infty} \frac{e_i}{10^i}.$$

Then $\langle x\rangle$ lies in $\left[M10^{-k},(M+1)10^{-k}\right] \subseteq (u-m, v-m)$ by the two displayed inequalities, so $x \in (u,v) \subseteq I$; and the string $c_1,\ldots,c_k,e_1,e_2,\ldots$ is a decimal representation of $\langle x\rangle$ that is not eventually 9, hence is the canonical one.
A fixed prefix does not move the limit superior: with $K = c_1+\cdots+c_k$ and $S_j = e_1+\cdots+e_j$, the average of the first $j+k$ digits of $\langle x\rangle$ is $(K+S_j)/(j+k)$, and

$$\frac{K+S_j}{j+k} - \frac{S_j}{j} = \frac{K}{j+k} - \frac{k\,S_j}{j(j+k)},$$

whose absolute value is at most $(K+9k)/(j+k) \to 0$ because $0 \le S_j \le 9j$. Hence $\varphi(x) = \limsup_j S_j/j$.
Step 2: every $t \in [0,9]$ is such a limit superior.
(i) $t = 9$: take $e_i = 0$ when i is a power of 2 and $e_i = 9$ otherwise. At most $1+\log_2 n$ powers of 2 lie in $\{1,\ldots,n\}$, so $9 \ge S_n/n \ge 9 - 9(1+\log_2 n)/n \to 9$. The string has infinitely many 0s, so it is not eventually 9.
(ii) $t \in [0,9)$: define (e_i) greedily by $e_n = 9$ if $(S_{n-1}+9)/n \le t$ and $e_n = 0$ otherwise, and put $A_n = S_n/n$. Then $A_n \le t$ for every n: if $e_n = 9$ this is the defining rule, and if $e_n = 0$ then $A_n = S_{n-1}/n \le S_{n-1}/(n-1) = A_{n-1} \le t$ for $n \ge 2$, while $A_1 = 0 \le t$. So $\limsup_n A_n \le t$. Also $e_n = 0$ for infinitely many n, since $e_n = 9$ from some point on would force $A_n \to 9 > t$; and each such n satisfies $A_n = S_{n-1}/n > t - \frac{9}{n}$, so $\limsup_n A_n \ge t$. Finally (e_n) is not eventually 9.
Given I and $t \in [0,9]$, Step 2 supplies a legal tail (e_i) and Step 1 converts it into $x \in I$ with $\varphi(x) = t$. Thus $\varphi(I) = [0,9]$ and $f(I) = h([0,9]) = \mathbb{R}$.

Problem 2D.24. For $A \subseteq \mathbb{R}$, the quantity

$$\sup\{|F| : F \text{ is a closed bounded subset of } \mathbb{R} \text{ and } F \subseteq A\}$$

is called the **inner measure** of A.
(a) Show that if A is a Lebesgue measurable subset of $\mathbb{R}$, then the inner measure of A equals the outer measure of A.
(b) Show that inner measure is not a measure on the σ-algebra of all subsets of $\mathbb{R}$.

Solution. Write $\mu_*(A)$ for the inner measure and $|A|$ for the outer measure. The empty set is an admissible F for every A, so $\mu_*(A) \in [0,\infty]$ is always defined; and every closed set is Borel, so each $|F|$ occurring is a Lebesgue measure.
(a). The inequality $\mu_*(A) \le |A|$ holds for every A: each closed bounded $F \subseteq A$ has $|F| \le |A|$ by 2.5. For the reverse, let $\varepsilon > 0$ and, A being Lebesgue measurable, use 2.71 ((a) implies (b)) to get a closed $F \subseteq A$ with $|A \setminus F| < \varepsilon$. The truncations $F_n = F \cap [-n,n]$ are closed, bounded, contained in A, and increase to F, so $\lim_n |F_n| = |F|$ by 2.59 applied to Lebesgue measure (a measure on the Lebesgue measurable sets by 2.72); as $|F_n| \le \mu_*(A)$ for every n, this gives $\mu_*(A) \ge |F|$. Meanwhile $A = F \cup (A\setminus F)$ and 2.8 give $|A| \le |F| + \varepsilon$. If $|A| = \infty$ this forces $|F| = \infty$, hence $\mu_*(A) = \infty = |A|$; otherwise $\mu_*(A) \ge |F| > |A| - \varepsilon$ for every $\varepsilon > 0$, so $\mu_*(A) \ge |A|$.
(b). For $r \in [0,1)$ and $S \subseteq [0,1)$ write

$$S \oplus r = \big((S \cap [0,1-r)) + r\big) \cup \big((S \cap [1-r,1)) + r - 1\big) \subseteq [0,1),$$

so that $x \mapsto x \oplus r$ is a bijection of $[0,1)$ onto itself. If S is Lebesgue measurable then so is $S \oplus r$, and since the two pieces are disjoint, additivity of Lebesgue measure (2.72) with translation invariance (2.7) gives $|S \oplus r| = |S|$.
By the axiom of choice let $V \subseteq [0,1)$ contain exactly one element of each class of the relation $x \sim y$ when $x - y \in \mathbb{Q}$ on $[0,1)$, let $r_1, r_2, \ldots$ enumerate $\mathbb{Q} \cap [0,1)$, and put $V_k = V \oplus r_k$.
The V_k are disjoint with union $[0,1)$. If $x \in V_j \cap V_k$, say $x = v \oplus r_j = w \oplus r_k$ with $v, w \in V$, then $x - (v + r_j)$ and $x - (w + r_k)$ lie in $\{0,-1\}$, so $v - w \in \mathbb{Q}$, forcing $v = w$ and then $r_j = r_k$ by injectivity of $r \mapsto v \oplus r$. Conversely, given $x \in [0,1)$ with representative $v \in V$, the number $r = x - v$ (or $x - v + 1$ if $x < v$) lies in $\mathbb{Q} \cap [0,1)$ and $x = v \oplus r$.
Each $\mu_*(V_k) = 0$. Since $x \oplus r_k \sim x$ for rational r_k and $\oplus r_k$ is injective, V_k meets each class exactly

once, so it suffices to show $\mu_*(S) = 0$ whenever $S \subseteq [0,1)$ meets each class at most once. Let $F \subseteq S$ be closed and bounded, hence Borel and Lebesgue measurable. The sets $F \oplus r$, $r \in \mathbb{Q} \cap [0,1)$, are disjoint: $x \oplus r = y \oplus r'$ with $x, y \in F$ and $r \neq r'$ forces $x - y \in \mathbb{Q}$ and $x \neq y$, two distinct points of S in one class. Each $F \oplus r$ is a measurable subset of $[0,1)$ with $|F \oplus r| = |F|$, so countable additivity (2.72) and 2.5 give

$$\sum_{r \in \mathbb{Q} \cap [0,1)} |F| = \Bigl| \bigcup_{r \in \mathbb{Q} \cap [0,1)} (F \oplus r) \Bigr| \leq |[0,1)| = 1,$$

a sum of countably infinitely many copies of $|F|$; hence $|F| = 0$ and $\mu_*(S) = 0$.
So $V_1, V_2, \ldots$ are disjoint subsets of $\mathbb{R}$ with $\sum_k \mu_*(V_k) = 0$, whereas by (a)

$$\mu_*\Bigl(\bigcup_{k=1}^{\infty} V_k\Bigr) = \mu_*\bigl([0,1)\bigr) = \bigl|[0,1)\bigr| = 1.$$

Thus μ_* is not countably additive on the σ-algebra of all subsets of $\mathbb{R}$.

2.5 Exercises 2E

Problem 2E.1. Suppose X is a finite set. Explain why a sequence of functions from X to $\mathbb{R}$ that converges pointwise on X also converges uniformly on X.

Solution. Take $n = \max\{n_1, \ldots, n_N\}$: a finite set of thresholds has a largest element, which is again in $\mathbb{Z}^+$.
The case $X = \emptyset$ is vacuous, so let $X = \{x_1, \ldots, x_N\}$ with $N \in \mathbb{Z}^+$, suppose $f_1, f_2, \ldots$ converges pointwise on X to f, and let $\varepsilon > 0$. By pointwise convergence (2.82) there is for each j an $n_j \in \mathbb{Z}^+$ with

$$|f_k(x_j) - f(x_j)| < \varepsilon \quad \text{for all integers } k \geq n_j.$$

With $n = \max\{n_1, \ldots, n_N\}$, every $k \geq n$ and every $x \in X$ satisfy $x = x_j$ for some j and $k \geq n_j$, so $|f_k(x) - f(x)| < \varepsilon$. This single n works for all $x \in X$ at once, so the convergence is uniform by 2.82.

Problem 2E.2. Give an example of a sequence of functions from $\mathbb{Z}^+$ to $\mathbb{R}$ that converges pointwise on $\mathbb{Z}^+$ but does not converge uniformly on $\mathbb{Z}^+$.

Solution. Take $f_k = \chi_{\{k\}}$.
For fixed $m \in \mathbb{Z}^+$ we have $f_k(m) = 0$ whenever $k > m$, so $f_k \to f \equiv 0$ pointwise. A uniform limit must agree with the pointwise limit, so uniform convergence would have to be to f; but for $\varepsilon = \frac{1}{2}$ and any $n \in \mathbb{Z}^+$, the index $k = n$ and the point $m = n$ give

$$|f_n(n) - f(n)| = 1 > \tfrac{1}{2},$$

so the condition in 2.82 fails. Hence the convergence is not uniform.

Problem 2E.3. Give an example of a sequence of continuous functions $f_1, f_2, \ldots$ from $[0,1]$ to $\mathbb{R}$ that converges pointwise to a function $f : [0,1] \to \mathbb{R}$ that is not a bounded function.

Solution. Take

$$f_k(x) = \begin{cases} k^2 x & \text{if } 0 \leq x \leq \frac{1}{k}, \\ \dfrac{1}{x} & \text{if } \frac{1}{k} < x \leq 1, \end{cases} \qquad f(x) = \begin{cases} 0 & \text{if } x = 0, \\ \dfrac{1}{x} & \text{if } 0 < x \leq 1. \end{cases}$$

Each f_k is continuous on $[0, \frac{1}{k}]$ and on $[\frac{1}{k}, 1]$ (where $x \geq \frac{1}{k} > 0$), and the two formulas agree at $\frac{1}{k}$ because $k^2 \cdot \frac{1}{k} = k$; hence f_k is continuous on $[0,1]$.
Pointwise convergence: $f_k(0) = 0 = f(0)$ for every k, while for $0 < x \leq 1$ any $k > 1/x$ gives $\frac{1}{k} < x$ and so $f_k(x) = \frac{1}{x} = f(x)$, an eventually constant sequence. Finally $f\bigl(\frac{1}{k}\bigr) = k$ for every k, so f is

unbounded.

Problem 2E.4. Prove or give a counterexample: If $A \subseteq \mathbb{R}$ and $f_1, f_2, \dots$ is a sequence of uniformly continuous functions from A to $\mathbb{R}$ that converges uniformly to a function $f : A \to \mathbb{R}$, then f is uniformly continuous on A.

Solution. The statement is true: f is uniformly continuous, by the $\varepsilon/3$ argument of 2.84 with the δ for f_n now independent of the point.
Suppose $\varepsilon > 0$. Because $f_1, f_2, \dots$ converges uniformly on A to f, there exists $n \in \mathbb{Z}^+$ such that

$$|f_n(z) - f(z)| < \frac{\varepsilon}{3} \quad \text{for all } z \in A.$$

Because f_n is uniformly continuous on A, there exists $\delta > 0$ such that

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \quad \text{for all } x, y \in A \text{ with } |x - y| < \delta.$$

Now suppose $x, y \in A$ and $|x - y| < \delta$. Then

$$\begin{aligned} |f(x) - f(y)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon. \end{aligned}$$

The δ works for all pairs of points of A at once, so f is uniformly continuous on A.

Problem 2E.5. Give an example to show that Egorov's Theorem can fail without the hypothesis that $\mu(X) < \infty$.

Solution. Take Lebesgue measure μ on the Borel subsets of $X = \mathbb{R}$ (so $\mu(X) = \infty$) and $f_k = \chi_{[k,\,k+1]}$, with $\varepsilon = 1$.
Each $x \in \mathbb{R}$ has $f_k(x) = 0$ for every $k > x$, so $f_k \to f \equiv 0$ pointwise. Now let E be any Borel set with $\mu(\mathbb{R} \setminus E) < 1$. For each $k \in \mathbb{Z}^+$, monotonicity applied to $[k, k+1] \setminus E \subseteq \mathbb{R} \setminus E$ gives

$$\big|[k, k+1] \setminus E\big| \le \big|\mathbb{R} \setminus E\big| < 1 = \big|[k, k+1]\big|,$$

so $[k, k+1] \cap E \neq \emptyset$; choose x_k there. Then for $\frac{1}{2}$ and any n, the index $k = n$ and the point $x_n \in E$ give $|f_n(x_n) - f(x_n)| = 1 > \frac{1}{2}$. Hence the convergence is uniform on no such E, and the conclusion of Egorov's Theorem (2.85) fails for $\varepsilon = 1$ although all its other hypotheses hold.

Problem 2E.6. Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) < \infty$. Suppose $f_1, f_2, \dots$ is a sequence of $\mathcal{S}$-measurable functions from X to $\mathbb{R}$ such that $\lim_{k\to\infty} f_k(x) = \infty$ for each $x \in X$. Prove that for every $\varepsilon > 0$, there exists a set $E \in \mathcal{S}$ such that $\mu(X \setminus E) < \varepsilon$ and $f_1, f_2, \dots$ converges uniformly to ∞ on E (meaning that for every $t > 0$, there exists $n \in \mathbb{Z}^+$ such that $f_k(x) > t$ for all integers $k \ge n$ and all $x \in E$).
[The exercise above is an Egorov-type theorem for sequences of functions that converge pointwise to ∞.]

Solution. Run the proof of Egorov's Theorem (2.85) with $\{x : |f_k(x) - f(x)| < \frac{1}{n}\}$ replaced by $\{x : f_k(x) > n\}$.
Suppose $\varepsilon > 0$ and fix $n \in \mathbb{Z}^+$. For $m \in \mathbb{Z}^+$ put

$$A_{m,n} = \bigcap_{k=m}^{\infty} \{x \in X : f_k(x) > n\} \in \mathcal{S}.$$

Then $A_{1,n} \subseteq A_{2,n} \subseteq \cdots$, and $\bigcup_m A_{m,n} = X$ because $\lim_{k\to\infty} f_k(x) = \infty$ gives, for each $x \in X$, an m with $f_k(x) > n$ for all $k \geq m$. Hence $\lim_{m\to\infty} \mu(A_{m,n}) = \mu(X)$ by 2.59, and since $\mu(X) < \infty$ we may choose m_n with $\mu(X) - \mu(A_{m_n,n}) < \varepsilon/2^n$; as $\mu(A_{m_n,n}) \leq \mu(X) < \infty$, 2.57(b) turns this into $\mu(X \setminus A_{m_n,n}) < \varepsilon/2^n$.
Let $E = \bigcap_{n=1}^{\infty} A_{m_n,n} \in \mathcal{S}$. By countable subadditivity (2.58),

$$\begin{aligned}\mu(X \setminus E) &\leq \sum_{n=1}^{\infty} \mu(X \setminus A_{m_n,n}) \\ &< \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.\end{aligned}$$

Given $t > 0$, choose $n \geq t$; then $E \subseteq A_{m_n,n}$ gives $f_k(x) > n \geq t$ for all $k \geq m_n$ and all $x \in E$, which is uniform convergence to ∞ on E.

Problem 2E.7. Suppose F is a closed bounded subset of $\mathbb{R}$ and $g_1, g_2, \ldots$ is an increasing sequence of continuous real-valued functions on F (thus $g_1(x) \leq g_2(x) \leq \cdots$ for all $x \in F$) such that $\sup\{g_1(x), g_2(x), \ldots\} < \infty$ for each $x \in F$. Define a real-valued function g on F by

$$g(x) = \lim_{k\to\infty} g_k(x).$$

Prove that g is continuous on F if and only if $g_1, g_2, \ldots$ converges uniformly on F to g.
[The result above is called Dini's Theorem.]

Solution. Each increasing bounded-above sequence $g_1(x) \leq g_2(x) \leq \cdots$ converges to $g(x) = \sup_k g_k(x) \in \mathbb{R}$, so $g_k \leq g$ on F for every k.
(i) If $g_1, g_2, \ldots$ converges uniformly to g on F, then g is continuous at each point of F by 2.84, since each g_k is.
(ii) Conversely, suppose g is continuous on F and set $h_k = g - g_k$, a continuous real-valued function on F with $h_k \geq 0$, $h_1 \geq h_2 \geq \cdots$, and $h_k(x) \to 0$ for each $x \in F$. Uniform convergence means: for each $\varepsilon > 0$ there is n with $h_k < \varepsilon$ on F for all $k \geq n$; by monotonicity in k it suffices to find one n with $h_n < \varepsilon$ on F.
If this fails for some $\varepsilon > 0$, pick $x_n \in F$ with $h_n(x_n) \geq \varepsilon$ for every n. As F is bounded, Bolzano-Weierstrass gives a subsequence $x_{n_j} \to x$, and $x \in F$ since F is closed. Fix m; for all j with $n_j \geq m$, monotonicity in the index gives $h_m(x_{n_j}) \geq h_{n_j}(x_{n_j}) \geq \varepsilon$, so continuity of h_m at x yields

$$h_m(x) = \lim_{j\to\infty} h_m(x_{n_j}) \geq \varepsilon.$$

This holds for every m, contradicting $h_m(x) \to 0$. Hence the convergence is uniform.

Problem 2E.8. Suppose μ is the measure on $(\mathbb{Z}^+, 2^{\mathbb{Z}^+})$ defined by

$$\mu(E) = \sum_{n\in E} \frac{1}{2^n}.$$

Prove that for every $\varepsilon > 0$, there exists a set $E \subseteq \mathbb{Z}^+$ with $\mu(\mathbb{Z}^+ \setminus E) < \varepsilon$ such that $f_1, f_2, \ldots$ converges uniformly on E for every sequence of functions $f_1, f_2, \ldots$ from $\mathbb{Z}^+$ to $\mathbb{R}$ that converges pointwise on $\mathbb{Z}^+$.
[This result does not follow from Egorov's Theorem because here we are asking for E to depend only on ε. In Egorov's Theorem, E depends on ε and on the sequence $f_1, f_2, \ldots$.]

Solution. Given $\varepsilon > 0$, choose N with $\frac{1}{2^N} < \varepsilon$ and take $E = \{1, \ldots, N\}$, a set depending only on ε.

Summing the geometric series,

$$\mu(\mathbb{Z}^+ \setminus E) = \sum_{n=N+1}^{\infty} \frac{1}{2^n} = \frac{1}{2^N} < \varepsilon.$$

Pointwise convergence on the finite set E is automatically uniform: if $f_k \to f$ pointwise on $\mathbb{Z}^+$ and $\delta > 0$, pick M_n for each $n \in E$ with $|f_k(n) - f(n)| < \delta$ for all $k \geq M_n$, and set $M = \max\{M_1, \dots, M_N\}$, a maximum over finitely many indices. Then $\sup_{n \in E} |f_k(n) - f(n)| \leq \delta$ for all $k \geq M$.

Problem 2E.9. Suppose $F_1, \dots, F_n$ are disjoint closed subsets of $\mathbb{R}$. Prove that if

$$g : F_1 \cup \dots \cup F_n \to \mathbb{R}$$

is a function such that $g|_{F_k}$ is a continuous function for each $k \in \{1, \dots, n\}$, then g is a continuous function.

Solution. Near any $x \in F = F_1 \cup \dots \cup F_n$ the other pieces are a positive distance away, so continuity of g at x reduces to continuity of the single restriction $g|_{F_j}$ with $x \in F_j$.
Fix $x \in F$ and $\varepsilon > 0$, and pick j with $x \in F_j$. The set $C = \bigcup_{k \neq j} F_k$ is closed (a finite union of closed sets) and misses x (disjointness), so $\mathbb{R} \setminus C$ is open about x and there is $\delta_1 > 0$ with

$$(x - \delta_1, x + \delta_1) \cap C = \varnothing.$$

Continuity of $g|_{F_j}$ at x gives $\delta_2 > 0$ with $|g(y) - g(x)| < \varepsilon$ for all $y \in F_j$ satisfying $|y - x| < \delta_2$. Put $\delta = \min\{\delta_1, \delta_2\}$. If $y \in F$ and $|y - x| < \delta$, then $y \notin C$, so $y \in F_j$, and hence $|g(y) - g(x)| < \varepsilon$. Thus g is continuous at x.

Problem 2E.10. Suppose $F \subseteq \mathbb{R}$ is such that every continuous function from F to $\mathbb{R}$ can be extended to a continuous function from $\mathbb{R}$ to $\mathbb{R}$. Prove that F is a closed subset of $\mathbb{R}$.

Solution. If F is not closed, then $g(x) = \frac{1}{x-b}$ is a continuous function on F with no continuous extension to $\mathbb{R}$; this contrapositive gives the result.
Not being closed, F omits some limit point $b \notin F$ of itself, so there are $x_j \in F$ with $x_j \to b$ and $x_j \neq b$. Since $b \notin F$ we have $F \subseteq \mathbb{R} \setminus \{b\}$, so g is the restriction to F of a function continuous on $\mathbb{R} \setminus \{b\}$, hence continuous. If $h : \mathbb{R} \to \mathbb{R}$ were continuous with $h|_F = g$, then $h(x_j) \to h(b) \in \mathbb{R}$, so $h(x_1), h(x_2), \dots$ would be bounded; but

$$|h(x_j)| = |g(x_j)| = \frac{1}{|x_j - b|} \to \infty$$

because $0 < |x_j - b| \to 0$, a contradiction.

Problem 2E.11. Prove or give a counterexample: If $F \subseteq \mathbb{R}$ is such that every bounded continuous function from F to $\mathbb{R}$ can be extended to a continuous function from $\mathbb{R}$ to $\mathbb{R}$, then F is a closed subset of $\mathbb{R}$.

Solution. The statement is true: the unbounded witness of Exercise 10 can be replaced by a bounded function oscillating between 0 and 1 along a sequence tending to the missing limit point.
We again prove the contrapositive, so let $b \notin F$ be a limit point of F. Then F meets $(b, b + \delta)$ for every $\delta > 0$ or meets $(b - \delta, b)$ for every $\delta > 0$ (otherwise the smaller of the two failing radii would isolate b from F, using $b \notin F$); replacing F, b by $-F, -b$ if necessary — the reflection $x \mapsto -x$ is a homeomorphism of $\mathbb{R}$, and transports a non-extendable bounded continuous function back — we may assume the first.
Recursively choose $x_1 \in F \cap (b, b + 1)$ and $x_{j+1} \in F \cap \big(b, b + \frac{x_j - b}{2}\big)$, nonempty at each stage. Then $x_1 > x_2 > \dots > b$ with $0 < x_j - b \leq 2^{1-j}$, so $x_j \to b$. Let $a_j = 1$ for j odd and $a_j = 0$ for j even, and

define $\varphi : \mathbb{R} \setminus \{b\} \to [0,1]$ to be 0 on $(-\infty, b)$, equal to 1 on $[x_1, \infty)$, and affine on each $[x_{j+1}, x_j]$ with

$$\varphi(x_j) = a_j \quad \text{for every } j \in \mathbb{Z}^+.$$

These intervals together with $[x_1, \infty)$ cover (b, ∞), since $x_i \to b < x$ makes $\{i : x \le x_i\}$ finite for each $x > b$, and consecutive pieces agree at the shared endpoints; so φ is well defined. It is constant on the open set $(-\infty, b)$, and each $x > b$ has a neighborhood (x_j, ∞) on which φ is pasted from the finitely many closed affine pieces $[x_j, x_{j-1}], \dots, [x_1, \infty)$, continuous by Exercise 9. Hence φ is continuous on $\mathbb{R} \setminus \{b\} \supseteq F$.
Set $g = \varphi|_F$, a continuous function on F with $0 \le g \le 1$. A continuous $h : \mathbb{R} \to \mathbb{R}$ with $h|_F = g$ would force

$$h(b) = \lim_{j\to\infty} h(x_j) = \lim_{j\to\infty} a_j,$$

but $1, 0, 1, 0, \dots$ does not converge. So g has no continuous extension, and F must be closed.

Problem 2E.12. Give an example of a Borel measurable function f from $\mathbb{R}$ to $\mathbb{R}$ such that there does not exist a set $B \subseteq \mathbb{R}$ such that $|\mathbb{R} \setminus B| = 0$ and $f|_B$ is a continuous function on B.

Solution. Take $f = \chi_A$ for a Borel set A with $|A \cap I| > 0$ and $|I \setminus A| > 0$ for every nonempty open interval I; we build such an A from nowhere dense sets of positive measure. (Note $f = \chi_{\mathbb{Q}}$ fails, since $B = \mathbb{R} \setminus \mathbb{Q}$ works for it.)
(i) Every nonempty bounded open interval $J = (c, d)$ contains a closed set K with empty interior and $|K| > 0$. Put $[c', d'] \subseteq J$ with $d' - c' = \frac{d-c}{2}$, let $\varepsilon = \frac{d'-c'}{2}$, enumerate $\mathbb{Q}$ as $q_1, q_2, \dots$, and let $U = \bigcup_j (q_j - \varepsilon 2^{-j-1}, q_j + \varepsilon 2^{-j-1})$, so $|U| \le \varepsilon$ by subadditivity (2.8) and 2.14. Then $K = [c', d'] \setminus U$ is closed, contains no rational and so has empty interior, and $[c', d'] \subseteq K \cup U$ gives

$$d' - c' \le |K| + |U| \le |K| + \tfrac{d'-c'}{2},$$

whence $|K| \ge \frac{d'-c'}{2} > 0$.
(ii) A finite union of closed sets with empty interior is closed with empty interior: for two such sets E, G and a nonempty open $V \subseteq E \cup G$, the open set $V \setminus E$ is nonempty (else $V \subseteq E$) and lies in G, a contradiction; induct.
(iii) Enumerate the open intervals with rational endpoints as $I_1, I_2, \dots$, and choose recursively disjoint closed sets K_n, L_n of positive measure and empty interior with $K_n \cup L_n \subseteq I_n$: given $C = K_1 \cup L_1 \cup \dots \cup L_{n-1}$, which is closed with empty interior by (ii), the open set $I_n \setminus C$ is nonempty, so it contains $(y - r, y + r)$ for some y, r, and (i) supplies $K_n \subseteq (y - \frac{r}{2}, y)$ and $L_n \subseteq (y, y + \frac{r}{2})$. Set $A = \bigcup_n K_n$, a Borel set disjoint from every L_n, and $f = \chi_A$.
Each nonempty open interval I contains some I_n, so $|A \cap I| \ge |K_n| > 0$ and $|I \setminus A| \ge |L_n| > 0$ by monotonicity. Now let $|\mathbb{R} \setminus B| = 0$. From $A \subseteq (A \cap B) \cup (\mathbb{R} \setminus B)$ and $|A| \ge |K_1| > 0$, subadditivity gives $|A \cap B| > 0$; pick $x \in A \cap B$, so $f(x) = 1$. If $f|_B$ were continuous at x, then taking $\varepsilon = \frac{1}{2}$ and using that f is $\{0,1\}$-valued, there is $\delta > 0$ with $B \cap I \subseteq A$ for $I = (x - \delta, x + \delta)$; hence

$$I \setminus A \subseteq I \setminus (B \cap I) \subseteq \mathbb{R} \setminus B,$$

forcing $|I \setminus A| = 0$, contrary to $|I \setminus A| > 0$.

Problem 2E.13. Prove or give a counterexample: If $f_t : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function for each $t \in \mathbb{R}$ and $f : \mathbb{R} \to (-\infty, \infty]$ is defined by

$$f(x) = \sup\{f_t(x) : t \in \mathbb{R}\},$$

then f is a Borel measurable function.

Solution. The statement is false: take a non-Borel set $A \subseteq \mathbb{R}$ (2.67) and set $f_t = \chi_{\{t\}}$ if $t \in A$, $f_t = 0$ if $t \notin A$. The countable analogue 2.53 holds only because σ-algebras are closed under countable unions. Each f_t is Borel measurable, being constant or the characteristic function of the closed set $\{t\}$. For $x \in A$ the choice $t = x$ gives $f_x(x) = 1$, and all f_t are $\{0,1\}$-valued, so $f(x) = 1$; for $x \notin A$ every $t \in A$ satisfies $t \neq x$ and so $f_t(x) = 0$, giving $f(x) = 0$. Thus $f = \chi_A$ and

$$f^{-1}\big((\tfrac{1}{2}, \infty]\big) = A,$$

which is not Borel; since $(\frac{1}{2}, \infty]$ is a Borel subset of $[-\infty, \infty]$ by 2.50, the definition 2.51 shows f is not Borel measurable.

Problem 2E.14. Suppose $b_1, b_2, \ldots$ is a sequence of real numbers. Define $f : \mathbb{R} \to [0, \infty]$ by

$$f(x) = \begin{cases} \displaystyle\sum_{k=1}^{\infty} \frac{1}{4^k |x - b_k|} & \text{if } x \notin \{b_1, b_2, \ldots\}, \\ \infty & \text{if } x \in \{b_1, b_2, \ldots\}. \end{cases}$$

Prove that $|\{x \in \mathbb{R} : f(x) < 1\}| = \infty$.
[This exercise is a variation of a problem originally considered by Borel. If $b_1, b_2, \ldots$ contains all the rational numbers, then it is not even obvious that $\{x \in \mathbb{R} : f(x) < \infty\} \neq \varnothing$.]

Solution. The set $N = \{x : f(x) \geq 1\}$ is covered by the intervals $A_k = [b_k - 2^{1-k}, b_k + 2^{1-k}]$, of total length 4, and a set whose complement has finite outer measure has infinite outer measure.
If $x \notin \bigcup_k A_k$, then $|x - b_k| > 2^{1-k} > 0$ for every k, so $x \notin \{b_1, b_2, \ldots\}$ and $f(x)$ is given by the series, with

$$\begin{aligned} f(x) = \sum_{k=1}^{\infty} \frac{1}{4^k |x - b_k|} &< \sum_{k=1}^{\infty} \frac{1}{2^{2k} \cdot 2^{1-k}} \\ &= \sum_{k=1}^{\infty} \frac{1}{2^{k+1}} = \frac{1}{2}. \end{aligned}$$

Hence $N \subseteq \bigcup_k A_k$, and countable subadditivity (2.8) with 2.14 gives

$$|N| \leq \sum_{k=1}^{\infty} |A_k| = \sum_{k=1}^{\infty} \frac{4}{2^k} = 4.$$

Writing $E = \{x : f(x) < 1\} = \mathbb{R} \setminus N$, for each $n \in \mathbb{Z}^+$ the inclusion $[-n, n] \subseteq (E \cap [-n, n]) \cup N$ and subadditivity give

$$2n = |[-n, n]| \leq |E \cap [-n, n]| + |N| \leq |E| + 4.$$

Letting $n \to \infty$ yields $|E| = \infty$.

Problem 2E.15. Suppose B is a Borel set and $f : B \to \mathbb{R}$ is a Lebesgue measurable function. Show that there exists a Borel measurable function $g : B \to \mathbb{R}$ such that

$$|\{x \in B : g(x) \neq f(x)\}| = 0.$$

Solution. Extend f by 0, apply 2.95 to the extension, and restrict the resulting Borel function back to B.
Define $\tilde{f} : \mathbb{R} \to \mathbb{R}$ by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in B, \\ 0 & \text{if } x \in \mathbb{R} \setminus B. \end{cases}$$

Then $\tilde{f}$ is Lebesgue measurable: for a Borel set $A \subseteq \mathbb{R}$,

$$\tilde{f}^{-1}(A) = \begin{cases} f^{-1}(A) & \text{if } 0 \notin A, \\ f^{-1}(A) \cup (\mathbb{R} \setminus B) & \text{if } 0 \in A, \end{cases}$$

a union of Lebesgue measurable sets, since f is Lebesgue measurable and the Borel set $\mathbb{R} \setminus B$ is Lebesgue measurable.
By 2.95 there is a Borel measurable $\tilde{g} : \mathbb{R} \to \mathbb{R}$ with $|\{x \in \mathbb{R} : \tilde{g}(x) \neq \tilde{f}(x)\}| = 0$. Put $g = \tilde{g}|_B$. For Borel A we have $g^{-1}(A) = B \cap \tilde{g}^{-1}(A)$, Borel because B is; so g is Borel measurable. Finally $\tilde{f} = f$ on B gives

$$\{x \in B : g(x) \neq f(x)\} \subseteq \{x \in \mathbb{R} : \tilde{g}(x) \neq \tilde{f}(x)\},$$

so $|\{x \in B : g(x) \neq f(x)\}| = 0$ by monotonicity of outer measure (2.5).

3 Integration

Contents

3.1 Exercises 3A

Problem 3A.1. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to [0, \infty]$ is an $\mathcal{S}$-measurable function such that $\int f \, d\mu < \infty$. Explain why

$$\inf_E f = 0$$

for each set $E \in \mathcal{S}$ with $\mu(E) = \infty$.

Solution. A positive lower bound on a set of infinite measure would already make one lower Lebesgue sum infinite.
Suppose $\mu(E) = \infty$ and $c = \inf_E f > 0$ (note $E \neq \varnothing$, since $\mu(\varnothing) = 0$). Taking the $\mathcal{S}$-partition $\mathcal{P}$ of X given by $E, X \setminus E$, the definition 3.2 and nonnegativity of the discarded term give

$$\mathcal{L}(f, \mathcal{P}) \geq \mu(E) \inf_E f = \infty \cdot c = \infty.$$

Hence $\int f \, d\mu = \infty$ by 3.3, contradicting $\int f \, d\mu < \infty$. As $\inf_E f \geq 0$ always, $\inf_E f = 0$.

Problem 3A.2. Suppose X is a set, $\mathcal{S}$ is a σ-algebra on X, and $c \in X$. Define the Dirac measure δ_c on $(X, \mathcal{S})$ by

$$\delta_c(E) = \begin{cases} 1 & \text{if } c \in E, \\ 0 & \text{if } c \notin E. \end{cases}$$

Prove that if $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable, then $\int f \, d\delta_c = f(c)$.
[Careful: $\{c\}$ may not be in $\mathcal{S}$.]

Solution. Every lower Lebesgue sum for δ_c collapses to $\inf_A f$ over the one partition piece A containing c, and the level sets $\{f > t\}$ supply pieces on which that infimum is nearly $f(c)$.
(i) $\int f \, d\delta_c \leq f(c)$. If $A_1, \ldots, A_m$ is an $\mathcal{S}$-partition of X, then c lies in exactly one piece A_{j_0}, so $\delta_c(A_{j_0}) = 1$ and the other terms vanish (with $0 \cdot \infty = 0$):

$$\mathcal{L}(f, \mathcal{P}) = \sum_{j=1}^{m} \delta_c(A_j) \inf_{A_j} f = \inf_{A_{j_0}} f \leq f(c).$$

Take the supremum over $\mathcal{P}$ and use 3.3.
(ii) $\int f \, d\delta_c \geq f(c)$. Assume $f(c) > 0$ and take any $t \in [0, f(c))$. Since $(t, \infty]$ is a Borel subset of $[-\infty, \infty]$ by 2.50 and f is $\mathcal{S}$-measurable, the set $A = \{x \in X : f(x) > t\}$ lies in $\mathcal{S}$ by 2.51, and $c \in A$; this replaces the partition by $\{c\}, X \setminus \{c\}$, which is unavailable since $\{c\}$ need not lie in $\mathcal{S}$. For the partition $A, X \setminus A$,

$$\mathcal{L}(f, \mathcal{P}) = \delta_c(A) \inf_A f = \inf_A f \geq t.$$

Letting $t \uparrow f(c)$ gives the inequality, also when $f(c) = \infty$.

Problem 3A.3. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to [0, \infty]$ is an $\mathcal{S}$-measurable function. Prove that

$$\int f \, d\mu > 0 \quad \text{if and only if} \quad \mu\big(\{x \in X : f(x) > 0\}\big) > 0.$$

Solution. Let $E = \{f > 0\}$ and $E_k = \{f > \frac{1}{k}\}$, all in $\mathcal{S}$ by 2.50 and 2.51.
(i) If $\mu(E) > 0$: the sets E_k increase to E (any x with $f(x) > 0$ has $f(x) > \frac{1}{k}$ for some k), so

$\mu(E_k) \to \mu(E) > 0$ by 2.59 and $\mu(E_k) > 0$ for some k. Since $\frac{1}{k}\chi_{E_k} \le f$ pointwise, 3.8 and 3.7 give

$$\int f\, d\mu \ge \int \tfrac{1}{k}\chi_{E_k}\, d\mu = \tfrac{1}{k}\mu(E_k) > 0.$$

(ii) If $\mu(E) = 0$: every term $\mu(A_j) \inf_{A_j} f$ of a lower Lebesgue sum vanishes. Either $A_j \subseteq E$, whence $\mu(A_j) = 0$ by 2.57(a), or A_j contains a point outside E, whence $\inf_{A_j} f = 0$; in both cases the convention $0 \cdot \infty = 0$ gives the term 0. So $\int f\, d\mu = 0$ by 3.3.

Problem 3A.4. Give an example of a Borel measurable function $f : [0,1] \to (0,\infty)$ such that $L(f,[0,1]) = 0$.
[Recall that $L(f,[0,1])$ denotes the lower Riemann integral, which was defined in Section 1A. If λ is Lebesgue measure on $[0,1]$, then the previous exercise states that $\int f\, d\lambda > 0$ for this function f, which is what we expect of a positive function. Thus even though both $L(f,[0,1])$ and $\int f\, d\lambda$ are defined by taking the supremum of approximations from below, Lebesgue measure captures the right behavior for this function f and the lower Riemann integral does not.]

Solution. List $\mathbb{Q} \cap [0,1]$ bijectively as $r_1, r_2, \ldots$ and take

$$f(x) = \begin{cases} \dfrac{1}{k} & \text{if } x = r_k, \\ 1 & \text{if } x \text{ is irrational,} \end{cases}$$

whose values lie in $(0,1]$. It is Borel measurable, since for Borel B the set $f^{-1}(B)$ is the union of the countable set $\{r_k : \frac{1}{k} \in B\}$ with either $[0,1] \setminus \mathbb{Q}$ or $\varnothing$.
Every nondegenerate subinterval $[x_{j-1}, x_j]$ contains r_k for infinitely many k, so $\inf_{[x_{j-1},x_j]} f = 0$. Hence every lower Riemann sum (1.3) of a partition P of $[0,1]$ is

$$L(f,P,[0,1]) = \sum_{j=1}^{n} (x_j - x_{j-1}) \cdot 0 = 0,$$

and taking the supremum over P gives $L(f,[0,1]) = 0$ by 1.7.

Problem 3A.5. Verify the assertion that integration with respect to counting measure is summation (Example 3.6).

Solution. The assertion is that $\int b\, d\mu = \sum_{k=1}^{\infty} b_k$ for counting measure μ on $\mathbb{Z}^+$ (2.55) and $b(k) = b_k \ge 0$; both inequalities come from 3.2 and 3.3, with $\sum_{k=1}^{\infty} b_k = \sup\{\sum_{k\in F} b_k : F \text{ finite}\}$.
(i) $\int b\, d\mu \ge \sum_{k=1}^{\infty} b_k$. Fix n and take the $\mathcal{S}$-partition $\mathcal{P}_n$ of $\mathbb{Z}^+$ given by

$$\{1\}, \{2\}, \ldots, \{n\}, \quad \{n+1, n+2, \ldots\}.$$

Since $\mu(\{k\}) = 1$ and the final term is nonnegative,

$$\mathcal{L}(b, \mathcal{P}_n) = \sum_{k=1}^{n} b_k + \mu(\{n+1, n+2, \ldots\}) \inf_{k>n} b_k \ge \sum_{k=1}^{n} b_k,$$

so $\int b\, d\mu \ge \sum_{k=1}^{n} b_k$ for every n; let $n \to \infty$.
(ii) $\int b\, d\mu \le \sum_{k=1}^{\infty} b_k$. For an $\mathcal{S}$-partition $A_1, \ldots, A_m$ and $c_j = \inf_{A_j} b$ we claim $\mu(A_j)c_j \le \sum_{k\in A_j} b_k$: for $A_j = \varnothing$ both sides are 0; for A_j finite with n_j elements each $b_k \ge c_j$ gives $n_j c_j \le \sum_{k\in A_j} b_k$; for A_j infinite with $c_j > 0$, finite subsets of any size N give $\sum_{k\in A_j} b_k \ge N c_j$, so the right side is ∞ (and if $c_j = 0$ the left side is $\infty \cdot 0 = 0$). Hence

$$\mathcal{L}(b, \mathcal{P}) \le \sum_{j=1}^{m} \sum_{k\in A_j} b_k \le \sum_{k=1}^{\infty} b_k,$$

the last step because approximating each $\sum_{k\in A_j} b_k$ from below by a finite $F_j \subseteq A_j$ and using disjointness makes $F_1 \cup \cdots \cup F_m$ a finite subset of $\mathbb{Z}^+$. Now apply 3.3.
Method (2): with $b^{(n)} = \sum_{k=1}^{n} b_k \chi_{\{k\}}$ we have $0 \le b^{(1)} \le b^{(2)} \le \cdots$ with $b^{(n)}(k) = b_k$ once $n \ge k$, so the Monotone Convergence Theorem 3.11 and 3.7 give

$$\int b\, d\mu = \lim_{n\to\infty} \sum_{k=1}^{n} b_k \mu(\{k\}) = \sum_{k=1}^{\infty} b_k.$$

Problem 3A.6. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $\mathcal{P}$ and $\mathcal{P}'$ are $\mathcal{S}$-partitions of X such that each set in $\mathcal{P}'$ is contained in some set in $\mathcal{P}$. Prove that $\mathcal{L}(f, \mathcal{P}) \le \mathcal{L}(f, \mathcal{P}')$.

Solution. Group the pieces of $\mathcal{P}'$ by the piece of $\mathcal{P}$ containing them; refining only raises the infima. Write $\mathcal{P} = A_1, \ldots, A_m$ and $\mathcal{P}' = B_1, \ldots, B_n$, and fix for each k an index $j(k)$ with $B_k \subseteq A_{j(k)}$. Then for each j,

$$A_j = \bigcup_{\{k\,:\, j(k)=j\}} B_k,$$

a disjoint union: $\supseteq$ is clear, and if $x \in A_j$ then the unique B_k containing x satisfies $x \in A_j \cap A_{j(k)}$, forcing $j(k) = j$ by disjointness of $\mathcal{P}$. Finite additivity of μ therefore gives $\mu(A_j) = \sum_{\{k:j(k)=j\}} \mu(B_k)$. Grouping the terms of $\mathcal{L}(f, \mathcal{P}')$ by j,

$$\begin{aligned}
\mathcal{L}(f, \mathcal{P}') &= \sum_{k=1}^{n} \mu(B_k) \inf_{B_k} f \\
&= \sum_{j=1}^{m} \sum_{\{k\,:\, j(k)=j\}} \mu(B_k) \inf_{B_k} f \\
&\ge \sum_{j=1}^{m} \sum_{\{k\,:\, j(k)=j\}} \mu(B_k) \inf_{A_j} f \\
&= \sum_{j=1}^{m} \Big(\inf_{A_j} f\Big) \sum_{\{k\,:\, j(k)=j\}} \mu(B_k) \\
&= \sum_{j=1}^{m} \mu(A_j) \inf_{A_j} f \\
&= \mathcal{L}(f, \mathcal{P}).
\end{aligned}$$

The third line uses $\inf_{B_k} f \ge \inf_{A_j} f$ for $B_k \subseteq A_j$ (multiplication by $\mu(B_k) \in [0, \infty]$ preserving it, with $0 \cdot \infty = 0$); the fourth uses the distributive law $\sum_i c\,t_i = c \sum_i t_i$ in $[0, \infty]$ (Check!), and the fifth the additivity above.

Problem 3A.7. Suppose X is a set, $\mathcal{S}$ is the σ-algebra of all subsets of X, and $w : X \to [0, \infty]$ is a function. Define a measure μ on $(X, \mathcal{S})$ by

$$\mu(E) = \sum_{x\in E} w(x)$$

for $E \subseteq X$. Prove that if $f : X \to [0, \infty]$ is a function, then

$$\int f\, d\mu = \sum_{x\in X} w(x) f(x),$$

where the infinite sums above are defined as the supremum of all sums over finite subsets of E (first sum) or X (second sum).

Solution. Write $S = \sum_{x\in X} w(x)f(x)$, each such sum being the supremum of its finite subsums, and $0\cdot\infty = \infty\cdot 0 = 0$; note $\mu(\{x\}) = w(x)$.
Two facts are used. (a) If $A_1,\dots,A_m \subseteq A$ are disjoint and $g : X \to [0,\infty]$, then $\sum_j \sum_{x\in A_j} g(x) \le \sum_{x\in A} g(x)$: choosing finite $F_j \subseteq A_j$ with $\sum_{x\in F_j} g$ close to (or, if infinite, large below) $\sum_{x\in A_j} g$ makes $F_1 \cup \cdots \cup F_m$ a finite subset of A, and taking suprema over such choices gives the claim. (b) If $g(x) \ge c\,w(x)$ on A for some $c \in [0,\infty]$, then $\sum_{x\in A} g(x) \ge c \sum_{x\in A} w(x)$, by the distributive law on finite subsums and $\sup_F (c\,t_F) = c \sup_F t_F$ (Check! in the cases $c = 0,\infty$).
(i) $\int f\,d\mu \ge S$. For a finite $F = \{x_1,\dots,x_p\} \subseteq X$ take the $\mathcal{S}$-partition $\{x_1\},\dots,\{x_p\}, X\setminus F$; discarding the last nonnegative term,

$$\mathcal{L}(f,\mathcal{P}) = \sum_{i=1}^{p} w(x_i)f(x_i) + \mu(X\setminus F)\inf_{X\setminus F} f \;\ge\; \sum_{x\in F} w(x)f(x).$$

Take suprema over F and apply 3.3.
(ii) $\int f\,d\mu \le S$. For an $\mathcal{S}$-partition $A_1,\dots,A_m$ and $c_j = \inf_{A_j} f$, we have $w(x)f(x) \ge c_j w(x)$ on A_j, so (b) gives

$$\sum_{x\in A_j} w(x)f(x) \;\ge\; c_j \sum_{x\in A_j} w(x) = \mu(A_j)\inf_{A_j} f.$$

Summing over j and applying (a) with $A = X$ gives $\mathcal{L}(f,\mathcal{P}) \le S$; now use 3.3.

Problem 3A.8. Suppose λ denotes Lebesgue measure on $\mathbb{R}$. Give an example of a sequence $f_1, f_2, \dots$ of simple Borel measurable functions from $\mathbb{R}$ to $[0,\infty)$ such that $\lim_{k\to\infty} f_k(x) = 0$ for every $x \in \mathbb{R}$ but $\lim_{k\to\infty} \int f_k\,d\lambda = 1$.

Solution. Take $f_k = \chi_{[k,\,k+1]}$, a simple Borel measurable function.
Each $x \in \mathbb{R}$ has $f_k(x) = 0$ for every $k > x$, so $f_k \to 0$ pointwise, while 3.4 gives

$$\int f_k\,d\lambda = \lambda\big([k,k+1]\big) = 1$$

for every k.

Problem 3A.9. Suppose μ is a measure on a measurable space $(X,\mathcal{S})$ and $f : X \to [0,\infty]$ is an $\mathcal{S}$-measurable function. Define $\nu : \mathcal{S} \to [0,\infty]$ by

$$\nu(A) = \int \chi_A f\,d\mu$$

for $A \in \mathcal{S}$. Prove that ν is a measure on $(X,\mathcal{S})$.

Solution. Countable additivity of ν is the Monotone Convergence Theorem 3.11 applied to the partial sums $\chi_{A_1\cup\cdots\cup A_n} f$; the rest is immediate.
Since $\chi_\emptyset f$ is identically 0, $\nu(\emptyset) = 0$. Let $A_1, A_2, \dots \in \mathcal{S}$ be disjoint with union A, and put

$$g_n = \sum_{k=1}^{n} \chi_{A_k} f = \chi_{A_1\cup\cdots\cup A_n} f,$$

the second equality by disjointness. Each g_n is $\mathcal{S}$-measurable, $0 \le g_1 \le g_2 \le \cdots$, and $g_n \to \chi_A f$ pointwise (each $x \in A$ lies in exactly one A_k, so $g_n(x) = f(x)$ for $n \ge k$; off A both sides are 0). Hence

3.11 and additivity 3.16 give

$$\nu(A) = \lim_{n\to\infty}\int g_n\,d\mu = \lim_{n\to\infty}\sum_{k=1}^{n}\nu(A_k) = \sum_{k=1}^{\infty}\nu(A_k).$$

So ν is a measure by 2.54.

Problem 3A.10. Suppose $(X,\mathcal{S},\mu)$ is a measure space and $f_1, f_2, \ldots$ is a sequence of nonnegative $\mathcal{S}$-measurable functions. Define $f : X \to [0,\infty]$ by $f(x) = \sum_{k=1}^{\infty} f_k(x)$. Prove that

$$\int f\,d\mu = \sum_{k=1}^{\infty}\int f_k\,d\mu.$$

Solution. Apply the Monotone Convergence Theorem 3.11 to the partial sums $g_n = \sum_{k=1}^{n} f_k$. Each g_n is $\mathcal{S}$-measurable, since for $a \ge 0$ the sum of two measurable $u, v : X \to [0,\infty]$ satisfies

$$(u+v)^{-1}\big((a,\infty]\big) = \bigcup_{q\in\mathbb{Q}}\Big(u^{-1}\big((q,\infty]\big)\cap v^{-1}\big((a-q,\infty]\big)\Big),$$

a countable union in $\mathcal{S}$ (for $\supseteq$ add the inequalities; for $\subseteq$ take rational q between $a - v(x)$ and $u(x)$, or $q > a$ if $u(x) = \infty$ and $q < 0$ if $v(x) = \infty$), so 2.52 applies; induct. Nonnegativity gives $0 \le g_1 \le g_2 \le \cdots$ with $g_n(x) \to f(x)$ for every x, so f is $\mathcal{S}$-measurable by 2.53, and 3.11 together with additivity 3.16 gives

$$\int f\,d\mu = \lim_{n\to\infty}\int g_n\,d\mu = \lim_{n\to\infty}\sum_{k=1}^{n}\int f_k\,d\mu = \sum_{k=1}^{\infty}\int f_k\,d\mu.$$

Problem 3A.11. Suppose $(X,\mathcal{S},\mu)$ is a measure space and $f_1, f_2, \ldots$ are $\mathcal{S}$-measurable functions from X to $\mathbb{R}$ such that $\sum_{k=1}^{\infty}\int |f_k|\,d\mu < \infty$. Prove that there exists $E \in \mathcal{S}$ such that $\mu(X\setminus E) = 0$ and $\lim_{k\to\infty} f_k(x) = 0$ for every $x \in E$.

Solution. Take $E = \{x \in X : g(x) < \infty\}$, where $g(x) = \sum_{k=1}^{\infty}|f_k(x)|$.
Each $|f_k|$ is $\mathcal{S}$-measurable, since for $a \ge 0$ we have $|f_k|^{-1}((a,\infty]) = f_k^{-1}((a,\infty)) \cup f_k^{-1}((-\infty,-a)) \in \mathcal{S}$, so 2.52 applies; hence g is $\mathcal{S}$-measurable as the pointwise supremum (2.53) of its partial sums, and $E \in \mathcal{S}$. By Exercise 3A.10 applied to $|f_1|, |f_2|, \ldots$,

$$\int g\,d\mu = \sum_{k=1}^{\infty}\int |f_k|\,d\mu < \infty.$$

For each $t \in (0,\infty)$ we have $g \ge t\chi_{X\setminus E}$ pointwise (the left side is ∞ on $X\setminus E$), so 3.8 with 3.4 gives

$$t\,\mu(X\setminus E) \le \int g\,d\mu < \infty;$$

letting $t \to \infty$ yields $\mu(X\setminus E) = 0$. Finally, for $x \in E$ the series $\sum_k |f_k(x)|$ converges in $\mathbb{R}$, so its terms tend to 0 and $\lim_{k\to\infty} f_k(x) = 0$.

Problem 3A.12. Show that there exists a Borel measurable function $f : \mathbb{R} \to (0, \infty)$ such that $\int \chi_I f \, d\lambda = \infty$ for every nonempty open interval $I \subseteq \mathbb{R}$, where λ denotes Lebesgue measure on $\mathbb{R}$.

Solution. Take an enumeration $r_1, r_2, \ldots$ of $\mathbb{Q}$, put $h_k(x) = 1/|x - r_k|$ with $h_k(r_k) = \infty$, and set

$$g = \sum_{k=1}^{\infty} 8^{-k} h_k, \qquad S = g^{-1}(\{\infty\}), \qquad f = g\chi_{\mathbb{R}\setminus S} + \chi_S.$$

Each h_k is Borel measurable (each $h_k^{-1}((a, \infty])$ is an open interval or $\mathbb{R}$), so g is Borel measurable as the supremum of its partial sums (2.53) and $g > 0$ everywhere; hence $f : \mathbb{R} \to (0, \infty)$ is Borel measurable by 2.52.
$\lambda(S) = 0$: with $B = \bigcap_n \bigcup_{k \geq n} (r_k - 4^{-k}, r_k + 4^{-k})$, monotonicity and countable subadditivity (2.57, 2.58) give $\lambda(B) \leq \sum_{k \geq n} 2 \cdot 4^{-k} = \frac{8}{3} 4^{-n}$ for every n, so $\lambda(B) = 0$; and if $x \notin B \cup \mathbb{Q}$, choosing N with $|x - r_k| \geq 4^{-k}$ for all $k \geq N$ gives $8^{-k} h_k(x) \leq 2^{-k}$ for $k \geq N$ while the finitely many earlier terms are finite, so $g(x) < \infty$. Thus $S \subseteq B \cup \mathbb{Q}$ and $\lambda(S) = 0$.
Fix a nonempty open interval I; by density of $\mathbb{Q}$ choose j with $r_j \in I$, and $\delta > 0$ with $(r_j - \delta, r_j + \delta) \subseteq I$. The disjoint sets $A_m = (r_j + \delta 2^{-m-2}, r_j + \delta 2^{-m-1}]$ lie in I, have $\lambda(A_m) = \delta 2^{-m-2}$, and carry $h_j \geq 2^{m+1}/\delta$, so 3.9 gives

$$\int \chi_I h_j \, d\lambda \ \geq \ \sum_{m=0}^{M} \frac{2^{m+1}}{\delta} \cdot \delta 2^{-m-2} = \frac{M+1}{2}$$

for every M, whence $\int \chi_I h_j \, d\lambda = \infty$. Since $\chi_I g = \sum_k 8^{-k} \chi_I h_k$ pointwise, 3A.10 and 3.20 give $\int \chi_I g \, d\lambda \geq 8^{-j} \int \chi_I h_j \, d\lambda = \infty$. Also $\int \chi_{I \cap S} \, g \, d\lambda = 0$: any $\sum_i c_i \chi_{A_i} \leq \chi_{I \cap S} \, g$ with the A_i disjoint Borel sets and $c_i > 0$ forces $A_i \subseteq S$ and hence $\lambda(A_i) = 0$, so every lower Lebesgue sum vanishes (3.3). Therefore additivity 3.16 gives

$$\infty = \int \chi_I g \, d\lambda = 0 + \int \chi_{I \setminus S} \, g \, d\lambda,$$

and $\chi_I f \geq \chi_{I \setminus S} \, g$ pointwise, so $\int \chi_I f \, d\lambda = \infty$ by 3.8.

Problem 3A.13. Give an example to show that the Monotone Convergence Theorem (3.11) can fail if the hypothesis that $f_1, f_2, \ldots$ are nonnegative functions is dropped.

Solution. Take $f_k = -\chi_{[k,\infty)}$ on $\mathbb{R}$ with Lebesgue measure λ and the Borel σ-algebra.
Each f_k is a Borel measurable simple function, and $f_1 \leq f_2 \leq \cdots$ because $[k+1, \infty) \subseteq [k, \infty)$; only the hypothesis $f_k \geq 0$ fails. Given x, we have $f_k(x) = 0$ for all $k > x$, so $f_k \to 0$ pointwise and $\int f \, d\lambda = 0$. But $f_k^+ = 0$ and $f_k^- = \chi_{[k,\infty)}$, so $\int f_k \, d\lambda$ is defined by 3.18 and 3.4 gives

$$\int f_k \, d\lambda = 0 - \lambda\big([k, \infty)\big) = -\infty$$

for every k. Hence $\lim_{k \to \infty} \int f_k \, d\lambda = -\infty \neq 0 = \int f \, d\lambda$.

Problem 3A.14. Give an example to show that the Monotone Convergence Theorem can fail if the hypothesis of an increasing sequence of functions is replaced by a hypothesis of a decreasing sequence of functions.
[This exercise shows that the Monotone Convergence Theorem should be called the Increasing Convergence Theorem. However, see Exercise 20.]

Solution. Take $f_k = \chi_{[k,\infty)}$ on $\mathbb{R}$ with Lebesgue measure λ and the Borel σ-algebra.
Each f_k is a nonnegative Borel measurable simple function, so every hypothesis of 3.11 holds except that $[k+1,\infty) \subseteq [k,\infty)$ makes the sequence decreasing: $f_1 \geq f_2 \geq \cdots \geq 0$. Given x, we have $f_k(x) = 0$ for all $k > x$, so $f_k \to 0$ pointwise and $\int f\,d\lambda = 0$. But 3.4 gives $\int f_k\,d\lambda = \lambda([k,\infty)) = \infty$ for every k, so

$$\lim_{k\to\infty} \int f_k\,d\lambda = \infty \neq 0 = \int f\,d\lambda.$$

Problem 3A.15. Suppose λ is Lebesgue measure on $\mathbb{R}$ and $f : \mathbb{R} \to [-\infty,\infty]$ is a Borel measurable function such that $\int f\,d\lambda$ is defined.
(a) For $t \in \mathbb{R}$, define $f_t : \mathbb{R} \to [-\infty,\infty]$ by $f_t(x) = f(x-t)$. Prove that $\int f_t\,d\lambda = \int f\,d\lambda$ for all $t \in \mathbb{R}$.
(b) For $t \in \mathbb{R}$, define $f_t : \mathbb{R} \to [-\infty,\infty]$ by $f_t(x) = f(tx)$. Prove that $\int f_t\,d\lambda = \frac{1}{|t|}\int f\,d\lambda$ for all $t \in \mathbb{R}\setminus\{0\}$.

Solution. Both parts follow from the fact that $A \mapsto t+A$ and $A \mapsto \frac{1}{t}A$ permute the Borel partitions of $\mathbb{R}$, multiplying lower Lebesgue sums by 1 and by $\frac{1}{|t|}$ respectively.
Write $\mathcal{B}$ for the Borel σ-algebra. Since $x \mapsto x-t$ and (for $t \neq 0$) $x \mapsto tx$ are continuous, hence Borel measurable (2.41), the maps $A \mapsto t+A$ and $A \mapsto \frac{1}{t}A$ send $\mathcal{B}$ into $\mathcal{B}$; each is a bijection of $\mathcal{B}$ onto $\mathcal{B}$ (inverses $A \mapsto -t+A$ and $A \mapsto tA$) preserving disjointness and unions, so each carries the Borel partitions of $\mathbb{R}$ bijectively onto themselves. Moreover $\lambda(t+A) = \lambda(A)$ by 2.7, and $\lambda(\frac{1}{t}A) = \frac{1}{|t|}\lambda(A)$ by Exercise 2 in Section 2A together with 2.68. Since $f_t^{-1}(B) = t + f^{-1}(B)$ in (a) and $f_t^{-1}(B) = \frac{1}{t}f^{-1}(B)$ in (b), the function f_t is Borel measurable in both cases.
(a) Assume first $f \geq 0$. For a Borel partition $P = A_1,\ldots,A_m$ put $P_t = t+A_1,\ldots,t+A_m$. As $x \in t+A_j$ if and only if $x-t \in A_j$, we get $\inf_{t+A_j} f_t = \inf_{A_j} f$, so

$$L(f_t,P_t) = \sum_{j=1}^{m} \lambda(t+A_j) \inf_{t+A_j} f_t = \sum_{j=1}^{m} \lambda(A_j) \inf_{A_j} f = L(f,P).$$

Because $P \mapsto P_t$ is a bijection of the Borel partitions, the two sets of lower Lebesgue sums coincide, and 3.3 gives $\int f_t\,d\lambda = \int f\,d\lambda$.
For general f note $(f_t)^{\pm} = (f^{\pm})_t$ pointwise, so the nonnegative case gives $\int (f_t)^{\pm} d\lambda = \int f^{\pm} d\lambda$; one of the latter is finite because $\int f\,d\lambda$ is defined, hence $\int f_t\,d\lambda$ is defined and 3.18 gives $\int f_t\,d\lambda = \int f\,d\lambda$.
(b) Identically, with $P^{(t)} = \frac{1}{t}A_1,\ldots,\frac{1}{t}A_m$ and $\inf_{\frac{1}{t}A_j} f_t = \inf_{A_j} f$,

$$L\big(f_t,P^{(t)}\big) = \sum_{j=1}^{m} \frac{1}{|t|}\lambda(A_j)\inf_{A_j} f = \frac{1}{|t|}L(f,P),$$

so the lower Lebesgue sums for f_t are exactly $\frac{1}{|t|}$ times those for f; multiplication by the positive constant $\frac{1}{|t|}$ commutes with suprema in $[0,\infty]$, so 3.3 gives $\int f_t\,d\lambda = \frac{1}{|t|}\int f\,d\lambda$ for $f \geq 0$. The passage to general f via $(f_t)^{\pm} = (f^{\pm})_t$ is as in (a), the final subtraction being legitimate because at most one of the two terms is infinite.

Problem 3A.16. Suppose $\mathcal{S}$ and $\mathcal{T}$ are σ-algebras on a set X and $\mathcal{S} \subseteq \mathcal{T}$. Suppose μ_1 is a measure on $(X,\mathcal{S})$, μ_2 is a measure on $(X,\mathcal{T})$, and $\mu_1(E) = \mu_2(E)$ for all $E \in \mathcal{S}$. Prove that if $f : X \to [0,\infty]$ is $\mathcal{S}$-measurable, then $\int f\,d\mu_1 = \int f\,d\mu_2$.

Solution. Both integrals are the limit of the same sequence of simple-function integrals.
Every $\mathcal{S}$-measurable function is $\mathcal{T}$-measurable, since its preimages lie in $\mathcal{S} \subseteq \mathcal{T}$; in particular $\int f\,d\mu_2$ makes sense. By 2.89 applied to $(X,\mathcal{S})$ there are simple $\mathcal{S}$-measurable $u_1,u_2,\ldots$ with $|u_k| \leq |u_{k+1}| \leq |f|$ pointwise and $u_k \to f$ pointwise; put $f_k = \max\{u_1^+,\ldots,u_k^+\}$. Each f_k takes finitely many values and $f_k^{-1}((a,\infty]) = \bigcup_{j\leq k}(u_j^+)^{-1}((a,\infty]) \in \mathcal{S}$, so f_k is a nonnegative simple $\mathcal{S}$-measurable function (2.52);

moreover $0 \leq f_1 \leq f_2 \leq \cdots \leq f$ because $u_j^+ \leq |u_j| \leq |f| = f$, and $f_k \geq u_k^+ \to f^+ = f$, so $f_k \to f$ pointwise.
If $h = \sum_{k=1}^n c_k \chi_{E_k}$ is the standard representation of a nonnegative simple $\mathcal{S}$-measurable function, then $E_1, \ldots, E_n$ are disjoint sets lying in $\mathcal{S} \subseteq \mathcal{T}$, so 3.7 applies in each of the two measure spaces and gives

$$\int h \, d\mu_1 = \sum_{k=1}^n c_k \mu_1(E_k) = \sum_{k=1}^n c_k \mu_2(E_k) = \int h \, d\mu_2,$$

the middle equality because each $E_k \in \mathcal{S}$. Applying this to each f_k and the Monotone Convergence Theorem 3.11 in $(X, \mathcal{S}, \mu_1)$ and in $(X, \mathcal{T}, \mu_2)$,

$$\int f \, d\mu_1 = \lim_{k\to\infty} \int f_k \, d\mu_1 = \lim_{k\to\infty} \int f_k \, d\mu_2 = \int f \, d\mu_2.$$

Problem 3A.17. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \ldots$ is a sequence of nonnegative $\mathcal{S}$-measurable functions on X. Define a function $f : X \to [0, \infty]$ by

$$f(x) = \liminf_{k\to\infty} f_k(x).$$

(a) Show that f is an $\mathcal{S}$-measurable function.
(b) Prove that

$$\int f \, d\mu \leq \liminf_{k\to\infty} \int f_k \, d\mu.$$

(c) Give an example showing that the inequality in (b) can be a strict inequality even when $\mu(X) < \infty$ and the family of functions $\{f_k\}_{k\in\mathbb{Z}^+}$ is uniformly bounded.
[The result in (b) is called Fatou's Lemma. Some textbooks prove Fatou's Lemma and then use it to prove the Monotone Convergence Theorem. Here we are taking the reverse approach—you should be able to use the Monotone Convergence Theorem to give a clean proof of Fatou's Lemma.]

Solution. Put $g_k = \inf\{f_j : j \geq k\}$, so that $f = \lim_{k\to\infty} g_k$.
(a) Each g_k is $\mathcal{S}$-measurable by 2.53, and $g_1 \leq g_2 \leq \cdots$ because the infimum is taken over a smaller set as k grows; hence $f = \liminf_k f_k = \sup_k g_k$ is $\mathcal{S}$-measurable, again by 2.53.
(b) By (a), $0 \leq g_1 \leq g_2 \leq \cdots$ are $\mathcal{S}$-measurable with $g_k \to f$ pointwise, so 3.11 gives $\int f \, d\mu = \lim_k \int g_k \, d\mu$. For every $j \geq k$ we have $g_k \leq f_j$, so 3.8 gives $\int g_k \, d\mu \leq \inf_{j\geq k} \int f_j \, d\mu$; letting $k \to \infty$,

$$\int f \, d\mu = \lim_{k\to\infty} \int g_k \, d\mu \leq \liminf_{k\to\infty} \int f_k \, d\mu.$$

(c) On $X = [0, 1]$ with Borel sets and Lebesgue measure (so $\mu(X) = 1$), take $f_k = \chi_{[0,1/2]}$ for k odd and $f_k = \chi_{(1/2,1]}$ for k even; then $0 \leq f_k \leq 1$, so the family is uniformly bounded. Each $x \in [0, 1]$ has $f_k(x) = 0$ for infinitely many k, so $g_k = 0$ for every k and $f = 0$, giving $\int f \, d\lambda = 0$; while 3.4 gives $\int f_k \, d\lambda = \frac{1}{2}$ for every k. Thus

$$0 = \int f \, d\lambda < \liminf_{k\to\infty} \int f_k \, d\lambda = \tfrac{1}{2}.$$

Problem 3A.18. Give an example of a sequence $x_1, x_2, \ldots$ of real numbers such that

$$\lim_{n\to\infty} \sum_{k=1}^n x_k \text{ exists in } \mathbb{R},$$

but $\int x \, d\mu$ is not defined, where μ is counting measure on $\mathbb{Z}^+$ and x is the function from $\mathbb{Z}^+$ to $\mathbb{R}$ defined by $x(k) = x_k$.

Solution. Take $x_k = \dfrac{(-1)^k}{k}$.
The sequence $1, \frac{1}{2}, \frac{1}{3}, \ldots$ decreases to 0, so by the alternating series test $\lim_{n\to\infty} \sum_{k=1}^{n} x_k$ exists in $\mathbb{R}$ (its value is $-\ln 2$). But integration with respect to counting measure on $\mathbb{Z}^+$ is summation (Example 3.6), and every subset of $\mathbb{Z}^+$ is measurable, so

$$\int x^+ \, d\mu = \sum_{m=1}^{\infty} \frac{1}{2m} = \infty, \qquad \int x^- \, d\mu = \sum_{m=1}^{\infty} \frac{1}{2m-1} = \infty,$$

both by divergence of the harmonic series. Since neither is finite, 3.18 assigns no value to $\int x \, d\mu$.

Problem 3A.19. Show that if $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to [0, \infty)$ is $\mathcal{S}$-measurable, then

$$\mu(X) \inf_X f \le \int f \, d\mu \le \mu(X) \sup_X f.$$

Solution. The one-set partition gives the left inequality, and $f \le (\sup_X f)\chi_X$ gives the right.
Assume $X \neq \varnothing$ (otherwise $\mu(X) = 0 = \int f \, d\mu$ and both inequalities are trivial), and recall the convention $0 \cdot \infty = \infty \cdot 0 = 0$ stated with 3.2. The single set $A_1 = X$ is an $\mathcal{S}$-partition P of X, so by 3.2 and the definition 3.3 of the integral as a supremum of lower Lebesgue sums,

$$\mu(X) \inf_X f = L(f, P) \le \int f \, d\mu.$$

For the other inequality put $c = \sup_X f \in [0, \infty]$, so $f \le c\chi_X$ pointwise; then 3.8 and 3.15 (with $n = 1$, $E_1 = X$, $c_1 = c$) give

$$\int f \, d\mu \le \int c\,\chi_X \, d\mu = c\,\mu(X) = \mu(X) \sup_X f.$$

Problem 3A.20. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \ldots$ is a monotone (meaning either increasing or decreasing) sequence of $\mathcal{S}$-measurable functions. Define $f : X \to [-\infty, \infty]$ by

$$f(x) = \lim_{k\to\infty} f_k(x).$$

Prove that if $\int |f_1| \, d\mu < \infty$, then

$$\lim_{k\to\infty} \int f_k \, d\mu = \int f \, d\mu.$$

Solution. The increasing case is 3.11 applied to $f_k - f_1$; the decreasing case follows by passing to $-f_k$.
Monotonicity makes $f(x) = \lim_k f_k(x)$ exist in $[-\infty, \infty]$ for each x, and f is $\mathcal{S}$-measurable by 2.53, being the pointwise supremum of the f_k in the increasing case and their pointwise infimum in the decreasing case.
Increasing case: assume $f_1 \le f_2 \le \cdots$. Then $f_k^- \le f_1^- \le |f_1|$ and likewise $f^- \le f_1^-$, so 3.8 gives $\int f_k^- \, d\mu \le \int |f_1| \, d\mu < \infty$ and $\int f^- \, d\mu < \infty$; hence every $\int f_k \, d\mu$ and $\int f \, d\mu$ is defined, with values in $(-\infty, \infty]$.
Reduction to f_1 real valued. Let $N = f_1^{-1}(\{-\infty, \infty\}) \in \mathcal{S}$ and $M = X \setminus N$. Since $|f_1| \ge n\chi_N$ pointwise for every n, 3.8 and 3.15 give $n\mu(N) \le \int |f_1| \, d\mu < \infty$, so $\mu(N) = 0$. For nonnegative $\mathcal{S}$-measurable h we then have $\int h\chi_N \, d\mu = 0$, because any $\mathcal{S}$-partition piece on which $h\chi_N$ has positive infimum lies in N and so has measure 0, making every lower Lebesgue sum vanish (3.3). Hence for $\mathcal{S}$-measurable g with $\int g \, d\mu$ defined, $g\chi_M$ is $\mathcal{S}$-measurable (its set $(g\chi_M)^{-1}((a, \infty])$ is $g^{-1}((a, \infty]) \cap M$ for $a \ge 0$ and that set union N for $a < 0$; apply 2.52), and $(g\chi_M)^{\pm} = g^{\pm}\chi_M$ with $g^{\pm} = g^{\pm}\chi_M + g^{\pm}\chi_N$, so additivity 3.16 gives $\int (g\chi_M)^{\pm} \, d\mu = \int g^{\pm} \, d\mu$ and therefore $\int g\chi_M \, d\mu = \int g \, d\mu$. Replacing each f_k by $f_k\chi_M$ and f by $f\chi_M$ thus changes no integral while preserving monotonicity, the pointwise limit, and $\int |f_1| \, d\mu < \infty$; so assume f_1 is real valued.

Now $g_k := f_k - f_1$ and $g := f - f_1$ are well defined with values in $[0, \infty]$, and they are $\mathcal{S}$-measurable by 2.52, since for $a \in \mathbb{R}$

$$g_k^{-1}\big((a, \infty]\big) = \bigcup_{r \in \mathbb{Q}} \Big(f_k^{-1}\big((r, \infty]\big) \cap f_1^{-1}\big([-\infty, r - a)\big)\Big)$$

(a rational r separates $a + f_1(x) < r < f_k(x)$ exactly when $g_k(x) > a$). As $0 \le g_1 \le g_2 \le \cdots$ with $g_k \to g$ pointwise, 3.11 gives $\int g_k \, d\mu \to \int g \, d\mu$.
Next, if $h : X \to (-\infty, \infty]$ is $\mathcal{S}$-measurable with $h \ge f_1$, then

$$\int h \, d\mu = \int (h - f_1) \, d\mu + \int f_1 \, d\mu.$$

Indeed $h - f_1$ is $\mathcal{S}$-measurable by the display above with h in place of f_k, and substituting $h - f_1 = h^+ - h^- - f_1^+ + f_1^-$ (legitimate pointwise, since f_1 is real valued and at most one of $h^\pm(x)$ is nonzero; both sides are ∞ where $h = \infty$) gives the identity between $[0, \infty]$-valued functions $h^+ + f_1^- = (h - f_1) + f_1^+ + h^-$, which integrates by 3.16 to

$$\int h^+ d\mu + \int f_1^- d\mu = \int (h - f_1) \, d\mu + \int f_1^+ d\mu + \int h^- d\mu.$$

Here $\int f_1^\pm d\mu \le \int |f_1| \, d\mu < \infty$, and $h \ge f_1$ forces $h^- \le f_1^-$ so $\int h^- d\mu < \infty$; subtracting these two finite numbers and applying 3.18 yields the claim.
Taking $h = f_k$ and $h = f$ there, and using that adding the real number $\int f_1 \, d\mu$ commutes with limits in $(-\infty, \infty]$,

$$\lim_{k \to \infty} \int f_k \, d\mu = \int g \, d\mu + \int f_1 \, d\mu = \int f \, d\mu.$$

Decreasing case: assume $f_1 \ge f_2 \ge \cdots$. Each $-f_k$ is $\mathcal{S}$-measurable (2.52), $-f_1 \le -f_2 \le \cdots$ has pointwise limit $-f$, and $\int | - f_1| \, d\mu < \infty$; also $f_k^+ \le f_1^+ \le |f_1|$ and $f^+ \le f_1^+$, so $\int f_k \, d\mu$ and $\int f \, d\mu$ are defined and 3.20 with $c = -1$ applies. The increasing case gives $\int (-f_k) \, d\mu \to \int (-f) \, d\mu$, that is $-\int f_k \, d\mu \to -\int f \, d\mu$, which is the assertion.

Problem 3A.21. Henri Lebesgue wrote the following about his method of integration:
I have to pay a certain sum, which I have collected in my pocket. I take the bills and coins out of my pocket and give them to the creditor in the order I find them until I have reached the total sum. This is the Riemann integral. But I can proceed differently. After I have taken all the money out of my pocket I order the bills and coins according to identical values and then I pay the several heaps one after the other to the creditor. This is my integral.
Use 3.15 to explain what Lebesgue meant and to explain why integration of a function with respect to a measure can be thought of as partitioning the range of the function, in contrast to Riemann integration, which depends on partitioning the domain of the function.
[The quote above is taken from page 796 of The Princeton Companion to Mathematics, edited by Timothy Gowers.]

Solution. Lebesgue's heaps are the sets $E_k = v^{-1}(\{c_k\})$ of 3.15, and his total is $\sum_k c_k \mu(E_k)$: a sum over the finitely many values in the range, not over the many items in the domain.
Model the pocket by the finite set X of bills and coins, $\mathcal{S}$ all subsets of X, μ counting measure, and $v : X \to [0, \infty)$ the face value. The sum owed is $\int v \, d\mu = \sum_{x \in X} v(x)$ (Example 3.6). Riemann's method is "in the order I find them": fix an enumeration of the domain and accumulate one item at a time. Lebesgue's method is 3.15 read from left to right: let $c_1, \ldots, c_n$ be the distinct denominations present and $E_k = v^{-1}(\{c_k\})$ the heap of items of value c_k, so that $v = \sum_{k=1}^n c_k \chi_{E_k}$ and

$$\int v \, d\mu = \sum_{k=1}^{n} c_k \, \mu(E_k) = \sum_{k=1}^{n} (\text{denomination}) \times (\text{size of heap}).$$

That this is a statement about the range is visible in what 3.15 asks of $E_1, \ldots, E_n$: only that they lie in $\mathcal{S}$, so that $\mu(E_k)$ makes sense. They need not be intervals, nor even disjoint. The c_k are values of the function; the E_k are whatever sets those values cut out. The same reading computes any $\mathcal{S}$-measurable $f : X \to [0, \infty)$: chopping the range gives $E_{j,k} = f^{-1}([k2^{-j}, (k+1)2^{-j}))$ for $0 \le k < j2^j$ and $E_{j,\infty} = f^{-1}([j, \infty))$, all in $\mathcal{S}$ precisely because f is $\mathcal{S}$-measurable, and the simple functions

$$s_j = \sum_{k=0}^{j2^j-1} k2^{-j}\chi_{E_{j,k}} + j\chi_{E_{j,\infty}}$$

satisfy $s_j = \min\{2^{-j}\lfloor 2^j f\rfloor, j\}$, hence increase to f pointwise; 3.15 evaluates each $\int s_j \, d\mu$ as a sum over the range, and 3.11 delivers $\int f \, d\mu$.
A Riemann sum, by contrast, starts from a partition $a = t_0 < \cdots < t_n = b$ of the domain and forms $\sum_j (\inf_{[t_{j-1}, t_j]} f)(t_j - t_{j-1})$ (1.3); its accuracy is governed by the oscillation of f on each subinterval, and nothing about the range enters the choice of partition. So $f = \chi_{\mathbb{Q}\cap[0,1]}$ has $L(f, [0,1]) = 0 \ne 1 = U(f, [0, 1])$, while sorting by value produces just two heaps and 3.15 gives at once

$$\int \chi_{\mathbb{Q}\cap[0,1]} \, d\lambda = 1 \cdot 0 + 0 \cdot 1 = 0.$$

The price is that the heaps must be measurable, which is why a theory of measure precedes the integral, and why the partitions in 3.3 range over arbitrary sets of $\mathcal{S}$ rather than intervals: once they do, a domain partition can follow the level sets of f, and 3.3 becomes the range-partition picture above.

3.2 Exercises 3B

Problem 3B.1. Give an example of a sequence $f_1, f_2, \ldots$ of functions from $\mathbb{Z}^+$ to $[0, \infty)$ such that

$$\lim_{k\to\infty} f_k(m) = 0$$

for every $m \in \mathbb{Z}^+$ but $\lim_{k\to\infty} \int f_k \, d\mu = 1$, where μ is counting measure on $\mathbb{Z}^+$.

Solution. Take $f_k = \chi_{\{k\}}$.
Every function on $\mathbb{Z}^+$ is measurable here. For fixed m we have $f_k(m) = 0$ whenever $k > m$, so $\lim_{k\to\infty} f_k(m) = 0$; and 3.4 gives $\int f_k \, d\mu = \mu(\{k\}) = 1$ for every k, so $\lim_{k\to\infty} \int f_k \, d\mu = 1$.

Problem 3B.2. Give an example of a sequence $f_1, f_2, \ldots$ of continuous functions from $\mathbb{R}$ to $[0, 1]$ such that

$$\lim_{k\to\infty} f_k(x) = 0$$

for every $x \in \mathbb{R}$ but $\lim_{k\to\infty} \int f_k \, d\lambda = \infty$, where λ is Lebesgue measure on $\mathbb{R}$.

Solution. Take the tent functions $f_k(x) = \max\bigl\{0,\, 1 - \frac{|x-k^2|}{k}\bigr\}$.
Each f_k is continuous with values in $[0, 1]$ (a maximum of two continuous functions), hence Borel measurable by 2.41, and it vanishes outside $[k^2 - k,\, k^2 + k]$. Given x, we have $k^2 - k > |x|$ for all large k, so $f_k(x) = 0$ eventually and $f_k \to 0$ pointwise. Being continuous, f_k is Riemann integrable on $[k^2 - k, k^2 + k]$, so 3.34 identifies its integral with the area of a triangle of base $2k$ and height 1:

$$\int f_k \, d\lambda = \int_{[k^2-k,\, k^2+k]} f_k \, d\lambda = \tfrac{1}{2} \cdot 2k \cdot 1 = k,$$

whence $\lim_{k\to\infty} \int f_k \, d\lambda = \infty$.

Problem 3B.3. Suppose λ is Lebesgue measure on $\mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function such that $\int |f|\, d\lambda < \infty$. Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \int_{(-\infty,\, x)} f\, d\lambda.$$

Prove that g is uniformly continuous on $\mathbb{R}$.

Solution. Uniform continuity comes straight from 3.28: integrals of $|f|$ over sets of small measure are small.
Each $g(x)$ is a real number: by 3.24 it is $\int \chi_{(-\infty,x)} f\, d\lambda$, and $|\chi_{(-\infty,x)} f| \le |f|$ with 3.8 makes both $(\chi_{(-\infty,x)} f)^{\pm}$ have finite integral, so 3.18 applies.
Let $\varepsilon > 0$. Since $|f|$ is nonnegative Borel measurable with $\int |f|\, d\lambda < \infty$, 3.28 supplies $\delta > 0$ such that $\int_B |f|\, d\lambda < \varepsilon$ for every Borel set B with $\lambda(B) < \delta$. Suppose $x \le y$ with $y - x < \delta$. Then $\chi_{(-\infty,y)} f = \chi_{(-\infty,x)} f + \chi_{[x,y)} f$, both summands having integrable absolute value, so 3.21 and then 3.23 and 3.8 give

$$|g(y) - g(x)| = \Big| \int_{[x,\, y)} f\, d\lambda \Big| \le \int_{[x,\, y)} |f|\, d\lambda < \varepsilon,$$

since $\lambda([x, y)) = y - x < \delta$. As δ depends only on ε, g is uniformly continuous on $\mathbb{R}$.

Problem 3B.4. (a) Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) < \infty$. Suppose that $f : X \to [0, \infty)$ is a bounded $\mathcal{S}$-measurable function. Prove that

$$\int f\, d\mu = \inf\Big\{ \sum_{j=1}^{m} \mu(A_j) \sup_{A_j} f \ : \ A_1, \dots, A_m \text{ is an } \mathcal{S}\text{-partition of } X \Big\}.$$

(b) Show that the conclusion of (a) can fail if the hypothesis that f is bounded is replaced by the hypothesis that $\int f\, d\mu < \infty$.
(c) Show that the conclusion of (a) can fail if the condition that $\mu(X) < \infty$ is deleted.
[Part (a) of this exercise shows that if we had defined an upper Lebesgue sum, then it could be used to define $\int f\, d\mu$ when f is bounded and $\mu(X) < \infty$. However, parts (b) and (c) show that the hypotheses that f is bounded and that $\mu(X) < \infty$ are needed if defining the integral via the equation above. The definition of the integral via the lower Lebesgue sum does not require these hypotheses, showing the advantage of using the lower Lebesgue sum.]

Solution. Write $U(f, P) = \sum_j \mu(A_j) \sup_{A_j} f$ for an $\mathcal{S}$-partition $P = A_1, \dots, A_m$, and $U(f) = \inf_P U(f, P)$; part (a) asserts $\int f\, d\mu = U(f)$.
(a) Assume $X \ne \varnothing$ (otherwise both sides are 0), and discard the empty sets from every partition: their terms carry the factor $\mu(\varnothing) = 0$, hence vanish by the convention $0 \cdot \infty = \infty \cdot 0 = 0$ stated before 3.2. All infima and suprema below are then real numbers, because f is bounded; fix $c \in [0, \infty)$ with $0 \le f \le c$, so that every $L(f, P)$ and $U(f, P)$ lies in $[0, c\mu(X)] \subseteq \mathbb{R}$.
Refinement. Given $\mathcal{S}$-partitions $P = A_1, \dots, A_m$ and $P' = B_1, \dots, B_n$, let Q consist of the nonempty sets $A_i \cap B_j$, again an $\mathcal{S}$-partition of X. Finite additivity of μ gives $\mu(A_i) = \sum_j \mu(A_i \cap B_j)$, and $\inf_{A_i} f \le \inf_{A_i \cap B_j} f$, so

$$\mu(A_i) \inf_{A_i} f \le \sum_{j=1}^{n} \mu(A_i \cap B_j) \inf_{A_i \cap B_j} f;$$

summing over i gives $L(f, P) \le L(f, Q)$, and the same computation with B_j in place of A_i and suprema in place of infima gives $U(f, Q) \le U(f, P')$. Since $L(f, Q) \le U(f, Q)$,

$$L(f, P) \le L(f, Q) \le U(f, Q) \le U(f, P'),$$

so taking the supremum over P (3.3) and then the infimum over P' yields $\int f\, d\mu \le U(f)$.

Conversely let $\varepsilon > 0$; we may assume $c > 0$, since $c = 0$ makes both quantities 0. Choose n with $\frac{c}{n}\mu(X) < \varepsilon$ (possible as $\mu(X) < \infty$) and let P be the nonempty sets among

$$A_j = f^{-1}\Big(\Big[\tfrac{(j-1)c}{n}, \tfrac{jc}{n}\Big)\Big) \ (j < n), \qquad A_n = f^{-1}\Big(\Big[\tfrac{(n-1)c}{n}, c\Big]\Big),$$

which lie in $\mathcal{S}$, are disjoint, and cover X because $f(X) \subseteq [0, c]$. On each of them f varies by at most c/n, so

$$U(f, P) - L(f, P) \leq \frac{c}{n}\sum_j \mu(A_j) = \frac{c}{n}\mu(X) < \varepsilon$$

(the subtraction legitimate as all terms are finite), whence $U(f) \leq U(f, P) < L(f, P) + \varepsilon \leq \int f\, d\mu + \varepsilon$ by 3.3. Let $\varepsilon \to 0$.

(b) Take $X = \mathbb{Z}^+$, $\mathcal{S}$ all subsets, $\mu(A) = \sum_{n\in A} 2^{-n}$ (countably additive, since nonnegative series may be summed in any order; $\mu(X) = 1 < \infty$), and $f(n) = n$, which is measurable and unbounded. Applying 3.11 to $f_N = \sum_{n\leq N} n\chi_{\{n\}} \uparrow f$ and 3.7,

$$\int f\, d\mu = \lim_{N\to\infty}\sum_{n=1}^{N}\frac{n}{2^n} = 2 < \infty.$$

But any $\mathcal{S}$-partition of the infinite set X has an infinite member A_k, for which $\sup_{A_k} f = \infty$ and $\mu(A_k) \geq 2^{-n_0} > 0$ for any $n_0 \in A_k$; so $U(f, P) = \infty$ for every P and $U(f) = \infty \neq 2$.

(c) Same X and $\mathcal{S}$, now with μ counting measure (so $\mu(X) = \infty$) and $f(n) = 2^{-n}$, which is measurable and bounded. As in (b), 3.11 and 3.7 give

$$\int f\, d\mu = \sum_{n=1}^{\infty} 2^{-n} = 1.$$

Again every $\mathcal{S}$-partition has an infinite member A_k, now with $\mu(A_k) = \infty$ and $\sup_{A_k} f \geq f(n_0) > 0$ for any $n_0 \in A_k$, so $U(f) = \infty \neq 1$.

Problem 3B.5. Let λ denote Lebesgue measure on $\mathbb{R}$. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function such that $\int |f|\, d\lambda < \infty$. Prove that

$$\lim_{k\to\infty}\int_{[-k,\, k]} f\, d\lambda = \int f\, d\lambda.$$

Solution. Apply the Dominated Convergence Theorem 3.31 to $f_k = \chi_{[-k,\, k]} f$, with dominating function $|f|$.

Each f_k is Borel measurable (2.46), $|f_k| \leq |f|$ pointwise with $\int |f|\, d\lambda < \infty$ by hypothesis, and for each x we have $f_k(x) = f(x)$ once $k \geq |x|$, so $f_k \to f$ pointwise. Hence 3.31 and the definition 3.24 of integration over a subset give

$$\lim_{k\to\infty}\int_{[-k,\, k]} f\, d\lambda = \lim_{k\to\infty}\int f_k\, d\lambda = \int f\, d\lambda.$$

Problem 3B.6. Let λ denote Lebesgue measure on $\mathbb{R}$. Give an example of a continuous function $f : [0, \infty) \to \mathbb{R}$ such that $\lim_{t\to\infty}\int_{[0,\, t]} f\, d\lambda$ exists (in $\mathbb{R}$) but $\int_{[0,\, \infty)} f\, d\lambda$ is not defined.

Solution. Take $f = \sum_{n=1}^{\infty} g_n$ with $g_n(x) = \frac{(-1)^{n+1}}{n}\varphi(x - n)$, where $\varphi(s) = \max\{0, 1 - |2s - 1|\}$ is the tent function supported on $(0, 1)$ with $\int_0^1 \varphi = \frac{1}{2}$.

The supports $(n, n + 1)$ are disjoint, so at most one term is nonzero at each x and f is real valued; moreover $f = \sum_{n=1}^{N} g_n$ on $[0, N + 1)$, a finite sum of continuous functions, and these sets cover $[0, \infty)$, so f is continuous, hence Borel measurable (2.41). Since g_n is continuous and vanishes off $[n, n+1]$, 3.34 gives

$$\int g_n \, d\lambda = \frac{(-1)^{n+1}}{n} \int_0^1 \varphi = \frac{(-1)^{n+1}}{2n}, \qquad \int |g_n| \, d\lambda = \frac{1}{2n}.$$

The limit exists. On $[0, N]$ we have $f = \sum_{n=1}^{N-1} g_n$, so 3.21 (legitimate since each $\int |g_n| \, d\lambda < \infty$) gives $\int_{[0,N]} f \, d\lambda = \sum_{n=1}^{N-1} \frac{(-1)^{n+1}}{2n}$, which converges to $\frac{\ln 2}{2}$ by the alternating series test. For real $t \geq 1$ put $N = \lfloor t \rfloor$; then $(N, t] \subseteq (N, N+1]$, where $|f| \leq \frac{1}{N}$, so 3.21 and 3.25 give

$$\Big| \int_{[0,\,t]} f \, d\lambda - \int_{[0,\,N]} f \, d\lambda \Big| \leq \lambda\big((N, t]\big) \sup_{(N,\,t]} |f| \leq \frac{1}{N},$$

and $N \to \infty$ as $t \to \infty$, so $\lim_{t\to\infty} \int_{[0,t]} f \, d\lambda = \frac{\ln 2}{2} \in \mathbb{R}$.
The integral is not defined. The g_n have disjoint supports, with $g_n \geq 0$ for n odd and $g_n \leq 0$ for n even, so $f^+ = \sum_{n \text{ odd}} g_n$ and $f^- = -\sum_{n \text{ even}} g_n$; hence 3.8 and 3.16 give, for every $K \in \mathbb{Z}^+$,

$$\int f^+ \, d\lambda \geq \sum_{k=1}^{K} \frac{1}{2(2k-1)}, \qquad \int f^- \, d\lambda \geq \sum_{k=1}^{K} \frac{1}{4k}.$$

Both sums diverge as $K \to \infty$ (comparison with the harmonic series), so $\int f^+ \, d\lambda = \int f^- \, d\lambda = \infty$ and 3.18 assigns no value to $\int_{[0,\infty)} f \, d\lambda$.

Problem 3B.7. Let λ denote Lebesgue measure on $\mathbb{R}$. Give an example of a continuous function $f : (0, 1) \to \mathbb{R}$ such that $\lim_{n\to\infty} \int_{(\frac{1}{n},\,1)} f \, d\lambda$ exists (in $\mathbb{R}$) but $\int_{(0,\,1)} f \, d\lambda$ is not defined.

Solution. Take $f = \sum_{n=1}^{\infty} h_n$ with $h_n(x) = (-1)^{n+1}\, 2(n+1)\, \varphi\big(n(n+1)x - n\big)$, where $\varphi(s) = \max\{0, 1 - |2s-1|\}$ is the tent function supported on $(0, 1)$.
Since $\varphi(n(n+1)x - n) \neq 0$ exactly when $\frac{1}{n+1} < x < \frac{1}{n}$, the function h_n is continuous on $(0, 1)$ and vanishes off $I_n = (\frac{1}{n+1}, \frac{1}{n})$, and the graph of $|h_n|$ over I_n is a triangle of base $\lambda(I_n) = \frac{1}{n(n+1)}$ and height $2(n+1)$; as h_n is continuous on $\overline{I_n}$, 3.34 gives

$$\int h_n \, d\lambda = (-1)^{n+1} \cdot \tfrac{1}{2} \cdot \tfrac{1}{n(n+1)} \cdot 2(n+1) = \frac{(-1)^{n+1}}{n}, \qquad \int |h_n| \, d\lambda = \frac{1}{n}.$$

The I_n are disjoint, so at most one term is nonzero at each x and f is real valued; moreover $f = \sum_{n=1}^{N} h_n$ on $(\frac{1}{N+1}, 1)$, a finite sum of continuous functions, and these sets cover $(0, 1)$, so f is continuous, hence Borel measurable (2.41).
The limit exists. For $n \geq 2$ we have $f = \sum_{k=1}^{n-1} h_k$ on $(\frac{1}{n}, 1)$, and $I_k \subseteq (\frac{1}{n}, 1)$ for $k \leq n-1$, so $\int_{(\frac{1}{n},1)} h_k \, d\lambda = \int h_k \, d\lambda$; since each $\int |h_k| \, d\lambda < \infty$, additivity 3.21 gives

$$\int_{(\frac{1}{n},\,1)} f \, d\lambda = \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k} \longrightarrow \ln 2 \in \mathbb{R}$$

by the alternating series test.
The integral is not defined. The h_k have disjoint supports, with $h_k \geq 0$ for k odd and $h_k \leq 0$ for k even, so $f^+ = \sum_{k \text{ odd}} h_k$ and $f^- = -\sum_{k \text{ even}} h_k$; hence 3.8 and 3.16 give, for every $K \in \mathbb{Z}^+$,

$$\int_{(0,\,1)} f^+ \, d\lambda \geq \sum_{k=1}^{K} \frac{1}{2k-1}, \qquad \int_{(0,\,1)} f^- \, d\lambda \geq \sum_{k=1}^{K} \frac{1}{2k}.$$

Both sums diverge as $K \to \infty$, so neither integral is finite and 3.18 assigns no value to $\int_{(0,1)} f \, d\lambda$.

Problem 3B.8. Verify the assertion in 3.38.

Solution. Both inclusions follow from the single formula $g_n(x) = \inf_{J_n(x)} f$, $h_n(x) = \sup_{J_n(x)} f$, where $J_n(x)$ is the union of the one or two intervals of P_n that contain x.
In the proof of 3.34, P_n cuts $[a, b]$ into 2^n closed intervals of length $\delta_n = (b-a)2^{-n}$; g_n and h_n are the corresponding lower and upper step functions, modified at the finitely many shared endpoints to take the infimum, respectively supremum, of f over the union of the two intervals meeting there; and $f_L = \lim_n g_n$, $f_U = \lim_n h_n$, these limits existing since f is bounded and $g_1 \le g_2 \le \cdots$, $h_1 \ge h_2 \ge \cdots$. With that modification the displayed formula holds at every x, and $J_n(x)$ is a closed interval containing x of length δ_n or $2\delta_n$. Hence $g_n(x) \le f(x) \le h_n(x)$, so $f_L \le f \le f_U$, and monotonicity of the two sequences gives

$$0 \le f_U(x) - f_L(x) \le h_n(x) - g_n(x) = \sup_{J_n(x)} f - \inf_{J_n(x)} f$$

for every $n \in \mathbb{Z}^+$.
(i) Continuity at x gives $f_L(x) = f_U(x)$. Let $\varepsilon > 0$; choose $\eta > 0$ with $|f(y) - f(x)| < \varepsilon$ for all $y \in [a, b]$ with $|y - x| < \eta$, then n with $2\delta_n < \eta$. Every $y \in J_n(x)$ satisfies $|y - x| < \eta$, so the display gives $f_U(x) - f_L(x) \le 2\varepsilon$; let $\varepsilon \to 0$.
(ii) $f_L(x) = f_U(x)$ gives continuity at x. Then $f_L(x) = f(x) = f_U(x)$, so for $\varepsilon > 0$ there is n with $h_n(x) - g_n(x) < \varepsilon$. Write $J_n(x) = [c, d]$. If $c > a$ then c is a partition point of P_n interior to $[a, b]$, hence lies in two of its intervals, so $x = c$ would put the interval left of c inside $J_n(x)$, contradicting $\min J_n(x) = c$; thus $c > a$ forces $x > c$, and symmetrically $d < b$ forces $x < d$. So x is interior to $[c, d]$ relative to $[a, b]$, and there is $\eta > 0$ with $\{y \in [a, b] : |y - x| < \eta\} \subseteq J_n(x)$. For such y both $f(y)$ and $f(x)$ lie in $[g_n(x), h_n(x)]$, so $|f(y) - f(x)| \le h_n(x) - g_n(x) < \varepsilon$.

Problem 3B.9. Verify the assertion in Example 3.41.

Solution. Both directions run through the identity $E_k = f^{-1}(\{a_k\})$ and 3.7.
Put $f = a_1\chi_{E_1} + \cdots + a_n\chi_{E_n}$. Disjointness makes $f = a_k$ on E_k and $f = 0$ off $E_1 \cup \cdots \cup E_n$; since the a_k are distinct and nonzero, this says exactly that

$$E_k = f^{-1}(\{a_k\}) \quad (1 \le k \le n), \qquad |f| = |a_1|\chi_{E_1} + \cdots + |a_n|\chi_{E_n}.$$

(Note the E_k are not assumed measurable; that is part of what is proved.)
If every $E_k \in \mathcal{S}$ with $\mu(E_k) < \infty$, then f is $\mathcal{S}$-measurable by 2.46, and 3.7 applied to $|f|$, whose sets E_k are disjoint and in $\mathcal{S}$, gives

$$\|f\|_1 = \int |f|\, d\mu = |a_1|\mu(E_1) + \cdots + |a_n|\mu(E_n) < \infty,$$

so $f \in L^1(\mu)$ with the asserted norm.
Conversely if $f \in L^1(\mu)$, then by 3.40 f is $\mathcal{S}$-measurable with $\|f\|_1 < \infty$, so $E_k = f^{-1}(\{a_k\}) \in \mathcal{S}$ by 2.35, the singleton $\{a_k\}$ being Borel. The same application of 3.7 gives the displayed formula for $\|f\|_1$; every term on the left lies in $[0, \infty]$ and their sum is finite, so $|a_k|\mu(E_k) < \infty$, and $a_k \ne 0$ forces $\mu(E_k) < \infty$.

Problem 3B.10. (a) Suppose $(X, \mathcal{S}, \mu)$ is a measure space such that $\mu(X) < \infty$. Suppose p, r are positive numbers with $p < r$. Prove that if $f : X \to [0, \infty)$ is an $\mathcal{S}$-measurable function such that $\int f^r\, d\mu < \infty$, then $\int f^p\, d\mu < \infty$.
(b) Give an example to show that the result in (a) can be false without the hypothesis that $\mu(X) < \infty$.

Solution. (a) The pointwise bound $f^p \le 1 + f^r$ does it; (b) take $f(x) = 1/x$ on $(1, \infty)$.
(a) Both f^p and f^r are $\mathcal{S}$-measurable, being compositions of f with the continuous map $s \mapsto s^t$ on $[0, \infty)$ (2.41 and 2.44). If $f(x) \le 1$ then $f(x)^p \le 1$; if $f(x) > 1$ then $f(x)^{r-p} \ge 1$, so $f(x)^p \le f(x)^r$.

Either way $f^p \le 1 + f^r$ pointwise, so 3.8, additivity 3.16, and 3.4 applied to χ_X give

$$\int f^p \, d\mu \le \int (1 + f^r) \, d\mu = \mu(X) + \int f^r \, d\mu < \infty.$$

(b) On $X = (1, \infty)$ with the Borel sets and Lebesgue measure, so $\mu(X) = \infty$, take $p = 1$, $r = 2$, and $f(x) = 1/x$, which is continuous hence Borel measurable. The functions $f^2 \chi_{(1,k)}$ and $f \chi_{(1,k)}$ increase pointwise to f^2 and f, so 3.11 together with 3.34 gives

$$\int f^2 \, d\mu = \lim_{k \to \infty} \left(1 - \frac{1}{k}\right) = 1 < \infty, \qquad \int f \, d\mu = \lim_{k \to \infty} \ln k = \infty.$$

Problem 3B.11. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \in L^1(\mu)$. Prove that the set $\{x \in X : f(x) \neq 0\}$ is the countable union of sets with finite μ-measure.

Solution. Take $A_k = \{x \in X : |f(x)| > \frac{1}{k}\}$; then $\{x : f(x) \neq 0\} = \bigcup_{k=1}^{\infty} A_k$, with every $\mu(A_k) < \infty$. Each $A_k = |f|^{-1}\big((\frac{1}{k}, \infty)\big) \in \mathcal{S}$, since $|f|$ is $\mathcal{S}$-measurable by 2.45. If $f(x) \neq 0$ then $|f(x)| > \frac{1}{k}$ for some k, so $x \in A_k$; conversely $x \in A_k$ forces $|f(x)| > 0$. This is the displayed union. Finally $\frac{1}{k}\chi_{A_k} \le |f|$ pointwise, so 3.8 and then 3.7 give

$$\tfrac{1}{k}\mu(A_k) = \int \tfrac{1}{k}\chi_{A_k} \, d\mu \le \int |f| \, d\mu = \|f\|_1,$$

and $\|f\|_1 < \infty$ because $f \in L^1(\mu)$; hence $\mu(A_k) \le k\|f\|_1 < \infty$.

Problem 3B.12. Suppose

$$f_k(x) = \frac{(1-x)^k \cos x^k}{\sqrt{x}}.$$

Prove that $\lim_{k \to \infty} \int_0^1 f_k = 0$.

Solution. Dominate by $g(x) = 1/\sqrt{x}$ and apply the Dominated Convergence Theorem 3.31 on $(0,1)$. Each f_k is continuous on $(0,1)$, hence Borel measurable (2.41). For $x \in (0,1)$ we have $0 < (1-x)^k \le 1$ and $|\cos x^k| \le 1$, so $|f_k| \le g$ for every k; and $(1-x)^k \to 0$ for each fixed x, so $f_k \to 0$ pointwise. The functions $g\chi_{(1/n,1)}$ increase pointwise to g, and g is continuous on $[1/n, 1]$, so 3.11 with 3.34 gives

$$\int_{(0,1)} g \, d\lambda = \lim_{n \to \infty} \int_{1/n}^{1} \frac{dx}{\sqrt{x}} = \lim_{n \to \infty} \left(2 - \frac{2}{\sqrt{n}}\right) = 2 < \infty.$$

Hence 3.31 applies, and with 3.39,

$$\lim_{k \to \infty} \int_0^1 f_k = \lim_{k \to \infty} \int_{(0,1)} f_k \, d\lambda = 0.$$

Problem 3B.13. Give an example of a sequence of nonnegative Borel measurable functions $f_1, f_2, \dots$ on $[0,1]$ such that both the following conditions hold.

- $\lim_{k \to \infty} \int_0^1 f_k = 0$;
- $\sup_{k \ge m} f_k(x) = \infty$ for every $m \in \mathbb{Z}^+$ and every $x \in [0,1]$.

Solution. Take the typewriter sequence: writing $k = \frac{(n-1)n}{2} + j$ with $n \in \mathbb{Z}^+$ and $j \in \{1, \dots, n\}$, a decomposition that exists and is unique because $\frac{(n-1)n}{2}$ is strictly increasing to ∞ in n, set

$$f_k = \sqrt{n}\, \chi_{[\frac{j-1}{n}, \frac{j}{n}]}.$$

Each f_k is a nonnegative multiple of the characteristic function of a closed interval, hence a nonnegative

Borel measurable function on $[0,1]$, and 3.7 gives

$$\int_0^1 f_k = \sqrt{n}\cdot\frac{1}{n} = \frac{1}{\sqrt{n}}.$$

Only the n indices $k \in \{\frac{(n-1)n}{2}+1,\dots,\frac{n(n+1)}{2}\}$ carry a given n, so $n\to\infty$ as $k\to\infty$, and the first condition holds.
For the second, fix $x\in[0,1]$ and $m\in\mathbb{Z}^+$, and let $n\in\mathbb{Z}^+$ be arbitrary. The intervals $[\frac{j-1}{n},\frac{j}{n}]$ for $1\le j\le n$ cover $[0,1]$, so some j has $x\in[\frac{j-1}{n},\frac{j}{n}]$, and then $k_n=\frac{(n-1)n}{2}+j$ satisfies $f_{k_n}(x)=\sqrt{n}$. Since $k_n>\frac{(n-1)n}{2}\to\infty$, we have $k_n\ge m$ for all large n, so $\sup_{k\ge m} f_k(x)\ge\sqrt{n}$ for all large n, that is $\sup_{k\ge m} f_k(x)=\infty$.

Problem 3B.14. Let λ denote Lebesgue measure on $\mathbb{R}$.
(a) Let $f(x)=1/\sqrt{x}$. Prove that $\int_{[0,1]} f\,d\lambda = 2$.
(b) Let $f(x)=1/(1+x^2)$. Prove that $\int_{\mathbb{R}} f\,d\lambda=\pi$.
(c) Let $f(x)=(\sin x)/x$. Show that the integral $\int_{(0,\infty)} f\,d\lambda$ is not defined but $\lim_{t\to\infty}\int_{(0,t)} f\,d\lambda$ exists in $\mathbb{R}$.

Solution. The values are 2 in (a) and π in (b); in (c) both $\int f^{\pm}\,d\lambda=\infty$, while integration by parts makes the truncated integrals converge.
(a) Set $f(0)=\infty$, which is immaterial since $\lambda(\{0\})=0$ and makes the approximation below converge everywhere. Then f is Borel measurable by 2.52, since $f^{-1}((a,\infty])$ is $[0,1]$ for $a<1$ and $[0,a^{-2})$ for $a\ge 1$. The functions $g_n=\min\{f,n\}$ are continuous on $[0,1]$ (the two formulas n on $[0,n^{-2}]$ and $x^{-1/2}$ on $[n^{-2},1]$ agree at n^{-2}), increase, and converge pointwise to f; as each is Riemann integrable, 3.34 gives

$$\int_{[0,1]} g_n\,d\lambda = n\cdot\frac{1}{n^2}+\int_{1/n^2}^1\frac{dx}{\sqrt{x}}=\frac{1}{n}+\left(2-\frac{2}{n}\right)=2-\frac{1}{n},$$

so 3.11 gives $\int_{[0,1]} f\,d\lambda=2$.
(b) Here f is continuous, hence Borel measurable (2.41), and $g_n=f\chi_{[-n,n]}$ increases pointwise to f, with $\int_{\mathbb{R}} g_n\,d\lambda=\int_{-n}^n\frac{dx}{1+x^2}=2\arctan n$ by 3.34. Hence 3.11 gives

$$\int_{\mathbb{R}} f\,d\lambda=\lim_{n\to\infty}2\arctan n=\pi.$$

(c) The function $f(x)=(\sin x)/x$ is continuous on $(0,\infty)$, so f and hence $f^{\pm}$ are Borel measurable (2.41, 2.44).
The integral is not defined. On $A_k=(2k\pi,(2k+1)\pi)$ we have $\sin x>0$, so $f^+=f\ge\frac{\sin x}{(2k+1)\pi}$ there; and 3.34 with the fundamental theorem of calculus gives $\int\chi_{A_k}\sin x\,d\lambda=[-\cos x]_{2k\pi}^{(2k+1)\pi}=2$, the two endpoints forming a λ-null set on which the integrand is bounded. Hence $\int f^+\chi_{A_k}\,d\lambda\ge\frac{2}{(2k+1)\pi}$, and since the A_k are disjoint, 3.8 and 3.16 give

$$\int f^+\,d\lambda\ge\sum_{k=0}^{n-1}\frac{2}{(2k+1)\pi}\qquad(n\in\mathbb{Z}^+),$$

whose right side diverges by comparison with the harmonic series; so $\int f^+\,d\lambda=\infty$. The same argument on $B_k=((2k+1)\pi,(2k+2)\pi)$, where $f^-=-f\ge\frac{-\sin x}{(2k+2)\pi}$ and $\int\chi_{B_k}(-\sin x)\,d\lambda=2$, gives $\int f^-\,d\lambda\ge\sum_{k=0}^{n-1}\frac{1}{(k+1)\pi}$ for every n, so $\int f^-\,d\lambda=\infty$ too, and 3.18 assigns no value to $\int_{(0,\infty)} f\,d\lambda$.
The limit exists. Extend f continuously to $[0,\infty)$ by $\tilde f(0)=1$. For $t>0$ the function $\tilde f$ is bounded on $[0,t]$ and $\{0,t\}$ is λ-null, so by 3.8, 3.7 and the last assertion of 3.34

$$\int_{(0,t)} f\,d\lambda=\int_{[0,t]}\tilde f\,d\lambda=\int_0^t\frac{\sin x}{x}\,dx\in\mathbb{R}.$$

For $t \geq 1$, integration by parts for Riemann integrals gives

$$\int_1^t \frac{\sin x}{x}\,dx = \cos 1 - \frac{\cos t}{t} - \int_1^t \frac{\cos x}{x^2}\,dx.$$

Now $h(x) = (\cos x)/x^2$ is continuous with $|h| \leq x^{-2}$, and 3.11 with 3.34 gives $\int_{(1,\infty)} x^{-2}\,d\lambda = 1 < \infty$, so $h \in L^1((1,\infty))$. For any sequence $t_n \to \infty$ in $(1,\infty)$ we have $h\chi_{(1,t_n)} \to h$ pointwise with $|h\chi_{(1,t_n)}| \leq x^{-2}$, so 3.31 (with 3.34 identifying $\int h\chi_{(1,t_n)}\,d\lambda = \int_1^{t_n} h$) gives that

$$L = \lim_{t\to\infty} \int_1^t \frac{\cos x}{x^2}\,dx = \int_{(1,\infty)} h\,d\lambda$$

exists in $\mathbb{R}$. Since $|\cos t/t| \leq 1/t \to 0$, letting $t \to \infty$ in the two displays above gives

$$\lim_{t\to\infty} \int_{(0,\,t)} f\,d\lambda = \int_0^1 \frac{\sin x}{x}\,dx + \cos 1 - L \in \mathbb{R}.$$

Problem 3B.15. Prove or give a counterexample: If G is an open subset of $(0,1)$, then χ_G is Riemann integrable on $[0,1]$.

Solution. False: with $r_1, r_2, \ldots$ an enumeration of $\mathbb{Q} \cap (0,1)$, take

$$G = (0,1) \cap \bigcup_{k=1}^{\infty} \big(r_k - 2^{-k-2},\ r_k + 2^{-k-2}\big).$$

Then G is open, and it contains every r_k, hence is dense in $(0,1)$; being open it is Borel, so $\lambda(G) = |G|$ (2.68, 2.69) and countable subadditivity of outer measure (2.8) with $|I| \leq \ell(I)$ (2.2) gives

$$\lambda(G) \leq \sum_{k=1}^{\infty} 2 \cdot 2^{-k-2} = \tfrac{1}{2}.$$

Put $f = \chi_G$ on $[0,1]$; it is bounded, so $L(f,[0,1])$ and $U(f,[0,1])$ of 1.7 make sense. Let $P : 0 = x_0 < \cdots < x_n = 1$ be any partition of $[0,1]$.
(i) Each $(x_{j-1}, x_j) \cap (0,1)$ is a nonempty open subset of $(0,1)$, so density gives $y \in G$ with $x_{j-1} < y < x_j$; hence $\sup_{[x_{j-1},x_j]} f = 1$ and, by 1.3,

$$U(f,P,[0,1]) = \sum_{j=1}^{n} (x_j - x_{j-1}) = 1.$$

(ii) Since f takes only the values 0 and 1, $\inf_{[x_{j-1},x_j]} f = 1$ exactly when $[x_{j-1},x_j] \subseteq G$. Those intervals have pairwise disjoint interiors, so with $J = \{j : [x_{j-1},x_j] \subseteq G\}$, finite additivity and monotonicity of λ (the overlaps being finitely many null points) give

$$L(f,P,[0,1]) = \sum_{j\in J} (x_j - x_{j-1}) = \lambda\Big(\bigcup_{j\in J} [x_{j-1},x_j]\Big) \leq \lambda(G) \leq \tfrac{1}{2}.$$

As P was arbitrary, $L(f,[0,1]) \leq \frac{1}{2} < 1 = U(f,[0,1])$, so χ_G is not Riemann integrable on $[0,1]$ by 1.9.

Problem 3B.16. Suppose $f \in \mathcal{L}^1(\mathbb{R})$.

- (a) For $t \in \mathbb{R}$, define $f_t \colon \mathbb{R} \to \mathbb{R}$ by $f_t(x) = f(x-t)$. Prove that
$$\lim_{t\to 0} \|f - f_t\|_1 = 0.$$
- (b) For $t > 0$, define $f_t \colon \mathbb{R} \to \mathbb{R}$ by $f_t(x) = f(tx)$. Prove that

$$\lim_{t\to 1}\|f - f_t\|_1 = 0.$$

Solution. Both parts are the $\varepsilon/3$ argument over the dense class supplied by 3.48, once we record how the $\mathcal{L}^1$-norm transforms.

Change of variables. If $h \in \mathcal{L}^1(\mathbb{R})$, $t \in \mathbb{R}$ and $c > 0$, then $x \mapsto h(x-t)$ and $x \mapsto h(cx)$ lie in $\mathcal{L}^1(\mathbb{R})$, with

$$\|h(\cdot - t)\|_1 = \|h\|_1, \qquad \|h(c\,\cdot)\|_1 = \tfrac{1}{c}\|h\|_1.$$

Indeed the σ-algebra $\mathcal{S}$ in question (Borel or Lebesgue, 3.45) is stable under $A \mapsto t + A$ and $A \mapsto cA$, since open sets are and null sets are, by 2.7 and 2.70; and $\lambda(t+A) = \lambda(A)$ by 2.7 while $\lambda(cA) = c\lambda(A)$ because scaling every interval of a cover in 2.2 by c scales its length by c. So $\{A_j\} \mapsto \{t + A_j\}$ and $\{A_j\} \mapsto \{\frac{1}{c}A_j\}$ are bijections of the set of $\mathcal{S}$-partitions of $\mathbb{R}$ (3.1) onto itself under which each lower Lebesgue sum (3.2) of $|h|$ is reproduced exactly, respectively multiplied by $\frac{1}{c}$; take suprema (3.3).

Part (a). Let $\varepsilon > 0$. By 3.48 choose g continuous and vanishing outside $[-M, M]$ with $\|f - g\|_1 < \frac{\varepsilon}{3}$; such a g is bounded and uniformly continuous on $\mathbb{R}$, being uniformly continuous on the compact interval $[-M-1, M+1]$ and zero outside $[-M, M]$. Put $g_t(x) = g(x-t)$. For $|t| \le 1$ both g and g_t vanish outside $[-M-1, M+1]$, so 3.8 and 3.4 give

$$\|g - g_t\|_1 \le (2M+2)\sup_{x\in\mathbb{R}}|g(x) - g(x-t)|.$$

Uniform continuity supplies $\delta \in (0,1]$ with that supremum at most $\frac{\varepsilon}{3(2M+2)}$ whenever $|t| < \delta$, whence $\|g - g_t\|_1 \le \frac{\varepsilon}{3}$. Since $f_t - g_t$ is the translate of $f - g$, the change of variables gives $\|f_t - g_t\|_1 = \|f - g\|_1 < \frac{\varepsilon}{3}$. Thus for $|t| < \delta$ the triangle inequality (3.43) gives

$$\|f - f_t\|_1 \le \|f - g\|_1 + \|g - g_t\|_1 + \|g_t - f_t\|_1 < \varepsilon.$$

Part (b). Let $\varepsilon > 0$ and choose g as above with $\|f - g\|_1 < \frac{\varepsilon}{4}$, and put $g_t(x) = g(tx)$. By the change of variables, $\|f_t - g_t\|_1 = \frac{1}{t}\|f - g\|_1 \le 2\|f - g\|_1 < \frac{\varepsilon}{2}$ for every $t \ge \frac{1}{2}$. For $\frac{1}{2} \le t \le 2$ and $|x| > 2M$ we have $|x| > M$ and $|tx| > M$, so $g - g_t$ vanishes outside $[-2M, 2M]$, where $|x - tx| \le 2M|t-1|$. Uniform continuity supplies $\eta > 0$ with $|g(u) - g(v)| < \frac{\varepsilon}{16M}$ whenever $|u - v| < \eta$; set $\rho = \min\{\frac{1}{2}, \frac{\eta}{2M}\}$. If $|t-1| < \rho$ then $|g - g_t| \le \frac{\varepsilon}{16M}\chi_{[-2M,2M]}$, so $\|g - g_t\|_1 \le \frac{\varepsilon}{4}$ by 3.8 and 3.4. Hence $\|f - f_t\|_1 < \varepsilon$ for $|t-1| < \rho$ by 3.43.

4 Differentiation

Contents

4.1 Exercises 4A

Problem 4A.1. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $h : X \to \mathbb{R}$ is an $\mathcal{S}$-measurable function. Prove that

$$\mu(\{x \in X : |h(x)| \geq c\}) \leq \frac{1}{c^p} \int |h|^p \, d\mu$$

for all positive numbers c and p.

Solution. Everything comes from the pointwise inequality $c^p \chi_E \leq |h|^p$ on X, where $E = \{x \in X : |h(x)| \geq c\}$: off E the left side is 0, while on E we have $|h(x)|^p \geq c^p$ because $t \mapsto t^p$ is increasing on $[0, \infty)$. Both sides are nonnegative and $\mathcal{S}$-measurable, so 3.8 and 3.15 give

$$c^p \, \mu\big(\{x \in X : |h(x)| \geq c\}\big) = \int c^p \chi_E \, d\mu \leq \int |h|^p \, d\mu.$$

Divide by $c^p > 0$.

Problem 4A.2. Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) = 1$ and $h \in \mathcal{L}^1(\mu)$. Prove that

$$\mu\left(\left\{x \in X : \left|h(x) - \int h \, d\mu\right| \geq c\right\}\right) \leq \frac{1}{c^2}\left(\int h^2 \, d\mu - \left(\int h \, d\mu\right)^2\right)$$

for all $c > 0$.
[The result above is called Chebyshev's inequality; it plays an important role in probability theory. Pafnuty Chebyshev (1821-1894) was Markov's thesis advisor.]

Solution. Write $m = \int h \, d\mu$, a real number because $h \in \mathcal{L}^1(\mu)$. Applying Exercise 1 in this section to the $\mathcal{S}$-measurable function $h - m$ with $p = 2$ gives

$$\mu\big(\{x \in X : |h(x) - m| \geq c\}\big) \leq \frac{1}{c^2} \int (h - m)^2 \, d\mu,$$

so it suffices to show $\int (h - m)^2 \, d\mu = \int h^2 \, d\mu - m^2$. If $\int h^2 \, d\mu = \infty$ the desired right side is ∞ and there is nothing to prove, so assume $h^2 \in \mathcal{L}^1(\mu)$. Then the constant function m is in $\mathcal{L}^1(\mu)$ too, since $\int |m| \, d\mu = |m|\mu(X) = |m| < \infty$, and so is $-2mh$; hence additivity and homogeneity of integration (3.21, 3.20) applied to the pointwise identity $(h - m)^2 = h^2 - 2mh + m^2$ yield

$$\begin{aligned}\int (h - m)^2 \, d\mu &= \int h^2 \, d\mu - 2m \int h \, d\mu + m^2 \mu(X) \\ &= \int h^2 \, d\mu - m^2,\end{aligned}$$

using $\mu(X) = 1$ again.

Problem 4A.3. Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Suppose $h \in \mathcal{L}^1(\mu)$ and $\|h\|_1 > 0$. Prove that there is at most one number $c \in (0, \infty)$ such that

$$\mu(\{x \in X : |h(x)| \geq c\}) = \frac{1}{c}\|h\|_1.$$

Solution. Call $c \in (0, \infty)$ an *equality point* if $\mu(E_c) = \frac{1}{c}\|h\|_1$, where $E_c = \{x \in X : |h(x)| \geq c\}$. The whole point is that an equality point pins down $|h|$ on E_c: we show $0 < \mu(E_c) < \infty$ and $|h| = c$ almost everywhere on E_c.
Since $|h| = \chi_{E_c}|h| + \chi_{X \setminus E_c}|h|$ pointwise and $\chi_{E_c}|h| \geq c\chi_{E_c}$, additivity for nonnegative functions (3.16)

and 3.8 give

$$\|h\|_1 = \int_{E_c} |h|\,d\mu + \int_{X\setminus E_c} |h|\,d\mu \ge c\,\mu(E_c) + \int_{X\setminus E_c} |h|\,d\mu,$$

whose first term is $\|h\|_1$; as $\|h\|_1 < \infty$ we may cancel it, leaving

$$\int_{E_c} |h|\,d\mu = c\,\mu(E_c) = \|h\|_1 \in (0,\infty),$$

so $0 < \mu(E_c) < \infty$. Applying 3.16 to $\chi_{E_c}|h| = \chi_{E_c}(|h| - c) + c\chi_{E_c}$ and subtracting the finite number $c\,\mu(E_c)$ gives $\int \chi_{E_c}(|h| - c)\,d\mu = 0$. A nonnegative $\mathcal{S}$-measurable g with $\int g\,d\mu = 0$ vanishes almost everywhere, since Exercise 1 in this section with $p = 1$ gives $\mu(\{g \ge \frac{1}{n}\}) = 0$ for every $n \in \mathbb{Z}^+$ and $\{g > 0\} = \bigcup_n \{g \ge \frac{1}{n}\}$ is then null by 2.58. Hence $|h| = c$ almost everywhere on E_c.
Now suppose $c_1 < c_2$ were both equality points. Then $E_{c_2} \subseteq E_{c_1}$, and no x can have $|h(x)|$ equal to both c_1 and c_2, so

$$E_{c_2} \subseteq \{x \in E_{c_1} : |h(x)| \ne c_1\} \cup \{x \in E_{c_2} : |h(x)| \ne c_2\},$$

a union of two sets of measure 0; thus $\mu(E_{c_2}) = 0$ by 2.58, contradicting $\mu(E_{c_2}) = \|h\|_1/c_2 > 0$.

Problem 4A.4. Show that the constant 3 in the Vitali Covering Lemma (4.4) cannot be replaced by a smaller positive constant.

Solution. Given $\alpha \in (0,3)$, choose ε with $0 < \varepsilon < \min\{2, 3-\alpha\}$ and take the two intervals

$$I_1 = (0,2), \qquad I_2 = (2-\varepsilon,\ 4-\varepsilon),$$

writing $\alpha * I$ for the open interval with the same center as I and α times its length (4.2 is the case $\alpha = 3$). Since $I_1 \cap I_2 = (2-\varepsilon, 2) \ne \emptyset$, the only disjoint sublists are the empty one and the two singletons, and $I_1 \cup I_2 = (0, 4-\varepsilon)$. The empty sublist fails because $I_1 \cup I_2 \ne \emptyset$. For the singletons,

$$\alpha * I_1 = (1-\alpha,\ 1+\alpha), \qquad \alpha * I_2 = (3-\varepsilon-\alpha,\ 3-\varepsilon+\alpha),$$

so $I_1 \cup I_2 \subseteq \alpha * I_1$ would force $4 - \varepsilon \le 1 + \alpha$ and $I_1 \cup I_2 \subseteq \alpha * I_2$ would force $3 - \varepsilon - \alpha \le 0$, each saying $\alpha \ge 3 - \varepsilon$, contrary to $\varepsilon < 3 - \alpha$. Hence no disjoint sublist has the covering property, so the conclusion of 4.4 fails with α in place of 3.

Problem 4A.5. Prove the assertion left as an exercise in the last sentence of the proof of the Vitali Covering Lemma (4.4).
[The last sentence of that proof asserts: if I and J are bounded nonempty open intervals of $\mathbb{R}$ with $I \cap J \ne \emptyset$ and $|I| \ge |J|$, then $J \subseteq 3 * I$. In the notation of the proof, $I = I_{k_L}$ and $J = I_j$.]

Solution. Write $I = (a-r, a+r)$ and $J = (b-s, b+s)$ with $r, s > 0$, so that the hypothesis $|I| \ge |J|$ reads $s \le r$, and $3*I = (a-3r,\ a+3r)$ by 4.2. Choosing $x \in I \cap J$ gives $|a-b| \le |a-x| + |x-b| < r+s$, so every $y \in J$ satisfies

$$|y-a| \le |y-b| + |b-a| < s + (r+s) = r + 2s \le 3r.$$

Thus $J \subseteq 3 * I$.

Problem 4A.6. Verify the formula in Example 4.7 for the Hardy-Littlewood maximal function of $\chi_{[0,1]}$.

[Example 4.7 asserts that, with $\chi_{[0,1]}$ the characteristic function of the interval $[0,1]$,

$$(\chi_{[0,1]})^*(b) = \begin{cases} \dfrac{1}{2(1-b)} & \text{if } b \le 0, \\ 1 & \text{if } 0 < b < 1, \\ \dfrac{1}{2b} & \text{if } b \ge 1. \end{cases}$$

as you should verify.]

Solution. Write $h = \chi_{[0,1]}$ and, following 4.6,

$$h^*(b) = \sup_{t>0} \psi_b(t), \qquad \psi_b(t) = \frac{\big|[0,1] \cap [b-t,\, b+t]\big|}{2t},$$

the numerator being $\int_{b-t}^{b+t} |h|$. Since that intersection lies inside $[b-t, b+t]$, we always have $\psi_b(t) \le 1$.
(i) $0 < b < 1$. For $0 < t \le \min\{b, 1-b\}$ we have $[b-t, b+t] \subseteq [0,1]$, so $\psi_b(t) = 1$, and therefore $h^*(b) = 1$.
(ii) $b \le 0$. Here $b - t < 0$ for every $t > 0$, so the overlap is $\max\{0, \min\{b+t, 1\}\}$ and

$$\psi_b(t) = \begin{cases} 0 & \text{if } 0 < t \le -b, \\ \frac{1}{2} + \frac{b}{2t} & \text{if } -b < t \le 1-b, \\ \frac{1}{2t} & \text{if } t > 1-b. \end{cases}$$

The middle branch is nondecreasing in t because $b \le 0$, with value $\frac{1}{2(1-b)}$ at $t = 1-b$, while the third branch is smaller than $\frac{1}{2(1-b)}$ throughout. Hence $h^*(b) = \frac{1}{2(1-b)}$.
(iii) $b \ge 1$. Here $b + t > 1$ for every $t > 0$, so the overlap is $\max\{0, 1 - \max\{b-t, 0\}\}$ and

$$\psi_b(t) = \begin{cases} 0 & \text{if } 0 < t \le b-1, \\ \frac{1}{2} + \frac{1-b}{2t} & \text{if } b-1 < t \le b, \\ \frac{1}{2t} & \text{if } t > b. \end{cases}$$

The middle branch is nondecreasing because $1 - b \le 0$, with value $\frac{1}{2b}$ at $t = b$, and the third branch stays below that. Hence $h^*(b) = \frac{1}{2b}$.
These three values are exactly the formula in Example 4.7.

Problem 4A.7. Find a formula for the Hardy-Littlewood maximal function of the characteristic function of $[0,1] \cup [2,3]$.

Solution. With $E = [0,1] \cup [2,3]$ and $h = \chi_E$,

$$h^*(b) = \begin{cases} \dfrac{1}{3-b} & \text{if } b \le -1, \\ \dfrac{1}{2(1-b)} & \text{if } -1 \le b \le 0, \\ 1 & \text{if } 0 < b < 1, \\ \dfrac{2b-1}{2b} & \text{if } 1 \le b \le \frac{3}{2}, \\ \dfrac{5-2b}{2(3-b)} & \text{if } \frac{3}{2} \le b \le 2, \\ 1 & \text{if } 2 < b < 3, \\ \dfrac{1}{2(b-2)} & \text{if } 3 \le b \le 4, \\ \dfrac{1}{b} & \text{if } b \ge 4 \end{cases}$$

(the two formulas agree at $b = -1, \frac{3}{2}, 4$). Put $g(y) = |E \cap (-\infty, y]|$, which is continuous, nondecreasing, and equal to $0, y, 1, y-1, 2$ on $(-\infty, 0], [0,1], [1,2], [2,3], [3,\infty)$ respectively. By additivity of Lebesgue measure and 4.6,

$$h^*(b) = \sup_{t>0} \psi_b(t), \qquad \psi_b(t) = \frac{g(b+t) - g(b-t)}{2t} \le 1,$$

the bound because $E \cap [b-t, b+t] \subseteq [b-t, b+t]$.

Reduction to finitely many t. Let $t_1 < \cdots < t_N$ be the positive numbers with $b+t$ or $b-t$ in $\{0,1,2,3\}$. Across any t-interval meeting none of them, both $t \mapsto g(b \pm t)$ are affine, so $\psi_b(t) = \frac{\alpha}{2} + \frac{\beta}{2t}$ is monotone there; ψ_b is continuous, $\lim_{t\downarrow 0} \psi_b(t)$ exists in $[0,1]$ by monotonicity and the bound $\psi_b \le 1$, and $\psi_b(t) \to 0$ as $t \to \infty$ since $g(b+t) - g(b-t) \le 2$, so on $[t_N, \infty)$ the monotone $\psi_b \ge 0$ is maximal at t_N. Hence

$$h^*(b) = \max\{\lim_{t\downarrow 0} \psi_b(t),\ \psi_b(t_1),\ \ldots,\ \psi_b(t_N)\}.$$

Also $y \mapsto 3-y$ maps E onto E and preserves Lebesgue measure, so $\psi_{3-b} = \psi_b$ and $h^*(3-b) = h^*(b)$; it therefore suffices to treat $b \le \frac{3}{2}$.

(i) $0 < b < 1$. For $0 < t \le \min\{b, 1-b\}$ we have $[b-t, b+t] \subseteq [0,1]$, so $\psi_b(t) = 1$ and $h^*(b) = 1$.

(ii) $b \le 0$. Then $g(b-t) = 0$ for all $t > 0$, the candidates are $t = -b, 1-b, 2-b, 3-b$, and

$$\psi_b(-b) = 0, \quad \psi_b(1-b) = \frac{1}{2(1-b)}, \quad \psi_b(2-b) = \frac{1}{2(2-b)}, \quad \psi_b(3-b) = \frac{1}{3-b},$$

while $\lim_{t\downarrow 0} \psi_b(t)$ is 0 if $b < 0$ and $\frac{1}{2} = \frac{1}{2(1-b)}$ if $b = 0$. As $\frac{1}{2(2-b)} < \frac{1}{2(1-b)}$, we get $h^*(b) = \max\{\frac{1}{2(1-b)}, \frac{1}{3-b}\}$, and $\frac{1}{2(1-b)} \ge \frac{1}{3-b}$ exactly when $b \ge -1$. This gives the first two branches.

(iii) $1 \le b \le \frac{3}{2}$. The candidates are $t = b-1, b, 2-b, 3-b$, and since then $2b-1 \in [1,2]$, $2b \in [2,3]$, $2b-2 \in [0,1]$, $2b-3 \le 0$,

$$\psi_b(b-1) = \frac{1-1}{2(b-1)} = 0,$$
$$\psi_b(b) = \frac{(2b-1)-0}{2b} =: A,$$
$$\psi_b(2-b) = \frac{1-(2b-2)}{2(2-b)} =: B,$$
$$\psi_b(3-b) = \frac{2-0}{2(3-b)} =: C,$$

while $\lim_{t\downarrow 0} \psi_b(t)$ is 0 for $b > 1$ and $\frac{1}{2} = A$ at $b = 1$. Now $A \ge C$ is equivalent to $(2b-3)(b-1) \le 0$, which holds on $[1, \frac{3}{2}]$; and $A = 1 - \frac{1}{2b}$ is increasing while $B = \frac{3-2b}{4-2b}$ is decreasing, with $A = B = \frac{1}{2}$ at $b = 1$, so $B \le A$. Hence $h^*(b) = A = \frac{2b-1}{2b}$.

(iv) $b \ge \frac{3}{2}$. By the symmetry $h^*(b) = h^*(3-b)$, cases (iii), (i), (ii) applied at $3-b$ give $\frac{5-2b}{2(3-b)}$ on $[\frac{3}{2}, 2]$, then 1 on $(2,3)$, then $\frac{1}{2(b-2)}$ on $[3,4]$, then $\frac{1}{b}$ for $b \ge 4$.

Problem 4A.8. Find a formula for the Hardy–Littlewood maximal function of the function $h\colon \mathbb{R} \to [0,\infty)$ defined by

$$h(x) = \begin{cases} x & \text{if } 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Solution. The answer is

$$h^*(b) = \begin{cases} \dfrac{1}{4(1-b)} & \text{if } b \le \frac{1}{2}, \\ b & \text{if } \frac{1}{2} \le b < 1, \\ \dfrac{b - \sqrt{b^2-1}}{2} = \dfrac{1}{2(b + \sqrt{b^2-1})} & \text{if } b \ge 1 \end{cases}$$

(the first two branches agree at $b = \frac{1}{2}$). Since $|h| = h$, writing $F(x) = \int_{-\infty}^{x} h$, so that $F = 0, \frac{x^2}{2}, \frac{1}{2}$ on $(-\infty, 0], [0,1], [1,\infty)$, Definition 4.6 reads

$$h^*(b) = \sup_{t>0} V_b(t), \qquad V_b(t) = \frac{F(b+t) - F(b-t)}{2t}.$$

(i) $b \le 0$. Then $F(b-t) = 0$ for every $t > 0$. For $0 < t \le -b$, $V_b(t) = 0$. For $-b \le t \le 1-b$, substituting $v = b + t \in [0,1]$,

$$V_b(t) = \frac{v^2}{4(v-b)}, \qquad \frac{d}{dv}\Big(\frac{v^2}{v-b}\Big) = \frac{v(v-2b)}{(v-b)^2} \ge 0,$$

so V_b increases to $\frac{1}{4(1-b)}$ at $t = 1-b$; for $t \ge 1-b$, $V_b(t) = \frac{1}{4t}$ decreases from that same value. Hence $h^*(b) = \frac{1}{4(1-b)}$.

(ii) $0 < b < 1$. For $0 < t \le \min\{b, 1-b\}$ both $b \pm t$ lie in $[0,1]$, so $V_b(t) = \frac{(b+t)^2 - (b-t)^2}{4t} = b$. If $b \le \frac{1}{2}$, then for $b \le t \le 1-b$ we get $V_b(t) = \frac{(b+t)^2}{4t}$ with derivative $\frac{(b+t)(t-b)}{4t^2} \ge 0$, so V_b increases to $\frac{1}{4(1-b)}$ at $t = 1-b$, beyond which $V_b(t) = \frac{1}{4t}$ decreases; since $1 - 4b(1-b) = (2b-1)^2 \ge 0$ gives $\frac{1}{4(1-b)} \ge b$, we get $h^*(b) = \frac{1}{4(1-b)}$. If $b \ge \frac{1}{2}$, then for $1 - b \le t \le b$, substituting $u = b - t \in [0, 2b-1]$,

$$V_b(t) = \frac{1-u^2}{4(b-u)}, \qquad \frac{d}{du}\Big(\frac{1-u^2}{b-u}\Big) = \frac{u^2 - 2bu + 1}{(b-u)^2} > 0$$

because that quadratic has discriminant $4b^2 - 4 < 0$; so V_b decreases in t from its value b at $t = 1-b$, while for $t \ge b$ we have $V_b(t) = \frac{1}{4t} \le \frac{1}{4b} \le b$. Hence $h^*(b) = b$.

(iii) $b \ge 1$. Then $F(b+t) = \frac{1}{2}$ for every $t > 0$, and $V_b(t) = 0$ for $0 < t \le b-1$. For $b - 1 \le t < b$, substituting $u = b - t$, the function $g(u) = \frac{1-u^2}{4(b-u)}$ has $g'(u) = \frac{u^2-2bu+1}{4(b-u)^2}$, whose numerator has roots $b \pm \sqrt{b^2-1}$ with $u_0 = b - \sqrt{b^2-1} \in (0,1]$ (as $(b-1)^2 \le (b-1)(b+1)$) and the other root at least 1; so g increases on $[0, u_0)$ and decreases on $(u_0, 1]$, and since $b - u_0 = \sqrt{b^2-1}$ and $1 - u_0^2 = 2\sqrt{b^2-1}\,(b - \sqrt{b^2-1})$,

$$g(u_0) = \frac{b - \sqrt{b^2-1}}{2}.$$

When $b > 1$ this value is attained at $t = \sqrt{b^2-1} > 0$; when $b = 1$ we have $u_0 = 1$, i.e. the excluded $t = 0$, but g increases on $[0,1)$ with $\lim_{u\uparrow 1} \frac{1+u}{4} = \frac{1}{2}$, again the same number. Finally $t \ge b$ gives $V_b(t) = \frac{1}{4t} \le \frac{1}{4b} \le g(u_0)$, because $b - \sqrt{b^2-1} = \frac{1}{b+\sqrt{b^2-1}} \ge \frac{1}{2b}$. Hence $h^*(b) = \frac{b - \sqrt{b^2-1}}{2}$.

Problem 4A.9. Suppose $h\colon \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable. Prove that

$$\{b \in \mathbb{R} : h^*(b) > c\}$$

is an open subset of $\mathbb{R}$ for every $c \in \mathbb{R}$.

Solution. Fix $c \in \mathbb{R}$ and b with $h^*(b) > c$; it suffices to produce $\delta > 0$ with $h^* > c$ on $(b-\delta, b+\delta)$. By 4.6 there is $t > 0$ with $\alpha > 2tc$, where $\alpha = \int_{b-t}^{b+t} |h| \in [0, \infty]$. Choose $t' > t$ with $2t'c < \alpha$: for $c \le 0$ any $t' > t$ serves (when $c = 0$ note $\alpha > 0$), while for $c > 0$ take $t' = t+1$ if $\alpha = \infty$ and any $t' \in (t, \frac{\alpha}{2c})$ otherwise, that interval being nonempty because $\frac{\alpha}{2c} > t$. Put $\delta = t' - t > 0$. If $|b' - b| < \delta$, then $b' - t' < b - t$ and $b' + t' > b + t$, so $(b-t, b+t) \subseteq (b'-t', b'+t')$ and monotonicity of the integral of $|h| \ge 0$ gives

$$h^*(b') \ge \frac{1}{2t'} \int_{b'-t'}^{b'+t'} |h| \ge \frac{\alpha}{2t'} > c.$$

Thus every point of $\{h^* > c\}$ is an interior point.

Problem 4A.10. Prove or give a counterexample: If $h\colon \mathbb{R} \to [0, \infty)$ is an increasing function, then h^* is an increasing function.

Solution. True: h^* is increasing. Suppose $b_1 < b_2$ and put $s = b_2 - b_1 > 0$ (note h is Borel measurable, each $\{h > c\}$ being an interval, so h^* is defined, and $|h| = h$). Translation invariance of Lebesgue measure (2.7) passes from characteristic functions through simple functions to all nonnegative Borel functions, giving $\int_{b_1-t}^{b_1+t} h(x+s)\,dx = \int_{b_2-t}^{b_2+t} h$ for every $t > 0$. Since h is increasing and $s > 0$, we have $h(x) \le h(x+s)$ pointwise, so monotonicity of the integral gives

$$\frac{1}{2t} \int_{b_1-t}^{b_1+t} |h| \le \frac{1}{2t} \int_{b_2-t}^{b_2+t} |h| \le h^*(b_2)$$

for every $t > 0$. Take the supremum over $t > 0$ on the left.

Problem 4A.11. Give an example of a Borel measurable function $h\colon \mathbb{R} \to [0, \infty)$ such that $h^*(b) < \infty$ for all $b \in \mathbb{R}$ but $\sup\{h^*(b) : b \in \mathbb{R}\} = \infty$.

Solution. Take

$$h = \sum_{n=1}^{\infty} n\, \chi_{[n,\, n+n^{-3}]},$$

a tall thin spike at each positive integer. The intervals are disjoint because $n + n^{-3} < n+1$, so h is $[0, \infty)$-valued, and it is Borel measurable as a pointwise limit of its partial sums (2.48). By the Monotone Convergence Theorem (3.11),

$$\|h\|_1 = \sum_{n=1}^{\infty} n \cdot n^{-3} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Finiteness of h^*: fix b, put $R = |b| + 1$, and note $M_R := \sup_{[-R,R]} h \le \lceil R \rceil < \infty$, only the spikes with $n \le R$ meeting $[-R, R]$. If $0 < t \le 1$ then $[b-t, b+t] \subseteq [-R, R]$, so the average of h over it is at most M_R; if $t \ge 1$ that average is at most $\frac{1}{2t}\|h\|_1 \le \frac{1}{2}\|h\|_1$. Hence $h^*(b) \le \max\{M_R, \frac{1}{2}\|h\|_1\} < \infty$. Unboundedness: taking $b = n$ and $t = n^{-3}$ gives

$$h^*(n) \ge \frac{n^3}{2} \int_{n-n^{-3}}^{n+n^{-3}} h \ge \frac{n^3}{2} \cdot n \cdot n^{-3} = \frac{n}{2} \longrightarrow \infty.$$

Problem 4A.12. Show that $|\{b \in \mathbb{R} : h^*(b) = \infty\}| = 0$ for every $h \in \mathcal{L}^1(\mathbb{R})$.

Solution. Write $E = \{b \in \mathbb{R} : h^*(b) = \infty\}$. For every positive integer n we have $E \subseteq \{b \in \mathbb{R} : h^*(b) > n\}$, so monotonicity of outer measure (2.5) and the Hardy–Littlewood maximal inequality (4.8), applicable since $h \in \mathcal{L}^1(\mathbb{R})$, give

$$|E| \le \bigl|\{b \in \mathbb{R} : h^*(b) > n\}\bigr| \le \frac{3}{n}\|h\|_1.$$

As $\|h\|_1 < \infty$, the right side tends to 0, so $|E| = 0$. (In fact $E = \bigcap_n \{h^* > n\}$ is Borel, each of those sets being open by Exercise 9 in this section.)

Problem 4A.13. Show that there exists $h \in \mathcal{L}^1(\mathbb{R})$ such that $h^*(b) = \infty$ for every $b \in \mathbb{Q}$.

Solution. Superimpose translated copies of one function with a single bad point. Define $g\colon \mathbb{R} \to [0, \infty)$ by

$$g(x) = \begin{cases} \dfrac{1}{x(\log x)^2} & \text{if } 0 < x < \frac{1}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

which is Borel measurable, being continuous on $(0, \frac{1}{2})$ and zero elsewhere. Since $\frac{d}{dx}\bigl(-\frac{1}{\log x}\bigr) = \frac{1}{x(\log x)^2}$ on $(0, 1)$ and $-\frac{1}{\log x} \to 0$ as $x \downarrow 0$, we get $\int_0^t g = \frac{1}{|\log t|}$ for $0 < t < \frac{1}{2}$; in particular $\|g\|_1 = \frac{1}{\log 2} < \infty$, and

$$\frac{1}{2t}\int_{-t}^{t} g = \frac{1}{2t\,|\log t|} \longrightarrow \infty \qquad (t \downarrow 0),$$

because $t|\log t| \to 0$; hence $g^*(0) = \infty$.
Let $r_1, r_2, \ldots$ enumerate $\mathbb{Q}$ and put $\tilde{h} = \sum_{k=1}^{\infty} 2^{-k} g(\cdot - r_k)$. Each summand is Borel measurable (a translate of a Borel set is Borel), so $\tilde{h}$ is Borel measurable as the supremum of the partial sums (2.53), and 3.11 together with translation invariance (2.7) gives

$$\int_{-\infty}^{\infty} \tilde{h} = \sum_{k=1}^{\infty} 2^{-k}\|g\|_1 = \frac{1}{\log 2} < \infty.$$

Hence $N = \{\tilde{h} = \infty\}$ is a null set (otherwise $\int \tilde{h} \ge c|N| \to \infty$), so $h = \tilde{h}\chi_{\mathbb{R}\setminus N}$ is a real-valued Borel function in $\mathcal{L}^1(\mathbb{R})$ agreeing with $\tilde{h}$ off N. The integral of a nonnegative function over a null set is 0, so $\int_{b-t}^{b+t} |h| = \int_{b-t}^{b+t} \tilde{h}$ for all b and $t > 0$, giving $h^* = \tilde{h}^*$.
Now fix $b \in \mathbb{Q}$, say $b = r_k$. Since $\tilde{h} \ge 2^{-k} g(\cdot - r_k)$ pointwise, and since $f_1 \le f_2$ implies $f_1^* \le f_2^*$, while $(\lambda f)^* = \lambda f^*$ for $\lambda > 0$ and $(g(\cdot - r))^*(b) = g^*(b - r)$ by 2.7 (all immediate from 4.6),

$$h^*(b) = \tilde{h}^*(b) \ge 2^{-k} g^*(b - r_k) = 2^{-k} g^*(0) = \infty.$$

Problem 4A.14. Suppose $h \in \mathcal{L}^1(\mathbb{R})$. Prove that

$$\bigl|\{b \in \mathbb{R} : h^*(b) \ge c\}\bigr| \le \frac{3}{c}\,\|h\|_1$$

for every $c > 0$.
[This result slightly strengthens the Hardy–Littlewood maximal inequality (4.8) because the set on the left side above includes those $b \in \mathbb{R}$ such that $h^*(b) = c$. A much deeper strengthening comes from replacing the constant 3 in the Hardy–Littlewood maximal inequality with a smaller constant. In 2003, Antonios Melas answered what had been an open question about the best constant. He proved that the smallest constant that can replace 3 in the Hardy–Littlewood maximal inequality is $(11 + \sqrt{61})/12 \approx 1.56752$; see Annals of Mathematics 157 (2003), 647–688.]

Solution. Approach c from below. For every integer $n > \frac{1}{c}$ the number $c - \frac{1}{n}$ is positive and $\{b : h^*(b) \geq c\} \subseteq \{b : h^*(b) > c - \frac{1}{n}\}$, so monotonicity of outer measure (2.5) and the Hardy–Littlewood maximal inequality (4.8), applicable because $h \in \mathcal{L}^1(\mathbb{R})$, give

$$\big|\{b \in \mathbb{R} : h^*(b) \geq c\}\big| \leq \frac{3}{c - \frac{1}{n}} \|h\|_1.$$

Since $\|h\|_1 < \infty$, letting $n \to \infty$ gives the bound $\frac{3}{c}\|h\|_1$.

4.2 Exercises 4B

Problem 4B.1. Suppose $f \in L^1(\mathbb{R})$. Prove that

$$\lim_{t \downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f - f_{[b-t,\, b+t]}| = 0$$

for almost every $b \in \mathbb{R}$.
[The book prefixes the exercises of this section with the following notation: for $f \in L^1(\mathbb{R})$ and I an interval of $\mathbb{R}$ with $0 < |I| < \infty$, let f_I denote the average of f on I. In other words, $f_I = \frac{1}{|I|} \int_I f$.]

Solution. The limit holds at every Lebesgue point of f, of which almost every $b \in \mathbb{R}$ is one by the first version of the Lebesgue Differentiation Theorem (4.10). So let E be the set of b with $\lim_{t \downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)| = 0$, fix $b \in E$ and $t > 0$, and write $I = [b-t, b+t]$. Since f and the constant $f(b)$ are both integrable on I, 3.21 and 3.23 give

$$|f_I - f(b)| = \Big|\frac{1}{2t} \int_{b-t}^{b+t} \big(f - f(b)\big)\Big| \leq \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)|,$$

so integrating the pointwise bound $|f - f_I| \leq |f - f(b)| + |f(b) - f_I|$ over I (3.21, 3.22) and dividing by $2t$ yields

$$\frac{1}{2t} \int_{b-t}^{b+t} |f - f_I| \leq 2 \cdot \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)| \xrightarrow[t \downarrow 0]{} 0.$$

Problem 4B.2. Suppose $f \in L^1(\mathbb{R})$. Prove that

$$\lim_{t \downarrow 0} \sup \left\{ \frac{1}{|I|} \int_I |f - f_I| \;:\; I \text{ is an interval of length } t \text{ containing } b \right\} = 0$$

for almost every $b \in \mathbb{R}$.
[As in the previous exercise, f_I denotes the average $\frac{1}{|I|} \int_I f$ of f over an interval I with $0 < |I| < \infty$.]

Solution. Again the limit holds at every b in the full-measure set E of 4.10, that is, where $\lim_{t \downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)| = 0$. Fix such a b and $t > 0$, and let I be any interval of length t containing b; then every $x \in I$ has $|x - b| \leq t$, so $I \subseteq [b-t, b+t]$. Exactly the estimate of Exercise 1 in this section, carried out on I, gives $\frac{1}{|I|} \int_I |f - f_I| \leq \frac{2}{|I|} \int_I |f - f(b)|$, whence

$$\frac{1}{|I|} \int_I |f - f_I| \leq \frac{2}{t} \int_{b-t}^{b+t} |f - f(b)| = 4 \cdot \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)|.$$

The right side is independent of I, hence bounds the supremum over all such I, and it tends to 0 as $t \downarrow 0$ because $b \in E$.

Problem 4B.3. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a Lebesgue measurable function such that $f^2 \in L^1(\mathbb{R})$. Prove that

$$\lim_{t \downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)|^2 = 0$$

for almost every $b \in \mathbb{R}$.

Solution. Approximate $f(b)$ by rationals, applying a localized form of 4.10 to each $(f-q)^2$.
Localization. If $h\colon \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable with $\int_{-n}^{n} |h| < \infty$ for every $n \in \mathbb{Z}^+$, then $\frac{1}{2t}\int_{b-t}^{b+t} h \to h(b)$ as $t \downarrow 0$ for almost every b. Indeed $h_n = h\chi_{(-n,n)}$ lies in $L^1(\mathbb{R})$, so 4.10 supplies a null set N_n off which $\frac{1}{2t}\int_{b-t}^{b+t} |h_n - h_n(b)| \to 0$; if $b \notin N = \bigcup_n N_n$ (null by 2.8) and $n > |b|$, then $h = h_n$ on $(b-t, b+t)$ for $0 < t < n - |b|$, and $|\frac{1}{2t}\int_{b-t}^{b+t} h - h(b)| \le \frac{1}{2t}\int_{b-t}^{b+t} |h - h(b)|$ by 3.23.
For $q \in \mathbb{Q}$ put $h_q = (f-q)^2$, Lebesgue measurable by 2.46. Since $(\alpha-\beta)^2 \le 2\alpha^2 + 2\beta^2$ and $f^2 \in L^1(\mathbb{R})$,

$$\int_{-n}^{n} |h_q| \le 2\int_{-\infty}^{\infty} f^2 + 4nq^2 < \infty \qquad (n \in \mathbb{Z}^+),$$

so localization gives a null set N_q with $\frac{1}{2t}\int_{b-t}^{b+t} h_q \to (f(b)-q)^2$ for $b \notin N_q$. Let $N = \bigcup_{q\in\mathbb{Q}} N_q$, null because $\mathbb{Q}$ is countable.
Fix $b \notin N$ and $\varepsilon > 0$, and choose $q \in \mathbb{Q}$ with $|f(b) - q| < \varepsilon$. Then $|f(x) - f(b)|^2 \le 2(f(x)-q)^2 + 2(q - f(b))^2$ for every x, so averaging over $(b-t, b+t)$ and letting $t \downarrow 0$ gives

$$\limsup_{t\downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} |f - f(b)|^2 \le 4\big(f(b) - q\big)^2 < 4\varepsilon^2.$$

As $\varepsilon > 0$ was arbitrary and the averages are nonnegative, the limit is 0.

Problem 4B.4. Prove that the Lebesgue Differentiation Theorem (4.19) still holds if the hypothesis that $\int_{-\infty}^{\infty} |f| < \infty$ is weakened to the requirement that $\int_{-\infty}^{x} |f| < \infty$ for all $x \in \mathbb{R}$.

Solution. Truncate on the right and use that differentiability at b is a local property. The hypothesis makes $g(x) = \int_{-\infty}^{x} f$ real valued, and for $n \in \mathbb{Z}^+$ the function $f_n = f\chi_{(-\infty,n)}$ satisfies $\int_{-\infty}^{\infty} |f_n| = \int_{-\infty}^{n} |f| < \infty$, so $f_n \in L^1(\mathbb{R})$. Hence 4.19 applied to f_n gives a set N_n with $|N_n| = 0$ such that $g_n'(b) = f_n(b)$ for $b \notin N_n$, where $g_n(x) = \int_{-\infty}^{x} f_n$.
For $x < n$ we have $f = f_n$ on $(-\infty, x)$, so $g_n = g$ on $(-\infty, n)$. Thus if $b < n$, the difference quotients of g and of g_n at b coincide for all $t \ne 0$ with $b + t < n$, so 4.16 gives $g'(b) = g_n'(b)$ whenever the latter exists; and $f_n(b) = f(b)$. Therefore

$$b \in (-\infty, n) \setminus N_n \implies g'(b) = f_n(b) = f(b).$$

Now $N = \bigcup_{n=1}^{\infty} N_n$ is null by 2.8, and every $b \notin N$ satisfies $b < n$ for some $n \in \mathbb{Z}^+$, so $g'(b) = f(b)$ for almost every $b \in \mathbb{R}$.

Problem 4B.5. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a Lebesgue measurable function. Prove that

$$|f(b)| \le f^*(b)$$

for almost every $b \in \mathbb{R}$.

Solution. Split $\mathbb{R}$ according to local integrability of $|f|$: let

$$\mathcal{F} = \Big\{(p,q) : p, q \in \mathbb{Q},\ p < q,\ \int_p^q |f| < \infty\Big\}, \qquad L = \bigcup_{(p,q)\in\mathcal{F}} (p,q),$$

so that $\mathcal{F}$ is countable.
(i) $b \notin L$. Then $\int_{b-t}^{b+t} |f| = \infty$ for every $t > 0$: were it finite, rationals p, q with $b-t < p < b < q < b+t$ would give $(p,q) \in \mathcal{F}$ and hence $b \in L$. So $f^*(b) = \infty > |f(b)|$, with no exceptional set.
(ii) $b \in L$. Fix $(p,q) \in \mathcal{F}$ and put $h = |f|\chi_{(p,q)}$, which lies in $L^1(\mathbb{R})$ since $\int h = \int_p^q |f| < \infty$. By 4.10 and 3.23 there is a null set $N_{p,q}$ off which $\frac{1}{2t}\int_{b-t}^{b+t} h \to h(b)$. If $b \in (p,q) \setminus N_{p,q}$ and $0 < t < \min\{b-p, q-b\}$, then $(b-t, b+t) \subseteq (p,q)$, so $h = |f|$ there and $h(b) = |f(b)|$; since every such average is at most $f^*(b)$ by 4.6,

$$|f(b)| = \lim_{t \downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} |f| \le f^*(b).$$

Since $\mathcal{F}$ is countable, $N = \bigcup_{(p,q)\in\mathcal{F}} N_{p,q}$ is null by 2.8, and every $b \notin N$ falls under (i) or (ii).

Problem 4B.6. Prove that if $h \in L^1(\mathbb{R})$ and $\int_{-\infty}^{s} h = 0$ for all $s \in \mathbb{R}$, then $h(s) = 0$ for almost every $s \in \mathbb{R}$.

Solution. The hypothesis says exactly that $g(x) = \int_{-\infty}^{x} h$ is the zero function, so $g'(b) = 0$ for every $b \in \mathbb{R}$. The Lebesgue Differentiation Theorem in its second version (4.19), applicable because $h \in L^1(\mathbb{R})$, gives $g'(b) = h(b)$ for almost every b. Hence $h(s) = 0$ for almost every $s \in \mathbb{R}$.

Problem 4B.7. Give an example of a Borel subset of $\mathbb{R}$ whose density at 0 is not defined.

Solution. Take

$$E = \bigcup_{k=0}^{\infty}\Big(\big[-4^{-2k},\, -4^{-2k-1}\big] \cup \big[4^{-2k-1},\, 4^{-2k}\big]\Big),$$

a Borel set symmetric about 0, for which the ratio $\frac{|E\cap(-t,t)|}{2t}$ oscillates between $\frac{4}{5}$ and $\frac{1}{5}$ as $t \downarrow 0$. Its intervals in $(0,\infty)$ are the disjoint $[4^{-2j-1}, 4^{-2j}]$, $j \ge 0$, of length $\frac{3}{4}16^{-j}$, and by symmetry the ratio equals $\frac{|E\cap(0,t)|}{t}$.
(i) $t_m = 4^{-2m}$. The intervals meeting $(0, t_m)$ are exactly those with $j \ge m$, the one with $j = m$ having right endpoint t_m; so countable additivity gives

$$|E \cap (0, t_m)| = \sum_{j \ge m} \tfrac{3}{4}16^{-j} = \tfrac{3}{4} \cdot \tfrac{16}{15} \cdot 16^{-m} = \tfrac{4}{5}\, t_m.$$

(ii) $s_m = 4^{-2m-1}$. This is the left endpoint of the interval with $j = m$, so only the intervals with $j \ge m+1$ meet $(0, s_m)$, and they lie inside it; by (i) with $m+1$ in place of m,

$$|E \cap (0, s_m)| = \tfrac{4}{5}\, 4^{-2m-2} = \tfrac{1}{5}\, s_m.$$

Both $t_m \downarrow 0$ and $s_m \downarrow 0$, yet the ratios are constantly $\frac{4}{5}$ and $\frac{1}{5}$, so the limit in 4.22 does not exist.

Problem 4B.8. Give an example of a Borel subset of $\mathbb{R}$ whose density at 0 is $\frac{1}{3}$.

Solution. Take $E = F \cup (-F)$, where $F = \bigcup_{k=1}^{\infty} J_k$ and

$$J_k = \Big[\tfrac{1}{k+1},\ \tfrac{1}{k+1} + \tfrac{1}{3k(k+1)}\Big)$$

is the left third of $\big[\frac{1}{k+1}, \frac{1}{k}\big)$, since $\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}$. These blocks sit on the harmonic scale rather than a geometric one, which is what makes the limit exist. As a countable union of intervals together with its reflection, E is Borel, and $F \subseteq (0,1)$, $-F \subseteq (-1,0)$.
Because Lebesgue measure is invariant under $x \mapsto -x$ (reflection carries a cover by open intervals to a cover of the same total length, as in 2.7), and $F \cap (0,t)$, $(-F) \cap (-t, 0)$ are disjoint Borel sets with

union $E \cap (-t, t)$,

$$\frac{|E \cap (-t,t)|}{2t} = \frac{|F \cap (0,t)|}{t},$$

so it suffices to show the right side tends to $\frac{1}{3}$. Since $J_k \subseteq \left[\frac{1}{k+1}, \frac{1}{k}\right)$, we have $F \cap (0, \frac{1}{n}) = \bigcup_{k \geq n} J_k$, a disjoint union, so 2.68 and 2.14 give

$$\left|F \cap (0, \tfrac{1}{n})\right| = \frac{1}{3} \sum_{k=n}^{\infty} \Big(\frac{1}{k} - \frac{1}{k+1}\Big) = \frac{1}{3n}$$

by telescoping. For $t \in (0, 1]$ put $n = \lfloor 1/t \rfloor$, so $\frac{1}{n+1} < t \leq \frac{1}{n}$; then $(0, \frac{1}{n+1}) \subseteq (0, t) \subseteq (0, \frac{1}{n}]$ and monotonicity of outer measure give

$$\frac{n}{3(n+1)} \leq \frac{|F \cap (0,t)|}{t} \leq \frac{n+1}{3n}.$$

As $t \downarrow 0$ we have $n \to \infty$ and both bounds tend to $\frac{1}{3}$, so the density of E at 0 is $\frac{1}{3}$.

Problem 4B.9. Prove that if $t \in [0, 1]$, then there exists a Borel set $E \subseteq \mathbb{R}$ such that the density of E at 0 is t.

Solution. Write the given number as $c \in [0, 1]$, to free the letter t for 4.22, and run the construction of Exercise 8 in this section with c in place of $\frac{1}{3}$: let $E = F \cup (-F)$, where $F = \bigcup_{k=1}^{\infty} J_k$ and

$$J_k = \Big[\tfrac{1}{k+1},\ \tfrac{1}{k+1} + \tfrac{c}{k(k+1)}\Big)$$

is the left-hand fraction c of $\left[\frac{1}{k+1}, \frac{1}{k}\right)$, of length $\frac{c}{k(k+1)}$ (empty if $c = 0$, the whole block if $c = 1$; both are allowed below). As before E is Borel, the blocks are disjoint, $F \subseteq (0, 1)$, $-F \subseteq (-1, 0)$, and reflection invariance of Lebesgue measure gives

$$\frac{|E \cap (-s,s)|}{2s} = \frac{|F \cap (0,s)|}{s}.$$

Exactly as in Exercise 8, $F \cap (0, \frac{1}{n}) = \bigcup_{k \geq n} J_k$ with

$$\left|F \cap (0, \tfrac{1}{n})\right| = c \sum_{k=n}^{\infty} \Big(\frac{1}{k} - \frac{1}{k+1}\Big) = \frac{c}{n}$$

by 2.68, 2.14 and telescoping, so taking $n = \lfloor 1/s \rfloor$ for $s \in (0, 1]$, whence $\frac{1}{n+1} < s \leq \frac{1}{n}$, monotonicity of outer measure gives

$$\frac{cn}{n+1} \leq \frac{|F \cap (0,s)|}{s} \leq \frac{c(n+1)}{n}.$$

Both bounds tend to c as $s \downarrow 0$, so the density of E at 0 equals c.

Problem 4B.10. Suppose E is a Lebesgue measurable subset of $\mathbb{R}$ such that the density of E equals 1 at every element of E and equals 0 at every element of $\mathbb{R} \setminus E$. Prove that $E = \varnothing$ or $E = \mathbb{R}$.

Solution. Build from E an everywhere-differentiable function with derivative exactly χ_E; Darboux's argument then forbids that derivative from taking both the values 1 and 0. Define

$$g(x) = |E \cap [0, x]| \quad \text{for } x \geq 0, \qquad g(x) = -|E \cap [x, 0]| \quad \text{for } x < 0.$$

Each $E \cap [u, v]$ is Lebesgue measurable with $|E \cap [u, v]| \leq v - u$ by 2.72 and 2.14, so g is real valued, and additivity of the measure supplied by 2.72 (endpoints being null) gives

$$g(v) - g(u) = |E \cap [u, v]| \in [0,\ v - u] \qquad (u < v),$$

so g is nondecreasing and 1-Lipschitz, hence continuous.
$g' = \chi_E$ at every point. Fix b and $t > 0$; the last display gives $\frac{g(b+t)-g(b)}{t} = \frac{|E\cap[b,b+t]|}{t}$ and $\frac{g(b)-g(b-t)}{t} = \frac{|E\cap[b-t,b]|}{t}$.
(i) $b \in E$, so E has density 1 at b. Since $[b, b+t] \subseteq [b-t, b+t]$, additivity and monotonicity give

$$0 \le 1 - \frac{|E \cap [b, b+t]|}{t} = \frac{|[b, b+t] \setminus E|}{t} \le \frac{|[b-t, b+t] \setminus E|}{t} = 2\Big(1 - \frac{|E \cap (b-t, b+t)|}{2t}\Big),$$

which tends to 0; the same bound holds with $[b-t, b]$ in place of $[b, b+t]$. Hence $g'(b) = 1$.
(ii) $b \notin E$, so E has density 0 at b. Then

$$0 \le \frac{|E \cap [b, b+t]|}{t} \le 2 \cdot \frac{|E \cap (b-t, b+t)|}{2t} \longrightarrow 0,$$

and likewise for $[b-t, b]$, so $g'(b) = 0$.
Now suppose $p \in E$ and $q \in \mathbb{R} \setminus E$. Replacing E by $-E$ if necessary, we may assume $p < q$: reflection invariance of Lebesgue measure makes the density of $-E$ at $-b$ equal to that of E at b, so $-E$ satisfies the same hypotheses and is empty (all of $\mathbb{R}$) exactly when E is. Put $\varphi(x) = g(x) - \frac{x}{2}$, continuous and everywhere differentiable with $\varphi' = \chi_E - \frac{1}{2}$, so $\varphi'(p) = \frac{1}{2} > 0$ and $\varphi'(q) = -\frac{1}{2} < 0$. The maximum of φ over the compact interval $[p, q]$ is attained at some c, not at p (as $\varphi'(p) > 0$ gives $\varphi(p+t) > \varphi(p)$ for some small $t > 0$) and not at q (as $\varphi'(q) < 0$ gives $\varphi(q+t) > \varphi(q)$ for some small $t < 0$); so $c \in (p, q)$, where the one-sided difference quotients force $\varphi'(c) \le 0$ and $\varphi'(c) \ge 0$. Then $g'(c) = \frac{1}{2}$, contradicting $g'(c) = \chi_E(c) \in \{0, 1\}$. Hence no such pair p, q exists, that is, $E = \varnothing$ or $E = \mathbb{R}$.

5 Product Measures

Contents

5.1 Exercises 5A

Problem 5A.1. Suppose $(X,\mathcal{S})$ and $(Y,\mathcal{T})$ are measurable spaces. Prove that if A is a nonempty subset of X and B is a nonempty subset of Y such that $A\times B\in\mathcal{S}\otimes\mathcal{T}$, then $A\in\mathcal{S}$ and $B\in\mathcal{T}$.

Solution. $A\in\mathcal{S}$ and $B\in\mathcal{T}$, because each is a cross section of $A\times B$.
Because A and B are nonempty, pick $a\in A$ and $b\in B$. As $A\times B\in\mathcal{S}\otimes\mathcal{T}$, 5.6 gives

$$[A\times B]^b\in\mathcal{S}\qquad\text{and}\qquad[A\times B]_a\in\mathcal{T}.$$

By Example 5.5, $[A\times B]^b=A$ (since $b\in B$) and $[A\times B]_a=B$ (since $a\in A$). Hence $A\in\mathcal{S}$ and $B\in\mathcal{T}$.

Problem 5A.2. Suppose $(X,\mathcal{S})$ is a measurable space. Prove that if $E\in\mathcal{S}\otimes\mathcal{S}$, then

$$\{x\in X:(x,x)\in E\}\in\mathcal{S}.$$

Solution. Write $\Delta(E)=\{x\in X:(x,x)\in E\}$ and let

$$\mathcal{E}=\{E\subseteq X\times X:\Delta(E)\in\mathcal{S}\};$$

the assertion is that $\mathcal{S}\otimes\mathcal{S}\subseteq\mathcal{E}$.
Every measurable rectangle lies in $\mathcal{E}$, since for $A,B\in\mathcal{S}$,

$$\Delta(A\times B)=A\cap B\in\mathcal{S}.$$

Moreover $\Delta(\varnothing)=\varnothing$ and

$$\Delta\big((X\times X)\setminus E\big)=X\setminus\Delta(E),$$

$$\Delta\Big(\bigcup_{k=1}^\infty E_k\Big)=\bigcup_{k=1}^\infty\Delta(E_k)$$

(Check!), so $\mathcal{E}$ is a σ-algebra on $X\times X$ because $\mathcal{S}$ is one on X. As $\mathcal{S}\otimes\mathcal{S}$ is the smallest σ-algebra containing the measurable rectangles, $\mathcal{S}\otimes\mathcal{S}\subseteq\mathcal{E}$.

Problem 5A.3. Let $\mathcal{B}$ denote the σ-algebra of Borel subsets of $\mathbb{R}$. Show that there exists a set $E\subseteq\mathbb{R}\times\mathbb{R}$ such that $[E]_a\in\mathcal{B}$ and $[E]^a\in\mathcal{B}$ for every $a\in\mathbb{R}$, but $E\notin\mathcal{B}\otimes\mathcal{B}$.

Solution. Take $E=\{(x,x):x\in A\}$, where $A\subseteq\mathbb{R}$ is a set that is not Borel (2.67).
For $a\in\mathbb{R}$, the relation $(a,y)\in E$ forces $y=a\in A$, and $(x,a)\in E$ forces $x=a\in A$; hence

$$[E]_a=[E]^a=\begin{cases}\{a\} & \text{if } a\in A,\\ \varnothing & \text{if } a\notin A,\end{cases}$$

which is closed and therefore in $\mathcal{B}$. But $\{x\in\mathbb{R}:(x,x)\in E\}=A\notin\mathcal{B}$, so $E\notin\mathcal{B}\otimes\mathcal{B}$ by Exercise 5A.2 (applied with $\mathcal{S}=\mathcal{B}$).

Problem 5A.4. Suppose $(X,\mathcal{S})$ and $(Y,\mathcal{T})$ are measurable spaces. Prove that if $f:X\to\mathbb{R}$ is $\mathcal{S}$-measurable and $g:Y\to\mathbb{R}$ is $\mathcal{T}$-measurable and $h:X\times Y\to\mathbb{R}$ is defined by $h(x,y)=f(x)\,g(y)$, then h is $(\mathcal{S}\otimes\mathcal{T})$-measurable.

Solution. $h=FG$, where $F(x,y)=f(x)$ and $G(x,y)=g(y)$, and both factors are $(\mathcal{S}\otimes\mathcal{T})$-measurable. Indeed, for every Borel set $D\subseteq\mathbb{R}$,

$$F^{-1}(D)=f^{-1}(D)\times Y,\qquad G^{-1}(D)=X\times g^{-1}(D),$$

which are measurable rectangles in $\mathcal{S} \otimes \mathcal{T}$ (5.2), because $f^{-1}(D) \in \mathcal{S}$ and $g^{-1}(D) \in \mathcal{T}$ by the measurability of f and of g. Hence FG is $(\mathcal{S} \otimes \mathcal{T})$-measurable by 2.46(a), and $(FG)(x, y) = f(x)g(y) = h(x, y)$.

Problem 5A.5. Verify the assertion in Example 5.11 that the collection of finite unions of intervals of $\mathbb{R}$ is closed under complementation.

Solution. A subset of $\mathbb{R}$ is an interval exactly when it is order convex (meaning that $x, z \in I$ and $x < y < z$ imply $y \in I$), and this makes the complement of an interval a union of two intervals; De Morgan's Laws then finish the proof.

Order convexity characterizes intervals. Each form (α, β), $[\alpha, \beta)$, $(\alpha, \beta]$, $[\alpha, \beta]$ (with $\alpha = -\infty$ or $\beta = \infty$ allowed at an open end), together with singletons and $\varnothing$, is cut out by inequalities that an intermediate point inherits. Conversely, let $I \neq \varnothing$ be order convex, $\alpha = \inf I$, $\beta = \sup I$. Then

$$(\alpha, \beta) \subseteq I \subseteq [\alpha, \beta] \cap \mathbb{R} :$$

the right inclusion is the definition of inf and sup, and if $\alpha < y < \beta$ then y is neither a lower nor an upper bound of I, so $x < y < z$ for some $x, z \in I$, whence $y \in I$. So I is (α, β) with possibly its finite endpoints adjoined, hence an interval (a single point when $\alpha = \beta$); and $\varnothing = (0, 0)$ is an interval.

Let $\mathcal{A}$ be the collection of finite unions of intervals.

(i) $\mathcal{A}$ is closed under finite intersections. The intersection of two order convex sets is order convex, hence an interval, and

$$\Big(\bigcup_{j=1}^{n} I_j\Big) \cap \Big(\bigcup_{k=1}^{m} J_k\Big) = \bigcup_{j=1}^{n} \bigcup_{k=1}^{m} (I_j \cap J_k) \in \mathcal{A};$$

induction extends this to finitely many elements of $\mathcal{A}$.

(ii) The complement of an interval I is a union of two intervals. Put

$$L = \{x \in \mathbb{R} : x < y \text{ for every } y \in I\},$$

$$U = \{x \in \mathbb{R} : x > y \text{ for every } y \in I\}.$$

Then $\mathbb{R} \setminus I = L \cup U$: a point of $I \cap L$ would satisfy $x < x$, and similarly for U, while if x lies in none of I, L, U, then some $y, z \in I$ satisfy $y < x < z$, forcing $x \in I$. Also L is closed downward and U closed upward, so both are order convex, hence intervals.

Now let $E = I_1 \cup \cdots \cup I_n \in \mathcal{A}$. By De Morgan's Laws,

$$\mathbb{R} \setminus E = \bigcap_{j=1}^{n} (\mathbb{R} \setminus I_j),$$

an intersection of finitely many elements of $\mathcal{A}$ by (ii), hence in $\mathcal{A}$ by (i).

Problem 5A.6. Verify the assertion in Example 5.12 that the collection of countable unions of intervals of $\mathbb{R}$ is not closed under complementation.

Solution. $\mathbb{Q}$ is a countable union of intervals but $\mathbb{R} \setminus \mathbb{Q}$ is not.

Let $\mathcal{A}$ be the collection of countable unions of intervals. Then $\mathbb{Q} = \bigcup_{q \in \mathbb{Q}} \{q\} \in \mathcal{A}$, each $\{q\}$ counting as an interval in Example 5.11.

An interval $I \subseteq \mathbb{R} \setminus \mathbb{Q}$ has at most one point: if $x, z \in I$ with $x < z$, density of $\mathbb{Q}$ gives $q \in \mathbb{Q}$ with $x < q < z$, and every interval contains the points lying between two of its points, so $q \in I \cap \mathbb{Q}$, a contradiction. So if $\mathbb{R} \setminus \mathbb{Q} = \bigcup_{k=1}^{\infty} I_k$ with each I_k an interval, then $\mathbb{R} \setminus \mathbb{Q}$ is a countable union of sets with at most one element, hence countable, making $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$ countable. This contradiction gives $\mathbb{R} \setminus \mathbb{Q} \notin \mathcal{A}$, so $\mathcal{A}$ is not closed under complementation.

Problem 5A.7. Suppose $\mathcal{A}$ is a nonempty collection of subsets of a set W. Show that $\mathcal{A}$ is an algebra on W if and only if $\mathcal{A}$ is closed under finite intersections and under complementation.

Solution. Both directions are De Morgan's Laws.
Suppose $\mathcal{A}$ is an algebra on W. Closure under complementation is the second condition of 5.10, and for $E, F \in \mathcal{A}$,

$$E \cap F = W \setminus \big((W \setminus E) \cup (W \setminus F)\big) \in \mathcal{A}$$

by closure under complementation and under unions of two elements; induction gives $E_1 \cap \cdots \cap E_n \in \mathcal{A}$.
Conversely, suppose $\mathcal{A}$ is nonempty and closed under finite intersections and under complementation. Choosing any $E \in \mathcal{A}$ (this is the only use of nonemptiness) gives

$$\varnothing = E \cap (W \setminus E) \in \mathcal{A},$$

and for $E, F \in \mathcal{A}$,

$$E \cup F = W \setminus \big((W \setminus E) \cap (W \setminus F)\big) \in \mathcal{A}.$$

The three conditions of 5.10 hold, so $\mathcal{A}$ is an algebra on W.

Problem 5A.8. Suppose μ is a measure on a measurable space $(X, \mathcal{S})$. Prove that the following are equivalent.
(a) The measure μ is σ-finite.
(b) There exists an increasing sequence $X_1 \subseteq X_2 \subseteq \cdots$ of sets in $\mathcal{S}$ such that $X = \bigcup_{k=1}^{\infty} X_k$ and $\mu(X_k) < \infty$ for every $k \in \mathbf{Z}^+$.
(c) There exists a disjoint sequence $X_1, X_2, X_3, \ldots$ of sets in $\mathcal{S}$ such that $X = \bigcup_{k=1}^{\infty} X_k$ and $\mu(X_k) < \infty$ for every $k \in \mathbf{Z}^+$.

Solution. (a) $\Rightarrow$ (b). By 5.18 there are $E_1, E_2, \ldots \in \mathcal{S}$ with $X = \bigcup_{k=1}^{\infty} E_k$ and $\mu(E_k) < \infty$. Put $X_k = E_1 \cup \cdots \cup E_k \in \mathcal{S}$. These increase, satisfy $\bigcup_{k=1}^{\infty} X_k = \bigcup_{k=1}^{\infty} E_k = X$, and by finite subadditivity (2.58 applied to $E_1, \ldots, E_k, \varnothing, \varnothing, \ldots$),

$$\mu(X_k) \leq \mu(E_1) + \cdots + \mu(E_k) < \infty.$$

(b) $\Rightarrow$ (c). Put $Y_1 = X_1$ and $Y_k = X_k \setminus X_{k-1}$ for $k \geq 2$, all in $\mathcal{S}$. If $j < k$ then $Y_j \subseteq X_j \subseteq X_{k-1}$ while $Y_k \cap X_{k-1} = \varnothing$, so the sequence is disjoint. Each $x \in X$ lies in Y_m for the smallest m with $x \in X_m$, so $X = \bigcup_{k=1}^{\infty} Y_k$. Finally $Y_k \subseteq X_k$ gives $\mu(Y_k) \leq \mu(X_k) < \infty$ by 2.57(a).
(c) $\Rightarrow$ (a). Such a sequence is in particular a sequence in $\mathcal{S}$ with union X and with every term of finite measure, which is the definition 5.18 of σ-finite.

Problem 5A.9. Suppose μ and ν are σ-finite measures. Prove that $\mu \times \nu$ is a σ-finite measure.

Solution. The sets $A_j \times B_k$ form a countable cover of $X \times Y$ by sets of finite $\mu \times \nu$ measure.
Say μ is a measure on $(X, \mathcal{S})$ and ν on $(Y, \mathcal{T})$; since both are σ-finite, $\mu \times \nu$ is defined on $\mathcal{S} \otimes \mathcal{T}$ by 5.25 and is a measure by 5.27. Choose $A_j \in \mathcal{S}$ with $X = \bigcup_{j=1}^{\infty} A_j$ and $\mu(A_j) < \infty$, and $B_k \in \mathcal{T}$ with $Y = \bigcup_{k=1}^{\infty} B_k$ and $\nu(B_k) < \infty$. Each $A_j \times B_k$ is a measurable rectangle (5.2), the index set $\mathbf{Z}^+ \times \mathbf{Z}^+$ is countable,

$$X \times Y = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} (A_j \times B_k),$$

and by 5.26,

$$(\mu \times \nu)(A_j \times B_k) = \mu(A_j)\, \nu(B_k) < \infty.$$

Thus $\mu \times \nu$ is σ-finite by 5.18.

Problem 5A.10. Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces. Prove that if ω is a measure on $\mathcal{S} \otimes \mathcal{T}$ such that $\omega(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{S}$ and all $B \in \mathcal{T}$, then $\omega = \mu \times \nu$.
[The exercise above means that $\mu \times \nu$ is the unique measure on $\mathcal{S} \otimes \mathcal{T}$ that behaves as we expect on measurable rectangles.]

Solution. The two measures agree on the algebra of finite unions of measurable rectangles, and the Monotone Class Theorem (5.17) propagates this to $\mathcal{S} \otimes \mathcal{T}$.
By Exercise 5A.8 choose increasing sequences $X_n \in \mathcal{S}$ and $Y_n \in \mathcal{T}$ with $X = \bigcup_n X_n$, $Y = \bigcup_n Y_n$, $\mu(X_n) < \infty$, $\nu(Y_n) < \infty$, and set $W_n = X_n \times Y_n$. Then $W_1 \subseteq W_2 \subseteq \cdots$ and $\bigcup_n W_n = X \times Y$ (for $(x, y) \in X_j \times Y_k$ take $n = \max\{j, k\}$), while the hypothesis on ω and 5.26 give

$$\omega(W_n) = \mu(X_n)\,\nu(Y_n) = (\mu \times \nu)(W_n) < \infty$$

(here $\mu \times \nu$ is defined by 5.25 and is a measure by 5.27, both spaces being σ-finite). Fix n and set

$$\mathcal{M}_n = \{E \in \mathcal{S} \otimes \mathcal{T} : \omega(E \cap W_n) = (\mu \times \nu)(E \cap W_n)\}.$$

(i) $\mathcal{M}_n$ contains the collection $\mathcal{A}$ of finite unions of measurable rectangles, an algebra by 5.13(a). For $A \in \mathcal{S}$ and $B \in \mathcal{T}$,

$$(A \times B) \cap W_n = (A \cap X_n) \times (B \cap Y_n),$$

so the hypothesis on ω and 5.26 give $A \times B \in \mathcal{M}_n$; a general $E \in \mathcal{A}$ is a disjoint union $E_1 \cup \cdots \cup E_m$ of measurable rectangles by 5.13(b), and finite additivity of both measures then gives $E \in \mathcal{M}_n$.
(ii) $\mathcal{M}_n$ is a monotone class. For $E_k \uparrow E$ in $\mathcal{M}_n$, apply 2.59 to $E_k \cap W_n \uparrow E \cap W_n$ for each measure. For $E_k \downarrow E$, we have $\omega(E_1 \cap W_n) \le \omega(W_n) < \infty$ and likewise for $\mu \times \nu$ by 2.57(a), so 2.60 applies to both and

$$\omega(E \cap W_n) = \lim_{k\to\infty} \omega(E_k \cap W_n) = \lim_{k\to\infty} (\mu \times \nu)(E_k \cap W_n) = (\mu \times \nu)(E \cap W_n).$$

By 5.17, $\mathcal{M}_n$ contains the smallest σ-algebra containing $\mathcal{A}$; that σ-algebra is $\mathcal{S} \otimes \mathcal{T}$, since $\mathcal{A}$ contains every measurable rectangle (5.2) and $\mathcal{A} \subseteq \mathcal{S} \otimes \mathcal{T}$. As $\mathcal{M}_n \subseteq \mathcal{S} \otimes \mathcal{T}$, we get $\mathcal{M}_n = \mathcal{S} \otimes \mathcal{T}$. Finally, for $E \in \mathcal{S} \otimes \mathcal{T}$ we have $E \cap W_n \uparrow E$, so 2.59 gives

$$\omega(E) = \lim_{n\to\infty} \omega(E \cap W_n) = \lim_{n\to\infty} (\mu \times \nu)(E \cap W_n) = (\mu \times \nu)(E).$$

Hence $\omega = \mu \times \nu$.

5.2 Exercises 5B

Problem 5B.1. (a) Let λ denote Lebesgue measure on $[0, 1]$. Show that

$$\int_{[0,1]} \int_{[0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2}\, d\lambda(y)\, d\lambda(x) = \frac{\pi}{4}$$

and

$$\int_{[0,1]} \int_{[0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2}\, d\lambda(x)\, d\lambda(y) = -\frac{\pi}{4}.$$

(b) Explain why (a) violates neither Tonelli's Theorem nor Fubini's Theorem.

Solution. (a) Everything comes from the pair of antiderivatives

$$\frac{\partial}{\partial y}\Big(\frac{y}{x^2 + y^2}\Big) = \frac{\partial}{\partial x}\Big(\frac{-x}{x^2 + y^2}\Big) = \frac{x^2 - y^2}{(x^2 + y^2)^2},$$

valid wherever $(x, y) \ne (0, 0)$. Write $f(x, y) = (x^2 - y^2)/(x^2 + y^2)^2$ off the origin and $f(0, 0) = 0$; note $f(y, x) = -f(x, y)$.

For $x \in (0,1]$ the map $y \mapsto f(x,y)$ is continuous on $[0,1]$ (the denominator is at least $x^4 > 0$), so 3.34 and the Fundamental Theorem of Calculus give

$$\int_{[0,1]} f(x,y)\, d\lambda(y) = \left[\frac{y}{x^2+y^2}\right]_{y=0}^{y=1} = \frac{1}{1+x^2}.$$

For $x = 0$ we have $f(0,\cdot)^+ = 0$ and $f(0,\cdot)^- = y^{-2}\chi_{(0,1]}$, whose integral is at least $\int_{[1/k,1]} y^{-2}\, dy = k-1$ for every $k \in \mathbb{Z}^+$ (3.34), hence is ∞; so by 3.18 the inner integral is $0 - \infty = -\infty$. Thus the inner integral is the function g with $g(0) = -\infty$ and $g(x) = 1/(1+x^2)$ for $x > 0$. Since $g^- = \infty \cdot \chi_{\{0\}}$ has integral $\infty \cdot \lambda(\{0\}) = 0$, and g^+ differs from the continuous function $x \mapsto 1/(1+x^2)$ only on the null set $\{0\}$, 3.18 and 3.34 give

$$\int_{[0,1]} \int_{[0,1]} f(x,y)\, d\lambda(y)\, d\lambda(x) = \int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4}.$$

Interchanging the two variables (which by $f(y,x) = -f(x,y)$ changes the sign) gives

$$\int_{[0,1]} \int_{[0,1]} f(x,y)\, d\lambda(x)\, d\lambda(y) = -\frac{\pi}{4}.$$

(b) Tonelli's Theorem (5.28) requires a nonnegative integrand, and f takes both signs; its other hypothesis holds, since λ on $[0,1]$ is finite, hence σ-finite. Fubini's Theorem (5.32) requires $f \in \mathcal{L}^1(\lambda \times \lambda)$, and this fails. Here f is $\mathcal{B} \otimes \mathcal{B}$-measurable: the coordinate maps are $\mathcal{B} \otimes \mathcal{B}$-measurable because their level sets are measurable rectangles (2.39), so each

$$f_k(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2 + \frac{1}{k}}$$

is measurable by 2.46, and $f_k \to f$ pointwise, so 2.48 applies. Applying 5.28 to $|f|$ and using $f(x,y) \geq 0$ for $0 \leq y \leq x$,

$$\int_{[0,1]} |f(x,y)|\, d\lambda(y) \geq \left[\frac{y}{x^2+y^2}\right]_{y=0}^{y=x} = \frac{1}{2x},$$

whence for every $k \in \mathbb{Z}^+$

$$\int_{[0,1]^2} |f|\, d(\lambda \times \lambda) \geq \frac{1}{2} \int_{[1/k,1]} \frac{dx}{x} = \frac{\ln k}{2},$$

so $\int_{[0,1]^2} |f|\, d(\lambda \times \lambda) = \infty$.

Problem 5B.2. (a) Give an example of a doubly indexed collection $\{x_{m,n} : m, n \in \mathbb{Z}^+\}$ of real numbers such that

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{m,n} = 0 \qquad \text{and} \qquad \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} x_{m,n} = \infty.$$

(b) Explain why (a) violates neither Tonelli's Theorem nor Fubini's Theorem.

Solution. (a) Take

$$x_{m,n} = \begin{cases} 1 & \text{if } n = m, \\ -1 & \text{if } n = 2m, \\ 0 & \text{otherwise.} \end{cases}$$

Row m has exactly the two nonzero terms $x_{m,m} = 1$ and $x_{m,2m} = -1$, distinct because $m \geq 1$, so every row sums to 0 and

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{m,n} = 0.$$

Column n receives $+1$ from $m = n$ and, when n is even, -1 from $m = n/2$ (a distinct index); so the column sums are 1 for n odd and 0 for n even, and the partial sums of $\sum_n \sum_m x_{m,n}$ count the odd integers up to n, giving

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} x_{m,n} = \infty.$$

(b) Let $\mu = \nu$ be counting measure on all subsets of $\mathbb{Z}^+$, which is σ-finite since $\mathbb{Z}^+ = \bigcup_{k=1}^{\infty}\{1,\dots,k\}$, and let $f(m,n) = x_{m,n}$. Every subset of the countable set $\mathbb{Z}^+ \times \mathbb{Z}^+$ is a countable union of the measurable rectangles $\{m\} \times \{n\}$, so f is $\mathcal{S} \otimes \mathcal{T}$-measurable; integration against counting measure is summation (3.6, and 3.42 in the real-valued case), so the two sums in (a) are the two iterated integrals of 5.28 and 5.32 (each inner function has at most two nonzero values, hence lies in $\mathcal{L}^1$).
Tonelli's Theorem (5.28), and likewise its corollary 5.31 on double sums, requires a nonnegative integrand, and f takes the value -1; the σ-finiteness hypothesis does hold. Fubini's Theorem (5.32) requires $f \in \mathcal{L}^1(\mu \times \nu)$, which fails: applying 5.28 to $|f|$, whose rows each have two entries of absolute value 1,

$$\int_{\mathbb{Z}^+ \times \mathbb{Z}^+} |f| \, d(\mu \times \nu) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |x_{m,n}| = \sum_{m=1}^{\infty} 2 = \infty.$$

Problem 5B.3. Suppose $(X, \mathcal{S})$ is a measurable space and $f : X \to [0,\infty]$ is a function. Let $\mathcal{B}$ denote the σ-algebra of Borel subsets of $(0,\infty)$. Prove that $U_f \in \mathcal{S} \otimes \mathcal{B}$ if and only if f is an $\mathcal{S}$-measurable function.

Solution. Both implications come from a rectangle decomposition of $U_f = \{(x,t) : 0 < t < f(x)\}$ and from 5.6.
Suppose f is $\mathcal{S}$-measurable. Then

$$U_f = \bigcup_{r \in \mathbb{Q} \cap (0,\infty)} f^{-1}\big((r,\infty]\big) \times (0,r] :$$

for $\supseteq$, such a point satisfies $0 < t \le r < f(x)$; for $\subseteq$, density of $\mathbb{Q}$ supplies a rational r with $t < r < f(x)$ (any rational $r > t$ if $f(x) = \infty$), and then $r > 0$. Each $f^{-1}((r,\infty])$ is in $\mathcal{S}$ by 2.51, since $(r,\infty]$ is Borel (2.50), and $(0,r] \in \mathcal{B}$, so each term is a measurable rectangle; the union is countable, so $U_f \in \mathcal{S} \otimes \mathcal{B}$.
Conversely, suppose $U_f \in \mathcal{S} \otimes \mathcal{B}$. By 5.6 each cross section $[U_f]^t$ lies in $\mathcal{S}$, and for $t \in (0,\infty)$ the condition $(x,t) \in U_f$ says exactly $f(x) > t$, so

$$f^{-1}\big((t,\infty]\big) = \{x \in X : f(x) > t\} = [U_f]^t \in \mathcal{S}.$$

Now let $a \in \mathbb{R}$. (i) If $a > 0$, then $f^{-1}((a,\infty]) = [U_f]^a \in \mathcal{S}$. (ii) If $a < 0$, then $f^{-1}((a,\infty]) = X \in \mathcal{S}$, as f takes values in $[0,\infty]$. (iii) If $a = 0$, then $f(x) > 0$ exactly when $f(x) > 1/k$ for some k, so

$$f^{-1}\big((0,\infty]\big) = \bigcup_{k=1}^{\infty} [U_f]^{1/k} \in \mathcal{S}.$$

Hence f is $\mathcal{S}$-measurable by 2.52.

Problem 5B.4. Suppose $(X, \mathcal{S})$ is a measurable space and $f : X \to \mathbb{R}$ is a function. Let $\operatorname{graph}(f) \subseteq X \times \mathbb{R}$ denote the graph of f:

$$\operatorname{graph}(f) = \Big\{\big(x, f(x)\big) : x \in X\Big\}.$$

Let $\mathcal{B}$ denote the σ-algebra of Borel subsets of $\mathbb{R}$. Prove that $\operatorname{graph}(f) \in \mathcal{S} \otimes \mathcal{B}$ if f is an $\mathcal{S}$-measurable function.

Solution. $\operatorname{graph}(f) = (g-h)^{-1}(\{0\})$, where $g(x,y) = f(x)$ and $h(x,y) = y$.
For every Borel set $B \subseteq \mathbb{R}$,

$$g^{-1}(B) = f^{-1}(B) \times \mathbb{R} \qquad \text{and} \qquad h^{-1}(B) = X \times B,$$

which are measurable rectangles in $\mathcal{S} \otimes \mathcal{B}$ because $f^{-1}(B) \in \mathcal{S}$ by the measurability of f. So g and h are $\mathcal{S} \otimes \mathcal{B}$-measurable, hence so is $g - h$ by 2.46(a), and $\{0\}$ is a Borel subset of $\mathbb{R}$. Thus

$$\operatorname{graph}(f) = \{(x,y) : f(x) - y = 0\} = (g-h)^{-1}(\{0\}) \in \mathcal{S} \otimes \mathcal{B}.$$

Method (2): build the graph from rectangles directly,

$$\operatorname{graph}(f) = \bigcap_{k=1}^{\infty} \bigcup_{j \in \mathbb{Z}} \Big(f^{-1}\big([\tfrac{j}{k}, \tfrac{j+1}{k})\big) \times [\tfrac{j}{k}, \tfrac{j+1}{k}) \Big).$$

Indeed if $y = f(x)$, then for each k the integer j with $j \le kf(x) < j+1$ puts both $f(x)$ and y in $[\frac{j}{k}, \frac{j+1}{k})$; conversely a point of the right side has $f(x)$ and y in a common interval of length $1/k$ for every k, so $|y - f(x)| < 1/k$ for all k and $y = f(x)$. Each term is a measurable rectangle, and a σ-algebra is closed under countable unions and countable intersections.

5.3 Exercises 5C

Problem 5C.1. Show that a set $G \subseteq \mathbb{R}^n$ is open in $\mathbb{R}^n$ if and only if for each $(b_1, \dots, b_n) \in G$, there exists $r > 0$ such that

$$\{(a_1, \dots, a_n) \in \mathbb{R}^n : \sqrt{(a_1 - b_1)^2 + \cdots + (a_n - b_n)^2} < r\} \subseteq G.$$

Solution. The two conditions are equivalent because $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty$.
Write $\|x\|_2 = (x_1^2 + \cdots + x_n^2)^{1/2}$ and $D(b,r) = \{a \in \mathbb{R}^n : \|a - b\|_2 < r\}$ for the set displayed in the exercise, and recall that G is open means every $x \in G$ has an open cube $B(x,\delta) = \{y : \|y - x\|_\infty < \delta\}$ inside G. The first inequality holds because $\|x\|_\infty^2 = x_j^2 \le \|x\|_2^2$ for an index j attaining the maximum; the second because $\|x\|_2^2 \le n\|x\|_\infty^2$. They say exactly that

$$D(b,\delta) \subseteq B(b,\delta) \quad \text{and} \quad B(b, r/\sqrt{n}\,) \subseteq D(b,r)$$

for all $b \in \mathbb{R}^n$ and $\delta, r > 0$. So if G is open and $b \in G$, then choosing $\delta > 0$ with $B(b,\delta) \subseteq G$ gives $D(b,\delta) \subseteq G$; and conversely, if $D(b,r) \subseteq G$, then $B(b, r/\sqrt{n}\,) \subseteq G$, so G is open.

Problem 5C.2. Show that there exists a set $E \subseteq \mathbb{R}^2$ (thinking of $\mathbb{R}^2$ as equal to $\mathbb{R} \times \mathbb{R}$) such that the cross sections $[E]_a$ and $[E]^a$ are open subsets of $\mathbb{R}$ for every $a \in \mathbb{R}$, but $E \notin \mathcal{B}_2$.

Solution. Take $E = E_A = \mathbb{R}^2 \setminus \Gamma_A$ for a suitable $A \subseteq (0,\infty)$, where Γ_A is the graph of

$$f_A(x) = \begin{cases} -x & \text{if } |x| \in A, \\ x & \text{if } |x| \notin A; \end{cases}$$

there are $2^{\mathfrak{c}}$ such sets but only $\mathfrak{c}$ Borel sets, so some E_A is not Borel. (Here $\mathfrak{c} = 2^{\aleph_0}$.)
(i) Every E_A has all cross sections open, and $A \mapsto E_A$ is injective. Since $|f_A(x)| = |x|$, we get $f_A(f_A(x)) = x$, so f_A is a bijection of $\mathbb{R}$ equal to its own inverse; hence $(a,y) \in \Gamma_A$ exactly when $y = f_A(a)$, and $(x,a) \in \Gamma_A$ exactly when $x = f_A(a)$, giving

$$[E_A]_a = [E_A]^a = \mathbb{R} \setminus \{f_A(a)\},$$

an open subset of $\mathbb{R}$. Injectivity holds because $f_A(t) = -t$ for $t \in A$ and $f_A(t) = t$ for $t \notin A$, with $-t \ne t$ when $t > 0$, so

$$A = \{t > 0 : (t, -t) \in \mathbb{R}^2 \setminus E_A\}.$$

(ii) $|\mathcal{B}_2| \le \mathfrak{c}$. By the proof of 5.38(a), every open subset of $\mathbb{R}^2$ is the union of a subcollection of the countable collection of open cubes with rational center and rational side length, so the collection Σ_0 of open sets satisfies $|\Sigma_0| \le 2^{\aleph_0} = \mathfrak{c}$. For $0 < \alpha < \omega_1$ (the first uncountable ordinal) let Σ_α be the collection of countable unions $\bigcup_k F_k$ in which each F_k or its complement lies in $\bigcup_{\beta<\alpha} \Sigma_\beta$, and put $\Sigma = \bigcup_{\alpha<\omega_1} \Sigma_\alpha$. Taking all F_k equal shows $\Sigma_\beta \subseteq \Sigma_\alpha$ for $\beta < \alpha$; Σ is closed under complementation, and under countable unions because a countable set of countable ordinals has countable supremum. So Σ is a σ-algebra containing the open sets, whence $\mathcal{B}_2 \subseteq \Sigma$. By transfinite induction $|\Sigma_\alpha| \le \mathfrak{c}$ for every $\alpha < \omega_1$: the sets available at stage α number at most $\aleph_0 \cdot \mathfrak{c} = \mathfrak{c}$, and each element of Σ_α is the union of a sequence of them, so

$$|\Sigma_\alpha| \le \mathfrak{c}^{\aleph_0} = \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0} = \mathfrak{c}.$$

Hence $|\mathcal{B}_2| \le |\Sigma| \le \aleph_1 \cdot \mathfrak{c} = \mathfrak{c}$, using $\aleph_1 \le \mathfrak{c}$.
By Cantor's Theorem $2^{\mathfrak{c}} > \mathfrak{c} \ge |\mathcal{B}_2|$, so by (i) the sets E_A cannot all be Borel; choose A with $E = E_A \notin \mathcal{B}_2$.

Problem 5C.3. Suppose $(X, \mathcal{S})$, $(Y, \mathcal{T})$, and $(Z, \mathcal{U})$ are measurable spaces. We can define $\mathcal{S} \otimes \mathcal{T} \otimes \mathcal{U}$ to be the smallest σ-algebra on $X \times Y \times Z$ that contains

$$\{A \times B \times C : A \in \mathcal{S},\ B \in \mathcal{T},\ C \in \mathcal{U}\}.$$

Prove that if we make the obvious identifications of the products $(X \times Y) \times Z$ and $X \times (Y \times Z)$ with $X \times Y \times Z$, then

$$\mathcal{S} \otimes \mathcal{T} \otimes \mathcal{U} = (\mathcal{S} \otimes \mathcal{T}) \otimes \mathcal{U} = \mathcal{S} \otimes (\mathcal{T} \otimes \mathcal{U}).$$

Solution. Both equalities follow from the fact that the generators of each side lie in the other, the reverse inclusion being proved by freezing one factor.
Identify $((x, y), z)$ and $(x, (y, z))$ with (x, y, z), so that $(A \times B) \times C$, $A \times (B \times C)$, and $A \times B \times C$ are the same set, and write $\mathcal{P} = \mathcal{S} \otimes \mathcal{T} \otimes \mathcal{U}$.
$\mathcal{P} \subseteq (\mathcal{S} \otimes \mathcal{T}) \otimes \mathcal{U}$: for $A \in \mathcal{S}$, $B \in \mathcal{T}$, $C \in \mathcal{U}$, the set $A \times B$ lies in $\mathcal{S} \otimes \mathcal{T}$, so $A \times B \times C = (A \times B) \times C$ is a measurable rectangle in $(\mathcal{S} \otimes \mathcal{T}) \otimes \mathcal{U}$; thus that σ-algebra contains all generators of $\mathcal{P}$, which is the smallest one that does.
For the reverse inclusion, fix $C \in \mathcal{U}$ and set

$$\mathcal{E} = \{D \subseteq X \times Y : D \times C \in \mathcal{P}\}.$$

Then $X \times Y \in \mathcal{E}$ (as $X \times Y \times C$ is a generator of $\mathcal{P}$); $\mathcal{E}$ is closed under countable unions because $(\bigcup_k D_k) \times C = \bigcup_k (D_k \times C)$; and $\mathcal{E}$ is closed under complementation because

$$\big((X \times Y) \setminus D\big) \times C = \big((X \times Y \times Z) \setminus (D \times C)\big) \cap \big(X \times Y \times C\big),$$

both sides consisting of the points (x, y, z) with $z \in C$ and $(x, y) \notin D$. So $\mathcal{E}$ is a σ-algebra on $X \times Y$ containing every measurable rectangle $A \times B$, whence $\mathcal{S} \otimes \mathcal{T} \subseteq \mathcal{E}$; that is, $D \times C \in \mathcal{P}$ for all $D \in \mathcal{S} \otimes \mathcal{T}$ and $C \in \mathcal{U}$. These sets generate $(\mathcal{S} \otimes \mathcal{T}) \otimes \mathcal{U}$, so $(\mathcal{S} \otimes \mathcal{T}) \otimes \mathcal{U} \subseteq \mathcal{P}$.
The proof of $\mathcal{P} = \mathcal{S} \otimes (\mathcal{T} \otimes \mathcal{U})$ is symmetric: fix $A \in \mathcal{S}$ and let $\mathcal{F} = \{F \subseteq Y \times Z : A \times F \in \mathcal{P}\}$, which is a σ-algebra by the same three verifications (for complementation use $A \times ((Y \times Z) \setminus F) = ((X \times Y \times Z) \setminus (A \times F)) \cap (A \times Y \times Z)$) and contains every $B \times C$; hence $\mathcal{T} \otimes \mathcal{U} \subseteq \mathcal{F}$, so the generators $A \times F$ of $\mathcal{S} \otimes (\mathcal{T} \otimes \mathcal{U})$ all lie in $\mathcal{P}$.

Problem 5C.4. Show that Lebesgue measure on $\mathbb{R}^n$ is translation invariant. More precisely, show that if $E \in \mathcal{B}_n$ and $a \in \mathbb{R}^n$, then $a + E \in \mathcal{B}_n$ and $\lambda_n(a + E) = \lambda_n(E)$, where

$$a + E = \{a + x : x \in E\}.$$

Solution. Fix $a \in \mathbb{R}^n$.
(i) $a + E \in \mathcal{B}_n$ for every $E \in \mathcal{B}_n$. The collection $\{E \in \mathcal{B}_n : a + E \in \mathcal{B}_n\}$ contains every open set, because $a+B(x,\delta) = B(a+x,\delta)$, and it is closed under complementation and countable unions because $x \mapsto a + x$ is a bijection of $\mathbb{R}^n$:

$$a + (\mathbb{R}^n \setminus E) = \mathbb{R}^n \setminus (a + E), \qquad a + \bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} (a + E_k).$$

So it is a σ-algebra containing the open sets, hence equals $\mathcal{B}_n$.
(ii) Translation invariance of λ_m forces translation invariance of the integral: if it holds in dimension m and $h : \mathbb{R}^m \to [0,\infty]$ is $\mathcal{B}_m$-measurable and $b \in \mathbb{R}^m$, then $x \mapsto h(x-b)$ is $\mathcal{B}_m$-measurable (its inverse images are $b + h^{-1}(B)$, measurable by (i)) and

$$\int_{\mathbb{R}^m} h(x-b)\, d\lambda_m(x) = \int_{\mathbb{R}^m} h\, d\lambda_m.$$

For $h = \chi_F$ this is $\lambda_m(b+F) = \lambda_m(F)$, since $\chi_F(x-b) = \chi_{b+F}(x)$; for nonnegative simple h it follows by 3.16 and 3.20; for general h take simple $s_k \uparrow h$ (2.89), note $s_k(\cdot - b) \uparrow h(\cdot - b)$, and apply the Monotone Convergence Theorem (3.11) twice.
(iii) Induction on n. For $n = 1$, λ_1 is the restriction of outer measure to $\mathcal{B}_1$, and outer measure is translation invariant by 2.7. Let $n > 1$ and assume the result in dimension $n-1$. View $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$, so $\mathcal{B}_n = \mathcal{B}_{n-1} \otimes \mathcal{B}_1$ by 5.39 and $\lambda_n = \lambda_{n-1} \times \lambda_1$; each λ_k is σ-finite, since $\mathbb{R}^k$ is the union of the cubes C_m centered at the origin with side m and $\lambda_k(C_m) = m^k < \infty$, so 5.20 and 5.25 apply. Write $a = (a', a_n)$. Since $(x,y) \in a + E$ exactly when $(x - a', y - a_n) \in E$,

$$[a+E]_x = a_n + [E]_{x-a'},$$

so by the case $n = 1$ we get $\lambda_1([a+E]_x) = g(x-a')$, where $g(x) = \lambda_1([E]_x)$ is $\mathcal{B}_{n-1}$-measurable by 5.20(a). Hence, by 5.25 and then (ii) with $m = n-1$ and $b = a'$,

$$\begin{aligned}
\lambda_n(a+E) &= \int_{\mathbb{R}^{n-1}} \lambda_1\big([a+E]_x\big)\, d\lambda_{n-1}(x) \\
&= \int_{\mathbb{R}^{n-1}} g(x-a')\, d\lambda_{n-1}(x) \\
&= \int_{\mathbb{R}^{n-1}} g\, d\lambda_{n-1} = \lambda_n(E).
\end{aligned}$$

Problem 5C.5. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{B}_n$-measurable and $t \in \mathbb{R} \setminus \{0\}$. Define $f_t : \mathbb{R}^n \to \mathbb{R}$ by $f_t(x) = f(tx)$.
(a) Prove that f_t is $\mathcal{B}_n$-measurable.
(b) Prove that if $\int_{\mathbb{R}^n} f\, d\lambda_n$ is defined, then

$$\int_{\mathbb{R}^n} f_t\, d\lambda_n = \frac{1}{|t|^n} \int_{\mathbb{R}^n} f\, d\lambda_n.$$

Solution. (a) $f_t^{-1}(B) = \frac{1}{t} f^{-1}(B) \in \mathcal{B}_n$ for every Borel $B \subseteq \mathbb{R}$, so f_t is $\mathcal{B}_n$-measurable by 2.35.
Indeed $f(tx) \in B$ exactly when $tx \in f^{-1}(B)$, that is, when $x \in \frac{1}{t} f^{-1}(B)$; and $f^{-1}(B) \in \mathcal{B}_n$ by the measurability of f, so the following lemma (applied to the nonzero number $1/t$) puts $\frac{1}{t} f^{-1}(B)$ in $\mathcal{B}_n$. The same argument applies to $[0,\infty]$-valued measurable functions.
Lemma. If $t \in \mathbb{R} \setminus \{0\}$ and $E \in \mathcal{B}_n$, then $tE \in \mathcal{B}_n$ and $\lambda_n(tE) = |t|^n \lambda_n(E)$.
For $t > 0$ this is 5.41, and $tE = -(|t|E)$ for $t < 0$, so only $x \mapsto -x$ needs treatment. The collection $\{F \in \mathcal{B}_n : -F \in \mathcal{B}_n\}$ contains the open sets ($-B(x,\delta) = B(-x,\delta)$) and is a σ-algebra ($x \mapsto -x$ is a bijection), hence equals $\mathcal{B}_n$. That $\lambda_n(-F) = \lambda_n(F)$ goes by induction on n. For $n = 1$, negating a cover of F by open intervals gives a cover of $-F$ by open intervals of the same lengths, so $|-F| = |F|$.

For $n > 1$, view $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$, so $\mathcal{B}_n = \mathcal{B}_{n-1} \otimes \mathcal{B}_1$ (5.39) and $\lambda_n = \lambda_{n-1} \times \lambda_1$, each λ_k being σ-finite ($\mathbb{R}^k$ is a union of cubes of finite measure), so 5.20 and 5.25 apply. The induction hypothesis gives

$$\int_{\mathbb{R}^{n-1}} h(-x)\, d\lambda_{n-1}(x) = \int_{\mathbb{R}^{n-1}} h\, d\lambda_{n-1}$$

for $\mathcal{B}_{n-1}$-measurable $h : \mathbb{R}^{n-1} \to [0, \infty]$, via $\chi_F(-x) = \chi_{-F}(x)$, then 3.16 and 3.20 for simple functions, then 2.89 and 3.11 in general. Since $[-F]_x = -[F]_{-x}$, the function $g(x) = \lambda_1([F]_x)$ ($\mathcal{B}_{n-1}$-measurable by 5.20(a)) satisfies $\lambda_1([-F]_x) = g(-x)$, so by 5.25

$$\lambda_n(-F) = \int_{\mathbb{R}^{n-1}} g(-x)\, d\lambda_{n-1}(x) = \int_{\mathbb{R}^{n-1}} g\, d\lambda_{n-1} = \lambda_n(F).$$

Finally, for $t < 0$, $\lambda_n(tE) = \lambda_n(|t|E) = |t|^n \lambda_n(E)$ by 5.41.
(b) Write $h_t(x) = h(tx)$. For $E \in \mathcal{B}_n$ we have $(\chi_E)_t = \chi_{(1/t)E}$, so the Lemma gives

$$\int_{\mathbb{R}^n} (\chi_E)_t\, d\lambda_n = \lambda_n\Big(\tfrac{1}{t}E\Big) = \frac{1}{|t|^n} \int_{\mathbb{R}^n} \chi_E\, d\lambda_n.$$

This extends to nonnegative simple functions by 3.16 and 3.20, and then to every $\mathcal{B}_n$-measurable $h : \mathbb{R}^n \to [0, \infty]$ by taking simple $s_k \uparrow h$ (2.89), so that $(s_k)_t \uparrow h_t$, and applying 3.11 twice. Since $(f_t)^\pm = (f^\pm)_t$,

$$\int_{\mathbb{R}^n} (f_t)^\pm\, d\lambda_n = \frac{1}{|t|^n} \int_{\mathbb{R}^n} f^\pm\, d\lambda_n.$$

Because $0 < |t|^{-n} < \infty$ and $\int_{\mathbb{R}^n} f\, d\lambda_n$ is defined, at least one of these is finite, so $\int_{\mathbb{R}^n} f_t\, d\lambda_n$ is defined (3.18) and subtracting gives

$$\int_{\mathbb{R}^n} f_t\, d\lambda_n = \frac{1}{|t|^n} \int_{\mathbb{R}^n} f\, d\lambda_n.$$

Problem 5C.6. Suppose λ denotes Lebesgue measure on $(\mathbb{R}, \mathcal{L})$, where $\mathcal{L}$ is the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}$. Show that there exist subsets E and F of $\mathbb{R}^2$ such that

- $F \in \mathcal{L} \otimes \mathcal{L}$ and $(\lambda \times \lambda)(F) = 0$;
- $E \subseteq F$ but $E \notin \mathcal{L} \otimes \mathcal{L}$.

[The measure space $(\mathbb{R}, \mathcal{L}, \lambda)$ has the property that every subset of a measurable set with measure 0 is measurable. This exercise asks you to show that the measure space $(\mathbb{R}^2, \mathcal{L} \otimes \mathcal{L}, \lambda \times \lambda)$ does not have this property.]

Solution. Take $F = \{0\} \times \mathbb{R}$ and $E = \{0\} \times V$, where $V \subseteq \mathbb{R}$ is not Lebesgue measurable.
Such a V exists: 2.18 gives disjoint $A, B \subseteq \mathbb{R}$ with $|A \cup B| \neq |A| + |B|$, while outer measure is a measure on $(\mathbb{R}, \mathcal{L})$ by 2.72(b) and so is additive on disjoint sets in $\mathcal{L}$; hence A and B are not both in $\mathcal{L}$. Also λ is σ-finite, since $\mathbb{R} = \bigcup_{k=1}^\infty (-k, k)$ with $\lambda((-k, k)) = 2k < \infty$, so $\lambda \times \lambda$ is defined by 5.25 and is a measure by 5.27.
Clearly $E \subseteq F$. Since $\{0\}$ and $\mathbb{R}$ lie in $\mathcal{L}$, the set F is a measurable rectangle, so $F \in \mathcal{L} \otimes \mathcal{L}$ and, by 5.26,

$$(\lambda \times \lambda)(F) = \lambda(\{0\})\, \lambda(\mathbb{R}) = 0 \cdot \infty = 0.$$

But $E \notin \mathcal{L} \otimes \mathcal{L}$: otherwise 5.6 would give $[E]_0 = V \in \mathcal{L}$, contradicting the choice of V.

Problem 5C.7. Suppose $m \in \mathbb{Z}^+$. Verify that the collection of sets $\mathcal{E}_m$ that appears in the proof of 5.41 is a monotone class.

Solution. $\mathcal{E}_m$ is closed under countable increasing unions and countable decreasing intersections, which by 5.15 is what must be checked.
In the proof of 5.41, $t > 0$ is fixed, C_m is the open cube in $\mathbb{R}^n$ centered at the origin with side length

m, and

$$\mathcal{E}_m = \{E \in \mathcal{B}_n : E \subseteq C_m \text{ and } \lambda_n(tE) = t^n \lambda_n(E)\},$$

the condition making sense because $tE \in \mathcal{B}_n$ (first paragraph of that proof). Since $x \mapsto tx$ is a bijection of $\mathbb{R}^n$, it carries unions to unions and intersections to intersections and preserves inclusions.
(i) Increasing unions. Let $E_k \uparrow E$ with $E_k \in \mathcal{E}_m$. Then $E \in \mathcal{B}_n$, $E \subseteq C_m$, and $tE_k \uparrow tE$, so two applications of 2.59 give

$$\lambda_n(tE) = \lim_{k\to\infty} \lambda_n(tE_k) = t^n \lim_{k\to\infty} \lambda_n(E_k) = t^n \lambda_n(E).$$

(ii) Decreasing intersections. Let $E_k \downarrow E$ with $E_k \in \mathcal{E}_m$. Then $E \in \mathcal{B}_n$, $E \subseteq C_m$, and $tE_k \downarrow tE$. The finiteness needed for 2.60 holds for both sequences, since by 2.57(a)

$$\lambda_n(E_1) \le \lambda_n(C_m) = m^n < \infty, \qquad \lambda_n(tE_1) \le \lambda_n(tC_m) = (tm)^n < \infty$$

(here $\lambda_n(C_m) = m^n$ by induction on n from $\lambda_n = \lambda_{n-1} \times \lambda_1$ and 5.26, the case $n = 1$ being the length of an interval; and tC_m is the cube of side tm). So 2.60 gives the same chain of equalities as in (i).
In both cases $E \in \mathcal{E}_m$, so $\mathcal{E}_m$ is a monotone class.

Problem 5C.8. Show that the open unit ball in $\mathbb{R}^n$ is an open subset of $\mathbb{R}^n$.

Solution. Given x in the open unit ball B_n, the open cube $B(x, \delta)$ with $\delta = (1 - \|x\|_2)/\sqrt{n} > 0$ is contained in B_n, where $\|y\|_2 = ({y_1}^2 + \cdots + {y_n}^2)^{1/2}$.
Indeed $\|y\|_2 \le \sqrt{n}\,\|y\|_\infty$, because $|y_j| \le \|y\|_\infty$ for each j; so $\|y - x\|_\infty < \delta$ gives

$$\|y - x\|_2 \le \sqrt{n}\,\|y - x\|_\infty < \sqrt{n}\,\delta = 1 - \|x\|_2.$$

The triangle inequality for $\|\cdot\|_2$, which follows from the Cauchy–Schwarz inequality since

$$\|u + v\|_2{}^2 = \|u\|_2{}^2 + 2\sum_{j=1}^{n} u_j v_j + \|v\|_2{}^2 \le (\|u\|_2 + \|v\|_2)^2,$$

then gives $\|y\|_2 \le \|x\|_2 + \|y - x\|_2 < 1$, so $y \in B_n$. Hence B_n is open.

Problem 5C.9. Suppose G_1 is a nonempty subset of $\mathbb{R}^m$ and G_2 is a nonempty subset of $\mathbb{R}^n$. Prove that $G_1 \times G_2$ is an open subset of $\mathbb{R}^m \times \mathbb{R}^n$ if and only if G_1 is an open subset of $\mathbb{R}^m$ and G_2 is an open subset of $\mathbb{R}^n$.
[One direction of this result was already proved (see 5.36); both directions are stated here to make the result look prettier and to be comparable to the next exercise, where neither direction has been proved.]

Solution. If G_1 and G_2 are open, then $G_1 \times G_2$ is open by 5.36 (nonemptiness is not needed here).
Conversely, suppose $G_1 \times G_2$ is open in $\mathbb{R}^{m+n}$, and recall $B(x, \delta) \times B(y, \delta) = B((x, y), \delta)$, which holds because $\|(u, v)\|_\infty = \max\{\|u\|_\infty, \|v\|_\infty\}$. Fix $y_0 \in G_2$, possible since $G_2 \ne \varnothing$. Given $x \in G_1$, openness supplies $\delta > 0$ with $B((x, y_0), \delta) \subseteq G_1 \times G_2$; then for every $u \in B(x, \delta)$,

$$(u, y_0) \in B(x, \delta) \times B(y_0, \delta) = B\big((x, y_0), \delta\big) \subseteq G_1 \times G_2,$$

so $u \in G_1$. Thus $B(x, \delta) \subseteq G_1$ and G_1 is open. The argument for G_2 is symmetric, using a fixed $x_0 \in G_1$.

Problem 5C.10. Suppose F_1 is a nonempty subset of $\mathbb{R}^m$ and F_2 is a nonempty subset of $\mathbb{R}^n$. Prove that $F_1 \times F_2$ is a closed subset of $\mathbb{R}^m \times \mathbb{R}^n$ if and only if F_1 is a closed subset of $\mathbb{R}^m$ and F_2 is a closed subset of $\mathbb{R}^n$.

Solution. If F_1 and F_2 are closed, then

$$\mathbb{R}^{m+n} \setminus (F_1 \times F_2) = \big[(\mathbb{R}^m \setminus F_1) \times \mathbb{R}^n\big] \cup \big[\mathbb{R}^m \times (\mathbb{R}^n \setminus F_2)\big],$$

since $(x, y) \notin F_1 \times F_2$ exactly when $x \notin F_1$ or $y \notin F_2$; each bracket is open by 5.36, so the complement of $F_1 \times F_2$ is open and $F_1 \times F_2$ is closed. (Nonemptiness is not needed here.)
Conversely, suppose $F_1 \times F_2$ is closed, and fix $y_0 \in F_2$, possible since $F_2 \neq \varnothing$. Given $x \in \mathbb{R}^m \setminus F_1$, we have $(x, y_0) \notin F_1 \times F_2$, so openness of the complement supplies $\delta > 0$ with

$$B\big((x, y_0), \delta\big) \cap (F_1 \times F_2) = \varnothing.$$

For $u \in B(x, \delta)$, the identity $B(x, \delta) \times B(y_0, \delta) = B((x, y_0), \delta)$ puts (u, y_0) outside $F_1 \times F_2$, and $y_0 \in F_2$ then forces $u \notin F_1$. Hence $B(x, \delta) \subseteq \mathbb{R}^m \setminus F_1$, so that complement is open and F_1 is closed. The argument for F_2 is symmetric, using a fixed $x_0 \in F_1$.

Problem 5C.11. Suppose E is a subset of $\mathbb{R}^m \times \mathbb{R}^n$ and

$$A = \{x \in \mathbb{R}^m : (x, y) \in E \text{ for some } y \in \mathbb{R}^n\}.$$

(a) Prove that if E is an open subset of $\mathbb{R}^m \times \mathbb{R}^n$, then A is an open subset of $\mathbb{R}^m$.
(b) Prove or give a counterexample: If E is a closed subset of $\mathbb{R}^m \times \mathbb{R}^n$, then A is a closed subset of $\mathbb{R}^m$.

Solution. (a) Yes. (b) False: for $m = n = 1$ the hyperbola $E = \{(x, y) \in \mathbb{R}^2 : xy = 1\}$ is closed while $A = \mathbb{R} \setminus \{0\}$ is not.
Throughout we identify $\mathbb{R}^m \times \mathbb{R}^n$ with $\mathbb{R}^{m+n}$ and use $B(x, \delta) \times B(y, \delta) = B((x, y), \delta)$.
(a) Let $x \in A$, so $(x, y) \in E$ for some y. Openness of E supplies $\delta > 0$ with

$$B\big((x, y), \delta\big) = B(x, \delta) \times B(y, \delta) \subseteq E,$$

so every $u \in B(x, \delta)$ has $(u, y) \in E$ and hence lies in A. Thus $B(x, \delta) \subseteq A$, and A is open.
(b) The hyperbola is closed: if $x_0 y_0 \neq 1$, put $\varepsilon = |x_0 y_0 - 1| > 0$ and choose $\delta > 0$ with $\delta \leq 1$ and $\delta(|x_0| + |y_0| + 1) < \varepsilon$; then $\|(x, y) - (x_0, y_0)\|_\infty < \delta$ gives $|y| < |y_0| + 1$ and

$$|xy - x_0 y_0| \leq |y|\,|x - x_0| + |x_0|\,|y - y_0| < \delta(|x_0| + |y_0| + 1) < \varepsilon,$$

so $xy \neq 1$. Every $x \neq 0$ has $(x, 1/x) \in E$, while $0 \cdot y = 0 \neq 1$, so $A = \mathbb{R} \setminus \{0\}$, whose complement $\{0\}$ is not open.
For general m and n, take

$$E = \{(x, y) : x_1 y_1 = 1,\ x_2 = \cdots = x_m = 0,\ y_2 = \cdots = y_n = 0\},$$

a finite intersection of closed sets: $\{(x, y) : x_1 y_1 = 1\}$ is closed by the estimate above (since $\|(x, y) - (x^0, y^0)\|_\infty < \delta$ controls $|x_1 - x_1{}^0|$ and $|y_1 - y_1{}^0|$), and each $\{(x, y) : x_j = 0\}$ is closed because a point with $x_j \neq 0$ has $B((x, y), |x_j|)$ in the complement, similarly for y_k. Then

$$A = \{x \in \mathbb{R}^m : x_1 \neq 0,\ x_2 = \cdots = x_m = 0\},$$

which is not closed: the origin is in $\mathbb{R}^m \setminus A$, yet every $B(0, \delta)$ contains $(\delta/2, 0, \ldots, 0) \in A$.

Problem 5C.12. (a) Prove that $\lim_{n\to\infty} \lambda_n(B_n) = 0$.
(b) Find the value of n that maximizes $\lambda_n(B_n)$.

Solution. (a) The limit is 0; (b) the maximum is at $n = 5$, where $\lambda_5(B_5) = 8\pi^2/15$.
Both parts run on the recursion established in the proof of 5.44, valid for every integer $n > 2$, together with the values 5.44 supplies:

$$\lambda_n(B_n) = \frac{2\pi}{n}\,\lambda_{n-2}(B_{n-2}),$$

$$\lambda_1(B_1) = 2, \quad \lambda_2(B_2) = \pi, \quad \lambda_3(B_3) = \frac{4\pi}{3},$$

$$\lambda_4(B_4) = \frac{\pi^2}{2}, \quad \lambda_5(B_5) = \frac{8\pi^2}{15}, \quad \lambda_6(B_6) = \frac{\pi^3}{6}.$$

(a) Put $r = 2\pi/7$, so $0 < r < 1$ because $\pi < 3.5$. If $n \geq 5$ then $n + 2 \geq 7$, so

$$\lambda_{n+2}(B_{n+2}) = \frac{2\pi}{n+2}\,\lambda_n(B_n) \leq r\,\lambda_n(B_n).$$

Iterating from $n = 5$ and from $n = 6$, and setting $M = \max\{\lambda_5(B_5), \lambda_6(B_6)\}$, we get for all $n \geq 5$

$$0 \leq \lambda_n(B_n) \leq M\,r^{(n-6)/2},$$

since $n = 5 + 2j$ or $n = 6 + 2j$ with $j = \lfloor (n-5)/2 \rfloor \geq (n-6)/2$. As $0 < r < 1$, the squeeze theorem gives $\lim_{n\to\infty} \lambda_n(B_n) = 0$.
(b) For $n \geq 5$ we have $n + 2 \geq 7 > 2\pi$, so the recursion and $\lambda_n(B_n) > 0$ give $\lambda_{n+2}(B_{n+2}) < \lambda_n(B_n)$; each parity chain therefore decreases strictly from its start at $n = 5$ and $n = 6$, and the supremum is the largest of the first six values. Using $3 < \pi < 3.2$, and abbreviating $\lambda_j = \lambda_j(B_j)$:

$$\frac{\lambda_2}{\lambda_1} = \frac{\pi}{2} > 1, \qquad \frac{\lambda_3}{\lambda_2} = \frac{4}{3} > 1, \qquad \frac{\lambda_4}{\lambda_3} = \frac{3\pi}{8} > \frac{9}{8} > 1,$$

$$\frac{\lambda_5}{\lambda_4} = \frac{16}{15} > 1, \qquad \frac{\lambda_5}{\lambda_6} = \frac{48}{15\pi} = \frac{16}{5\pi} > 1 \text{ since } 5\pi < 16.$$

Thus $\lambda_5(B_5) = 8\pi^2/15$ is the unique maximum.

Problem 5C.13. For readers familiar with the gamma function Γ: Prove that

$$\lambda_n(B_n) = \frac{\pi^{n/2}}{\Gamma\big(\frac{n}{2} + 1\big)}$$

for every positive integer n.

Solution. Induct in steps of 2, using the recursion $\lambda_n(B_n) = \frac{2\pi}{n}\lambda_{n-2}(B_{n-2})$ for $n > 2$ from the proof of 5.44 and the gamma identities $\Gamma(x+1) = x\Gamma(x)$, $\Gamma(1) = 1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
The two base cases, $n = 1$ and $n = 2$, start the two parity chains that together reach every positive integer:

$$\frac{\pi^{1/2}}{\Gamma\big(\frac{3}{2}\big)} = \frac{\sqrt{\pi}}{\frac{1}{2}\sqrt{\pi}} = 2 = \lambda_1(B_1), \qquad \frac{\pi}{\Gamma(2)} = \frac{\pi}{1} = \pi = \lambda_2(B_2).$$

For $n > 2$, the recursion, the induction hypothesis for $n - 2$, and $\Gamma(\frac{n}{2} + 1) = \frac{n}{2}\Gamma(\frac{n}{2})$ give

$$\lambda_n(B_n) = \frac{2\pi}{n} \cdot \frac{\pi^{(n-2)/2}}{\Gamma\big(\frac{n}{2}\big)} = \frac{2\,\pi^{n/2}}{n\,\Gamma\big(\frac{n}{2}\big)} = \frac{\pi^{n/2}}{\Gamma\big(\frac{n}{2} + 1\big)},$$

completing the induction.

Problem 5C.14. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

(a) Prove that $D_1(D_2 f)$ and $D_2(D_1 f)$ exist everywhere on $\mathbb{R}^2$.
(b) Show that $\big(D_1(D_2 f)\big)(0,0) \neq \big(D_2(D_1 f)\big)(0,0)$.
(c) Explain why (b) does not violate 5.48.

Solution. (a) On the open set $G = \mathbb{R}^2 \setminus \{(0,0)\}$ the formula $f(x,y) = (x^3y - xy^3)/(x^2+y^2)$ is a quotient of polynomials with nonvanishing denominator, hence infinitely differentiable there; so both mixed partials exist at every point of G, and the quotient rule gives

$$(D_1 f)(x,y) = \frac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}, \qquad (D_2 f)(x,y) = \frac{x^5 - 4x^3y^2 - xy^4}{(x^2+y^2)^2}$$

on G. (Check!) Since $f(t,0) = f(0,t) = 0$ for all t, the difference quotients at the origin give $(D_1 f)(0,0) = (D_2 f)(0,0) = 0$, so $D_1 f$ and $D_2 f$ exist on all of $\mathbb{R}^2$, and the displayed formulas evaluated on the axes yield

$$(D_1 f)(0,y) = -y \quad (y \in \mathbb{R}), \qquad (D_2 f)(x,0) = x \quad (x \in \mathbb{R}),$$

the value 0 at the origin being consistent with both. Hence the two mixed partials also exist at the origin:

$$\big(D_1(D_2 f)\big)(0,0) = \lim_{t\to 0} \frac{(D_2 f)(t,0) - 0}{t} = \lim_{t\to 0} \frac{t}{t} = 1,$$

$$\big(D_2(D_1 f)\big)(0,0) = \lim_{t\to 0} \frac{(D_1 f)(0,t) - 0}{t} = \lim_{t\to 0} \frac{-t}{t} = -1.$$

(b) Those two limits are 1 and -1, so the mixed partials differ at the origin.
(c) Result 5.48 requires the four functions $D_1 f$, $D_2 f$, $D_1(D_2 f)$, $D_2(D_1 f)$ to be continuous on the open set, not merely to exist there. All four are continuous on G, where f is smooth, so 5.48 does apply on G and gives $D_1(D_2 f) = D_2(D_1 f)$ there; if both mixed partials were also continuous at $(0,0)$, letting $(x,y) \to (0,0)$ within G in that identity would force equality of the two values at the origin, contradicting (b). So the continuity hypothesis of 5.48 fails on every open set containing $(0,0)$.

6 Banach Spaces

Contents

6.1 Exercises 6A

Problem 6A.1. Verify that each of the claimed metrics in Example 6.2 is indeed a metric.

Solution. In each of the five examples the first three conditions of 6.1 follow at once from the corresponding property of $|\cdot|$, coordinatewise or termwise. (Check!) What remains is finiteness of d and the triangle inequality.
(i) *Discrete metric on a nonempty set* V. The values lie in $\{0,1\}$. If $f = h$ then $d(f,h) = 0 \leq d(f,g) + d(g,h)$; if $f \neq h$ then g differs from at least one of f, h, so $d(f,g) + d(g,h) \geq 1 = d(f,h)$.
(ii) $\mathbb{R}$ *with* $d(x,y) = |x - y|$.

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

(iii) $\mathbb{R}^n$ *with the max metric.* Finite, as a maximum of n nonnegative reals. For each j,

$$|x_j - z_j| \leq |x_j - y_j| + |y_j - z_j| \leq d(x,y) + d(y,z),$$

and the right side is independent of j, so the maximum of the left side over j obeys the same bound.
(iv) $C([0,1])$ *with the sup metric.* Finite because $|f - g|$ is continuous on the closed bounded interval $[0,1]$, hence bounded. For each $t \in [0,1]$,

$$|f(t) - h(t)| \leq |f(t) - g(t)| + |g(t) - h(t)| \leq d(f,g) + d(g,h),$$

so $d(f,g) + d(g,h)$ is an upper bound for the set whose supremum is $d(f,h)$.
(v) ℓ^1. Finite because $|a_k - b_k| \leq |a_k| + |b_k|$ termwise gives $d(a,b) \leq \sum_k |a_k| + \sum_k |b_k| < \infty$. For every $n \in \mathbb{Z}^+$,

$$\sum_{k=1}^{n} |a_k - c_k| \leq \sum_{k=1}^{n} |a_k - b_k| + \sum_{k=1}^{n} |b_k - c_k| \leq d(a,b) + d(b,c),$$

the last step because a nonnegative series dominates its partial sums; let $n \to \infty$ to get $d(a,c) \leq d(a,b) + d(b,c)$.

Problem 6A.2. Prove that every finite subset of a metric space is closed.

Solution. For $E = \{f_1, \ldots, f_n\}$ finite in a metric space (V,d) and $g \in V \setminus E$, the witness is

$$r = \min\{d(g,f_1), \ldots, d(g,f_n)\},$$

which is positive because $g \neq f_j$ forces $d(g,f_j) > 0$ for each j (second bullet of 6.1, contrapositive, with $d \geq 0$) and the minimum is over finitely many terms. If some $h \in B(g,r)$ lay in E, say $h = f_j$, then $d(g,f_j) = d(g,h) < r \leq d(g,f_j)$, a contradiction; hence $B(g,r) \subseteq V \setminus E$. So $V \setminus E$ is open by 6.4 (for $E = \emptyset$ take $r = 1$), and E is closed by 6.6.

Problem 6A.3. Prove that every closed ball in a metric space is closed.

Solution. For h outside the closed ball $\overline{B}(f,r) = \{g : d(f,g) \leq r\}$ of 6.3, the witness is $s = d(f,h) - r > 0$. If $g \in B(h,s)$, then the triangle inequality and the symmetry of d (6.1) give

$$d(f,g) \geq d(f,h) - d(g,h) > d(f,h) - s = r,$$

so $g \notin \overline{B}(f,r)$. Thus $B(h,s) \subseteq V \setminus \overline{B}(f,r)$, the complement is open by 6.4, and $\overline{B}(f,r)$ is closed by 6.6.

Problem 6A.4. Suppose V is a metric space.
(a) Prove that the union of each collection of open subsets of V is an open subset of V.
(b) Prove that the intersection of each finite collection of open subsets of V is an open subset of V.

Solution. Both parts produce the ball required by 6.4 at an arbitrary point of the set.
(a) Let $G = \bigcup_{A \in \mathcal{A}} A$ with each $A \in \mathcal{A}$ open, and let $f \in G$. Choose $A \in \mathcal{A}$ with $f \in A$; openness of A gives $r > 0$ with

$$B(f, r) \subseteq A \subseteq G.$$

Hence G is open.
(b) Let $G = G_1 \cap \cdots \cap G_n$ with each G_j open, and let $f \in G$. Openness of G_j gives $r_j > 0$ with $B(f, r_j) \subseteq G_j$, and $r = \min\{r_1, \ldots, r_n\} > 0$ because the collection is finite. If $g \in B(f, r)$ and $j \in \{1, \ldots, n\}$, then $d(f, g) < r \leq r_j$, so

$$B(f, r) \subseteq B(f, r_j) \subseteq G_j \quad \text{for every } j,$$

whence $B(f, r) \subseteq G$ and G is open.

Problem 6A.5. Suppose V is a metric space.
(a) Prove that the intersection of each collection of closed subsets of V is a closed subset of V.
(b) Prove that the union of each finite collection of closed subsets of V is a closed subset of V.

Solution. Both parts are Exercise 6A.4 read through complements, using 6.6 (a set is closed exactly when its complement is open) and De Morgan's laws.
(a) For $\mathcal{F}$ a collection of closed subsets of V and $F = \bigcap_{A \in \mathcal{F}} A$,

$$V \setminus F = \bigcup_{A \in \mathcal{F}} (V \setminus A),$$

a union of open sets, hence open by 6A.4(a). So F is closed.
(b) For closed $F_1, \ldots, F_n$ and $F = F_1 \cup \cdots \cup F_n$,

$$V \setminus F = \bigcap_{j=1}^{n} (V \setminus F_j),$$

an intersection of finitely many open sets, hence open by 6A.4(b). So F is closed.

Problem 6A.6. (a) Prove that if V is a metric space, $f \in V$, and $r > 0$, then $\overline{B(f, r)} \subseteq \overline{B}(f, r)$.
(b) Give an example of a metric space V, $f \in V$, and $r > 0$ such that $\overline{B(f, r)} \neq \overline{B}(f, r)$.

Solution. (a) The closed ball $\overline{B}(f, r)$ is a closed set containing $B(f, r)$ (closed by 6A.3, containing it because $d(f, g) < r$ implies $d(f, g) \leq r$), and by 6.9(b) the closure $\overline{B(f, r)}$ is the intersection of all such sets; hence $\overline{B(f, r)} \subseteq \overline{B}(f, r)$.
Method (2): If $g \in \overline{B(f, r)}$ and $\varepsilon > 0$, pick $h \in B(g, \varepsilon) \cap B(f, r)$ (6.7); then the triangle inequality and symmetry (6.1) give

$$d(f, g) \leq d(f, h) + d(h, g) < r + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ gives $d(f, g) \leq r$.
(b) Take $V = \{0, 1\}$ with the discrete metric of 6.2, $f = 0$, $r = 1$. Then $d(0, 1) = 1$ is not less than 1 but is at most 1, so

$$B(0, 1) = \{0\}, \qquad \overline{B}(0, 1) = \{0, 1\}.$$

The finite set $\{0\}$ is closed by 6A.2, hence equals its own closure by 6.9(d), so $\overline{B(0, 1)} = \{0\} \neq \overline{B}(0, 1)$.

Problem 6A.7. Show that each sequence in a metric space has at most one limit.

Solution. Two limits of the same sequence are at distance 0, hence equal. Suppose $f_k \to f$ and $f_k \to g$ in a metric space (V, d), and let $\varepsilon > 0$. By 6.8 both $d(f_k, f) \to 0$ and $d(f_k, g) \to 0$, so some single index n (the larger of the two the definitions supply) satisfies $d(f_n, f) < \varepsilon/2$ and $d(f_n, g) < \varepsilon/2$; the triangle inequality with the symmetry of d (6.1) then gives

$$d(f, g) \le d(f, f_n) + d(f_n, g) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Thus $0 \le d(f, g) < \varepsilon$ for every $\varepsilon > 0$, forcing $d(f, g) = 0$ and hence $f = g$ by the second bullet of 6.1.

Problem 6A.8. Prove 6.9.
[6.9 states: Suppose V is a metric space and $E \subseteq V$. Then
(a) $\overline{E} = \{g \in V : \text{there exist } f_1, f_2, \dots \text{ in } E \text{ such that } \lim_{k\to\infty} f_k = g\}$;
(b) $\overline{E}$ is the intersection of all closed subsets of V that contain E;
(c) $\overline{E}$ is a closed subset of V;
(d) E is closed if and only if $E = \overline{E}$;
(e) E is closed if and only if E contains the limit of every convergent sequence of elements of E.]

Solution. Throughout, $\overline{E} = \{g \in V : B(g, \varepsilon) \cap E \neq \emptyset \text{ for every } \varepsilon > 0\}$ is the closure as defined in 6.7, and $E \subseteq \overline{E}$ because $f \in B(f, \varepsilon) \cap E$ for every $f \in E$ and $\varepsilon > 0$. The parts are proved in the order (a), (c), (b), (d), (e).
(a) If $g \in \overline{E}$, choose $f_k \in B(g, 1/k) \cap E$ for each $k \in \mathbb{Z}^+$; then $d(f_k, g) < 1/k \to 0$, so $f_k \to g$ by 6.8. Conversely, if $f_1, f_2, \dots$ lie in E with $f_k \to g$, then each $\varepsilon > 0$ admits a k with $d(f_k, g) < \varepsilon$, so $f_k \in B(g, \varepsilon) \cap E$ and $g \in \overline{E}$.
(c) Suppose $g \in V \setminus \overline{E}$, and take $\varepsilon > 0$ with $B(g, \varepsilon) \cap E = \emptyset$. For $h \in B(g, \varepsilon)$ put $\delta = \varepsilon - d(g, h) > 0$; the triangle inequality gives

$$d(g, u) \le d(g, h) + d(h, u) < d(g, h) + \delta = \varepsilon \qquad (u \in B(h, \delta)),$$

so $B(h, \delta) \subseteq B(g, \varepsilon)$ misses E and $h \notin \overline{E}$. Hence $B(g, \varepsilon) \subseteq V \setminus \overline{E}$, so $V \setminus \overline{E}$ is open and $\overline{E}$ is closed by 6.6.
(b) First, $(*)$: every closed $F \supseteq E$ contains $\overline{E}$. Indeed, for $g \in V \setminus F$ the openness of $V \setminus F$ gives $\varepsilon > 0$ with $B(g, \varepsilon) \subseteq V \setminus F \subseteq V \setminus E$, so $g \notin \overline{E}$. Thus $\overline{E}$ is contained in the intersection C of all closed sets containing E; and $\overline{E}$ is itself one of those sets, by (c) and $E \subseteq \overline{E}$, so $C \subseteq \overline{E}$. Hence $C = \overline{E}$.
(d) If $E = \overline{E}$, then E is closed by (c). If E is closed, then $(*)$ with $F = E$ gives $\overline{E} \subseteq E \subseteq \overline{E}$.
(e) If E is closed and $f_1, f_2, \dots \in E$ converge to g, then $g \in \overline{E} = E$ by (a) and (d). Conversely, if E contains every such limit, then (a) gives $\overline{E} \subseteq E$, so $E = \overline{E}$ and E is closed by (d).

Problem 6A.9. Prove that each open subset of a metric space V is the union of some sequence of closed subsets of V.

Solution. For G open in V, take the sets of points at distance at least $1/k$ from the complement:

$$F_k = V \setminus \bigcup_{h \in V \setminus G} B\bigl(h, \tfrac{1}{k}\bigr) = \bigl\{f \in V : d(f, h) \ge \tfrac{1}{k} \text{ for all } h \in V \setminus G\bigr\}$$

for $k \in \mathbb{Z}^+$ (so $F_k = V$ when $G = V$, the union being empty).
Each F_k is closed, since each $B(h, 1/k)$ is open by 6.5, their union is open by 6A.4(a), and 6.6 applies. Each $F_k \subseteq G$, since $f \notin G$ gives $f \in B(f, 1/k)$ and hence $f \notin F_k$. Conversely, if $f \in G$, choose $r > 0$ with $B(f, r) \subseteq G$ and then k with $1/k \le r$: were $f \in B(h, 1/k)$ for some $h \in V \setminus G$, then $d(f, h) < 1/k \le r$ would put $h \in B(f, r) \subseteq G$, a contradiction, so $f \in F_k$. Hence $G = \bigcup_{k=1}^{\infty} F_k$.

Problem 6A.10. Prove or give a counterexample: If V is a metric space and U, W are subsets of V, then $\overline{U} \cup \overline{W} = \overline{U \cup W}$.

Solution. The statement is true.
Closure is order preserving: if $E_1 \subseteq E_2$ and $g \in \overline{E_1}$, then $\emptyset \neq B(g,\varepsilon) \cap E_1 \subseteq B(g,\varepsilon) \cap E_2$ for every $\varepsilon > 0$, so $g \in \overline{E_2}$ by 6.7. Applying this to $U \subseteq U \cup W$ and $W \subseteq U \cup W$ gives $\overline{U} \cup \overline{W} \subseteq \overline{U \cup W}$.
For the reverse inclusion, suppose $g \notin \overline{U} \cup \overline{W}$ and pick $\varepsilon_1, \varepsilon_2 > 0$ with $B(g,\varepsilon_1) \cap U = \emptyset$ and $B(g,\varepsilon_2) \cap W = \emptyset$. With $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\} > 0$,

$$B(g,\varepsilon) \cap (U \cup W) = \big(B(g,\varepsilon) \cap U\big) \cup \big(B(g,\varepsilon) \cap W\big) = \emptyset,$$

so $g \notin \overline{U \cup W}$. Hence $\overline{U \cup W} \subseteq \overline{U} \cup \overline{W}$, and the two inclusions give equality.
Method (2) for the reverse inclusion: $\overline{U} \cup \overline{W}$ is closed, being a union of two closed sets (6.9(c) and 6A.5(b)), and it contains $U \cup W$; now apply 6.9(b).

Problem 6A.11. Prove or give a counterexample: If V is a metric space and U, W are subsets of V, then $\overline{U} \cap \overline{W} = \overline{U \cap W}$.

Solution. False: in $V = \mathbb{R}$ with $d(x,y) = |x-y|$, take

$$U = (0,1), \qquad W = (1,2).$$

Here $U \cap W = \emptyset$, and $\overline{\emptyset} = \emptyset$ because no ball meets $\emptyset$ (6.7), so $\overline{U \cap W} = \emptyset$. But $1 \in \overline{U} \cap \overline{W}$: given $\varepsilon > 0$, the points

$$t = \max\left\{\tfrac{1}{2},\, 1 - \tfrac{\varepsilon}{2}\right\} \in U, \qquad s = \min\left\{\tfrac{3}{2},\, 1 + \tfrac{\varepsilon}{2}\right\} \in W$$

both lie in $B(1,\varepsilon)$, since $|1-t| \leq \varepsilon/2 < \varepsilon$ and $|s-1| \leq \varepsilon/2 < \varepsilon$. (Check!) Hence $1 \in \overline{U} \cap \overline{W}$ while $1 \notin \overline{U \cap W}$.

Problem 6A.12. Suppose (U, d_U), (V, d_V), and (W, d_W) are metric spaces. Suppose also that $T : U \to V$ and $S : V \to W$ are continuous functions.
(a) Using the definition of continuity, show that $S \circ T : U \to W$ is continuous.
(b) Using the equivalence of 6.11(a) and 6.11(b), show that $S \circ T : U \to W$ is continuous.
(c) Using the equivalence of 6.11(a) and 6.11(c), show that $S \circ T : U \to W$ is continuous.

Solution. (a) Fix $f \in U$ and $\varepsilon > 0$. Continuity of S at $T(f)$ supplies $\gamma > 0$ with $d_W(S(T(f)), S(v)) < \varepsilon$ whenever $d_V(T(f), v) < \gamma$, and continuity of T at f then supplies (with γ as the tolerance) $\delta > 0$ with $d_V(T(f), T(g)) < \gamma$ whenever $d_U(f,g) < \delta$. Taking $v = T(g)$, we get for all $g \in U$ with $d_U(f,g) < \delta$

$$d_W\big((S \circ T)(f), (S \circ T)(g)\big) = d_W\big(S(T(f)), S(T(g))\big) < \varepsilon,$$

so $S \circ T$ is continuous at f by 6.10, and f was arbitrary.
(b) If $f_k \to f$ in U, then 6.11(a) $\Rightarrow$ (b) applied to the continuous T gives $T(f_k) \to T(f)$ in V, and the same implication applied to the continuous S with the convergent sequence $T(f_1), T(f_2), \ldots$ gives

$$\lim_{k\to\infty} (S \circ T)(f_k) = \lim_{k\to\infty} S\big(T(f_k)\big) = S\big(T(f)\big) = (S \circ T)(f).$$

So $S \circ T$ satisfies 6.11(b), hence is continuous by 6.11(b) $\Rightarrow$ (a).
(c) For every $G \subseteq W$ we have $(S \circ T)^{-1}(G) = T^{-1}(S^{-1}(G))$, since $S(T(f)) \in G$ exactly when $T(f) \in S^{-1}(G)$. If G is open, then 6.11(a) $\Rightarrow$ (c) applied to S makes $S^{-1}(G)$ open in V, and the same implication applied to T with that open set makes

$$(S \circ T)^{-1}(G) = T^{-1}\big(S^{-1}(G)\big)$$

open in U. So $S \circ T$ satisfies 6.11(c), hence is continuous by 6.11(c) $\Rightarrow$ (a).

Problem 6A.13. Prove the parts of 6.11 that were not proved in the text.
[6.11 states: Suppose V and W are metric spaces and $T : V \to W$ is a function. Then the following are equivalent.
(a) T is continuous.
(b) $\lim_{k\to\infty} f_k = f$ in V implies $\lim_{k\to\infty} T(f_k) = T(f)$ in W.
(c) $T^{-1}(G)$ is an open subset of V for every open set $G \subseteq W$.
(d) $T^{-1}(F)$ is a closed subset of V for every closed set $F \subseteq W$.
The text proves that (b) implies (d), and that (c) and (d) are equivalent.]

Solution. Since the text gives (b) $\Rightarrow$ (d) and (c) $\Leftrightarrow$ (d), only (a) $\Rightarrow$ (b) and (c) $\Rightarrow$ (a) are missing; with them the cycle (a) $\Rightarrow$ (b) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (a) makes all four equivalent. Write d_V, d_W for the two metrics.
(a) $\Rightarrow$ (b). Suppose T is continuous, $f_k \to f$ in V, and $\varepsilon > 0$. Continuity at f gives $\delta > 0$ with $d_W(T(f), T(g)) < \varepsilon$ whenever $d_V(f, g) < \delta$ (6.10), and $d_V(f_k, f) \to 0$ (6.8) gives $n \in \mathbb{Z}^+$ with $d_V(f_k, f) < \delta$ for all $k \geq n$. Hence

$$d_W\big(T(f_k), T(f)\big) < \varepsilon \quad \text{for all } k \geq n,$$

so $d_W(T(f_k), T(f)) \to 0$, that is, $T(f_k) \to T(f)$.
(c) $\Rightarrow$ (a). Let $f \in V$ and $\varepsilon > 0$. The ball $B(T(f), \varepsilon)$ is open in W by 6.5, so

$$G = T^{-1}\big(B(T(f), \varepsilon)\big)$$

is open in V by (c), and $f \in G$. By 6.4 there is $\delta > 0$ with $B(f, \delta) \subseteq G$; then $d_V(f, g) < \delta$ forces $T(g) \in B(T(f), \varepsilon)$, that is, $d_W(T(f), T(g)) < \varepsilon$. So T is continuous at each $f \in V$.

Problem 6A.14. Suppose a Cauchy sequence in a metric space has a convergent subsequence. Prove that the Cauchy sequence converges.

Solution. The Cauchy sequence converges to the limit f of its subsequence $f_{k_1}, f_{k_2}, \ldots$, where $k_1 < k_2 < \cdots$; note $k_m \geq m$ for every m (induction: $k_1 \geq 1$, and $k_{m+1} > k_m \geq m$ forces $k_{m+1} \geq m + 1$).
Let $\varepsilon > 0$. Being Cauchy (6.12) gives $n \in \mathbb{Z}^+$ with $d(f_p, f_q) < \varepsilon/2$ for all $p, q \geq n$, and $d(f_{k_m}, f) \to 0$ (6.8) gives M with $d(f_{k_m}, f) < \varepsilon/2$ for all $m \geq M$. Fix an integer $m \geq \max\{M, n\}$ and put $j = k_m$, so that $j \geq m \geq n$ and $d(f_j, f) < \varepsilon/2$. Then for every $k \geq n$,

$$d(f_k, f) \leq d(f_k, f_j) + d(f_j, f) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon,$$

the first term being bounded by the Cauchy estimate since $k, j \geq n$. Hence $d(f_k, f) \to 0$, which by 6.8 says $f_k \to f$.

Problem 6A.15. Verify that all five of the metric spaces in Example 6.2 are complete metric spaces.

Solution. Each space is complete; (ii) is proved from the least-upper-bound property of $\mathbb{R}$ and then drives (iii), (iv), (v).
(i) *Discrete metric.* Taking $\varepsilon = 1$ in 6.12 gives n with $d(f_j, f_k) < 1$, hence $d(f_j, f_k) = 0$ and $f_j = f_k$, for all $j, k \geq n$. The sequence is eventually constant, so $f_k \to f_n \in V$.
(ii) $\mathbb{R}$. A Cauchy sequence $x_1, x_2, \ldots$ is bounded: $\varepsilon = 1$ in 6.12 gives n with $|x_k| \leq |x_n| + 1$ for $k \geq n$, so $|x_k| \leq M$ for all k, where $M = \max\{|x_1|, \ldots, |x_{n-1}|, |x_n| + 1\}$. Hence $s_m = \sup\{x_k : k \geq m\}$ exists for each m, with $s_1 \geq s_2 \geq \cdots \geq -M$, so $L = \inf_m s_m$ exists. Given $\varepsilon > 0$, take n with $|x_p - x_q| < \varepsilon/2$ for $p, q \geq n$. Then $x_k \leq x_n + \varepsilon/2$ for $k \geq n$ gives $L \leq s_n \leq x_n + \varepsilon/2$, while $s_m \geq x_m > x_n - \varepsilon/2$ for $m \geq n$ together with the monotonicity of (s_m) gives $L = \inf_{m\geq n} s_m \geq x_n - \varepsilon/2$; so $|x_n - L| \leq \varepsilon/2$ and

$$|x_j - L| \leq |x_j - x_n| + |x_n - L| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon \qquad (j \geq n).$$

(iii) $\mathbb{R}^n$, *max metric.* Each coordinate sequence is Cauchy, since $|x_i^{(j)} - x_i^{(k)}| \le d(x^{(j)}, x^{(k)})$, so converges to some y_i by (ii). Given $\varepsilon > 0$, let N be the largest of the n thresholds m_i beyond which $|x_i^{(k)} - y_i| < \varepsilon$; then for $k \ge N$,

$$d\big(x^{(k)}, y\big) = \max_{1 \le i \le n} |x_i^{(k)} - y_i| < \varepsilon.$$

(iv) $C([0,1])$, *sup metric.* Since $|f_j(t) - f_k(t)| \le d(f_j, f_k)$, each $f_1(t), f_2(t), \ldots$ is Cauchy in $\mathbb{R}$; put $f(t) = \lim_k f_k(t)$, which exists by (ii). Given $\varepsilon > 0$, take n with $d(f_j, f_k) < \varepsilon/3$ for $j, k \ge n$; letting $j \to \infty$ in $|f_j(t) - f_k(t)| < \varepsilon/3$ gives

$$|f(t) - f_k(t)| \le \tfrac{\varepsilon}{3} \qquad (t \in [0,1],\ k \ge n).$$

Moreover f is continuous: given t, continuity of f_n at t gives $\delta > 0$ with $|f_n(t) - f_n(s)| < \varepsilon/3$ whenever $|s - t| < \delta$, and then

$$|f(t) - f(s)| \le |f(t) - f_n(t)| + |f_n(t) - f_n(s)| + |f_n(s) - f(s)| < \varepsilon.$$

So $f \in C([0,1])$, and the previous display gives $d(f_k, f) \le \varepsilon/3 < \varepsilon$ for $k \ge n$.
(v) ℓ^1. Since $|a_k^{(j)} - a_k^{(m)}| \le d(a^{(j)}, a^{(m)})$, each coordinate sequence is Cauchy and converges by (ii) to some b_k; set $b = (b_1, b_2, \ldots)$. Given $\varepsilon > 0$, take n with $d(a^{(j)}, a^{(m)}) < \varepsilon/2$ for $j, m \ge n$. For $m \ge n$ and $N \in \mathbb{Z}^+$, letting $j \to \infty$ in the finite sum $\sum_{k=1}^N |a_k^{(j)} - a_k^{(m)}| \le \varepsilon/2$ gives $\sum_{k=1}^N |b_k - a_k^{(m)}| \le \varepsilon/2$, and taking the supremum over N gives

$$d\big(a^{(m)}, b\big) = \sum_{k=1}^{\infty} |b_k - a_k^{(m)}| \le \tfrac{\varepsilon}{2} < \varepsilon \qquad (m \ge n).$$

With $m = n$ this puts $b - a^{(n)}$ in ℓ^1, so $b \in \ell^1$ because $a^{(n)} \in \ell^1$; the display then says $a^{(m)} \to b$ in ℓ^1.

Problem 6A.16. Suppose (U, d) is a metric space. Let W denote the set of all Cauchy sequences of elements of U.
(a) For $(f_1, f_2, \ldots)$ and $(g_1, g_2, \ldots)$ in W, define $(f_1, f_2, \ldots) \equiv (g_1, g_2, \ldots)$ to mean that

$$\lim_{k \to \infty} d(f_k, g_k) = 0.$$

Show that $\equiv$ is an equivalence relation on W.
(b) Let V denote the set of equivalence classes of elements of W under the equivalence relation above. For $(f_1, f_2, \ldots) \in W$, let $(f_1, \widehat{f_2, \ldots})$ denote the equivalence class of $(f_1, f_2, \ldots)$. Define $d_V : V \times V \to [0, \infty)$ by

$$d_V\big((f_1, \widehat{f_2, \ldots}), (g_1, \widehat{g_2, \ldots})\big) = \lim_{k \to \infty} d(f_k, g_k).$$

Show that this definition of d_V makes sense and that d_V is a metric on V.
(c) Show that (V, d_V) is a complete metric space.
(d) Show that the map from U to V that takes $f \in U$ to $(f, \widehat{f, f, \ldots})$ preserves distances, meaning that

$$d(f, g) = d_V\big((f, \widehat{f, f, \ldots}), (g, \widehat{g, g, \ldots})\big)$$

for all $f, g \in U$.
(e) Explain why (d) shows that every metric space is a subset of some complete metric space.

Solution. Completeness of $\mathbb{R}$ (6.15, verified in 6A.15) is used throughout.
(a) The three properties hold termwise, from the first, third, and fourth bullets of 6.1: $d(f_k, f_k) = 0$; $d(g_k, f_k) = d(f_k, g_k)$; and

$$0 \le d(f_k, h_k) \le d(f_k, g_k) + d(g_k, h_k) \to 0$$

when $(f_k) \equiv (g_k) \equiv (h_k)$.
(b) Everything follows from the quadrilateral inequality, which two applications of the triangle inequality give for $f, f', g, g' \in U$:

$$\left|d(f,g) - d(f',g')\right| \leq d(f,f') + d(g,g').$$

Existence. Taking $f = f_j$, $f' = f_k$, $g = g_j$, $g' = g_k$ and choosing n beyond which $d(f_j, f_k) < \varepsilon/2$ and $d(g_j, g_k) < \varepsilon/2$ (both sequences being Cauchy) shows $d(f_1, g_1), d(f_2, g_2), \ldots$ is Cauchy in $\mathbb{R}$, hence convergent, with limit in $[0, \infty)$.
Representatives. Taking $f = f_k$, $f' = f'_k$, $g = g_k$, $g' = g'_k$ makes the right side tend to 0 when $(f_k) \equiv (f'_k)$ and $(g_k) \equiv (g'_k)$, so the two limits agree and d_V is well defined.
Metric axioms. Write F, G, H for the classes of $(f_k), (g_k), (h_k)$. Then $d_V(F,F) = \lim_k d(f_k, f_k) = 0$; $d_V(F,G) = 0$ says precisely $(f_k) \equiv (g_k)$, that is, $F = G$; symmetry is termwise; and since all three limits exist, letting $k \to \infty$ in $d(f_k, h_k) \leq d(f_k, g_k) + d(g_k, h_k)$ gives

$$d_V(F,H) \leq d_V(F,G) + d_V(G,H).$$

(c) Write $\iota(f) = (\widehat{f, f, f, \ldots})$, so $d_V(\iota(f), \iota(g)) = d(f,g)$ by (d). Let $F^{(1)}, F^{(2)}, \ldots$ be Cauchy in V, with representatives $(f_1^{(m)}, f_2^{(m)}, \ldots) \in W$. For each m pick k_m with $d(f_j^{(m)}, f_k^{(m)}) < 1/m$ for all $j, k \geq k_m$ and set $h_m = f_{k_m}^{(m)}$; since $d(f_k^{(m)}, h_m) < 1/m$ for $k \geq k_m$,

$$d_V\big(F^{(m)}, \iota(h_m)\big) = \lim_{k\to\infty} d\big(f_k^{(m)}, h_m\big) \leq \tfrac{1}{m}.$$

Then $h_1, h_2, \ldots$ is Cauchy in U: by (d) and the triangle inequality in V,

$$d(h_j, h_m) = d_V\big(\iota(h_j), \iota(h_m)\big) \leq \tfrac{1}{j} + d_V\big(F^{(j)}, F^{(m)}\big) + \tfrac{1}{m},$$

which is $< \varepsilon$ for $j, m \geq n$ once $n \geq n_0$ (where $d_V(F^{(j)}, F^{(m)}) < \varepsilon/3$ for $j, m \geq n_0$) and $n > 3/\varepsilon$. Put $H = (\widehat{h_1, h_2, \ldots})$. Given $\varepsilon > 0$, take n_1 with $d(h_m, h_k) < \varepsilon/3$ for $m, k \geq n_1$; letting $k \to \infty$ gives $d_V(\iota(h_m), H) \leq \varepsilon/3$ for $m \geq n_1$, so for every integer $m \geq n_1$ with $m > 3/\varepsilon$,

$$d_V\big(F^{(m)}, H\big) \leq d_V\big(F^{(m)}, \iota(h_m)\big) + d_V\big(\iota(h_m), H\big) \leq \tfrac{1}{m} + \tfrac{\varepsilon}{3} < \varepsilon.$$

Hence $F^{(m)} \to H$, and (V, d_V) is complete by 6.14.
(d) The constant sequence $(f, f, f, \ldots)$ is Cauchy, so ι is defined on all of U, and the real sequence with k-th term $d(f,g)$ is constant, so

$$d_V\big((\widehat{f, f, f, \ldots}), (\widehat{g, g, g, \ldots})\big) = \lim_{k\to\infty} d(f,g) = d(f,g).$$

(e) Distance preservation makes ι injective (if $\iota(f) = \iota(g)$ then $d(f,g) = 0$, so $f = g$ by the second bullet of 6.1), hence a bijection of U onto $\iota(U) \subseteq V$ carrying d to the metric $\iota(U)$ inherits from d_V. Identifying each f with $\iota(f)$ exhibits U, with its own metric, as a subset of the complete space V of (c).

6.2 Exercises 6B

Problem 6B.1. Show that if $a, b \in \mathbb{R}$ with $a + bi \neq 0$, then

$$\frac{1}{a+bi} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\, i.$$

Solution. Set $w = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\, i$, which makes sense because $a + bi \neq 0$ forces $a^2 + b^2 > 0$. The multiplication rule of 6.17 gives

$$(a+bi)w = \frac{a^2+b^2}{a^2+b^2} + \frac{-ab+ab}{a^2+b^2}\, i = 1,$$

and the multiplicative inverse of a nonzero element of the field $\mathbb{C}$ is unique, so $1/(a+bi) = w$.
Method (2): By 6.24, $z\bar{z} = |z|^2$; taking $z = a+bi$, so $|z|^2 = a^2+b^2 \neq 0$, and dividing by $z|z|^2$ gives

$$\frac{1}{a+bi} = \frac{\overline{a+bi}}{|a+bi|^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\,i.$$

Problem 6B.2. Suppose $z \in \mathbb{C}$. Prove that

$$\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq \sqrt{2}\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\}.$$

Solution. Both inequalities are one squared chain. Write $z = a+bi$ with $a, b \in \mathbb{R}$, so $|z| = \sqrt{a^2+b^2}$ by 6.18, and put $M = \max\{|a|, |b|\}$. Then

$$M^2 = \max\{a^2, b^2\} \leq a^2+b^2 = |z|^2 \leq 2M^2 = \left(\sqrt{2}\,M\right)^2,$$

the first equality because $t \mapsto t^2$ is increasing on $[0,\infty)$ and the last inequality because $a^2, b^2 \leq M^2$. All the quantities squared are nonnegative, so taking square roots preserves the order and gives $M \leq |z| \leq \sqrt{2}\,M$.

Problem 6B.3. Suppose $z \in \mathbb{C}$. Prove that

$$\frac{|\operatorname{Re} z| + |\operatorname{Im} z|}{\sqrt{2}} \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|.$$

Solution. Write $z = a+bi$ with $a, b \in \mathbb{R}$, so $|z| = \sqrt{a^2+b^2}$ by 6.18. Since $2|a|\,|b| \geq 0$, and since $(|a|-|b|)^2 \geq 0$ expands to $2|a|\,|b| \leq a^2+b^2$,

$$|z|^2 = a^2+b^2 \leq \left(|a|+|b|\right)^2 = a^2+2|a|\,|b|+b^2 \leq 2\left(a^2+b^2\right) = \left(\sqrt{2}\,|z|\right)^2.$$

All the quantities squared are nonnegative, so taking square roots gives $|z| \leq |a|+|b| \leq \sqrt{2}\,|z|$, which is the assertion.

Problem 6B.4. Suppose $w, z \in \mathbb{C}$. Prove that $|wz| = |w|\,|z|$ and $|w+z| \leq |w|+|z|$.

Solution. Write $w = a+bi$ and $z = c+di$ with $a, b, c, d \in \mathbb{R}$. By the multiplication rule 6.17 and the formula $|u+vi|^2 = u^2+v^2$ of 6.18,

$$\begin{aligned}
|wz|^2 &= (ac-bd)^2 + (ad+bc)^2 \\
&= a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2 \\
&= \left(a^2+b^2\right)\left(c^2+d^2\right) = (|w|\,|z|)^2,
\end{aligned}$$

the cross terms $\mp 2abcd$ cancelling; both sides are nonnegative, so $|wz| = |w|\,|z|$.
For the triangle inequality, 6.24 supplies $u\bar{u} = |u|^2$, $u+\bar{u} = 2\operatorname{Re} u$, $|\bar{u}| = |u|$, and the additivity and multiplicativity of conjugation, so

$$\begin{aligned}
|w+z|^2 = (w+z)(\bar{w}+\bar{z}) &= |w|^2+|z|^2+w\bar{z}+\overline{w\bar{z}} \\
&= |w|^2+|z|^2+2\operatorname{Re}(w\bar{z}) \\
&\leq |w|^2+2|w\bar{z}|+|z|^2 = (|w|+|z|)^2,
\end{aligned}$$

using $\overline{w\bar{z}} = \bar{w}z$, then $\operatorname{Re} u \leq |\operatorname{Re} u| \leq |u|$ (6B.2), then $|w\bar{z}| = |w|\,|\bar{z}| = |w|\,|z|$ by the first part. Taking square roots of nonnegative quantities gives $|w+z| \leq |w|+|z|$.

Problem 6B.5. Suppose $(X, \mathcal{S})$ is a measurable space and $f\colon X \to \mathbb{C}$ is a complex-valued function. For conditions (b) and (c) below, identify $\mathbb{C}$ with $\mathbb{R}^2$. Prove that the following are equivalent.
(a) f is $\mathcal{S}$-measurable.
(b) $f^{-1}(G) \in \mathcal{S}$ for every open set G in $\mathbb{R}^2$.
(c) $f^{-1}(B) \in \mathcal{S}$ for every Borel set $B \in \mathcal{B}_2$.

Solution. Write $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$, so that $f(x) = (u(x), v(x))$ under the identification and, by 6.19, condition (a) says exactly that u and v are $\mathcal{S}$-measurable. We prove (a) $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ (a).
(a) $\Rightarrow$ (c). The collection $\mathcal{T} = \{B \subseteq \mathbb{R}^2 : f^{-1}(B) \in \mathcal{S}\}$ is a σ-algebra on $\mathbb{R}^2$, since $f^{-1}(\varnothing) = \varnothing$ and inverse images commute with complements and countable unions (2.33). It contains every open cube $C = I_1 \times I_2$ (two bounded open intervals of equal length), because

$$f^{-1}(C) = u^{-1}(I_1) \cap v^{-1}(I_2) \in \mathcal{S}:$$

the intervals are Borel sets, so the two inverse images lie in $\mathcal{S}$ by 2.35, and $\mathcal{S}$ is closed under finite intersections (2.25(b)). Since $\mathcal{B}_2$ is the smallest σ-algebra containing the open cubes (5.38(b)), $\mathcal{B}_2 \subseteq \mathcal{T}$.
(c) $\Rightarrow$ (b). Every open subset of $\mathbb{R}^2$ lies in $\mathcal{B}_2$ by 5.37.
(b) $\Rightarrow$ (a). For each $a \in \mathbb{R}$ the sets $(a, \infty) \times \mathbb{R}$ and $\mathbb{R} \times (a, \infty)$ are open in $\mathbb{R}^2$, and

$$u^{-1}\big((a,\infty)\big) = f^{-1}\big((a,\infty) \times \mathbb{R}\big), \qquad v^{-1}\big((a,\infty)\big) = f^{-1}\big(\mathbb{R} \times (a,\infty)\big),$$

both in $\mathcal{S}$ by (b). Hence u and v are $\mathcal{S}$-measurable by 2.39, and f is $\mathcal{S}$-measurable by 6.19.

Problem 6B.6. Suppose $(X, \mathcal{S})$ is a measurable space and $f, g\colon X \to \mathbb{C}$ are $\mathcal{S}$-measurable. Prove that
(a) $f + g$, $f - g$, and fg are $\mathcal{S}$-measurable functions;
(b) if $g(x) \neq 0$ for all $x \in X$, then $\frac{f}{g}$ is an $\mathcal{S}$-measurable function.

Solution. Write $f = u_1 + iv_1$ and $g = u_2 + iv_2$ with real and imaginary parts as named; by 6.19 the hypothesis says u_1, v_1, u_2, v_2 are $\mathcal{S}$-measurable, and the same result reduces each claim to the measurability of two real-valued functions. Sums, differences, products, and quotients with nowhere-vanishing denominators of such functions are $\mathcal{S}$-measurable by 2.46.
(a) The definitions of addition and multiplication in $\mathbb{C}$ (6.17) give, pointwise on X,

$$\operatorname{Re}(f \pm g) = u_1 \pm u_2, \qquad \operatorname{Im}(f \pm g) = v_1 \pm v_2,$$

$$\operatorname{Re}(fg) = u_1u_2 - v_1v_2, \qquad \operatorname{Im}(fg) = u_1v_2 + v_1u_2,$$

and every right side is $\mathcal{S}$-measurable by 2.46(a).
(b) Put $h = |g|^2 = u_2^2 + v_2^2$, which is $\mathcal{S}$-measurable by 2.45 and 2.46(a) and never 0, since $h(x) = 0$ would force $u_2(x) = v_2(x) = 0$ and hence $g(x) = 0$. By 6.24 we have $g\bar{g} = h$, so $f/g = f\bar{g}/h$ pointwise, and $\bar{g} = u_2 - iv_2$ (6.23) with the rule 6.17 gives $f\bar{g} = (u_1u_2 + v_1v_2) + (v_1u_2 - u_1v_2)i$. Dividing a complex number by the nonzero real h divides both parts by h, so

$$\operatorname{Re}\frac{f}{g} = \frac{u_1u_2 + v_1v_2}{u_2^2 + v_2^2}, \qquad \operatorname{Im}\frac{f}{g} = \frac{v_1u_2 - u_1v_2}{u_2^2 + v_2^2},$$

both $\mathcal{S}$-measurable by 2.46(a) for the numerators and 2.46(b) for the quotients.

Problem 6B.7. Suppose $(X, \mathcal{S})$ is a measurable space and $f_1, f_2, \ldots$ is a sequence of $\mathcal{S}$-measurable functions from X to $\mathbb{C}$. Suppose $\lim_{k\to\infty} f_k(x)$ exists for each $x \in X$. Define $f\colon X \to \mathbb{C}$ by

$$f(x) = \lim_{k\to\infty} f_k(x).$$

Prove that f is an $\mathcal{S}$-measurable function.

Solution. Split into real and imaginary parts: each $\operatorname{Re} f_k$ and $\operatorname{Im} f_k$ is $\mathcal{S}$-measurable by 6.19. Fix $x \in X$ and put $z_k = f_k(x)$, $L = f(x)$, so $|z_k - L| \to 0$ by 6.18. Since Re and Im are additive, Exercise 6B.2 gives

$$\left|\operatorname{Re} z_k - \operatorname{Re} L\right| \le |z_k - L| \quad \text{and} \quad \left|\operatorname{Im} z_k - \operatorname{Im} L\right| \le |z_k - L|,$$

so $(\operatorname{Re} f_k)(x) \to (\operatorname{Re} f)(x)$ and $(\operatorname{Im} f_k)(x) \to (\operatorname{Im} f)(x)$ for every $x \in X$. Thus $\operatorname{Re} f$ and $\operatorname{Im} f$ are everywhere-existing pointwise limits of $\mathcal{S}$-measurable real-valued functions, hence $\mathcal{S}$-measurable by 2.48. Therefore f is $\mathcal{S}$-measurable by 6.19.

Problem 6B.8. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to \mathbb{C}$ is an $\mathcal{S}$-measurable function such that $\int |f| \, d\mu < \infty$. Prove that if $\alpha \in \mathbb{C}$, then

$$\int \alpha f \, d\mu = \alpha \int f \, d\mu.$$

Solution. Both sides equal $(aI_u - bI_v) + i(aI_v + bI_u)$, where $f = u + iv$ and $\alpha = a + bi$ with $u = \operatorname{Re} f$, $v = \operatorname{Im} f$ $\mathcal{S}$-measurable by 6.19, and $I_u = \int u \, d\mu$, $I_v = \int v \, d\mu$.
Because $|u| \le |f|$ and $|v| \le |f|$ pointwise, 3.8 gives $\int |u| \, d\mu < \infty$ and $\int |v| \, d\mu < \infty$, so I_u and I_v are real numbers and $\int f \, d\mu = I_u + iI_v$ by 6.21. Multiplying out with 6.17,

$$\operatorname{Re}(\alpha f) = au - bv \quad \text{and} \quad \operatorname{Im}(\alpha f) = av + bu,$$

while $|\alpha f| = |\alpha| \, |f|$ (Exercise 6B.4) and 3.20 give $\int |\alpha f| \, d\mu = |\alpha| \int |f| \, d\mu < \infty$, so $\int \alpha f \, d\mu$ is defined. Each of au, bv, av, bu has absolute value with finite integral by 3.20, so additivity 3.21 and homogeneity 3.20 apply termwise:

$$\begin{aligned}\int \alpha f \, d\mu &= \int (au - bv) \, d\mu + i \int (av + bu) \, d\mu \\ &= (aI_u - bI_v) + i(aI_v + bI_u),\end{aligned}$$

the first equality by 6.21. Expanding with 6.17,

$$\alpha \int f \, d\mu = (a + bi)(I_u + iI_v) = (aI_u - bI_v) + i(aI_v + bI_u).$$

Problem 6B.9. Suppose V is a vector space. Show that the intersection of every collection of subspaces of V is a subspace of V.

Solution. The intersection inherits each of the three conditions of 6.31 memberwise. Let $\mathcal{A}$ be a collection of subspaces of V and set

$$U = \bigcap_{U' \in \mathcal{A}} U',$$

so $U \subseteq V$ (with the convention that the empty intersection inside V is V itself). Every $U' \in \mathcal{A}$ contains 0 by 6.31, so $0 \in U$. If $f, g \in U$ and $\alpha \in \mathbf{F}$, then $f, g \in U'$ for each $U' \in \mathcal{A}$, whence $f + g \in U'$ and $\alpha f \in U'$ by 6.31; as U' was arbitrary, $f + g \in U$ and $\alpha f \in U$. Hence U is a subspace of V by 6.31.

Problem 6B.10. Suppose V and W are vector spaces. Define $V \times W$ by

$$V \times W = \{(f, g) : f \in V \text{ and } g \in W\}.$$

Define addition and scalar multiplication on $V \times W$ by

$$(f_1, g_1) + (f_2, g_2) = (f_1 + f_2,\ g_1 + g_2) \quad \text{and} \quad \alpha(f, g) = (\alpha f,\ \alpha g).$$

Prove that $V \times W$ is a vector space with these operations.

Solution. Every axiom of 6.27 holds coordinatewise, inherited from V and W (vector spaces over the same $\mathbf{F}$), because two ordered pairs are equal exactly when both pairs of coordinates are.
The operations land in $V \times W$, since $f_1 + f_2 \in V$, $g_1 + g_2 \in W$, $\alpha f \in V$, $\alpha g \in W$. A typical axiom check reads

$$\begin{aligned}\alpha\big((f_1, g_1) + (f_2, g_2)\big) &= \big(\alpha(f_1 + f_2),\ \alpha(g_1 + g_2)\big) \\ &= (\alpha f_1 + \alpha f_2,\ \alpha g_1 + \alpha g_2) \\ &= \alpha(f_1, g_1) + \alpha(f_2, g_2),\end{aligned}$$

the middle equality being distributivity in V and in W; commutativity, associativity of addition, $(\alpha\beta)(f, g) = \alpha(\beta(f, g))$, $1(f, g) = (f, g)$, and $(\alpha + \beta)(f, g) = \alpha(f, g) + \beta(f, g)$ go identically. (Check!) The additive identity is $(0_V, 0_W)$, and (f, g) has additive inverse (f', g'), where $f + f' = 0_V$ in V and $g + g' = 0_W$ in W. Hence $V \times W$ is a vector space over $\mathbf{F}$.

6.3 Exercises 6C

Problem 6C.1. Show that the map $f \mapsto \|f\|$ from a normed vector space V to $\mathbf{F}$ is continuous (where the norm on $\mathbf{F}$ is the usual absolute value).

Solution. The map $N(f) = \|f\|$ is Lipschitz with constant 1, hence continuous. Indeed, the triangle inequality of 6.33 gives

$$\|f\| \le \|f - g\| + \|g\| \quad \text{and} \quad \|g\| \le \|g - f\| + \|f\|,$$

while homogeneity (6.33, with $\alpha = -1$) gives $\|g - f\| = \|f - g\|$; since $\big|\|f\| - \|g\|\big|$ is one of $\|f\| - \|g\|$, $\|g\| - \|f\|$, the two inequalities yield

$$\big|N(f) - N(g)\big| = \big|\|f\| - \|g\|\big| \le \|f - g\|$$

for all $f, g \in V$. Given $f \in V$ and $\varepsilon > 0$, take $\delta = \varepsilon$ in 6.10 (the metrics being $\|f - g\|$ on V and $|s - t|$ on $\mathbf{F}$, by 6.36): $\|g - f\| < \delta$ forces $|N(g) - N(f)| < \varepsilon$.

Problem 6C.2. Prove that if V is a normed vector space, $f \in V$, and $r > 0$, then

$$\overline{B(f, r)} = \overline{B}(f, r).$$

Solution. Both inclusions follow from the sequential description 6.9(a) of the closure, taken in the metric $d(g, h) = \|g - h\|$ of 6.36.
(i) $\overline{B(f, r)} \subseteq \overline{B}(f, r)$. If $g_k \in B(f, r)$ and $g_k \to g$, then for each k

$$\|f - g\| \le \|f - g_k\| + \|g_k - g\| < r + \|g_k - g\|,$$

and letting $k \to \infty$ gives $\|f - g\| \le r$.
(ii) $\overline{B}(f, r) \subseteq \overline{B(f, r)}$. Given $\|f - g\| \le r$, put $g_k = f + \big(1 - \frac{1}{k}\big)(g - f)$. Homogeneity of the norm gives

$$\begin{aligned}\|g_k - f\| &= (1 - \tfrac{1}{k})\|g - f\| \le (1 - \tfrac{1}{k})r < r, \\ \|g_k - g\| &= \tfrac{1}{k}\|g - f\| \le \tfrac{r}{k},\end{aligned}$$

so each $g_k \in B(f, r)$ and $g_k \to g$.

Problem 6C.3. Show that the functions defined in the last two bullet points of Example 6.35 are not norms.

Solution. One-point vectors in $\mathbf{F}^n$ break a condition of 6.33 in each case.
(i) For $\|(a_1,\dots,a_n)\| = |a_1|^{1/2} + \cdots + |a_n|^{1/2}$, homogeneity fails: with $f = (1,0,\dots,0)$ and $\alpha = 4$,

$$\|\alpha f\| = \|(4,0,\dots,0)\| = |4|^{1/2} = 2 \neq 4 = |\alpha|\,\|f\|.$$

(ii) For $\|(a_1,\dots,a_n)\|_{1/2} = \big(|a_1|^{1/2} + \cdots + |a_n|^{1/2}\big)^2$, the triangle inequality fails whenever $n > 1$ (for $n = 1$ this is the absolute value): with $f = (1,0,\dots,0)$ and $g = (0,1,0,\dots,0)$, so that $f + g = (1,1,0,\dots,0)$,

$$\begin{aligned}\|f+g\|_{1/2} &= \big(|1|^{1/2} + |1|^{1/2}\big)^2 = 4\\ &> 2 = \|f\|_{1/2} + \|g\|_{1/2}.\end{aligned}$$

Problem 6C.4. Prove that each Cauchy sequence in a normed vector space is bounded (meaning that there is a real number that is greater than the norm of every element in the Cauchy sequence).

Solution. Take $M = \max\big\{\|f_1\|,\dots,\|f_{n-1}\|,\, 1 + \|f_n\|\big\} + 1$, where $n \in \mathbf{Z}^+$ is supplied by the Cauchy condition 6.12 with $\varepsilon = 1$, so that $\|f_j - f_k\| < 1$ for all $j,k \geq n$. (The maximum is over a finite nonempty set of reals, so $M \in \mathbf{R}$.)
If $k \geq n$, then the triangle inequality gives

$$\|f_k\| = \|(f_k - f_n) + f_n\| \leq \|f_k - f_n\| + \|f_n\| < 1 + \|f_n\| \leq M - 1,$$

and if $k < n$, then $\|f_k\| \leq M - 1$ by the definition of M. Hence $\|f_k\| < M$ for every $k \in \mathbf{Z}^+$.

Problem 6C.5. Show that if $n \in \mathbf{Z}^+$, then $\mathbf{F}^n$ is a Banach space with both the norms used in the first bullet point of Example 6.34.

Solution. Both $\|\cdot\|_1$ and $\|\cdot\|_\infty$ reduce convergence in $\mathbf{F}^n$ to convergence of each coordinate in the complete space $\mathbf{F}$, so $\mathbf{F}^n$ is a Banach space for each.
Each is a norm: positive definiteness and homogeneity follow from $|\alpha a_m| = |\alpha|\,|a_m|$, and $|a_m + b_m| \leq |a_m| + |b_m|$ gives the triangle inequality on summing over m, respectively on maximizing over m. (Check!) They are comparable: for $a = (a_1,\dots,a_n) \in \mathbf{F}^n$ and each m,

$$|a_m| \leq \|a\|_\infty \leq \|a\|_1 \leq n\|a\|_\infty,$$

since the maximum is one of the n summands of $\|a\|_1$, each of which is at most $\|a\|_\infty$.
Now let $\|\cdot\|$ be either norm and let $f_k = (a_{k,1},\dots,a_{k,n})$ be Cauchy for $\|\cdot\|$. Fix m; the comparison gives

$$|a_{j,m} - a_{k,m}| \leq \|f_j - f_k\|_\infty \leq \|f_j - f_k\|$$

(the last step an equality when $\|\cdot\| = \|\cdot\|_\infty$), so $a_{1,m}, a_{2,m},\dots$ is Cauchy in $\mathbf{F}$ and hence converges to some $b_m \in \mathbf{F}$ by completeness of $\mathbf{F}$. Put $b = (b_1,\dots,b_n)$. Given $\varepsilon > 0$, pick N_m with $|a_{k,m} - b_m| < \varepsilon/n$ for all $k \geq N_m$ and set $N = \max\{N_1,\dots,N_n\}$; then for every $k \geq N$,

$$\|f_k - b\|_\infty \leq \|f_k - b\|_1 = \sum_{m=1}^{n} |a_{k,m} - b_m| < \varepsilon.$$

Hence $f_k \to b$ in $(\mathbf{F}^n, \|\cdot\|)$ for both choices of $\|\cdot\|$, so each is a Banach space by 6.37.

Problem 6C.6. Suppose X is a nonempty set and $b(X)$ is the vector space of bounded functions from X to $\mathbf{F}$. Prove that if $\|\cdot\|$ is defined on $b(X)$ by $\|f\| = \sup_X |f|$, then $b(X)$ is a Banach space.

Solution. The limit of a $\|\cdot\|$-Cauchy sequence is its pointwise limit, which exists because $\mathbf{F}$ is complete and is bounded because Cauchy sequences are bounded.

That $\|\cdot\|$ is a norm on $b(X)$ is the third bullet point of 6.34: $\|f\| = 0$ forces $|f(x)| \le 0$ hence $f = 0$; $\|\alpha f\| = \sup_X |\alpha|\,|f| = |\alpha|\,\|f\|$; and $|(f+g)(x)| \le \|f\| + \|g\|$ for each $x \in X$ gives $\|f+g\| \le \|f\| + \|g\|$ on taking the supremum. (Check!)
Let $f_1, f_2, \ldots$ be Cauchy in $b(X)$. For each $x \in X$ we have $|f_j(x) - f_k(x)| \le \|f_j - f_k\|$, so $f_1(x), f_2(x), \ldots$ is Cauchy in $\mathbf{F}$; set $f(x) = \lim_{k\to\infty} f_k(x)$, which exists by completeness of $\mathbf{F}$. By 6C.4 there is $M \in \mathbf{R}$ with $\|f_k\| < M$ for every k, so $|f(x)| = \lim_{k\to\infty} |f_k(x)| \le M$ for each $x \in X$ and hence $f \in b(X)$.
Given $\varepsilon > 0$, pick n with $\|f_j - f_k\| < \varepsilon/2$ for all $j, k \ge n$. Fixing $k \ge n$ and $x \in X$ and letting $j \to \infty$ in $|f_j(x) - f_k(x)| < \varepsilon/2$,

$$|f(x) - f_k(x)| = \lim_{j\to\infty} |f_j(x) - f_k(x)| \le \frac{\varepsilon}{2},$$

so the supremum over $x \in X$ gives $\|f - f_k\| \le \varepsilon/2 < \varepsilon$ for all $k \ge n$. Thus $f_k \to f$ in $b(X)$, which is therefore a Banach space by 6.37.

Problem 6C.7. Show that ℓ^1 with the norm defined by $\|(a_1, a_2, \ldots)\|_\infty = \sup_{k\in\mathbf{Z}^+} |a_k|$ is not a Banach space.

Solution. The truncated harmonic sequences

$$f_k = \left(1, \tfrac{1}{2}, \tfrac{1}{3}, \ldots, \tfrac{1}{k}, 0, 0, \ldots\right) \in \ell^1$$

form a $\|\cdot\|_\infty$-Cauchy sequence with no limit in ℓ^1. (Each f_k has finitely many nonzero coordinates, so lies in ℓ^1; write $h(m)$ for the m^{th} coordinate of h.)
Cauchy: for $j > k$ the coordinates of $f_j - f_k$ vanish except in positions $k+1, \ldots, j$, where they are $\frac{1}{k+1}, \ldots, \frac{1}{j}$, so $\|f_j - f_k\|_\infty = \frac{1}{k+1}$; hence $\|f_j - f_k\|_\infty \le \frac{1}{n+1} < \varepsilon$ for all $j, k \ge n$ once $n > 1/\varepsilon$ (the cases $j = k$ and $j < k$ being trivial and symmetric).
No limit: suppose $\|f_k - g\|_\infty \to 0$ for some $g = (b_1, b_2, \ldots) \in \ell^1$. Fix m; for every $k \ge m$ we have $f_k(m) = \frac{1}{m}$, so

$$\left|\tfrac{1}{m} - b_m\right| = |f_k(m) - b_m| \le \|f_k - g\|_\infty \to 0,$$

and the left side is independent of k, forcing $b_m = \frac{1}{m}$ for every m. Then $\sum_{m=1}^\infty |b_m| = \sum_{m=1}^\infty \frac{1}{m} = \infty$, so $g \notin \ell^1$, a contradiction. Hence $(\ell^1, \|\cdot\|_\infty)$ is not complete.

Problem 6C.8. Show that ℓ^1 with the norm defined by $\|(a_1, a_2, \ldots)\|_1 = \sum_{k=1}^\infty |a_k|$ is a Banach space.

Solution. The coordinatewise limit of an $\|\cdot\|_1$-Cauchy sequence lies in ℓ^1 and is its limit there.
First, $\|\cdot\|_1$ is a norm: $\sum_{k=1}^\infty |a_k| = 0$ forces every $a_k = 0$, homogeneity is $\sum_k |\alpha a_k| = |\alpha| \sum_k |a_k|$, and for $a = (a_1, a_2, \ldots)$, $b = (b_1, b_2, \ldots) \in \ell^1$ and each $n \in \mathbf{Z}^+$,

$$\sum_{k=1}^n |a_k + b_k| \le \sum_{k=1}^n |a_k| + \sum_{k=1}^n |b_k| \le \|a\|_1 + \|b\|_1,$$

so $n \to \infty$ gives $\|a + b\|_1 \le \|a\|_1 + \|b\|_1$ (and $a + b \in \ell^1$).
Now let $f^n = (a_1^n, a_2^n, \ldots)$ be Cauchy in ℓ^1. For fixed k we have $|a_k^n - a_k^m| \le \|f^n - f^m\|_1$, so $a_k^1, a_k^2, \ldots$ is Cauchy in the complete space $\mathbf{F}$ and converges to some a_k; put $f = (a_1, a_2, \ldots)$. Given $\varepsilon > 0$, choose N with $\|f^n - f^m\|_1 < \varepsilon$ for all $m, n \ge N$. Fix $n \ge N$ and $M \in \mathbf{Z}^+$; the finite sum $\sum_{k=1}^M |a_k^n - a_k^m| \le \|f^n - f^m\|_1 < \varepsilon$ may be passed to the limit $m \to \infty$ term by term, giving $\sum_{k=1}^M |a_k^n - a_k| \le \varepsilon$, and the partial sums increase in M, so

$$\sum_{k=1}^\infty |a_k^n - a_k| \le \varepsilon \quad \text{for all } n \ge N.$$

Hence $f^N - f \in \ell^1$, so $f = f^N - (f^N - f) \in \ell^1$, and the display reads $\|f^n - f\|_1 \le \varepsilon$ for all $n \ge N$. Thus $f^n \to f$ in ℓ^1, which is therefore a Banach space by 6.37.

Problem 6C.9. Show that the vector space $C([0,1])$ of continuous functions from $[0,1]$ to $\mathbf{F}$ with the norm defined by $\|f\| = \int_0^1 |f|$ is not a Banach space.

Solution. The ramps $f_k \in C([0,1])$, $k \geq 3$, defined by

$$f_k(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \frac{1}{2}, \\ k(x - \frac{1}{2}) & \text{if } \frac{1}{2} \leq x \leq \frac{1}{2} + \frac{1}{k}, \\ 1 & \text{if } \frac{1}{2} + \frac{1}{k} \leq x \leq 1, \end{cases}$$

are Cauchy for $\|f\| = \int_0^1 |f|$ but have no limit in $C([0,1])$. (The formulas agree at both junctions, so each f_k is continuous, and $0 \leq f_k \leq 1$.)
Used twice below: if g is continuous on $[a,b]$ with $a < b$ and $\int_a^b |g| = 0$, then $g = 0$ on $[a,b]$, since otherwise continuity puts $|g| > c/2$ on an interval of positive length δ about a point where $|g| = c > 0$, forcing $\int_a^b |g| \geq c\delta/2 > 0$.
Cauchy: for $j,k \geq 3$ with $m = \min\{j,k\}$, both f_j and f_k vanish on $[0, \frac{1}{2}]$ and both equal 1 on $[\frac{1}{2} + \frac{1}{m}, 1]$, so $f_j - f_k$ vanishes off an interval of length $\frac{1}{m}$ on which $|f_j - f_k| \leq 1$; hence

$$\|f_j - f_k\| \leq \frac{1}{\min\{j,k\}} < \varepsilon \quad \text{for all } j,k \geq n > 1/\varepsilon.$$

No limit: suppose $f \in C([0,1])$ with $\|f_k - f\| \to 0$. Since $f_k = 0$ on $[0, \frac{1}{2}]$,

$$\int_0^{1/2} |f| = \int_0^{1/2} |f - f_k| \leq \|f_k - f\| \longrightarrow 0,$$

and the left side is independent of k, so $f = 0$ on $[0, \frac{1}{2}]$. Fixing $\delta \in (0, \frac{1}{2})$ and using $f_k = 1$ on $[\frac{1}{2} + \delta, 1]$ for $k > 1/\delta$,

$$\int_{1/2+\delta}^1 |f - 1| = \int_{1/2+\delta}^1 |f - f_k| \leq \|f_k - f\| \longrightarrow 0,$$

so $f = 1$ on $[\frac{1}{2} + \delta, 1]$ and hence, δ being arbitrary, on $(\frac{1}{2}, 1]$. Then $\lim_{x \downarrow 1/2} f(x) = 1 \neq 0 = f(\frac{1}{2})$, contradicting continuity at $\frac{1}{2}$.

Problem 6C.10. Suppose U is a subspace of a normed vector space V such that some open ball of V is contained in U. Prove that $U = V$.

Solution. Shrink each $f \in V$ into the ball, then scale back. Suppose $B(h,r) = \{g \in V : \|g - h\| < r\} \subseteq U$. If $\|g\| < r$, then $h + g$ and h both lie in $B(h,r) \subseteq U$, so $g = (h+g) - h \in U$; thus $B(0,r) \subseteq U$.
Now let $f \in V$ with $f \neq 0$ (the case $f = 0$ holding since U is a subspace), so $\|f\| > 0$ by positive definiteness, and set $\alpha = r/(2\|f\|) > 0$. Homogeneity gives $\|\alpha f\| = r/2 < r$, so $\alpha f \in B(0,r) \subseteq U$ and therefore

$$f = \frac{1}{\alpha}(\alpha f) \in U.$$

Hence $V \subseteq U$, and $U = V$.

Problem 6C.11. Prove that the only subsets of a normed vector space V that are both open and closed are $\emptyset$ and V.

Solution. The segment joining a point of E to a point outside E cannot leave E. Both $\emptyset$ and V are open, and each is the complement of the other, so both are also closed.
Suppose $E \subseteq V$ is open and closed with $f \in E$ and $g \in V \setminus E$, so $\|g - f\| > 0$. Put $\gamma(t) = f + t(g - f)$

for $t \in [0,1]$, so homogeneity of the norm gives

$$\|\gamma(s) - \gamma(t)\| = \|(s-t)(g-f)\| = |s-t|\,\|g-f\|,$$

and let $c = \sup A$, where $A = \{t \in [0,1] : \gamma(t) \in E\}$ contains 0 and is bounded above by 1.
Then $c \in A$: choosing $t_k \in A$ with $t_k \to c$, the display gives $\gamma(t_k) \to \gamma(c)$, and $\gamma(c) \in E$ by 6.9(e) since E is closed. Also $c < 1$, because $\gamma(1) = g \notin E$. But E is open, so $B(\gamma(c), \varepsilon) \subseteq E$ for some $\varepsilon > 0$, and any

$$t \in \Bigl(c,\ \min\Bigl\{1,\ c + \frac{\varepsilon}{\|g-f\|}\Bigr\}\Bigr)$$

(a nonempty interval, as $c < 1$) satisfies $\|\gamma(t) - \gamma(c)\| = (t-c)\|g-f\| < \varepsilon$, so $\gamma(t) \in E$ and $t \in A$ with $t > \sup A$, a contradiction.

Problem 6C.12. Suppose V is a normed vector space. Prove that the closure of each subspace of V is a subspace of V.

Solution. Limits of sums are sums of limits. Let U be a subspace of V; by 6.9(a), $\overline{U}$ is exactly the set of limits of sequences in U, and $0 \in U \subseteq \overline{U}$.
Suppose $f, g \in \overline{U}$ and $\alpha \in \mathbf{F}$; choose $f_k \to f$ and $g_k \to g$ with all $f_k, g_k \in U$, so $f_k + g_k \in U$ and $\alpha f_k \in U$. The triangle inequality and homogeneity give

$$\|(f_k + g_k) - (f+g)\| \le \|f_k - f\| + \|g_k - g\| \to 0 \quad \text{and} \quad \|\alpha f_k - \alpha f\| = |\alpha|\,\|f_k - f\| \to 0,$$

so $f + g \in \overline{U}$ and $\alpha f \in \overline{U}$ by 6.9(a). Hence $\overline{U}$ is a subspace of V by 6.31.

Problem 6C.13. Suppose U is a normed vector space. Let d be the metric on U defined by $d(f,g) = \|f - g\|$ for $f, g \in U$. Let V be the complete metric space constructed in Exercise 16 in Section 6A.
(a) Show that the set V is a vector space under natural operations of addition and scalar multiplication.
(b) Show that there is a natural way to make V into a normed vector space and that with this norm, V is a Banach space.
(c) Explain why (b) shows that every normed vector space is a subspace of some Banach space.

Solution. Everything is defined termwise on Cauchy sequences:

$$\begin{aligned} (f_k)^\wedge + (g_k)^\wedge &= (f_k + g_k)^\wedge, \qquad \alpha (f_k)^\wedge = (\alpha f_k)^\wedge, \\ \bigl\|(f_k)^\wedge\bigr\|_V &= \lim_{k\to\infty} \|f_k\|, \end{aligned}$$

where $(f_k)^\wedge$ denotes the class of the Cauchy sequence $(f_1, f_2, \ldots)$ in the construction of 6A.16, equivalence meaning $\lim_{k\to\infty} d(f_k, g_k) = 0$ with $d(f,g) = \|f-g\|$.
(a) The two operations produce Cauchy sequences and respect equivalence, by

$$\|(f_j + g_j) - (f_k + g_k)\| \le \|f_j - f_k\| + \|g_j - g_k\|, \qquad \|\alpha f_j - \alpha f_k\| = |\alpha|\,\|f_j - f_k\|,$$

read first for large j, k, then with $(f_k'), (g_k')$ equivalent to $(f_k), (g_k)$ in place of the second entries. Each axiom of 6.27 then holds because it holds at every index in U; the zero is $(0, 0, \ldots)^\wedge$ and $(f_k)^\wedge$ has additive inverse $(-f_k)^\wedge$. (Check!)
(b) The limit defining $\|\cdot\|_V$ exists because $\bigl|\|f_j\| - \|f_k\|\bigr| \le \|f_j - f_k\|$ makes $\|f_1\|, \|f_2\|, \ldots$ Cauchy in the complete space $\mathbf{R}$, and the same inequality applied to equivalent sequences shows it depends only on the class. It is a norm: $\lim_k \|f_k\| = 0$ says exactly that $(f_k) \equiv (0, 0, \ldots)$, i.e. $(f_k)^\wedge$ is the zero of V, while homogeneity and the triangle inequality pass to the limit from U. Moreover

$$\begin{aligned} \bigl\|(f_k)^\wedge - (g_k)^\wedge\bigr\|_V &= \bigl\|(f_k - g_k)^\wedge\bigr\|_V = \lim_{k\to\infty} \|f_k - g_k\| \\ &= d_V\bigl((f_k)^\wedge, (g_k)^\wedge\bigr), \end{aligned}$$

so the metric that 6.36 attaches to $\|\cdot\|_V$ is d_V, which is complete by 6A.16(c). Hence V is a Banach space by 6.37.
(c) The map $\varphi(f) = (f, f, f, \dots)^\wedge$ is linear by the definitions in (a) and norm preserving, since $\|\varphi(f)\|_V = \lim_{k\to\infty} \|f\| = \|f\|$; in particular φ is injective, and $\varphi(U)$ is a subspace of V. Thus U is isometrically isomorphic to a subspace of the Banach space V.

Problem 6C.14. Suppose U is a subspace of a normed vector space V. Suppose also that W is a Banach space and $S : U \to W$ is a bounded linear map.
(a) Prove that there exists a unique continuous function $T : \overline{U} \to W$ such that $T|_U = S$.
(b) Prove that the function T in (a) is a bounded linear map from $\overline{U}$ to W and $\|T\| = \|S\|$.
(c) Give an example to show that (a) can fail if the assumption that W is a Banach space is replaced by the assumption that W is a normed vector space.

Solution. The extension is $Tf = \lim_{k\to\infty} Sf_k$ for any $f_k \in U$ with $f_k \to f \in \overline{U}$. Throughout, $\overline{U}$ is a subspace of V by 6C.12, and $\|Sh\| \le \|S\|\,\|h\|$ for $h \in U$ by 6C.16(c), with $\|S\| < \infty$ as S is bounded.
(a) Uniqueness: if T_1, T_2 are continuous on $\overline{U}$ with $T_i|_U = S$, pick $f_k \in U$ with $f_k \to f$ (6.9(a)); then 6.11 gives

$$T_1 f = \lim_{k\to\infty} Sf_k = T_2 f.$$

Existence: such a sequence is Cauchy (6.13), and

$$\|Sf_j - Sf_k\| = \|S(f_j - f_k)\| \le \|S\|\,\|f_j - f_k\|$$

makes $Sf_1, Sf_2, \dots$ Cauchy in the Banach space W, hence convergent. The limit does not depend on the sequence: for another $g_k \to f$ in U,

$$\|Sf_k - Sg_k\| \le \|S\|\big(\|f_k - f\| + \|f - g_k\|\big) \to 0.$$

Constant sequences give $T|_U = S$. Finally, for $f, g \in \overline{U}$ with $f_k \to f$ and $g_k \to g$ in U we have $Sf_k - Sg_k \to Tf - Tg$ (triangle inequality), so letting $k \to \infty$ in $\|Sf_k - Sg_k\| \le \|S\|\,\|f_k - g_k\|$ and using continuity of the norm (6C.1),

$$\|Tf - Tg\| \le \|S\|\,\|f - g\|;$$

thus T is Lipschitz, hence continuous.
(b) Linearity: $f_k + g_k \to f + g$ and $\alpha f_k \to \alpha f$ in U (as in 6C.12), so evaluating T along these sequences and using linearity of S,

$$T(f+g) = \lim_{k\to\infty} (Sf_k + Sg_k) = Tf + Tg, \qquad T(\alpha f) = \lim_{k\to\infty} \alpha S f_k = \alpha T f.$$

Putting $g = 0$ in the Lipschitz estimate (with $T0 = S0 = 0$) gives $\|Tf\| \le \|S\|\,\|f\|$ on $\overline{U}$, so $\|T\| \le \|S\| < \infty$; conversely $\|S\| = \sup\{\|Tf\| : f \in U,\ \|f\| \le 1\} \le \|T\|$, the supremum for $\|T\|$ being over the larger set $\overline{U}$. Hence $\|T\| = \|S\|$.
(c) Take $V = \ell^1$ with $\|\cdot\|_1$, let U be the subspace of sequences with only finitely many nonzero coordinates, let $W = U$ with the inherited norm, and let $S : U \to W$ be the identity, so $\|S\| = 1$. Truncations $a^{(n)} = (a_1, \dots, a_n, 0, 0, \dots) \in U$ satisfy

$$\|a - a^{(n)}\|_1 = \sum_{k=n+1}^{\infty} |a_k| \to 0,$$

so $\overline{U} = \ell^1$ by 6.9(a), while $f = (1/2, 1/4, 1/8, \dots) \in \ell^1 \setminus U$. If $T : \ell^1 \to W$ were continuous with $T|_U = S$, then 6.11 applied to the truncations $f_k \to f$ would give $Tf = \lim_{k\to\infty} f_k$ in W, hence in ℓ^1 since the norm of W is the restriction of $\|\cdot\|_1$, so $Tf = f$ by uniqueness of limits. That is impossible: $Tf \in W = U$ but $f \notin U$.

Problem 6C.15. For readers familiar with the quotient of a vector space and a subspace: Suppose V is a normed vector space and U is a subspace of V. Define $\|\cdot\|$ on V/U by

$$\|f+U\| = \inf\{\|f+g\| : g \in U\}.$$

(a) Prove that $\|\cdot\|$ is a norm on V/U if and only if U is a closed subspace of V.
(b) Prove that if V is a Banach space and U is a closed subspace of V, then V/U (with the norm defined above) is a Banach space.
(c) Prove that if U is a Banach space (with the norm it inherits from V) and V/U is a Banach space (with the norm defined above), then V is a Banach space.

Solution. The quotient norm is the distance to U: replacing g by $-g$ in the infimum gives $\|f+U\| = \operatorname{dist}(f,U)$, which depends only on the coset and satisfies $0 \le \|f+U\| \le \|f\|$ (take $g=0$).
(a) Homogeneity and the triangle inequality hold for any subspace U, so positive definiteness carries the whole statement. For $\alpha \neq 0$ the map $g \mapsto \alpha g$ is a bijection of U, so

$$\|\alpha(f+U)\| = \inf_{h\in U} \|\alpha f + \alpha h\| = |\alpha|\,\|f+U\|,$$

while $\|0(f+U)\| = \|0+U\| = 0$; and for $f,h \in V$, $g_1,g_2 \in U$,

$$\|(f+h)+U\| \le \|(f+h)+(g_1+g_2)\| \le \|f+g_1\| + \|h+g_2\|,$$

whose infimum over g_1 and then g_2 is the triangle inequality. Since $\|0+U\| = 0$ always, $\|\cdot\|$ is a norm exactly when $\operatorname{dist}(f,U)=0$ forces $f \in U$; and $\operatorname{dist}(f,U)=0$ says precisely that f is a limit of a sequence in U, i.e. $f \in \overline{U}$. So $\|\cdot\|$ is a norm if and only if $\overline{U} \subseteq U$, i.e. if and only if U is closed.
(b) By (a), V/U is a normed vector space; apply the series criterion 6.41. Suppose $\sum_{k=1}^{\infty} \|f_k+U\| < \infty$ and choose $g_k \in U$ with $\|f_k+g_k\| \le \|f_k+U\| + 2^{-k}$, so $\sum_{k=1}^{\infty} \|f_k+g_k\| < \infty$. Since V is a Banach space, 6.41 gives $f \in V$ with $\sum_{k=1}^{n}(f_k+g_k) \to f$; as $\sum_{k=1}^{n} g_k \in U$, the cosets agree and $\|h+U\| \le \|h\|$ yields

$$\Big\|\sum_{k=1}^{n}(f_k+U) - (f+U)\Big\| \le \Big\|\sum_{k=1}^{n}(f_k+g_k) - f\Big\| \longrightarrow 0.$$

Thus every absolutely convergent series in V/U converges, so V/U is a Banach space by 6.41.
(c) First U is closed: if $g_k \in U$ and $g_k \to f$ in V, then (g_k) is Cauchy by 6.13, so completeness of U gives $g_k \to g \in U$, and $f = g \in U$ by uniqueness of limits; hence $\|\cdot\|$ is a norm on V/U by (a). Now let $f_1, f_2, \dots$ be Cauchy in V. Since

$$\|(f_j+U)-(f_k+U)\| = \|(f_j-f_k)+U\| \le \|f_j-f_k\|,$$

the cosets are Cauchy in V/U and so converge to some $f+U$, that is, $\|(f_k-f)+U\| \to 0$. Choose $h_k \in U$ with $\|f_k - f + h_k\| \le \|(f_k-f)+U\| + \frac{1}{k}$ and put $u_k = f_k + h_k$, so $u_k \to f$ in V. Then $h_k = u_k - f_k$ obeys $\|h_j - h_k\| \le \|u_j-u_k\| + \|f_j-f_k\|$, and both sequences on the right are Cauchy (6.13 for u_k), so (h_k) is Cauchy in U and converges to some $h \in U$. Therefore

$$\|f_k - (f-h)\| = \|(u_k-h_k)-(f-h)\| \le \|u_k - f\| + \|h_k - h\| \longrightarrow 0,$$

so $f_k \to f-h$ in V, and V is a Banach space.

Problem 6C.16. Suppose V and W are normed vector spaces with $V \neq \{0\}$ and $T : V \to W$ is a linear map.
(a) Show that $\|T\| = \sup\{\|Tf\| : f \in V \text{ and } \|f\| < 1\}$.
(b) Show that $\|T\| = \sup\{\|Tf\| : f \in V \text{ and } \|f\| = 1\}$.
(c) Show that $\|T\| = \inf\{c \in [0,\infty) : \|Tf\| \le c\|f\| \text{ for all } f \in V\}$.

(d) Show that

$$\|T\| = \sup\Big\{\frac{\|Tf\|}{\|f\|} : f \in V \text{ and } f \neq 0\Big\}.$$

Solution. Rescaling moves each defining set onto the closed unit ball. Write $A = \|T\| = \sup\{\|Tf\| : \|f\| \leq 1\}$ (6.43) and let B, C, D, E be the quantities on the right in (a), (b), (c), (d); no boundedness is assumed, so all lie in $[0, \infty]$, with $\inf \emptyset = \infty$. Used throughout: $T0 = 0$ and $\|T(\alpha f)\| = |\alpha|\,\|Tf\|$, and $V \neq \{0\}$ makes the sets for C and E nonempty.
(a) $B \leq A$ since $\{\|f\| < 1\} \subseteq \{\|f\| \leq 1\}$. Conversely, if $\|f\| \leq 1$ and $t \in (0,1)$, then $\|tf\| \leq t < 1$, so $t\|Tf\| = \|T(tf)\| \leq B$; this gives $\|Tf\| \leq B$ trivially when $B = \infty$, and when $B < \infty$ by letting $t \uparrow 1$ in $\|Tf\| \leq B/t$. Hence $A \leq B$.
(b) $C \leq A$ since $\{\|f\| = 1\} \subseteq \{\|f\| \leq 1\}$. Conversely, let $\|f\| \leq 1$; then $\|Tf\| = 0 \leq C$ if $f = 0$, and otherwise $u = f/\|f\|$ has $\|u\| = 1$ and

$$\|Tf\| = \big\|T(\|f\|u)\big\| = \|f\|\,\|Tu\| \leq \|Tu\| \leq C.$$

Hence $A \leq C$.
(c) Put $S = \{c \in [0, \infty) : \|Tf\| \leq c\|f\| \text{ for all } f \in V\}$, so $D = \inf S$. If $c \in S$ and $\|f\| \leq 1$, then $\|Tf\| \leq c\|f\| \leq c$, so $A \leq c$ and $S \subseteq [A, \infty)$. Also $\|Tf\| \leq A\|f\|$ for every $f \in V$, both sides vanishing at $f = 0$ and otherwise

$$\|Tf\| = \|f\|\,\|T(f/\|f\|)\| \leq A\|f\|.$$

So if $A < \infty$, then $A \in S$ and $D = \inf S = A$; and if $A = \infty$, then $S \subseteq [\infty, \infty)$ is empty, so $D = \inf \emptyset = \infty = A$.
(d) The two sets $\{\|Tf\|/\|f\| : f \neq 0\}$ and $\{\|Tu\| : \|u\| = 1\}$ are equal: $\|u\| = 1$ gives $\|Tu\| = \|Tu\|/\|u\|$, and $f \neq 0$ gives $u = f/\|f\|$ of norm 1 with $\|Tu\| = \|Tf\|/\|f\|$. Hence $E = C = A$ by (b).

Problem 6C.17. Suppose U, V, and W are normed vector spaces and $T : U \to V$ and $S : V \to W$ are linear. Prove that $\|S \circ T\| \leq \|S\|\,\|T\|$.

Solution. Apply $\|Rf\| \leq \|R\|\,\|f\|$ (6C.16(c)) twice, to S at Tf and to T at f. The composition $S \circ T$ is linear, so $\|S \circ T\|$ is defined by 6.43.
(i) $\|S\| \neq 0$ and $\|T\| \neq 0$. If either is ∞, then $\|S\|\,\|T\| = \infty$ and there is nothing to prove; so let $\|S\|, \|T\| \in (0, \infty)$ and $f \in U$ with $\|f\| \leq 1$. Then

$$\|(S \circ T)f\| = \|S(Tf)\| \leq \|S\|\,\|Tf\| \leq \|S\|\,\|T\|\,\|f\| \leq \|S\|\,\|T\|,$$

and the supremum over such f gives $\|S \circ T\| \leq \|S\|\,\|T\|$.
(ii) $\|S\| = 0$ or $\|T\| = 0$, where the convention $0 \cdot \infty = 0$ makes the claim $\|S \circ T\| = 0$. If $\|T\| = 0$, the cited inequality forces $Tf = 0$ for every $f \in U$, so $(S \circ T)f = S0 = 0$; if $\|S\| = 0$, it forces $Sg = 0$ for every $g \in V$, so again $(S \circ T)f = 0$. Either way $\|S \circ T\| = 0 \leq \|S\|\,\|T\|$.

Problem 6C.18. Suppose V and W are normed vector spaces and $T : V \to W$ is a linear map. Prove that the following are equivalent.
(a) T is bounded.
(b) There exists $f \in V$ such that T is continuous at f.
(c) T is uniformly continuous (which means that for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\|Tf - Tg\| < \varepsilon$ for all $f, g \in V$ with $\|f - g\| < \delta$).
(d) $T^{-1}\big(B(0, r)\big)$ is an open subset of V for some $r > 0$.

Solution. We run (a) $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ (a) and (a) $\Rightarrow$ (d) $\Rightarrow$ (a), with V and W metrized by 6.36 and $T^{-1}\big(B(0,r)\big) = \{f \in V : \|Tf\| < r\}$. The engine for both returns to (a): if $\|Th\| \leq M < \infty$ for all $\|h\| < \delta$, then $\|T\| \leq 2M/\delta$, since $\|f\| \leq 1$ gives $\|(\delta/2)f\| < \delta$ and hence $(\delta/2)\|Tf\| = \|T((\delta/2)f)\| \leq$

M.
(a) $\Rightarrow$ (c). By 6C.16(c), $\|Th\| \le \|T\|\,\|h\|$; given $\varepsilon > 0$, set $\delta = \varepsilon/(1+\|T\|) > 0$. If $\|f-g\| < \delta$, then

$$\|Tf - Tg\| = \|T(f-g)\| \le \|T\|\,\|f-g\| \le \|T\|\,\delta = \frac{\|T\|}{1+\|T\|}\,\varepsilon < \varepsilon$$

(the second $\le$ rather than $<$ because $\|T\|$ may be 0).
(c) $\Rightarrow$ (b). Uniform continuity gives, for each $\varepsilon > 0$, a $\delta > 0$ with $\|Tg - T0\| < \varepsilon$ whenever $\|g-0\| < \delta$; by 6.10 this is continuity at $f = 0$.
(b) $\Rightarrow$ (a). Continuity at f with $\varepsilon = 1$ in 6.10 gives $\delta > 0$ with $\|Tg - Tf\| < 1$ whenever $\|g - f\| < \delta$. For $\|h\| < \delta$, take $g = f + h$:

$$\|Th\| = \|T(g-f)\| = \|Tg - Tf\| < 1,$$

so the engine with $M = 1$ gives $\|T\| \le 2/\delta < \infty$.
(a) $\Rightarrow$ (d). A bounded T is continuous by 6.48, and $B(0,1)$ is open in W by 6.5, so $T^{-1}\big(B(0,1)\big)$ is open in V by 6.11(a) $\Leftrightarrow$ 6.11(c); thus (d) holds with $r = 1$.
(d) $\Rightarrow$ (a). Since $T0 = 0$, the open set $T^{-1}\big(B(0,r)\big)$ contains 0, so 6.4 supplies $\delta > 0$ with $\|Th\| < r$ whenever $\|h\| < \delta$; the engine with $M = r$ gives $\|T\| \le 2r/\delta < \infty$.

6.4 Exercises 6D

Problem 6D.1. Suppose V is a normed vector space and φ is a linear functional on V. Suppose $\alpha \in \mathbb{F} \setminus \{0\}$. Prove that the following are equivalent.
(a) φ is a bounded linear functional.
(b) $\varphi^{-1}(\alpha)$ is a closed subset of V.
(c) $\overline{\varphi^{-1}(\alpha)} \ne V$.

Solution. Write $S = \varphi^{-1}(\alpha)$, a translate of $\operatorname{null}\varphi$. Translation $\tau_h(f) = h + f$ is a bijection of V with $\|\tau_h(f) - \tau_h(g)\| = \|f - g\|$, hence a homeomorphism, so $\overline{h + A} = h + \overline{A}$ for every $A \subseteq V$.
(i) φ identically 0. Then $S = \varnothing$ because $\alpha \ne 0$, and all three statements hold: $\|\varphi\| = 0$, $\varnothing$ is closed, and $\overline{\varnothing} = \varnothing \ne V$.
(ii) φ not identically 0, so 6.52 applies. Pick g with $\varphi(g) = \beta \ne 0$ and set $h = (\alpha/\beta)g$, so $\varphi(h) = \alpha$ and, since $\varphi(f) = \alpha$ exactly when $f - h \in \operatorname{null}\varphi$,

$$S = h + \operatorname{null}\varphi.$$

(a) $\Rightarrow$ (b): a bounded φ is continuous by 6.48, and $\{\alpha\}$ is closed in $\mathbb{F}$, so $S = \varphi^{-1}(\{\alpha\})$ is closed by 6.11(d).
(b) $\Rightarrow$ (c): $\varphi(0) = 0 \ne \alpha$, so $0 \notin S = \overline{S}$ and hence $\overline{S} \ne V$.
(c) $\Rightarrow$ (a), contrapositively: if φ is unbounded, then statement (a) of 6.52 fails, hence so does its equivalent statement (d), giving $\overline{\operatorname{null}\varphi} = V$ and

$$\overline{S} = h + \overline{\operatorname{null}\varphi} = h + V = V.$$

Problem 6D.2. Suppose φ is a linear functional on a vector space V. Prove that if U is a subspace of V such that $\operatorname{null}\varphi \subseteq U$, then $U = \operatorname{null}\varphi$ or $U = V$.

Solution. If $U \ne \operatorname{null}\varphi$, a single vector of U outside $\operatorname{null}\varphi$ sweeps out V. Pick $u \in U$ with $\beta = \varphi(u) \ne 0$, and let $g \in V$. Then

$$\varphi\Big(g - \frac{\varphi(g)}{\beta}u\Big) = \varphi(g) - \frac{\varphi(g)}{\beta}\,\varphi(u) = 0,$$

so $g - \frac{\varphi(g)}{\beta}u \in \operatorname{null}\varphi \subseteq U$, while $\frac{\varphi(g)}{\beta}u \in U$ because U is a subspace containing u; adding gives $g \in U$. Hence $V \subseteq U$, so $U = V$.

Problem 6D.3. Suppose φ and ψ are linear functionals on the same vector space. Prove that

$$\operatorname{null}\varphi \subseteq \operatorname{null}\psi$$

if and only if there exists $\alpha \in \mathbb{F}$ such that $\psi = \alpha\varphi$.

Solution. The scalar is $\alpha = \psi(h)$ for any h with $\varphi(h) = 1$. Let V be the common domain.
If $\psi = \alpha\varphi$ and $\varphi(f) = 0$, then $\psi(f) = \alpha \cdot 0 = 0$, so $\operatorname{null}\varphi \subseteq \operatorname{null}\psi$.
Conversely, suppose $\operatorname{null}\varphi \subseteq \operatorname{null}\psi$.
(i) φ identically 0. Then $V = \operatorname{null}\varphi \subseteq \operatorname{null}\psi$, so ψ is identically 0 and $\psi = 0 \cdot \varphi$.
(ii) $\varphi(g) \neq 0$ for some $g \in V$. Put $h = g/\varphi(g)$, so $\varphi(h) = 1$, and set $\alpha = \psi(h)$. For $f \in V$,

$$\varphi\big(f - \varphi(f)h\big) = \varphi(f) - \varphi(f)\varphi(h) = 0,$$

so $f - \varphi(f)h \in \operatorname{null}\varphi \subseteq \operatorname{null}\psi$, and linearity of ψ gives

$$0 = \psi\big(f - \varphi(f)h\big) = \psi(f) - \varphi(f)\psi(h) = \psi(f) - \alpha\varphi(f).$$

Hence $\psi = \alpha\varphi$.

Problem 6D.4. [For this exercise and the next, $\mathbb{F}^n$ should be endowed with the norm $\|\cdot\|_\infty$ as defined in Example 6.34.]
Suppose $n \in \mathbb{Z}^+$ and V is a normed vector space. Prove that every linear map from $\mathbb{F}^n$ to V is continuous.

Solution. Every linear map $T : \mathbb{F}^n \to V$ is bounded, hence continuous by 6.48.
Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{F}^n$ and put $c = \sum_{k=1}^n \|Te_k\| < \infty$. For $x = (a_1, \ldots, a_n) \in \mathbb{F}^n$ we have $x = \sum_{k=1}^n a_k e_k$ and $|a_k| \le \|x\|_\infty$, so

$$\begin{aligned}\|Tx\| &= \Big\|\sum_{k=1}^n a_k\, Te_k\Big\| \le \sum_{k=1}^n |a_k|\,\|Te_k\| \\ &\le \Big(\max_{1\le k\le n} |a_k|\Big) \sum_{k=1}^n \|Te_k\| = c\,\|x\|_\infty.\end{aligned}$$

Thus $\|T\| \le c < \infty$, so T is bounded and therefore continuous by 6.48.

Problem 6D.5. [For this exercise and the preceding one, $\mathbb{F}^n$ should be endowed with the norm $\|\cdot\|_\infty$ as defined in Example 6.34.]
Suppose $n \in \mathbb{Z}^+$, V is a normed vector space, and $T : \mathbb{F}^n \to V$ is a linear map that is one-to-one and onto V.
(a) Show that

$$\inf\{\|Tx\| : x \in \mathbb{F}^n \text{ and } \|x\|_\infty = 1\} > 0.$$

(b) Prove that $T^{-1} : V \to \mathbb{F}^n$ is a bounded linear map.

Solution. (a) Were the infimum 0, a Bolzano-Weierstrass extraction would produce a nonzero element of $\operatorname{null} T$.
Write $S = \{x \in \mathbb{F}^n : \|x\|_\infty = 1\}$ and $c = \inf\{\|Tx\| : x \in S\} \ge 0$, and suppose $c = 0$. Choose $x_k \in S$ with $\|Tx_k\| < \frac{1}{k}$. Every coordinate of every x_k has absolute value at most 1, so applying Bolzano-Weierstrass successively in each of the n coordinates (to real and imaginary parts when $\mathbb{F} = \mathbb{C}$) yields a subsequence converging coordinatewise to some $x \in \mathbb{F}^n$; coordinatewise convergence of finitely many coordinates is convergence in $\|\cdot\|_\infty$, so $\|x_{k_m} - x\|_\infty \to 0$, and the reverse triangle inequality then forces

$\|x\|_\infty = 1$. By 6D.4 the map T is continuous, so

$$\|Tx\| = \lim_{m\to\infty} \|Tx_{k_m}\| \le \lim_{m\to\infty} \tfrac{1}{k_m} = 0.$$

Thus x is a nonzero element of null T, contradicting that T is one-to-one. Hence $c > 0$.
(b) T^{-1} is bounded with $\|T^{-1}\| \le 1/c$, where $c > 0$ is the infimum from (a).
Linearity of T^{-1}: applying T^{-1} to $T(T^{-1}g + \alpha T^{-1}h) = g + \alpha h$ gives $T^{-1}(g + \alpha h) = T^{-1}g + \alpha T^{-1}h$.
Next, $\|Tx\| \ge c\|x\|_\infty$ for every $x \in \mathbb{F}^n$: this is trivial for $x = 0$, and for $x \ne 0$ the vector $x/\|x\|_\infty$ lies in S, so homogeneity gives $c \le \|Tx\|/\|x\|_\infty$. Hence for $g \in V$, taking $x = T^{-1}g$,

$$\|T^{-1}g\|_\infty \le \tfrac{1}{c}\|Tx\| = \tfrac{1}{c}\|g\|,$$

so T^{-1} is a bounded linear map with $\|T^{-1}\| \le 1/c$.

Problem 6D.6. Suppose $n \in \mathbb{Z}^+$.
(a) Prove that all norms on $\mathbb{F}^n$ have the same convergent sequences, the same open sets, and the same closed sets.
(b) Prove that all norms on $\mathbb{F}^n$ make $\mathbb{F}^n$ into a Banach space.

Solution. Both parts follow from the comparison: for every norm $\|\cdot\|$ on $\mathbb{F}^n$ there are $c, C \in (0,\infty)$ with

$$c\,\|x\|_\infty \le \|x\| \le C\,\|x\|_\infty \quad \text{for all } x \in \mathbb{F}^n.$$

For the upper bound, take $C = 1 + \sum_{k=1}^n \|e_k\|$ with $e_1, \dots, e_n$ the standard basis; the computation of 6D.4, applied to the identity map I from $(\mathbb{F}^n, \|\cdot\|_\infty)$ to $(\mathbb{F}^n, \|\cdot\|)$, gives $\|x\| \le C\|x\|_\infty$. For the lower bound, I is linear, one-to-one, and onto the normed vector space $(\mathbb{F}^n, \|\cdot\|)$, so 6D.5(a) gives

$$c := \inf\{\|x\| : \|x\|_\infty = 1\} > 0;$$

then $c\|x\|_\infty \le \|x\|$ for all x, trivially at $x = 0$ and by applying the definition of c to $x/\|x\|_\infty$ otherwise. Consequently any two norms $\|\cdot\|_a, \|\cdot\|_b$ on $\mathbb{F}^n$ satisfy $\lambda\|x\|_b \le \|x\|_a \le \Lambda\|x\|_b$ for all x, with $\Lambda = C_a/c_b$ and $\lambda = c_a/C_b$.
(a) Because $\lambda\|x_k - x\|_b \le \|x_k - x\|_a \le \Lambda\|x_k - x\|_b$, one side tends to 0 exactly when the other does, so the two norms have the same convergent sequences with the same limits. For open sets, the same bounds give

$$B_b(y, r/\Lambda) \subseteq B_a(y, r) \quad \text{and} \quad B_a(y, \lambda r) \subseteq B_b(y, r),$$

so a set open for one norm is open for the other. Complements then give the same closed sets.
(b) Let $\|\cdot\|$ be a norm on $\mathbb{F}^n$ with c, C as above, and let $x_1, x_2, \dots$ be $\|\cdot\|$-Cauchy. From $\|x_j - x_k\|_\infty \le \frac{1}{c}\|x_j - x_k\|$ the sequence is $\|\cdot\|_\infty$-Cauchy, hence each of its n coordinate sequences $(a_{k,m})_{k\in\mathbb{Z}^+}$ is Cauchy in $\mathbb{F}$ and so converges to some a_m, the field $\mathbb{F}$ being complete. With $x = (a_1, \dots, a_n)$,

$$\begin{aligned}\|x_k - x\| &\le C\,\|x_k - x\|_\infty \\ &= C \max_{1\le m\le n} |a_{k,m} - a_m| \longrightarrow 0.\end{aligned}$$

Thus $(\mathbb{F}^n, \|\cdot\|)$ is a Banach space.

Problem 6D.7. Suppose V and W are normed vector spaces and V is finite-dimensional. Prove that every linear map from V to W is continuous.

Solution. Every linear map $T : V \to W$ is bounded, hence continuous by 6.48.
Assume $V \ne \{0\}$ (otherwise $T = 0$), let $e_1, \dots, e_n$ be a basis of V, and define $S : \mathbb{F}^n \to V$ by $S(a_1, \dots, a_n) = a_1e_1 + \cdots + a_ne_n$, where $\mathbb{F}^n$ carries $\|\cdot\|_\infty$. Because $e_1, \dots, e_n$ spans V and is linearly independent, S is linear, one-to-one, and onto V; hence S^{-1} is bounded by 6D.5(b), and the linear

map $T \circ S$ is bounded by 6D.4 together with 6.48. Set $M = \|S^{-1}\|$ and $K = \|T \circ S\|$. For $f \in V$, applying $\|Rg\| \le \|R\| \, \|g\|$ (immediate from 6.43) twice,

$$\|Tf\| = \|(T \circ S)(S^{-1}f)\| \le K \, \|S^{-1}f\|_\infty \le KM \, \|f\|.$$

Hence $\|T\| \le KM < \infty$, so T is bounded and therefore continuous by 6.48.

Problem 6D.8. Prove that every finite-dimensional normed vector space is a Banach space.

Solution. Suppose V is a finite-dimensional normed vector space; completeness transfers to V from $\mathbf{F}^n$ through a coordinate map.

If $V = \{0\}$ then V is trivially complete. Otherwise 6.54, after discarding from a finite spanning family each element lying in the span of the others, supplies a basis $e_1, \ldots, e_n$ of V; define $T \colon \mathbf{F}^n \to V$ by $T(a_1, \ldots, a_n) = a_1 e_1 + \cdots + a_n e_n$, with $\mathbf{F}^n$ carrying $\| \cdot \|_\infty$. Then T is linear, and one-to-one and onto V because $e_1, \ldots, e_n$ is a basis. By 6D.4 the map T is bounded; set $M = \|T\|$. By 6D.5(b) the map T^{-1} is bounded; set $N = \|T^{-1}\|$.

Let $f_1, f_2, \ldots$ be Cauchy in V and put $x_j = T^{-1} f_j$. Then

$$\|x_j - x_k\|_\infty \le N \|f_j - f_k\|,$$

so $x_1, x_2, \ldots$ is $\| \cdot \|_\infty$-Cauchy; each of its n coordinate sequences is then Cauchy in the complete field $\mathbf{F}$, so $\|x_j - x\|_\infty \to 0$ for the coordinatewise limit x. Hence

$$\|f_j - Tx\| = \|T(x_j - x)\| \le M \|x_j - x\|_\infty \to 0,$$

so $f_j \to Tx \in V$ and V is a Banach space.

Problem 6D.9. Prove that every finite-dimensional subspace of each normed vector space is closed.

Solution. Suppose U is a finite-dimensional subspace of a normed vector space V; with the inherited norm, U is a Banach space by Exercise 8 in this section.

Suppose $f \in \overline{U}$. By 6.7 the ball of radius $\frac{1}{j}$ about f meets U, so there is a sequence $f_1, f_2, \ldots$ in U with $\|f_j - f\| \to 0$. A convergent sequence is Cauchy, so completeness of U supplies $g \in U$ with $\|f_j - g\| \to 0$, whence

$$\|f - g\| \le \|f - f_j\| + \|f_j - g\| \to 0.$$

Thus $f = g \in U$, so $\overline{U} = U$ and U is closed in V.

Problem 6D.10. Give a concrete example of an infinite-dimensional normed vector space and a basis of that normed vector space.

Solution. Take

$$V = \{(a_1, a_2, \ldots) \in \ell^\infty : a_k = 0 \text{ for all but finitely many } k\},$$

with the norm $\| \cdot \|_\infty$ inherited from ℓ^∞, and let $e_k \in V$ be the sequence with 1 in slot k and 0 elsewhere; then $\{e_k\}_{k \in \mathbf{Z}^+}$ is a basis of V in the sense of 6.54.

V is a subspace of ℓ^∞, since the coordinates of $a + \alpha b$ vanish off the union of the two finite sets where those of a and b are nonzero, so V is a normed vector space. The family $\{e_k\}$ spans V: if $a \in V$ has $a_k = 0$ for $k > n$, then $a = \sum_{k=1}^n a_k e_k$. It is linearly independent: if Ω is finite and $\sum_{k \in \Omega} \alpha_k e_k = 0$, reading off the j^{th} coordinate gives $\alpha_j = 0$ for each $j \in \Omega$.

Finally V is infinite-dimensional: given $f_1, \ldots, f_m \in V$, all their coordinates past some index N vanish, as then do those of every element of $\operatorname{span}(f_1, \ldots, f_m)$, whereas $e_{N+1} \in V$ has $(N+1)^{\text{st}}$ coordinate 1.

Problem 6D.11. Show that the collection $\mathcal{A} = \{k\mathbf{Z} : k = 2, 3, 4, \ldots\}$ of subsets of $\mathbf{Z}$ satisfies the hypothesis of Zorn's Lemma (6.60).

Solution. The union of a chain $\mathcal{C} \subseteq \mathcal{A}$ is its largest member $m\mathbf{Z} \in \mathcal{A}$, where m is the smallest k with $k\mathbf{Z} \in \mathcal{C}$. (Chains are understood nonempty, the only reading under which 6.60 can apply to a collection omitting $\emptyset$.)
For $j, k \in \{2, 3, 4, \ldots\}$ we have the divisibility criterion

$$k\mathbf{Z} \subseteq j\mathbf{Z} \iff j \text{ divides } k :$$

if $k\mathbf{Z} \subseteq j\mathbf{Z}$ then $k \in j\mathbf{Z}$, so $k = jm$ for some $m \in \mathbf{Z}$; conversely $k = jm$ gives $kn = j(mn) \in j\mathbf{Z}$.
Now let $\mathcal{C} \subseteq \mathcal{A}$ be a chain, put $S = \{k \geq 2 : k\mathbf{Z} \in \mathcal{C}\}$, and let $m = \min S$, which exists by the well-ordering of $\mathbf{Z}^+$. Then $m\mathbf{Z} \subseteq \bigcup_{E\in\mathcal{C}} E$. Conversely let $k \in S$; since $\mathcal{C}$ is a chain, either $k\mathbf{Z} \subseteq m\mathbf{Z}$, or $m\mathbf{Z} \subseteq k\mathbf{Z}$, in which case k divides m, so $k \leq m$ and minimality forces $k = m$. Either way $k\mathbf{Z} \subseteq m\mathbf{Z}$, so

$$\bigcup_{E\in\mathcal{C}} E = \bigcup_{k\in S} k\mathbf{Z} = m\mathbf{Z} \in \mathcal{A},$$

which is the hypothesis of 6.60.

Problem 6D.12. Prove that every linearly independent family in a vector space can be extended to a basis of the vector space.

Solution. Apply Zorn's Lemma (6.60) to the collection $\mathcal{A}$ of all linearly independent subsets of V containing $A = \{f_j : j \in \Lambda\}$, where $\{f_j\}_{j\in\Lambda}$ is the given linearly independent family.
Linear independence forces $j \mapsto f_j$ to be one-to-one (otherwise $f_i - f_j = 0$ is a nontrivial vanishing combination), so, as in the convention preceding 6.57, the family may be identified with the set A; note $A \in \mathcal{A}$.
Hypothesis of 6.60: let $\mathcal{C} \subseteq \mathcal{A}$ be a chain (nonempty, as in Exercise 11; when $A = \emptyset$ the empty union $\emptyset$ lies in $\mathcal{A}$ too) and let $U = \bigcup_{E\in\mathcal{C}} E$, so $A \subseteq U$. Every finite subset of U lies in a single element of $\mathcal{C}$: choosing $E_k \in \mathcal{C}$ with $u_k \in E_k$, an induction on n using the chain condition makes one of $E_1, \ldots, E_n$ contain the others. Since linear independence involves only finitely many elements at a time, a relation $\sum_{u\in\Omega} \alpha_u u = 0$ over a finite $\Omega \subseteq U$ takes place inside one linearly independent $E \in \mathcal{C}$, so every $\alpha_u = 0$. Hence $U \in \mathcal{A}$.
By 6.60 there is a maximal $\Gamma \in \mathcal{A}$, and Γ is maximal among all linearly independent subsets of V: any linearly independent $\Gamma' \supsetneq \Gamma$ contains A and so lies in $\mathcal{A}$. Therefore 6.57 makes Γ a basis of V, and $A \subseteq \Gamma$ says it extends the given family.

Problem 6D.13. Suppose V is a normed vector space, U is a subspace of V, and $\psi \colon U \to \mathbf{R}$ is a bounded linear functional. Prove that ψ has a unique extension to a bounded linear functional φ on V with $\|\varphi\| = \|\psi\|$ if and only if

$$\sup_{f\in U} \bigl(-\|\psi\|\,\|f + h\| - \psi(f) \bigr) = \inf_{g\in U} \bigl(\|\psi\|\,\|g + h\| - \psi(g) \bigr)$$

for every $h \in V \setminus U$.

Solution. The bridge is that the norm-preserving extensions of ψ to $U + \mathbf{R}h$ are exactly the φ_c with $s(h) \leq c \leq t(h)$, so uniqueness on V is equivalent to $s(h) = t(h)$ throughout $V \setminus U$. Here $\mathbf{F} = \mathbf{R}$, since ψ is real-valued and over $\mathbf{C}$ the identity $i\psi(f) = \psi(if) \in \mathbf{R}$ would force $\psi = 0$; for $h \in V \setminus U$ write

$$s(h) = \sup_{f\in U} \bigl(-\|\psi\|\,\|f + h\| - \psi(f) \bigr), \qquad t(h) = \inf_{g\in U} \bigl(\|\psi\|\,\|g + h\| - \psi(g) \bigr).$$

Taking $f = g = 0$ and invoking 6.66 gives $-\|\psi\|\,\|h\| \leq s(h) \leq t(h) \leq \|\psi\|\,\|h\|$, so $s(h)$ and $t(h)$ are real and the displayed condition says exactly $s(h) = t(h)$.

Claim. For $c \in \mathbf{R}$, the linear functional $\varphi_c(f+\alpha h) = \psi(f) + \alpha c$ on $U + \mathbf{R}h$ (well defined since $h \notin U$) extends ψ with $\|\varphi_c\| = \|\psi\|$ if and only if $s(h) \le c \le t(h)$. Indeed $\varphi_c|_U = \psi$ gives $\|\varphi_c\| \ge \|\psi\|$, so $\|\varphi_c\| = \|\psi\|$ is exactly 6.64; as in the proof of 6.63, 6.64 is equivalent to 6.65, since 6.64 at $\alpha = 1$ is 6.65 while 6.64 is trivial at $\alpha = 0$ and follows for $\alpha \ne 0$ from 6.65 applied to f/α after multiplying by $|\alpha|$. Finally 6.65 reads

$$\big(-\|\psi\|\,\|f+h\| - \psi(f)\big) \le c \le \big(\|\psi\|\,\|f+h\| - \psi(f)\big)$$

for all $f \in U$, whose supremum and infimum over U are $s(h) \le c \le t(h)$.
(i) Suppose $s(h) = t(h)$ for all $h \in V \setminus U$. At least one norm-preserving extension exists by the Hahn-Banach Theorem (6.69). If φ_1, φ_2 are two of them and $h \in V \setminus U$, then $c_j = \varphi_j(h)$ satisfies

$$|\psi(f) + c_j| = |\varphi_j(f+h)| \le \|\psi\|\,\|f+h\| \quad (f \in U),$$

which is 6.65, so the claim's computation gives $s(h) \le c_j \le t(h)$ and hence $c_1 = c_2$. As $\varphi_1 = \psi = \varphi_2$ on U, we get $\varphi_1 = \varphi_2$.
(ii) Conversely, suppose $s(h) \ne t(h)$ for some $h \in V \setminus U$, so $s(h) < t(h)$ by 6.66. By the claim, $\varphi_{s(h)}$ and $\varphi_{t(h)}$ are norm-$\|\psi\|$ extensions of ψ to $U + \mathbf{R}h$ taking different values at h; applying 6.69 to each yields Φ_1, Φ_2 on V with $\|\Phi_j\| = \|\psi\|$, both restricting to ψ on U, and $\Phi_1(h) = s(h) \ne t(h) = \Phi_2(h)$. So the extension is not unique.

Problem 6D.14. Show that there exists a linear functional $\varphi\colon \ell^\infty \to \mathbf{F}$ such that

$$|\varphi(a_1, a_2, \ldots)| \le \|(a_1, a_2, \ldots)\|_\infty$$

for all $(a_1, a_2, \ldots) \in \ell^\infty$ and

$$\varphi(a_1, a_2, \ldots) = \lim_{k\to\infty} a_k$$

for all $(a_1, a_2, \ldots) \in \ell^\infty$ such that the limit above on the right exists.

Solution. Take for φ a norm-preserving Hahn-Banach extension of the limit functional on the subspace of convergent sequences.
Let $U = \{a \in \ell^\infty : \lim_{k\to\infty} a_k \text{ exists}\}$ and $\psi(a) = \lim_{k\to\infty} a_k$. Every convergent sequence is bounded and the limit laws are exactly additivity and homogeneity, so U is a subspace of ℓ^∞ and ψ is a linear functional on it. Continuity of the absolute value gives

$$|\psi(a)| = \lim_{k\to\infty} |a_k| \le \sup_{k\in\mathbf{Z}^+} |a_k| = \|a\|_\infty,$$

so $\|\psi\| \le 1$, with equality because $\psi(1,1,1,\ldots) = 1$ and $\|(1,1,1,\ldots)\|_\infty = 1$.
By the Hahn-Banach Theorem (6.69), applied to the bounded linear functional ψ on the subspace U of ℓ^∞, there is a bounded linear functional φ on ℓ^∞ with $\varphi|_U = \psi$ and $\|\varphi\| = 1$. Hence $|\varphi(a)| \le \|\varphi\|\,\|a\|_\infty = \|a\|_\infty$ for every $a \in \ell^\infty$, while $\varphi(a) = \psi(a) = \lim_{k\to\infty} a_k$ whenever that limit exists.

Problem 6D.15. Suppose B is an open ball in a normed vector space V such that $0 \notin B$. Prove that there exists $\varphi \in V'$ such that

$$\operatorname{Re}\varphi(f) > 0$$

for all $f \in B$.

Solution. Take $\varphi \in V'$ with $\|\varphi\| = 1$ and $\varphi(g) = \|g\|$, where g is the center of $B = \{f \in V : \|f-g\| < r\}$ and $r > 0$ its radius.

Such a φ is supplied by 6.72, since $0 \notin B$ forces $\|g\| \geq r > 0$ and in particular $g \neq 0$. For $f \in B$,

$$\begin{aligned}
\operatorname{Re}\varphi(f) &= \operatorname{Re}\varphi(g) + \operatorname{Re}\varphi(f-g) \\
&= \|g\| + \operatorname{Re}\varphi(f-g) \\
&\geq \|g\| - |\varphi(f-g)| \\
&\geq \|g\| - \|\varphi\|\,\|f-g\| \\
&= \|g\| - \|f-g\| \\
&> \|g\| - r \\
&\geq 0,
\end{aligned}$$

using $\varphi(g) = \|g\| \in \mathbb{R}$, then $\operatorname{Re} z \geq -|z|$, then the definition of $\|\varphi\|$, then $\|f-g\| < r$, and finally $\|g\| \geq r$.

Problem 6D.16. Show that the dual space of each infinite-dimensional normed vector space is infinite-dimensional.

Solution. For each $n \in \mathbb{Z}^+$ the dual space V' of an infinite-dimensional V contains a linearly independent list of length n, which is impossible for a space spanned by fewer than n elements; hence V' is infinite-dimensional.

First, V contains linearly independent $f_1, \dots, f_n$: starting from any $f_1 \neq 0$, if $f_1, \dots, f_k$ is linearly independent then $\operatorname{span}\{f_1, \dots, f_k\} \neq V$ (else V would be finite-dimensional by 6.54), and any f_{k+1} outside that span extends the list, since a vanishing combination must have $\alpha_{k+1} = 0$ and then all $\alpha_j = 0$.

Let $W = \operatorname{span}\{f_1, \dots, f_n\}$ carry the norm inherited from V, and let $\psi_j : W \to \mathbb{F}$ be the j^{th} coordinate functional for the basis $f_1, \dots, f_n$ of W, so $\psi_j(f_k) = 1$ if $j = k$ and 0 otherwise. Each ψ_j is bounded because W is finite-dimensional (Exercise 7 in Section 6D), so 6.69 extends it to some $\varphi_j \in V'$ with $\varphi_j(f_k) = \psi_j(f_k)$.

If $\alpha_1\varphi_1 + \cdots + \alpha_n\varphi_n = 0$, evaluating at f_k gives $\alpha_k = 0$ for each k. Thus $\varphi_1, \dots, \varphi_n$ is linearly independent in V'.

Problem 6D.17. Suppose V is a separable normed vector space. Explain how the Hahn-Banach Theorem (6.69) for V can be proved without using any results (such as Zorn's Lemma) that depend on the Axiom of Choice.

Solution. Replace the maximal element that Zorn's Lemma supplies in the proof of 6.69 by an explicitly defined recursion: extend one dimension at a time along a fixed enumeration of a countable dense set, then extend by continuity. Every step is given by a formula (a least index, a supremum, a limit), so no choices are made.

Fix an enumeration $g_1, g_2, \dots$ of a countable dense subset of V, indexed by all of $\mathbb{Z}^+$ (repeat elements if it is finite; if it is empty then $V = \{0\}$ and 6.69 is trivial). Fixing one such enumeration is a single existential instantiation, not an application of the Axiom of Choice. Let ψ be a bounded linear functional on a subspace U of V.

Case (i): $\mathbb{F} = \mathbb{R}$. The Extension Lemma (6.63) becomes choice-free once c is specified: for a bounded linear functional ρ on a subspace W and $h \in V \setminus W$, put

$$c = \sup_{f \in W} \bigl(-\|\rho\|\,\|f+h\| - \rho(f)\bigr),$$

a real number, since $f = 0$ bounds it below by $-\|\rho\|\,\|h\|$ and 6.66 bounds it above by $\|\rho\|\,\|h\|$; with this c, the computation in the proof of 6.63 makes $\rho_h(f + \alpha h) = \rho(f) + \alpha c$ a norm-preserving extension of ρ to $W + \mathbb{R}h$.

Recursively set $U_0 = U$, $\varphi_0 = \psi$, and given (U_n, φ_n): if every g_k lies in U_n, keep $(U_{n+1}, \varphi_{n+1}) = (U_n, \varphi_n)$; otherwise let m_n be the least index with $g_{m_n} \notin U_n$, put $U_{n+1} = U_n + \mathbb{R}g_{m_n}$, and let φ_{n+1}

be the explicit extension above with $h = g_{m_n}$. Each ingredient is determined by (U_n, φ_n) and the enumeration, so the recursion theorem produces the sequence, and $\|\varphi_n\| = \|\psi\|$ throughout.
Every g_k lies in some U_n: otherwise m_n is defined for all n and strictly increases (as $g_{m_n} \in U_{n+1}$), so $m_n \ge m_0 + n \to \infty$. Hence $W = \bigcup_n U_n$, a subspace because the U_n increase, is dense in V, and $\varphi(f) = \varphi_n(f)$ for any n with $f \in U_n$ is a well-defined linear functional on W extending ψ with $|\varphi(f)| \le \|\psi\| \, \|f\|$.
For $f \in V$ let $k_n(f)$ be the least k with $\|g_k - f\| < \frac{1}{n}$, which exists by density. Then

$$\big|\varphi(g_{k_n(f)}) - \varphi(g_{k_m(f)})\big| \le \|\psi\| \Big(\frac{1}{n} + \frac{1}{m}\Big),$$

so $\tilde{\varphi}(f) = \lim_{n\to\infty} \varphi(g_{k_n(f)})$ exists by completeness of $\mathbb{R}$ (itself choice-free, the limit being $\sup_n \inf_{j \ge n}$ of the sequence). The same estimate shows $\varphi(h_j) \to \tilde{\varphi}(f)$ for every sequence $h_j \to f$ in W; applying this along the explicit sequences $h_j = g_{k_j(f)}$ and $h'_j = g_{k_j(f')}$ gives

$$\tilde{\varphi}(f + \alpha f') = \lim_{j\to\infty} \big(\varphi(h_j) + \alpha\varphi(h'_j)\big) = \tilde{\varphi}(f) + \alpha\tilde{\varphi}(f').$$

Taking $h_j = f$ for $f \in W$ shows $\tilde{\varphi}|_W = \varphi$, so $\tilde{\varphi}$ extends ψ; and $|\varphi(g_{k_n(f)})| \le \|\psi\| \, \|g_{k_n(f)}\|$ with $g_{k_n(f)} \to f$ gives $\|\tilde{\varphi}\| \le \|\psi\|$, hence $\|\tilde{\varphi}\| = \|\psi\|$.
Case (ii): $\mathbb{F} = \mathbb{C}$. The reduction in the proof of 6.69 uses no choice: set $\psi_1 = \operatorname{Re}\psi$, apply case (i) to the real normed vector space V (same norm, hence the same countable dense set, so separable) to get φ_1 extending ψ_1 with $\|\varphi_1\| = \|\psi_1\|$, and put $\varphi(f) = \varphi_1(f) - i\varphi_1(if)$; the verifications in 6.69 that φ is $\mathbb{C}$-linear, extends ψ, and has $\|\varphi\| \le \|\psi\|$ are computations.

Problem 6D.18. Suppose V is a normed vector space such that the dual space V' is a separable Banach space. Prove that V is separable.

Solution. Take a countable dense $\{\varphi_1, \varphi_2, \dots\}$ in V' and vectors $f_n \in V$ with $\|f_n\| \le 1$ and $|\varphi_n(f_n)| > \frac{1}{2}\|\varphi_n\|$ (possible by the definition of $\|\varphi_n\|$ as a supremum when $\varphi_n \ne 0$; put $f_n = 0$ otherwise); then the rational span of $\{f_n\}$ is a countable dense subset of V.
Let $W = \operatorname{span}\{f_n : n \in \mathbb{Z}^+\}$ and let D consist of the finite combinations $\sum_{n\in\Omega} \alpha_n f_n$ with all $\alpha_n \in \mathbb{Q}$ (in $\mathbb{Q} + i\mathbb{Q}$ when $\mathbb{F} = \mathbb{C}$), a countable set. Then $W \subseteq \overline{D}$, because

$$\Big\|\sum_{n\in\Omega} (\alpha_n - \beta_n) f_n\Big\| \le \sum_{n\in\Omega} |\alpha_n - \beta_n| \, \|f_n\|$$

is arbitrarily small for suitable rational β_n; so it suffices to show $\overline{W} = V$.
If not, take $h \in V \setminus \overline{W}$. Since W is a subspace, 6.73 gives $\varphi \in V'$ with $\varphi|_W = 0$ and $\varphi(h) \ne 0$, and dividing by $\|\varphi\| \ne 0$ we may take $\|\varphi\| = 1$. By density choose n with $\|\varphi - \varphi_n\| < \frac{1}{4}$, so $\|\varphi_n\| \ge \|\varphi\| - \|\varphi - \varphi_n\| > \frac{3}{4}$; in particular $\varphi_n \ne 0$, so f_n is as displayed above. But $f_n \in W$ forces $\varphi(f_n) = 0$, whence

$$\tfrac{1}{2}\|\varphi_n\| < |\varphi_n(f_n) - \varphi(f_n)| \le \|\varphi_n - \varphi\| \, \|f_n\| < \tfrac{1}{4},$$

contradicting $\|\varphi_n\| > \frac{3}{4}$. Hence $\overline{D} \supseteq \overline{W} = V$ and V is separable.

Problem 6D.19. Prove that the dual of the Banach space $C([0,1])$ is not separable; here the norm on $C([0,1])$ is defined by $\|f\| = \sup_{[0,1]} |f|$.

Solution. The evaluation functionals $\varphi_t(f) = f(t)$, for $t \in [0,1]$, form an uncountable family in $C([0,1])'$ that is pairwise at distance 2, leaving no room for a countable dense subset.
Each φ_t is linear with $|\varphi_t(f)| = |f(t)| \le \|f\|$, so $\varphi_t \in C([0,1])'$ with $\|\varphi_t\| = 1$, the constant function 1 attaining the supremum. For $s \ne t$ put $\delta = |s - t| > 0$ and

$$f(x) = \max\Big\{-1,\ 1 - \frac{2|x - s|}{\delta}\Big\},$$

continuous with $\|f\| \le 1$, $f(s) = 1$, and $f(t) = -1$; hence

$$2 = |f(s) - f(t)| \le \|\varphi_s - \varphi_t\| \le \|\varphi_s\| + \|\varphi_t\| = 2.$$

Suppose $\{\mu_1, \mu_2, \ldots\}$ were dense in $C([0,1])'$, and let $\kappa(t)$ be the least j with $\|\mu_j - \varphi_t\| < 1$ (such j exists by density). If $\kappa(s) = \kappa(t) = j$ then

$$\|\varphi_s - \varphi_t\| \le \|\varphi_s - \mu_j\| + \|\mu_j - \varphi_t\| < 2,$$

so $s = t$ by the previous paragraph. Thus κ is a one-to-one map from the uncountable set $[0,1]$ into $\mathbb{Z}^+$, which is impossible.

Problem 6D.20. Define $\Phi : V \to V''$ by

$$(\Phi f)(\varphi) = \varphi(f)$$

for $f \in V$ and $\varphi \in V'$. Show that $\|\Phi f\| = \|f\|$ for every $f \in V$.
[The map Φ defined above is called the canonical isometry of V into V''.]

Solution. Each Φf is linear on V' by the definition of the vector space operations on V', and the definition of $\|\varphi\|$ gives

$$|(\Phi f)(\varphi)| = |\varphi(f)| \le \|\varphi\|\,\|f\| \quad \text{for all } \varphi \in V',$$

so $\Phi f \in V''$ with $\|\Phi f\| \le \|f\|$.
For the reverse inequality, both sides vanish when $f = 0$; and when $f \neq 0$, 6.72 supplies $\varphi \in V'$ with $\|\varphi\| = 1$ and $\varphi(f) = \|f\|$, whence

$$\|\Phi f\| \ge |(\Phi f)(\varphi)| = |\varphi(f)| = \|f\|.$$

Thus $\|\Phi f\| = \|f\|$ for every $f \in V$.

Problem 6D.21. Suppose V is an infinite-dimensional normed vector space. Show that there is a convex subset U of V such that $\overline{U} = V$ and such that the complement $V \setminus U$ is also a convex subset of V with $\overline{V \setminus U} = V$.
[See 8.25 for the definition of a convex set. This exercise should stretch your geometric intuition because this behavior cannot happen in finite dimensions.]

Solution. Take $U = \{f \in V : \rho(f) > 0\}$, where $\rho : V \to \mathbb{R}$ is a discontinuous $\mathbb{R}$-linear functional.
Convexity (8.25) involves only real scalars, so regard V as a real normed vector space; it is still infinite-dimensional, since a finite real spanning list would span V over $\mathbb{C}$ as well. Hence 6.62 supplies a discontinuous linear functional $\rho : V \to \mathbb{R}$, which is unbounded by 6.48 and in particular is not identically 0.
$\operatorname{null} \rho$ is dense in V (this is (d) $\Rightarrow$ (a) of 6.52 in contrapositive form): unboundedness gives h_j with $\|h_j\| \le 1$ and $|\rho(h_j)| > j$, obtained by normalizing any f_j with $|\rho(f_j)| > j\|f_j\|$, and then for $g \in V$ the vectors $k_j = g - \frac{\rho(g)}{\rho(h_j)} h_j$ satisfy $\rho(k_j) = 0$ and

$$\|k_j - g\| = \frac{|\rho(g)|}{|\rho(h_j)|}\|h_j\| \le \frac{|\rho(g)|}{j} \longrightarrow 0.$$

Both U and $V \setminus U = \{f : \rho(f) \le 0\}$ are convex: for $t \in [0,1]$,

$$\rho\big((1-t)f + tg\big) = (1-t)\rho(f) + t\rho(g),$$

which is > 0 when $\rho(f), \rho(g) > 0$ (both terms are ≥ 0 and one of t, $1-t$ is nonzero) and is ≤ 0 when $\rho(f), \rho(g) \le 0$.

Density of $V \setminus U$ is immediate from $\operatorname{null} \rho \subseteq V \setminus U$. For U, fix $h = h_0/\rho(h_0)$ with $\rho(h_0) \neq 0$, so $\rho(h) = 1$ and $h \neq 0$; given $g \in V$ and $\varepsilon > 0$, pick $k \in \operatorname{null} \rho$ with $\|k - g\| < \varepsilon/2$ and set $\delta = \frac{\varepsilon}{2\|h\|+2} > 0$ and $u = k + \delta h$. Then $\rho(u) = \delta > 0$, so $u \in U$, while

$$\|u - g\| \leq \|k - g\| + \delta\|h\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $\overline{U} = V = \overline{V \setminus U}$.

6.5 Exercises 6E

Problem 6E.1. Suppose U is a subset of a metric space V. Show that U is dense in V if and only if every nonempty open subset of V contains at least one element of U.

Solution. Both directions are 6.7, which says $\overline{U}$ consists of the $g \in V$ with $B(g, \varepsilon) \cap U \neq \emptyset$ for every $\varepsilon > 0$.
(i) Suppose $\overline{U} = V$ and let G be a nonempty open subset of V. Pick $f \in G$ and, by 6.4, an $r > 0$ with $B(f, r) \subseteq G$. Then $f \in \overline{U}$ gives $B(f, r) \cap U \neq \emptyset$, and any of its elements lies in $G \cap U$.
(ii) Conversely, let $f \in V$ and $\varepsilon > 0$. The ball $B(f, \varepsilon)$ is open (6.5) and contains f, so by hypothesis it meets U; as ε was arbitrary, 6.7 gives $f \in \overline{U}$. Hence $\overline{U} = V$.

Problem 6E.2. Suppose U is a subset of a metric space V. Show that U has an empty interior if and only if $V \setminus U$ is dense in V.

Solution. Both directions run through Exercise 1: $V \setminus U$ is dense exactly when every nonempty open subset of V meets $V \setminus U$, that is, when no nonempty open set is contained in U.
(i) Suppose $\operatorname{int} U = \emptyset$ and some nonempty open G satisfies $G \subseteq U$. Any $f \in G$ then has $B(f, r) \subseteq G \subseteq U$ for some $r > 0$ by 6.4, so $f \in \operatorname{int} U$ by 6.74, a contradiction. Hence every nonempty open set meets $V \setminus U$, and Exercise 1 gives density.
(ii) Conversely, suppose $V \setminus U$ is dense and $f \in \operatorname{int} U$, so $B(f, r) \subseteq U$ for some $r > 0$. The ball $B(f, r)$ is nonempty and open (6.5), so by Exercise 1 it contains an element of $V \setminus U$, a contradiction. Hence $\operatorname{int} U = \emptyset$.

Problem 6E.3. Prove or give a counterexample: If V is a metric space and U, W are subsets of V, then $(\operatorname{int} U) \cup (\operatorname{int} W) = \operatorname{int}(U \cup W)$.

Solution. False: in $V = \mathbb{R}$ take $U = \mathbb{Q}$ and $W = \mathbb{R} \setminus \mathbb{Q}$. Each is dense in $\mathbb{R}$ and is the complement of the other, so Exercise 2 gives $\operatorname{int} U = \operatorname{int} W = \emptyset$, whereas

$$\operatorname{int}(U \cup W) = \operatorname{int} \mathbb{R} = \mathbb{R} \neq \emptyset = (\operatorname{int} U) \cup (\operatorname{int} W).$$

Only the inclusion $\subseteq$ holds in general: if $f \in \operatorname{int} U$ then $B(f, r) \subseteq U \subseteq U \cup W$ for some $r > 0$, so $f \in \operatorname{int}(U \cup W)$ by 6.74, and symmetrically with W in place of U.

Problem 6E.4. Prove or give a counterexample: If V is a metric space and U, W are subsets of V, then $(\operatorname{int} U) \cap (\operatorname{int} W) = \operatorname{int}(U \cap W)$.

Solution. True; by 6.74, $f \in \operatorname{int} E$ means $f \in E$ and $B(f, r) \subseteq E$ for some $r > 0$.
(i) If $f \in (\operatorname{int} U) \cap (\operatorname{int} W)$, choose $r_1, r_2 > 0$ with $B(f, r_1) \subseteq U$ and $B(f, r_2) \subseteq W$; then $f \in U \cap W$ and $B(f, \min\{r_1, r_2\}) \subseteq U \cap W$, so $f \in \operatorname{int}(U \cap W)$.
(ii) If $f \in \operatorname{int}(U \cap W)$, choose $r > 0$ with $B(f, r) \subseteq U \cap W$; since $U \cap W$ is contained in each of U and W, this gives $f \in \operatorname{int} U$ and $f \in \operatorname{int} W$.

Problem 6E.5. Suppose
$$X = \{0\} \cup \bigcup_{k=1}^{\infty} \{\tfrac{1}{k}\}$$
and $d(x,y) = |x-y|$ for $x, y \in X$.
(a) Show that (X,d) is a complete metric space.
(b) Each set of the form $\{x\}$ for $x \in X$ is a closed subset of $\mathbb{R}$ that has an empty interior as a subset of $\mathbb{R}$. Clearly X is a countable union of such sets. Explain why this does not violate the statement of Baire's Theorem that a complete metric space is not the countable union of closed subsets with empty interior.

Solution. (a) X is a closed subset of the complete metric space $\mathbb{R}$, hence complete.
Closedness: since $X \subseteq [0,1]$ and every $x \in (0,1)$ not of the form $1/k$ satisfies $\frac{1}{k+1} < x < \frac{1}{k}$ for exactly one $k \in \mathbb{Z}^+$,
$$\mathbb{R} \setminus X = (-\infty, 0) \cup (1, \infty) \cup \bigcup_{k=1}^{\infty} \left(\tfrac{1}{k+1}, \tfrac{1}{k}\right),$$
a union of open sets, so X is closed by 6.6. A Cauchy sequence in X is Cauchy in $\mathbb{R}$, hence converges to some $x \in \mathbb{R}$, and $x \in X$ by 6.9(e); as d is the restriction of the metric of $\mathbb{R}$, this is convergence in (X,d).
(b) Baire's Theorem [6.76(a)] concerns empty interior in the complete space at hand, here X rather than $\mathbb{R}$, and each singleton $\{\frac{1}{k}\}$ has nonempty interior in X.
Indeed, with $r = \frac{1}{k(k+1)}$ the only element of X within r of $\frac{1}{k}$ is $\frac{1}{k}$ itself: $|\frac{1}{k} - 0| = \frac{1}{k} > r$, while $|\frac{1}{k} - \frac{1}{j}| \geq \frac{1}{k} - \frac{1}{k+1} = r$ for $j > k$ and $|\frac{1}{k} - \frac{1}{j}| \geq \frac{1}{k-1} - \frac{1}{k} > r$ for $j < k$. So $B(\frac{1}{k}, r) = \{\frac{1}{k}\}$, which by 6.74 gives $\operatorname{int}\{\frac{1}{k}\} = \{\frac{1}{k}\}$ in X. (Only $\{0\}$ has empty interior in X, since every $B(0,r)$ contains some $\frac{1}{k} \neq 0$.) Thus the displayed decomposition is not a union of closed sets with empty interior in X, and 6.76(a) is not contradicted.

Problem 6E.6. Give an example of a metric space that is the countable union of closed subsets with empty interior.
[This exercise shows that the completeness hypothesis in Baire's Theorem cannot be dropped.]

Solution. Take $\mathbb{Q}$ with $d(x,y) = |x-y|$ and an enumeration $r_1, r_2, \dots$ of its elements, so that
$$\mathbb{Q} = \bigcup_{j=1}^{\infty} \{r_j\}.$$
Each $\{r_j\}$ is closed in $\mathbb{Q}$ by 6.6, since $B(x, |x - r_j|) \subseteq \mathbb{Q} \setminus \{r_j\}$ for every rational $x \neq r_j$. Each has empty interior in $\mathbb{Q}$: given $r > 0$, density of $\mathbb{Q}$ in $\mathbb{R}$ puts a rational $q \neq r_j$ in $(r_j, r_j + r)$, so $B(r_j, r) \not\subseteq \{r_j\}$ and 6.74 applies.

Problem 6E.7. (a) Define $f : \mathbb{R} \to \mathbb{R}$ as follows:
$$f(a) = \begin{cases} 0 & \text{if } a \text{ is irrational,} \\ \frac{1}{n} & \text{if } a \text{ is rational and } n \text{ is the smallest} \\ & \text{positive integer such that } a = \frac{m}{n} \text{ for some integer } m. \end{cases}$$
At which numbers in $\mathbb{R}$ is f continuous?
(b) Show that there does not exist a countable collection of open subsets of $\mathbb{R}$ whose intersection equals $\mathbb{Q}$.
(c) Show that there does not exist a function $f : \mathbb{R} \to \mathbb{R}$ such that f is continuous at each element of $\mathbb{Q}$ and discontinuous at each element of $\mathbb{R} \setminus \mathbb{Q}$.

Solution. (a) f is continuous exactly at the irrational numbers. Write $n(a)$ for the least positive integer with $a = m/n(a)$ for some $m \in \mathbb{Z}$, so $f(a) = 1/n(a) > 0$ for rational a and $f = 0$ on the irrationals.
Discontinuity at rational a: take $\varepsilon = f(a) > 0$. Every interval $(a-\delta, a+\delta)$ contains an irrational x, by density of $\mathbb{R} \setminus \mathbb{Q}$, and then $|f(x) - f(a)| = f(a) = \varepsilon$, so no δ works in 6.10.
Continuity at irrational a: here $f(a) = 0$. Given $\varepsilon > 0$, choose $N \in \mathbb{Z}^+$ with $\frac{1}{N} < \varepsilon$ and let

$$S = \{x \in (a-1, a+1) : f(x) \geq \varepsilon\}.$$

Each $x \in S$ has $f(x) > 0$, hence is rational with $\frac{1}{n(x)} \geq \varepsilon > \frac{1}{N}$, so $n(x) < N$; and for each fixed $n \leq N$ the condition $\frac{m}{n} \in (a-1, a+1)$ confines m to an interval of length $2n \leq 2N$, so at most $2N+1$ integers. Hence S is finite, and $a \notin S$ because a is irrational. Put

$$\delta = \min\left(\{1\} \cup \{|a-s| : s \in S\}\right) > 0.$$

If $|x - a| < \delta$ then $x \in (a-1, a+1) \setminus S$, so $|f(x) - f(a)| = f(x) < \varepsilon$.
(b) Suppose $G_1, G_2, \ldots$ are open with $\bigcap_k G_k = \mathbb{Q}$. Each G_k contains $\mathbb{Q}$, hence is dense. Enumerating $\mathbb{Q} = \{r_1, r_2, \ldots\}$, each $H_j = \mathbb{R} \setminus \{r_j\}$ is open, and dense because every nonempty open subset of $\mathbb{R}$ is infinite (Exercise 1). But then $G_1, H_1, G_2, H_2, \ldots$ are countably many dense open subsets of the complete metric space $\mathbb{R}$ with

$$\Big(\bigcap_{k=1}^{\infty} G_k\Big) \cap \Big(\bigcap_{j=1}^{\infty} H_j\Big) = \mathbb{Q} \cap (\mathbb{R} \setminus \mathbb{Q}) = \emptyset,$$

contradicting Baire's Theorem [6.76(b)].
(c) The set of continuity points of any $f : \mathbb{R} \to \mathbb{R}$ is a countable intersection of open sets, so by (b) it is never $\mathbb{Q}$.
For $a \in \mathbb{R}$ and $\delta > 0$ let $\omega(a, \delta) = \sup\{|f(x) - f(y)| : x, y \in (a-\delta, a+\delta)\} \in [0, \infty]$ and $\omega(a) = \inf_{\delta > 0} \omega(a, \delta)$. Then f is continuous at a if and only if $\omega(a) = 0$: continuity gives δ with $|f(x) - f(a)| < \varepsilon/3$ on $(a-\delta, a+\delta)$, so $\omega(a) \leq \omega(a, \delta) \leq 2\varepsilon/3 < \varepsilon$ for every $\varepsilon > 0$; conversely $\omega(a, \delta) < \varepsilon$ gives $|f(x) - f(a)| \leq \omega(a, \delta) < \varepsilon$ for $|x - a| < \delta$, taking $y = a$ in the supremum.
Each $G_n = \{a : \omega(a) < \frac{1}{n}\}$ is open: choosing δ with $\omega(a, \delta) < \frac{1}{n}$, any b with $|b - a| < \delta/2$ has $(b - \frac{\delta}{2}, b + \frac{\delta}{2}) \subseteq (a - \delta, a + \delta)$ and hence $\omega(b) \leq \omega(b, \delta/2) \leq \omega(a, \delta) < \frac{1}{n}$, so $B(a, \delta/2) \subseteq G_n$ (6.4). The set of continuity points is therefore $\bigcap_{n=1}^{\infty} G_n$, which by (b) cannot equal $\mathbb{Q}$.

Problem 6E.8. Suppose (X, d) is a complete metric space and $G_1, G_2, \ldots$ is a sequence of dense open subsets of X. Prove that $\bigcap_{k=1}^{\infty} G_k$ is a dense subset of X.

Solution. By Exercise 1 in Section 6E, density of $\bigcap_{k=1}^{\infty} G_k$ amounts to: every nonempty open $U \subseteq X$ meets it. So fix such a U, pick $f \in U$, and use the fact proved just before 6.76 to get $r > 0$ with

$$B(f, r) \subseteq Y := \overline{B(f, r)} \subseteq U.$$

Then Y is nonempty and closed in X, so (Y, d) is a complete metric space by 6.16(b).
Set $H_k = G_k \cap Y$, open in Y because G_k is open in X. Each H_k is dense in Y: given $y \in Y$ and $\varepsilon > 0$, the point y lies in the closure of $B(f, r)$, so there is $x \in B(f, r)$ with $d(x, y) < \varepsilon/2$; the set $B(f, r) \cap B(x, \varepsilon/2)$ is open and nonempty, so density of G_k supplies a point z of it lying in G_k, and then $z \in G_k \cap B(f, r) \subseteq H_k$ with

$$d(z, y) \leq d(z, x) + d(x, y) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Baire's Theorem [6.76(b)], applied to the dense open subsets $H_1, H_2, \ldots$ of the complete metric space Y, produces

$$g \in \bigcap_{k=1}^{\infty} H_k = Y \cap \bigcap_{k=1}^{\infty} G_k \subseteq U \cap \bigcap_{k=1}^{\infty} G_k,$$

as required.

Problem 6E.9. Prove that there does not exist an infinite-dimensional Banach space with a countable basis.
[This exercise implies, for example, that there is not a norm that makes the vector space of polynomials with coefficients in $\mathbf{F}$ into a Banach space.]

Solution. Suppose V is an infinite-dimensional Banach space with a countable basis in the sense of 6.54. The basis cannot be finite (else V would be finite-dimensional), so relabel it as $e_1, e_2, \ldots$ and put $V_n = \operatorname{span}\{e_1, \ldots, e_n\}$.
(i) $V = \bigcup_{n=1}^{\infty} V_n$, because every $f \in V$ is a *finite* linear combination $\sum_{j \in \Omega} \alpha_j e_j$, so $f \in V_n$ for $n = \max \Omega$.
(ii) Each V_n is closed in V by Exercise 9 in Section 6D.
(iii) Each V_n has empty interior. Otherwise $B(g, \rho) \subseteq V_n$ for some $g \in V_n$ and $\rho > 0$; then for $f \in V$ with $f \neq 0$,

$$\left\| \left(g + \tfrac{\rho}{2\|f\|} f \right) - g \right\| = \frac{\rho}{2} < \rho,$$

so $g + \frac{\rho}{2\|f\|} f \in V_n$, and subtracting $g \in V_n$ and rescaling gives $f \in V_n$. Thus $V = V_n$ is finite-dimensional, a contradiction.
Because V is a Banach space, (V, d) with $d(f, g) = \|f - g\|$ is a complete metric space, so (i)–(iii) exhibit V as a countable union of closed sets with empty interior, contradicting Baire's Theorem [6.76(a)].

Problem 6E.10. Give an example of a Banach space V, a normed vector space W, a bounded linear map T of V onto W, and an open subset G of V such that $T(G)$ is not an open subset of W.
[This exercise shows that the hypothesis in the Open Mapping Theorem that W is a Banach space cannot be relaxed to the hypothesis that W is a normed vector space.]

Solution. Take $V = \ell^1$ with $\|a\|_1 = \sum_{k=1}^{\infty} |a_k|$, take $W = \ell^1$ with $\|a\|_\infty = \sup_k |a_k|$, let $T : V \to W$ be the identity map, and let $G = \{a \in \ell^1 : \|a\|_1 < 1\}$ be the open unit ball of V.
Then V is a Banach space (second bullet of Example 6.38) and W is a normed vector space (second bullet of Example 6.39). The map T is linear and onto W, being the identity on the common underlying vector space, and $|a_k| \leq \sum_j |a_j|$ for each k gives

$$\|Ta\|_\infty \leq \|a\|_1,$$

so T is bounded with $\|T\| \leq 1$.
But $T(G) = G$ contains no $\|\cdot\|_\infty$-ball about 0: given $\varepsilon > 0$, choose $n > 2/\varepsilon$ and let a have n entries equal to $\varepsilon/2$ followed by zeros. Then

$$\|a\|_\infty = \frac{\varepsilon}{2} < \varepsilon, \qquad \|a\|_1 = n \cdot \frac{\varepsilon}{2} > \frac{2}{\varepsilon} \cdot \frac{\varepsilon}{2} = 1,$$

so $a \notin T(G)$. Hence $T(G)$ is not an open subset of W.

Problem 6E.11. Show that there exists a normed vector space V, a Banach space W, a bounded linear map T of V onto W, and an open subset G of V such that $T(G)$ is not an open subset of W.
[This exercise shows that the hypothesis in the Open Mapping Theorem that V is a Banach space cannot be relaxed to the hypothesis that V is a normed vector space.]

Solution. Take $W = \ell^1$ with $\|f\|_1 = \sum_{k=1}^{\infty} |f_k|$, a Banach space by the second bullet of Example 6.38 and infinite-dimensional (the coordinate vectors $e_1, e_2, \ldots$ form an infinite linearly independent family, so no finite list spans ℓ^1). By 6.62 there is a discontinuous linear functional $\varphi : W \to \mathbf{F}$, which by 6.48 means: no $c \in (0, \infty)$ satisfies $|\varphi(f)| \leq c\|f\|_1$ for all $f \in W$. Now take

$$V = \ell^1 \text{ with } \|f\|_V = \|f\|_1 + |\varphi(f)|$$

(a norm, by linearity of φ and because $\|f\|_V = 0$ forces $\|f\|_1 = 0$), let $T : V \to W$ be the identity map,

and let $G = \{f \in V : \|f\|_V < 1\}$.
Then T is linear and onto W, and bounded with $\|T\| \le 1$ since $\|Tf\|_1 \le \|f\|_V$.
Suppose $T(G) = \{f \in \ell^1 : \|f\|_1 + |\varphi(f)| < 1\}$ were open in W. Since $0 \in T(G)$, there would be $\varepsilon > 0$ with $\|f\|_1 < \varepsilon \implies |\varphi(f)| < 1$. Applying this to $g = \frac{\varepsilon}{2\|f\|_1} f$ for $f \neq 0$, which has $\|g\|_1 = \varepsilon/2 < \varepsilon$, and using homogeneity of φ gives

$$|\varphi(f)| \le \frac{2}{\varepsilon}\|f\|_1 \quad \text{for all } f \in W$$

(the case $f = 0$ being trivial), contradicting the unboundedness of φ. Hence $T(G)$ is not open in W.

Problem 6E.12. Suppose $T : V \to W$ is a bounded linear map from a Banach space V to a Banach space W. Prove that T is bounded below if and only if T is injective and the range of T is a closed subspace of W.

Solution. Write $R = T(V)$, a subspace of W by linearity of T.
(i) Suppose $\|f\| \le c\|Tf\|$ for all $f \in V$, where $c \in (0, \infty)$. Then $Tf = 0$ forces $f = 0$, so T is injective. For closedness of R, let $g_k = Tf_k \to g$ in W. The sequence $g_1, g_2, \ldots$ is Cauchy (6.13), and

$$\|f_j - f_k\| \le c\|T(f_j - f_k)\| = c\|g_j - g_k\|$$

makes $f_1, f_2, \ldots$ Cauchy in the Banach space V, say $f_k \to f$. Since T is bounded it is continuous (6.48), so $g = \lim_k Tf_k = Tf \in R$; thus R is closed by 6.9(e).
(ii) Conversely, suppose T is injective with R closed. Being a closed subset of the Banach space W, R is complete [6.16(b)], hence itself a Banach space under the inherited norm. Viewed as a map $V \to R$, T is a bounded linear bijection between Banach spaces, so the Bounded Inverse Theorem (6.83) gives $b \in [0, \infty)$ with $\|T^{-1}g\| \le b\|g\|$ for all $g \in R$. Taking $g = Tf$ yields

$$\|f\| \le b\|Tf\| \le c\|Tf\|, \qquad c := \max\{b, 1\} \in (0, \infty),$$

so T is bounded below.

Problem 6E.13. Give an example of a Banach space V, a normed vector space W, and a one-to-one bounded linear map T of V onto W such that T^{-1} is not a bounded linear map of W onto V.
[This exercise shows that the hypothesis in the Bounded Inverse Theorem (6.83) that W is a Banach space cannot be relaxed to the hypothesis that W is a normed vector space.]

Solution. Take the example of Exercise 10: $V = \ell^1$ with $\|a\|_1 = \sum_{k=1}^{\infty} |a_k|$, a Banach space by the second bullet of Example 6.38; $W = \ell^1$ with $\|a\|_\infty = \sup_k |a_k|$, a normed vector space by the second bullet of Example 6.39; and $T : V \to W$ the identity map, which is linear, one-to-one, onto, and bounded with $\|T\| \le 1$ since $\|Ta\|_\infty \le \|a\|_1$.
Its inverse $T^{-1} : W \to V$ is again the identity, and it is unbounded: for $a^{(n)} = (1, \ldots, 1, 0, 0, \ldots)$ with n ones,

$$\|a^{(n)}\|_\infty = 1, \qquad \|T^{-1}a^{(n)}\|_1 = n,$$

so $\|T^{-1}g\|_1 \le c\|g\|_\infty$ for all $g \in W$ would force $n \le c$ for every n.

Problem 6E.14. Show that there exists a normed space V, a Banach space W, and a one-to-one bounded linear map T of V onto W such that T^{-1} is not a bounded linear map of W onto V.
[This exercise shows that the hypothesis in the Bounded Inverse Theorem (6.83) that V is a Banach space cannot be relaxed to the hypothesis that V is a normed vector space.]

Solution. Take the construction of Exercise 11: $W = \ell^1$ with $\|\cdot\|_1$, a Banach space by the second bullet of Example 6.38 and infinite-dimensional, so 6.62 and 6.48 supply an unbounded linear functional $\varphi : W \to \mathbf{F}$; $V = \ell^1$ with the norm $\|f\|_V = \|f\|_1 + |\varphi(f)|$; and $T : V \to W$ the identity map, which is

linear, one-to-one, onto, and bounded with $\|T\| \le 1$ since $\|Tf\|_1 \le \|f\|_V$.
Its inverse $T^{-1} : W \to V$ is again the identity, and it is unbounded: were $\|T^{-1}g\|_V \le c\|g\|_1$ for all $g \in W$, then

$$|\varphi(f)| \le \|f\|_1 + |\varphi(f)| = \|f\|_V \le c\|f\|_1$$

for all $f \in \ell^1$, making φ bounded.

Problem 6E.15. Prove 6.84.

Solution. Give $V \times W$ the coordinatewise operations and $\|(f,g)\| = \max\{\|f\|, \|g\|\}$.
Norm. Here $\|(f,g)\| = 0$ forces $\|f\| = \|g\| = 0$, and $\|\lambda(f,g)\| = |\lambda| \max\{\|f\|, \|g\|\}$ because multiplication by $|\lambda| \ge 0$ commutes with maxima. For the triangle inequality,

$$\|f_1 + f_2\| \le \|f_1\| + \|f_2\| \le \|(f_1, g_1)\| + \|(f_2, g_2)\|,$$

and the same bound holds for $\|g_1 + g_2\|$, hence for their maximum.
Coordinatewise convergence. Because $(f_k, g_k) - (f, g) = (f_k - f, g_k - g)$,

$$\|(f_k, g_k) - (f, g)\| = \max\{\|f_k - f\|, \|g_k - g\|\},$$

and $a, b \le \max\{a, b\} \le a + b$ for $a, b \ge 0$ shows this tends to 0 precisely when $\|f_k - f\| \to 0$ and $\|g_k - g\| \to 0$.
Completeness. If $(f_1, g_1), (f_2, g_2), \dots$ is Cauchy in $V \times W$, the same display makes $f_1, f_2, \dots$ Cauchy in V and $g_1, g_2, \dots$ Cauchy in W; as V and W are Banach spaces, $f_k \to f$ and $g_k \to g$, so $(f_k, g_k) \to (f, g)$ by the previous paragraph.

Problem 6E.16. Suppose V is a Banach space with norm $\|\cdot\|$ and that $\varphi : V \to \mathbf{F}$ is a linear functional. Define another norm $\|\cdot\|_\varphi$ on V by

$$\|f\|_\varphi = \|f\| + |\varphi(f)|.$$

Prove that if V is a Banach space with the norm $\|\cdot\|_\varphi$, then φ is a continuous linear functional on V (with the original norm).

Solution. Write V_φ for V with the norm $\|\cdot\|_\varphi$, a Banach space by hypothesis, and let $I : V_\varphi \to V$ be the identity map. It is a one-to-one linear map onto V, bounded with $\|I\| \le 1$ because

$$\|If\| = \|f\| \le \|f\| + |\varphi(f)| = \|f\|_\varphi.$$

Both V_φ and V are Banach spaces, so the Bounded Inverse Theorem (6.83) makes I^{-1} bounded; with $c = \|I^{-1}\| < \infty$ this says $\|f\|_\varphi \le c\|f\|$ for all $f \in V$. Hence

$$|\varphi(f)| \le \|f\| + |\varphi(f)| = \|f\|_\varphi \le c\|f\|,$$

so φ is bounded on $(V, \|\cdot\|)$ with $\|\varphi\| \le c$, and therefore continuous by 6.48.
Remark. The converse is also true and easy: if φ is continuous on V with $\|\varphi\| = c'$, then $\|f\| \le \|f\|_\varphi \le (1 + c')\|f\|$, so the two norms are equivalent and V_φ is complete. Thus V_φ is a Banach space if and only if φ is continuous.

Problem 6E.17. Suppose V is a Banach space, W is a normed vector space, and $T_1, T_2, \dots$ is a sequence of bounded linear maps from V to W such that $\lim_{k\to\infty} T_k f$ exists for each $f \in V$. Define $T : V \to W$ by

$$Tf = \lim_{k\to\infty} T_k f$$

for $f \in V$. Prove that T is a bounded linear map from V to W.

[This result states that the pointwise limit of a sequence of bounded linear maps on a Banach space is a bounded linear map.]

Solution. T is linear because limits in a normed vector space respect sums and scalar multiples (the triangle inequality and homogeneity of the norm): $T_k(f+h) = T_kf + T_kh \to Tf + Th$ and $T_k(\lambda f) = \lambda T_k f \to \lambda Tf$, while these same sequences converge to $T(f+h)$ and $T(\lambda f)$, and limits in a metric space are unique.
For boundedness, fix $f \in V$. The convergent sequence $T_1 f, T_2 f, \dots$ is bounded: choosing n with $\|T_k f - Tf\| \le 1$ for $k \ge n$ gives

$$\sup_k \|T_k f\| \le \max\bigl(\{\|T_j f\| : j < n\} \cup \{\|Tf\| + 1\}\bigr) < \infty,$$

a maximum over a finite set. So $\{T_k : k \in \mathbf{Z}^+\}$ is a pointwise bounded family of bounded linear maps from the Banach space V to W, and the Principle of Uniform Boundedness (6.86) gives $M := \sup_k \|T_k\| < \infty$. Since $\|T_k f\| \le M\|f\|$ for each k and $\bigl|\, \|T_k f\| - \|Tf\| \,\bigr| \le \|T_k f - Tf\| \to 0$,

$$\|Tf\| = \lim_{k\to\infty} \|T_k f\| \le M\|f\|,$$

so T is bounded with $\|T\| \le \sup_k \|T_k\|$.

Problem 6E.18. Suppose that V is a normed vector space and B is a subset of V such that $\sup_{f\in B} |\varphi(f)| < \infty$ for every $\varphi \in V'$. Prove that $\sup_{f\in B} \|f\| < \infty$.

Solution. Run the Principle of Uniform Boundedness on V' rather than on V: by 6.71 and 6.47, $V' = \mathcal{B}(V, \mathbf{F})$ is a Banach space because $\mathbf{F}$ is complete, whereas V is only normed.
For $f \in V$ define $\widehat{f} : V' \to \mathbf{F}$ by $\widehat{f}(\varphi) = \varphi(f)$, which is linear because the operations on V' are pointwise. Then $|\widehat{f}(\varphi)| = |\varphi(f)| \le \|\varphi\|\,\|f\|$ puts $\widehat{f} \in (V')'$ with $\|\widehat{f}\| \le \|f\|$; conversely $\|\widehat{f}\| \ge \|f\|$, trivially for $f = 0$ and otherwise by taking the $\varphi \in V'$ that 6.72 supplies with $\|\varphi\| = 1$ and $\varphi(f) = \|f\|$. Hence $\|\widehat{f}\| = \|f\|$.
The hypothesis says exactly that the family $\{\widehat{f} : f \in B\}$ of bounded linear maps from V' to $\mathbf{F}$ is pointwise bounded, since

$$\sup_{f\in B} |\widehat{f}(\varphi)| = \sup_{f\in B} |\varphi(f)| < \infty$$

for each $\varphi \in V'$. So 6.86 gives $\sup_{f\in B} \|\widehat{f}\| < \infty$, that is, $\sup_{f\in B} \|f\| < \infty$.

Problem 6E.19. Suppose $T : V \to W$ is a linear map from a Banach space V to a Banach space W such that

$$\varphi \circ T \in V' \quad \text{for all } \varphi \in W'.$$

Prove that T is a bounded linear map.

Solution. Apply the Closed Graph Theorem (6.85): V and W are Banach spaces, and $\operatorname{graph}(T) = \{(f, Tf) : f \in V\}$ is a subspace of $V \times W$ by 6.68(a), so it suffices to show the graph is closed in the norm $\|(f,g)\| = \max\{\|f\|, \|g\|\}$ of 6.84.
Suppose $(f_k, Tf_k) \to (f, g)$, which by the convergence criterion of 6.84 means $f_k \to f$ in V and $Tf_k \to g$ in W. Fix $\varphi \in W'$. Being bounded, φ is continuous (6.48), so $\varphi(Tf_k) \to \varphi(g)$; and $\varphi \circ T \in V'$ is likewise continuous, so $\varphi(Tf_k) \to \varphi(Tf)$. Uniqueness of limits in $\mathbf{F}$ gives

$$\varphi(g - Tf) = 0 \qquad \text{for every } \varphi \in W'.$$

Were $g - Tf \ne 0$, then 6.72 would supply $\varphi \in W'$ with $\varphi(g - Tf) = \|g - Tf\| > 0$. Hence $g = Tf$, so $(f, g) \in \operatorname{graph}(T)$ and the graph is closed by 6.9(e); 6.85 now makes T bounded.

7 L^p Spaces

Contents

7.1 Exercises 7A

Problem 7A.1. Suppose μ is a measure. Prove that

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty \qquad \text{and} \qquad \|\alpha f\|_\infty = |\alpha|\,\|f\|_\infty$$

for all $f, g \in L^\infty(\mu)$ and all $\alpha \in \mathbf{F}$. Conclude that with the usual operations of addition and scalar multiplication of functions, $L^\infty(\mu)$ is a vector space.

Solution. Everything follows from the fact that the infimum in 7.1 is attained. Indeed, put $c = \|f\|_\infty < \infty$; for each n there is $t_n < c + \frac{1}{n}$ with $\mu(\{|f| > t_n\}) = 0$, whence $\mu(\{|f| > c + \frac{1}{n}\}) = 0$, and countable subadditivity (2.58) applied to

$$\{|f| > c\} = \bigcup_{n=1}^{\infty} \{|f| > c + \tfrac{1}{n}\}$$

gives $\mu(\{|f| > \|f\|_\infty\}) = 0$; that is, $|f| \le \|f\|_\infty$ almost everywhere.
Triangle inequality. For $f, g \in L^\infty(\mu)$ the set

$$E = \{|f| > \|f\|_\infty\} \cup \{|g| > \|g\|_\infty\}$$

has $\mu(E) = 0$ by the above and subadditivity, and off E we have $|f+g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty$. So every $t > \|f\|_\infty + \|g\|_\infty$ has $\{|f+g| > t\} \subseteq E$ and hence measure 0; taking the infimum over such t,

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty < \infty.$$

Since $f+g$ is $\mathcal{S}$-measurable (2.46, applied to real and imaginary parts when $\mathbf{F} = \mathbf{C}$), $f + g \in L^\infty(\mu)$.
Homogeneity. For $\alpha = 0$ both sides are 0. For $\alpha \ne 0$ and every $t > 0$ we have $\{|\alpha f| > t\} = \{|f| > t/|\alpha|\}$, so writing $A_f = \{t > 0 : \mu(\{|f| > t\}) = 0\}$ we get $A_{\alpha f} = |\alpha| A_f$, and multiplication by $|\alpha| > 0$ commutes with infima:

$$\|\alpha f\|_\infty = \inf A_{\alpha f} = |\alpha| \inf A_f = |\alpha|\,\|f\|_\infty$$

(both sides being ∞ when $A_f = \varnothing$). With αf measurable (2.46), $\alpha f \in L^\infty(\mu)$.
As $0 \in L^\infty(\mu)$, these two closure properties make $L^\infty(\mu)$ a subspace of $\mathbf{F}^X$, hence a vector space.

Problem 7A.2. Suppose $a \ge 0$, $b \ge 0$, and $1 < p < \infty$. Prove that

$$ab = \frac{a^p}{p} + \frac{b^{p'}}{p'}$$

if and only if $a^p = b^{p'}$ [compare to Young's inequality (7.8)].

Solution. By 7.6, $p' = \frac{p}{p-1}$, so $\frac{p'}{p} = p' - 1 = \frac{1}{p-1}$.
(i) $b = 0$. The equation reads $0 = \frac{a^p}{p}$ and the condition reads $a^p = 0$; both say $a = 0$.
(ii) $b > 0$. As in the proof of Young's inequality (7.8), set

$$f(a) = \frac{a^p}{p} + \frac{b^{p'}}{p'} - ab,$$

so the equation in question says $f(a) = 0$. Here $f'(a) = a^{p-1} - b$, and since $a \mapsto a^{p-1}$ is strictly increasing on $(0, \infty)$ this is negative for $0 < a < a_0$ and positive for $a > a_0$, where $a_0 = b^{1/(p-1)} > 0$. Hence $f(a) > f(a_0)$ for all $a \ge 0$ with $a \ne a_0$. Because $a_0^p = b^{p/(p-1)} = b^{p'}$ and $a_0 b = b^{\frac{1}{p-1}+1} = b^{p'}$,

$$f(a_0) = b^{p'}\Big(\frac{1}{p} + \frac{1}{p'}\Big) - b^{p'} = 0,$$

so $f(a) = 0$ if and only if $a = a_0$. Finally $a = b^{1/(p-1)}$ gives $a^p = b^{p/(p-1)} = b^{p'}$, and conversely $a^p = b^{p'} > 0$ gives $a = b^{p'/p} = b^{1/(p-1)}$.

Problem 7A.3. Suppose $a_1, \dots, a_n$ are nonnegative numbers. Prove that

$$(a_1 + \cdots + a_n)^5 \le n^4\big(a_1^5 + \cdots + a_n^5\big).$$

Solution. Apply Hölder's inequality (7.9) with $p = 5$ and $p' = 5/4$ (by 7.6) on $X = \{1, \dots, n\}$ with counting measure, where every function is measurable and integration is summation (3.6 and 3.7). Take $f(k) = a_k$ and $h(k) = 1$, so that

$$\|fh\|_1 = \sum_{k=1}^n a_k, \qquad \|h\|_{5/4} = n^{4/5}, \qquad \|f\|_5 = \Big(\sum_{k=1}^n a_k^5\Big)^{1/5},$$

all finite. Hölder's inequality then gives

$$\sum_{k=1}^n a_k \le n^{4/5}\Big(\sum_{k=1}^n a_k^5\Big)^{1/5},$$

and raising both nonnegative sides to the fifth power yields the claim.
Method (2): $t \mapsto t^5$ is convex on $[0, \infty)$ (its second derivative $20t^3$ is nonnegative), so the finite form of Jensen's inequality with equal weights $\frac{1}{n}$ gives

$$\Big(\frac{a_1 + \cdots + a_n}{n}\Big)^5 \le \frac{a_1^5 + \cdots + a_n^5}{n};$$

multiply by n^5.

Problem 7A.4. Prove Hölder's inequality (7.9) in the cases $p = 1$ and $p = \infty$.

Solution. By 7.7 we have $1' = \infty$ and $\infty' = 1$, and $|fh| = |hf|$, so the case $p = \infty$ is the case $p = 1$ with f and h interchanged. Thus it suffices to prove $\|fh\|_1 \le \|f\|_1\|h\|_\infty$. Two facts recur: $\mu(E) = 0$ forces $\int u\,d\mu = 0$ for any $u \ge 0$ vanishing off E, because $u \le \infty \cdot \chi_E$ and 3.8 with 3.7 give $\int u\,d\mu \le \infty \cdot \mu(E) = 0$; and $\mu(\{|h| > \|h\|_\infty\}) = 0$ whenever $\|h\|_\infty < \infty$, by Exercise 1 in this section.
(i) $\|f\|_1 = 0$. With $E_n = \{|f| > \frac{1}{n}\}$, the bound $\frac{1}{n}\chi_{E_n} \le |f|$ and 3.8, 3.7 give $\frac{1}{n}\mu(E_n) \le 0$, so $\mu(E_n) = 0$; as $\{f \ne 0\} = \bigcup_n E_n$, countable subadditivity (2.58) makes $\mu(\{f \ne 0\}) = 0$. Since $|fh|$ vanishes off that set, $\|fh\|_1 = 0$.
(ii) $\|h\|_\infty = 0$. Then $\mu(\{|h| > 0\}) = 0$, and $|fh|$ vanishes off that set, so $\|fh\|_1 = 0$.
(iii) $\|f\|_1 > 0$, $\|h\|_\infty > 0$, and one of them is ∞. Then the right side is ∞.
(iv) $0 < \|f\|_1 < \infty$ and $0 < \|h\|_\infty < \infty$. Put $E = \{|h| > \|h\|_\infty\}$, so $\mu(E) = 0$. Pointwise

$$|fh| \le \|h\|_\infty |f| + \infty \cdot \chi_E$$

(on E the right side is ∞; off E we have $|h| \le \|h\|_\infty$), so 3.8, additivity 3.16, homogeneity 3.20, and 3.7 give

$$\|fh\|_1 \le \|h\|_\infty \|f\|_1 + \infty \cdot \mu(E) = \|f\|_1\|h\|_\infty.$$

Problem 7A.5. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $1 < p < \infty$, $f \in L^p(\mu)$, and $h \in L^{p'}(\mu)$. Prove that Hölder's inequality (7.9) is an equality if and only if there exist nonnegative numbers a and b, not both 0, such that

$$a|f(x)|^p = b|h(x)|^{p'}$$

for almost every $x \in X$.

Solution. Both $\|f\|_p$ and $\|h\|_{p'}$ are finite, and three facts from Chapter 3 are used throughout.
(F1) If $u \ge 0$ is measurable with $\int u\,d\mu = 0$, then $u = 0$ almost everywhere: with $E_n = \{u > \frac{1}{n}\}$, the

bound $\frac{1}{n}\chi_{E_n} \leq u$ and 3.8, 3.7 give $\mu(E_n) = 0$, and $\{u > 0\} = \bigcup_n E_n$ is null by 2.58.
(F2) Nonnegative measurable functions agreeing outside a null set E have equal integrals, since $u \leq v + \infty \cdot \chi_E$ with 3.8, 3.16, 3.7 gives $\int u\, d\mu \leq \int v\, d\mu$, and symmetrically.
(F3) Hence $\|f\|_p = 0$ if and only if $f = 0$ almost everywhere.
(i) Suppose $a|f|^p = b|h|^{p'}$ almost everywhere, with $a, b \geq 0$ not both 0. If $b = 0$ then $a > 0$, so $f = 0$ almost everywhere and both sides of Hölder's inequality vanish by (F2) and (F3); the case $a = 0$ is symmetric. So let $a, b > 0$ and $c = b/a$, giving $|f|^p = c|h|^{p'}$ almost everywhere and hence $\|f\|_p^p = c\|h\|_{p'}^{p'}$ by (F2). If $\|h\|_{p'} = 0$ both sides vanish again, so assume $\|h\|_{p'} > 0$; then $\|f\|_p > 0$ too. Put $F = |f|/\|f\|_p$ and $H = |h|/\|h\|_{p'}$, so that $\int F^p\, d\mu = \int H^{p'}\, d\mu = 1$ and, almost everywhere,

$$F^p = \frac{|f|^p}{\|f\|_p^p} = \frac{c\,|h|^{p'}}{c\,\|h\|_{p'}^{p'}} = H^{p'}.$$

By Exercise 2 in this section, Young's inequality is therefore an equality almost everywhere, and integrating

$$FH = \frac{F^p}{p} + \frac{H^{p'}}{p'}$$

via (F2), 3.16, and 3.20 gives $\int FH\, d\mu = \frac{1}{p} + \frac{1}{p'} = 1$. Since $|fh| = \|f\|_p\|h\|_{p'}FH$, homogeneity yields $\|fh\|_1 = \|f\|_p\|h\|_{p'}$.
(ii) Conversely, suppose $\|fh\|_1 = \|f\|_p\|h\|_{p'}$. If $\|f\|_p = 0$ take $a = 1, b = 0$, and if $\|h\|_{p'} = 0$ take $a = 0, b = 1$; in each case (F3) makes both sides vanish almost everywhere. Otherwise both norms lie in $(0, \infty)$; with F, H as above, $\int F^p\, d\mu = \int H^{p'}\, d\mu = 1$ and homogeneity turns the assumed equality into $\int FH\, d\mu = 1$. The function

$$g = \frac{F^p}{p} + \frac{H^{p'}}{p'} - FH$$

is nonnegative by Young's inequality (7.8), and $g + FH = \frac{F^p}{p} + \frac{H^{p'}}{p'}$ with 3.16 and 3.20 gives $\int g\, d\mu + \int FH\, d\mu = \frac{1}{p} + \frac{1}{p'} = 1$; subtracting the finite quantity $\int FH\, d\mu = 1$ leaves $\int g\, d\mu = 0$, so $g = 0$ almost everywhere by (F1). By Exercise 2 this forces $F^p = H^{p'}$ almost everywhere, that is,

$$a|f|^p = b|h|^{p'} \quad \text{almost everywhere, where } a = \|h\|_{p'}^{p'} > 0,\ b = \|f\|_p^p > 0.$$

Problem 7A.6. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f \in L^1(\mu)$, and $h \in L^\infty(\mu)$. Prove that $\|fh\|_1 = \|f\|_1\|h\|_\infty$ if and only if

$$|h(x)| = \|h\|_\infty$$

for almost every $x \in X$ such that $f(x) \neq 0$.

Solution. Write $M = \|h\|_\infty < \infty$ and $N = \|f\|_1 < \infty$, and use facts (F1) and (F2) from Exercise 5 in this section.
Reduction. The set $E_0 = \{|h| > M\}$ is null by Exercise 1 in this section, so replacing h by $h\chi_{X\setminus E_0}$ changes neither $\|h\|_\infty$ (the essential supremum is unaffected by modification on a null set; see the discussion following 7.1), nor $\|fh\|_1$ (by (F2)), nor the measure of $\{f \neq 0 \text{ and } |h| \neq M\}$. So assume $|h| \leq M$ everywhere.
Then $u := (M - |h|)|f|$ is nonnegative and measurable with $u + |fh| = M|f|$ pointwise, so additivity (3.16) and homogeneity (3.20) give

$$\int u\, d\mu + \|fh\|_1 = M\|f\|_1 = MN.$$

Hölder's inequality (7.9) for $p = 1$, proved in Exercise 4 of this section, makes $\|fh\|_1 \leq MN < \infty$, so that term may be subtracted:

$$\int u\, d\mu = \|f\|_1\|h\|_\infty - \|fh\|_1.$$

Hence $\|fh\|_1 = \|f\|_1\|h\|_\infty$ if and only if $\int u\,d\mu = 0$, which by (F1) holds if and only if $u = 0$ almost everywhere. Because $|h| \le M$,

$$\{u > 0\} = \{f \ne 0 \text{ and } |h| < M\} = \{f \ne 0 \text{ and } |h| \ne M\},$$

so that condition says exactly that $|h(x)| = \|h\|_\infty$ for almost every $x \in X$ with $f(x) \ne 0$.

Problem 7A.7. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, h : X \to \mathbf{F}$ are $\mathcal{S}$-measurable. Prove that

$$\|fh\|_r \le \|f\|_p\,\|h\|_q$$

for all positive numbers p, q, r such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$.

Solution. Apply Hölder's inequality (7.9) to $|f|^r$ and $|h|^r$ with exponent $P = p/r$. Here $1 < P < \infty$, since $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} > \frac{1}{p}$ forces $r < p$, and $Q := q/r$ is the dual exponent of P because

$$\frac{1}{P} + \frac{1}{Q} = r\Big(\frac{1}{p} + \frac{1}{q}\Big) = 1,$$

so $Q = P'$ by 7.6. As $|f|^r$ and $|h|^r$ are measurable by 6.20, 7.9 gives

$$\big\|\, |f|^r |h|^r \,\big\|_1 \le \big\|\, |f|^r \,\big\|_P \,\big\|\, |h|^r \,\big\|_Q.$$

By 7.1 the left side is $\int |fh|^r\,d\mu = \|fh\|_r^r$, while $rP = p$ gives

$$\big\|\, |f|^r \,\big\|_P = \Big(\int |f|^p\,d\mu\Big)^{r/p} = \|f\|_p^r,$$

and likewise $\big\|\, |h|^r \,\big\|_Q = \|h\|_q^r$ from $rQ = q$. Hence

$$\|fh\|_r^r \le \big(\|f\|_p\,\|h\|_q\big)^r$$

(using the convention $0 \cdot \infty = 0$ under which 7.9 is stated), and taking r-th roots, which preserves order on $[0, \infty]$, completes the proof.

Problem 7A.8. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $n \in \mathbb{Z}^+$. Prove that

$$\|f_1 f_2 \cdots f_n\|_1 \le \|f_1\|_{p_1}\|f_2\|_{p_2} \cdots \|f_n\|_{p_n}$$

for all positive numbers $p_1, \dots, p_n$ such that $\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n} = 1$ and all $\mathcal{S}$-measurable functions $f_1, f_2, \dots, f_n : X \to \mathbf{F}$.

Solution. Induct on n, with Exercise 7 in this section as the two-factor step. The statement to prove is: if $p_1, \dots, p_n, r$ are positive numbers with $\sum_{k=1}^n \frac{1}{p_k} = \frac{1}{r}$, then

$$\|f_1 f_2 \cdots f_n\|_r \le \|f_1\|_{p_1}\|f_2\|_{p_2} \cdots \|f_n\|_{p_n}.$$

For $n = 1$ we have $r = p_1$ and the two sides agree. Assume the statement for n, and let $p_1, \dots, p_{n+1}, r$ be positive with $\sum_{k=1}^{n+1} \frac{1}{p_k} = \frac{1}{r}$. Define s by $\frac{1}{s} = \sum_{k=1}^n \frac{1}{p_k}$, a positive finite number, so that $\frac{1}{s} + \frac{1}{p_{n+1}} = \frac{1}{r}$. The induction hypothesis bounds $\|f_1 \cdots f_n\|_s$, and Exercise 7 applied to $f_1 \cdots f_n$ and f_{n+1} gives

$$\|f_1 \cdots f_n f_{n+1}\|_r \le \|f_1 \cdots f_n\|_s\,\|f_{n+1}\|_{p_{n+1}} \le \prod_{k=1}^{n+1} \|f_k\|_{p_k},$$

the second inequality because multiplication by a fixed element of $[0, \infty]$ preserves order under the convention $0 \cdot \infty = 0$ used in 7.9. Taking $r = 1$ yields the exercise.

Problem 7A.9. Show that the formula in 7.12 holds for $p = \infty$ if μ is a σ-finite measure.

Solution. With $p = \infty$ and $p' = 1$ (7.6 and 7.7), the formula in 7.12 reads

$$\|f\|_\infty = \sup\Big\{\int fh\,d\mu : h \in \mathcal{L}^1(\mu) \text{ and } \|h\|_1 \le 1\Big\}.$$

For any admissible h, Hölder's inequality (7.9) with $p = \infty$, $p' = 1$ gives

$$\int fh\,d\mu \le \int |fh|\,d\mu = \|fh\|_1 \le \|f\|_\infty \|h\|_1 \le \|f\|_\infty,$$

so the supremum is at most $\|f\|_\infty$; as $h = 0$ is admissible the supremum is also at least 0, settling the case $\|f\|_\infty = 0$. Assume now $\|f\|_\infty > 0$ and fix $t \in (0, \|f\|_\infty)$.
Then $E = \{|f| > t\} \in \mathcal{S}$ has $\mu(E) > 0$, since t lies below the infimum in 7.1. Writing $X = \bigcup_{j=1}^\infty X_j$ with each $\mu(X_j) < \infty$ by σ-finiteness, countable subadditivity (2.58) rules out $\mu(E \cap X_j) = 0$ for every j, so $A := E \cap X_j$ satisfies $0 < \mu(A) < \infty$ for some j. Put

$$h = \frac{\overline{f}}{\mu(A)\,|f|}\,\chi_A,$$

which is defined because $|f| > t > 0$ on A. Then $|h| = \frac{1}{\mu(A)}\chi_A$, so $\|h\|_1 = 1$ by 3.15 and h is admissible; moreover $fh = \frac{|f|}{\mu(A)}\chi_A \ge \frac{t}{\mu(A)}\chi_A$ pointwise, so 3.8 and 3.15 give

$$\int fh\,d\mu \ge \frac{t\,\mu(A)}{\mu(A)} = t.$$

Hence the supremum is at least t for every $t < \|f\|_\infty$, and the two bounds give equality.

Problem 7A.10. Suppose $0 < p < q \le \infty$.
(a) Prove that $\ell^p \subseteq \ell^q$.
(b) Prove that $\|(a_1, a_2, \ldots)\|_p \ge \|(a_1, a_2, \ldots)\|_q$ for every sequence $a_1, a_2, \ldots$ of elements of $\mathbf{F}$.

Solution. Here μ is counting measure on $\mathbf{Z}^+$, so by 7.2 and 7.4 we have $\|a\|_p = (\sum_k |a_k|^p)^{1/p}$ for $p < \infty$, $\|a\|_\infty = \sup_k |a_k|$, and $\ell^p = \{a : \|a\|_p < \infty\}$.
(a) Immediate from (b): if $a \in \ell^p$ then $\|a\|_q \le \|a\|_p < \infty$, so $a \in \ell^q$.
(b) If $\|a\|_p = \infty$ there is nothing to prove, and if $\|a\|_p = 0$ then every $a_k = 0$ and both sides vanish. So assume $0 < \|a\|_p < \infty$ and put $b = a/\|a\|_p$, which has $\|b\|_p = 1$ by homogeneity (7.5(b)); thus $\sum_{k=1}^\infty |b_k|^p = 1$, and in particular $|b_k| \le 1$ for every k.
(i) $q = \infty$. Then $\|b\|_\infty = \sup_k |b_k| \le 1$.
(ii) $q < \infty$. Then $0 \le |b_k| \le 1$ and $q > p$ give $|b_k|^q \le |b_k|^p$, so

$$\|b\|_q^q = \sum_{k=1}^\infty |b_k|^q \le \sum_{k=1}^\infty |b_k|^p = 1.$$

Either way $\|b\|_q \le 1$, so homogeneity of the q-norm (7.5(b), or Exercise 1 in this section when $q = \infty$) gives

$$\|a\|_q = \big\|\, \|a\|_p\, b \,\big\|_q = \|a\|_p\, \|b\|_q \le \|a\|_p.$$

Problem 7A.11. Show that

$$\bigcap_{p>1} \ell^p \ne \ell^1.$$

Solution. Take $a = (a_1, a_2, \ldots)$ with $a_k = \frac{1}{k}$. Since $\ell^1 \subseteq \ell^p$ for every $p > 1$ by Exercise 10(a) in this section, it suffices to show a lies in every ℓ^p with $p > 1$ but not in ℓ^1.
Not in ℓ^1. The block $\{2^m, \ldots, 2^{m+1} - 1\}$ has 2^m terms, each at least $2^{-(m+1)}$, so it contributes at least $\frac{1}{2}$; summing the blocks $m = 0, \ldots, M - 1$,

$$\sum_{k=1}^{2^M - 1} \frac{1}{k} \geq \frac{M}{2}.$$

The partial sums are unbounded, so $\|a\|_1 = \infty$.
In ℓ^p for $p > 1$. For $p = \infty$ this is clear, as $\sup_k \frac{1}{k} = 1$. For $1 < p < \infty$ the same block has 2^m terms, each at most 2^{-mp}, so it contributes at most $(2^{1-p})^m$; since $0 < 2^{1-p} < 1$, summing the geometric series bounds the partial sums of the nonnegative series:

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \leq \sum_{m=0}^{\infty} \left(2^{1-p}\right)^m = \frac{1}{1 - 2^{1-p}} < \infty.$$

Problem 7A.12. Show that

$$\bigcap_{p<\infty} \mathcal{L}^p([0,1]) \neq \mathcal{L}^\infty([0,1]).$$

Solution. Take $f = \sum_{n=1}^{\infty} n\, \chi_{A_n}$ on $[0,1]$ (with $f(0) = 0$), where $A_n = (2^{-n}, 2^{-n+1}]$ are disjoint Borel sets with $\lambda(A_n) = 2^{-n}$; here f is Borel measurable because $f^{-1}((t, \infty))$ is the union of those A_n with $n > t$, together with $\{0\}$ when $t < 0$.
The containment. If $g \in \mathcal{L}^\infty([0,1])$ and $0 < p < \infty$, then $N = \{|g| > \|g\|_\infty\}$ is null by Exercise 1 in this section, so $|g|^p \leq \|g\|_\infty^p \chi_{[0,1]} + \infty \cdot \chi_N$ pointwise and 3.8, 3.16, 3.15 give

$$\int |g|^p \, d\lambda \leq \|g\|_\infty^p \lambda([0,1]) + \infty \cdot \lambda(N) = \|g\|_\infty^p < \infty.$$

Thus $\mathcal{L}^\infty([0,1]) \subseteq \bigcap_{p<\infty} \mathcal{L}^p([0,1])$, and it suffices to place f in the right side but not the left.
$f \notin \mathcal{L}^\infty([0,1])$. Given $t > 0$, any integer $n > t$ gives $\lambda(\{|f| > t\}) \geq \lambda(A_n) > 0$, so the set whose infimum defines $\|f\|_\infty$ in 7.1 is empty and $\|f\|_\infty = \infty$.
$f \in \mathcal{L}^p([0,1])$ for $0 < p < \infty$. Disjointness gives $|f|^p = \sum_n n^p \chi_{A_n}$, whose increasing partial sums are simple with integrals $\sum_{n \leq N} n^p 2^{-n}$ by 3.15, so the Monotone Convergence Theorem (3.11) gives

$$\int |f|^p \, d\lambda = \sum_{n=1}^{\infty} \frac{n^p}{2^n} < \infty,$$

the series converging because $c_n = n^p 2^{-n}$ satisfies $\frac{c_{n+1}}{c_n} = \frac{1}{2}(1 + \frac{1}{n})^p \to \frac{1}{2}$, hence $c_{n+1} \leq \frac{3}{4} c_n$ for large n, which dominates the tail by a geometric series.

Problem 7A.13. Show that

$$\bigcup_{p>1} \mathcal{L}^p([0,1]) \neq \mathcal{L}^1([0,1]).$$

Solution. Take $f = \sum_{n=1}^{\infty} \frac{2^n}{n^2} \chi_{A_n}$ on $[0,1]$, with $A_n = (2^{-n}, 2^{-n+1}]$ as in Exercise 12 in this section: disjoint Borel sets with $\lambda(A_n) = 2^{-n}$, and $f(0) = 0$; as there, f is Borel measurable because $f^{-1}((t, \infty))$ is a union of A_n's together with $\{0\}$ when $t < 0$.
The containment $\bigcup_{p>1} \mathcal{L}^p([0,1]) \subseteq \mathcal{L}^1([0,1])$ holds by 7.10 when $1 < p < \infty$, since $\lambda_{[0,1]}$ is a finite measure, and by the first paragraph of Exercise 12 when $p = \infty$. So it suffices that f lie in $\mathcal{L}^1([0,1])$ but in no $\mathcal{L}^p([0,1])$ with $p > 1$. Exactly as in Exercise 12 (disjointness, 3.15, and the Monotone

Convergence Theorem 3.11), for $0 < p < \infty$,

$$\int |f|^p \, d\lambda = \sum_{n=1}^{\infty} \Big(\frac{2^n}{n^2}\Big)^p \lambda(A_n) = \sum_{n=1}^{\infty} \frac{2^{n(p-1)}}{n^{2p}}.$$

(i) $p = 1$. The sum is $\sum_n \frac{1}{n^2} < \infty$, since $\frac{1}{n^2} \le \frac{1}{n-1} - \frac{1}{n}$ for $n \ge 2$ bounds the partial sums by 2. Hence $f \in \mathcal{L}^1([0,1])$.
(ii) $1 < p < \infty$. Choose an integer $m > 2p$; then $e^t \ge t^m/m!$ with $t = n(p-1)\ln 2$ gives

$$\frac{2^{n(p-1)}}{n^{2p}} \ge \frac{((p-1)\ln 2)^m}{m!} n^{m-2p} \longrightarrow \infty,$$

so the terms of the series do not tend to 0 and $\int |f|^p \, d\lambda = \infty$.
(iii) $p = \infty$. The same device with $m = 3$ and $t = n \ln 2$ gives $2^n \ge (\ln 2)^3 n^3/6$, so $2^n/n^2 \to \infty$ and every $t > 0$ admits n with $\lambda(\{|f| > t\}) \ge \lambda(A_n) > 0$; hence $\|f\|_\infty = \infty$.

Problem 7A.14. Suppose $p, q \in (0, \infty]$, with $p \ne q$. Prove that neither of the sets $\mathcal{L}^p(\mathbb{R})$ and $\mathcal{L}^q(\mathbb{R})$ is a subset of the other.

Solution. By symmetry assume $p < q$, so $p \in (0, \infty)$; put $r = q$ if $q < \infty$ and $r = p + 1$ if $q = \infty$, and take

$$f = x^{-1/r} \chi_{(0,1)}, \qquad g = x^{-1/p} \chi_{(1,\infty)}.$$

Each is continuous and nonnegative on the interval carrying it, so its Lebesgue integrals are the improper Riemann integrals

$$\int |f|^s \, d\lambda = \int_0^1 x^{-s/r} \, dx, \qquad \int |g|^s \, d\lambda = \int_1^\infty x^{-s/p} \, dx,$$

the first finite exactly when $s < r$, the second exactly when $s > p$.
(i) $f \in \mathcal{L}^p(\mathbb{R}) \setminus \mathcal{L}^q(\mathbb{R})$. Since $p < r$, $\|f\|_p < \infty$. If $q < \infty$, then $r = q$ and $\int |f|^q \, d\lambda = \int_0^1 dx/x = \infty$. If $q = \infty$, then for every $t > 0$ the set $\{|f| > t\}$ contains an interval $(0, \delta)$ of positive measure, so $\|f\|_\infty = \infty$.
(ii) $g \in \mathcal{L}^q(\mathbb{R}) \setminus \mathcal{L}^p(\mathbb{R})$. Here $\int |g|^p \, d\lambda = \int_1^\infty dx/x = \infty$. If $q < \infty$, then $q > p$ gives $\|g\|_q < \infty$; if $q = \infty$, then $|g| \le 1$.
Thus $\mathcal{L}^p(\mathbb{R}) \not\subseteq \mathcal{L}^q(\mathbb{R})$ and $\mathcal{L}^q(\mathbb{R}) \not\subseteq \mathcal{L}^p(\mathbb{R})$.

Problem 7A.15. Show that there exists $f \in \mathcal{L}^2(\mathbb{R})$ such that $f \notin \mathcal{L}^p(\mathbb{R})$ for all $p \in (0, \infty] \setminus \{2\}$.

Solution. Take $f(x) = 1/\big(\sqrt{x}\,|\ln x|\big)$ for $x \in \big(0, \frac{1}{2}\big) \cup (2, \infty)$ and $f(x) = 0$ otherwise; on each of those intervals f is continuous and nonnegative, so its Lebesgue integrals are the improper Riemann integrals below.
(i) $p = 2$. The substitution $u = \ln x$, $du = dx/x$, gives

$$\|f\|_2^2 = \int_0^{1/2} \frac{dx}{x(\ln x)^2} + \int_2^\infty \frac{dx}{x(\ln x)^2} = \int_{-\infty}^{-\ln 2} \frac{du}{u^2} + \int_{\ln 2}^\infty \frac{du}{u^2} = \frac{2}{\ln 2},$$

so $f \in \mathcal{L}^2(\mathbb{R})$.
(ii) $0 < p < 2$; put $q = p/2 \in (0,1)$. On $(2, \infty)$ we have $|f|^p = x^{-q} (\ln x)^{-p}$, and $(\ln x)^p \le x^{(1-q)/2}$ for all $x \ge M$, some $M \ge 2$, whence

$$|f(x)|^p \ge x^{-(1+q)/2} \quad (x \ge M), \qquad \tfrac{1+q}{2} < 1,$$

so $\int |f|^p \, d\lambda = \infty$.
(iii) $2 < p < \infty$; put $q = p/2 > 1$ and fix $r \in (1, q)$. On $\big(0, \frac{1}{2}\big)$ we have $|f|^p = x^{-q} |\ln x|^{-p}$, and

$x^{q-r}|\ln x|^p \to 0$ as $x \downarrow 0$ (a positive power of x beats any power of $|\ln x|$), so $|\ln x|^p \le x^{-(q-r)}$ on some $(0,\delta)$; there $|f(x)|^p \ge x^{-r}$ with $r>1$, so $\int |f|^p \, d\lambda = \infty$.
(iv) $p=\infty$. Since $\sqrt{x}\,|\ln x| \to 0$ as $x \downarrow 0$, each set $\{|f|>t\}$ contains an interval $(0,\delta_t)$ of positive measure, so the infimum in 7.1 is over the empty set and $\|f\|_\infty = \infty$.

Problem 7A.16. Suppose $(X,\mathcal{S},\mu)$ is a finite measure space. Prove that

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$$

for every $\mathcal{S}$-measurable function $f\colon X \to \mathbf{F}$.

Solution. The two bounds $\|f\|_p \le \|f\|_\infty \mu(X)^{1/p}$ and $\|f\|_p \ge t\,\mu(\{|f|>t\})^{1/p}$ give the limit, since both measure factors tend to 1. In detail: if $\mu(X)=0$, then every set is null, so $\|f\|_p = 0 = \|f\|_\infty$ for all p; assume henceforth $0<\mu(X)<\infty$.
(i) $\limsup_{p\to\infty} \|f\|_p \le \|f\|_\infty$. Only $\|f\|_\infty < \infty$ needs proof. Then $|f| \le \|f\|_\infty$ almost everywhere, because

$$\{|f| > \|f\|_\infty\} = \bigcup_{n=1}^{\infty} \{|f| > \|f\|_\infty + \tfrac{1}{n}\}$$

and each set on the right is null by the definition of the infimum in 7.1; countable subadditivity finishes it. Hence $\int |f|^p \, d\mu \le \|f\|_\infty^p \, \mu(X)$, so $\|f\|_p \le \|f\|_\infty \mu(X)^{1/p}$, and $\mu(X)^{1/p} \to 1$ because $0<\mu(X)<\infty$.
(ii) $\liminf_{p\to\infty} \|f\|_p \ge \|f\|_\infty$. Only $\|f\|_\infty > 0$ needs proof. Fix $t \in (0, \|f\|_\infty)$; then t fails the defining condition in 7.1, so $c := \mu(\{|f|>t\})$ satisfies $0 < c \le \mu(X) < \infty$, and

$$\int |f|^p \, d\mu \ge t^p c \qquad \text{for every } p \in (0,\infty).$$

Thus $\|f\|_p \ge t\,c^{1/p} \to t$, giving $\liminf_{p\to\infty} \|f\|_p \ge t$; now let $t \uparrow \|f\|_\infty$.

Problem 7A.17. Suppose μ is a measure, $0 < p \le \infty$, and $f \in \mathcal{L}^p(\mu)$. Prove that for every $\varepsilon > 0$, there exists a simple function $g \in \mathcal{L}^p(\mu)$ such that $\|f-g\|_p < \varepsilon$.
[This exercise extends 3.44.]

Solution. Take g from the simple approximations supplied by 2.89, applied to $\operatorname{Re} f$ and $\operatorname{Im} f$. Precisely, 2.89(b) and (c) give simple $\mathcal{S}$-measurable $u_k \to \operatorname{Re} f$ and $v_k \to \operatorname{Im} f$ pointwise on X with $|u_k| \le |\operatorname{Re} f|$ and $|v_k| \le |\operatorname{Im} f|$; then $g_k := u_k + i v_k$ is simple and $\mathcal{S}$-measurable, with

$$|g_k| \le |\operatorname{Re} f| + |\operatorname{Im} f| \le 2|f| \qquad \text{and} \qquad g_k \to f \ \text{ pointwise},$$

so $\|g_k\|_p \le 2\|f\|_p < \infty$ and $g_k \in \mathcal{L}^p(\mu)$ for every $p \in (0,\infty]$.
(i) $0 < p < \infty$. Then $|f-g_k|^p \le 3^p |f|^p$, which is integrable because $f \in \mathcal{L}^p(\mu)$, and $|f-g_k|^p \to 0$ pointwise, so the Dominated Convergence Theorem 3.31 gives $\|f-g_k\|_p \to 0$; take $g = g_k$ for k large.
(ii) $p=\infty$. Put $M = \|f\|_\infty < \infty$, $E = \{|f| \le M\}$, and $\tilde{f} = f\chi_E$; as in 7A.16(i), $\mu(X \setminus E) = 0$, so $\|f-\tilde{f}\|_\infty = 0$. Since $\operatorname{Re} \tilde{f}$ and $\operatorname{Im} \tilde{f}$ are bounded, 2.89(d) supplies simple $\mathcal{S}$-measurable a, b with $\sup_X |\operatorname{Re} \tilde{f} - a| < \varepsilon/4$ and $\sup_X |\operatorname{Im} \tilde{f} - b| < \varepsilon/4$. Then $g = a + ib$ is simple with $|g| \le M + \varepsilon/2$, so $g \in \mathcal{L}^\infty(\mu)$, and by the triangle inequality of 7A.1,

$$\|f-g\|_\infty \le \|f-\tilde{f}\|_\infty + \|\tilde{f}-g\|_\infty \le 0 + \frac{\varepsilon}{2} < \varepsilon.$$

Problem 7A.18. Suppose $0 < p < \infty$ and $f \in \mathcal{L}^p(\mathbb{R})$. Prove that for every $\varepsilon > 0$, there exists a step function $g \in \mathcal{L}^p(\mathbb{R})$ such that $\|f-g\|_p < \varepsilon$.
[This exercise extends 3.47.]

Solution. Approximate f first by a simple function using 7A.17, then replace each of its level sets by a finite union of intervals; with $q = \min\{p, 1\}$, the errors add through the q-triangle inequality

$$\Big\|\sum_{k=1}^{m} f_k\Big\|_p^q \leq \sum_{k=1}^{m} \|f_k\|_p^q.$$

That inequality is Minkowski 7.14 with induction on m when $p \geq 1$ (so $q = 1$); when $0 < p < 1$ it is $q = p$ and follows by integrating $\big(\sum_k |f_k|\big)^p \leq \sum_k |f_k|^p$, which iterates $(a+b)^p \leq a^p + b^p$ for $a, b \geq 0$ (divide by $(a+b)^p$ and use $s^p \geq s$ for $s \in [0, 1]$).
Interval approximation. If $A \subseteq \mathbb{R}$ is Lebesgue measurable with $|A| < \infty$ and $\delta > 0$, there is a union E of finitely many disjoint bounded open intervals with $\|\chi_A - \chi_E\|_p^p = |A \setminus E| + |E \setminus A| < \delta$. Indeed, 2.71(e) gives an open $G \supseteq A$ with $|G \setminus A| < \delta/2$, so $|G| < \infty$; writing $G = \bigcup_{j=1}^{\infty} J_j$ with the J_j disjoint open intervals (bounded, as $|G| < \infty$), choose N with $\sum_{j>N} |J_j| < \delta/2$ and set $E = J_1 \cup \cdots \cup J_N$; then

$$A \setminus E \subseteq \bigcup_{j>N} J_j \quad \text{and} \quad E \setminus A \subseteq G \setminus A.$$

Now let $\varepsilon > 0$. By 7A.17 there is a simple function $h \in \mathcal{L}^p(\mathbb{R})$ with

$$\|f - h\|_p^q < \frac{\varepsilon^q}{2}.$$

If $h = 0$, take $g = 0$. Otherwise write $h = a_1\chi_{A_1} + \cdots + a_n\chi_{A_n}$ in the standard representation following 2.88, the a_k distinct and nonzero and the A_k disjoint and measurable; disjointness gives $\|h\|_p^p = \sum_k |a_k|^p |A_k| < \infty$, so each $|A_k| < \infty$. Choose $\delta_k > 0$ with $|a_k|^q \delta_k^{q/p} < \varepsilon^q/(2n)$ and then E_k as above with $\|\chi_{A_k} - \chi_{E_k}\|_p^p < \delta_k$, and set

$$g = a_1\chi_{E_1} + \cdots + a_n\chi_{E_n}.$$

Expanding each χ_{E_k} over its finitely many bounded intervals exhibits g as a step function; it is bounded with bounded support, so $g \in \mathcal{L}^p(\mathbb{R})$. By the q-triangle inequality and the homogeneity 7.5(b),

$$\|h - g\|_p^q \leq \sum_{k=1}^{n} |a_k|^q \, \|\chi_{A_k} - \chi_{E_k}\|_p^q < \sum_{k=1}^{n} |a_k|^q \delta_k^{q/p} < \frac{\varepsilon^q}{2}.$$

Therefore

$$\|f - g\|_p^q \leq \|f - h\|_p^q + \|h - g\|_p^q < \frac{\varepsilon^q}{2} + \frac{\varepsilon^q}{2} = \varepsilon^q,$$

which gives $\|f - g\|_p < \varepsilon$.

Problem 7A.19. Suppose $0 < p < \infty$ and $f \in \mathcal{L}^p(\mathbb{R})$. Prove that for every $\varepsilon > 0$, there exists a continuous function $g \colon \mathbb{R} \to \mathbf{F}$ such that $\|f - g\|_p < \varepsilon$ and the set $\{x \in \mathbb{R} : g(x) \neq 0\}$ is bounded.
[This exercise extends 3.48.]

Solution. Round off the corners of the step function supplied by 7A.18: replace each χ_I by a continuous trapezoid. Throughout, $q = \min\{p, 1\}$ and errors add through the q-triangle inequality $\|\sum_k f_k\|_p^q \leq \sum_k \|f_k\|_p^q$ proved in 7A.18.
Let $\varepsilon > 0$. By 7A.18 there is a step function $h \in \mathcal{L}^p(\mathbb{R})$ with

$$\|f - h\|_p^q < \frac{\varepsilon^q}{2}.$$

If $h = 0$, take $g = 0$. Otherwise write $h = a_1\chi_{I_1} + \cdots + a_n\chi_{I_n}$ with the a_k nonzero and the I_k intervals, which may be taken disjoint: the finitely many endpoints cut $\mathbb{R}$ into finitely many pieces on each of which h is constant, and we discard the pieces where $h = 0$ and relabel. Disjointness then gives $\|h\|_p^p = \sum_k |a_k|^p |I_k| < \infty$, so each I_k is bounded, say with endpoints $b_k \leq c_k$.

Trapezoids. For a bounded interval I with endpoints $b \le c$ and for $\delta > 0$, let $\varphi = 0$ if $c - b \le 2\delta$, and otherwise let φ be the continuous piecewise-linear function that equals 1 on $[b+\delta, c-\delta]$, equals 0 off (b, c), and is linear on $[b, b+\delta]$ and on $[c-\delta, c]$. In the first case $\|\chi_I - \varphi\|_p^p = c - b \le 2\delta$; in the second, $\chi_I - \varphi$ is bounded by 1 and vanishes off $[b, b+\delta] \cup [c-\delta, c]$, a set of measure 2δ. Either way φ is continuous, vanishes off $[b, c]$, and

$$\|\chi_I - \varphi\|_p^p \le 2\delta.$$

Choose $\delta > 0$ small enough that

$$|a_k|^q (2\delta)^{q/p} < \frac{\varepsilon^q}{2n} \qquad \text{for } k = 1, \dots, n,$$

possible since $(2\delta)^{q/p} \to 0$ as $\delta \downarrow 0$; let φ_k be the trapezoid for I_k and this δ, and set

$$g = a_1\varphi_1 + \cdots + a_n\varphi_n.$$

Then g is continuous and vanishes outside the bounded set $[b_1, c_1] \cup \cdots \cup [b_n, c_n]$, hence is bounded and lies in $\mathcal{L}^p(\mathbb{R})$. By the q-triangle inequality and 7.5(b),

$$\|h - g\|_p^q \le \sum_{k=1}^{n} |a_k|^q \, \|\chi_{I_k} - \varphi_k\|_p^q \le \sum_{k=1}^{n} |a_k|^q (2\delta)^{q/p} < \frac{\varepsilon^q}{2},$$

so $\|f - g\|_p^q \le \|f - h\|_p^q + \|h - g\|_p^q < \varepsilon^q$.

Problem 7A.20. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 < p < \infty$, and $f, g \in \mathcal{L}^p(\mu)$. Prove that Minkowski's inequality (7.14) is an equality if and only if there exist nonnegative numbers a and b, not both 0, such that

$$af(x) = bg(x)$$

for almost every $x \in X$.

Solution. Normalize and use strict convexity of $t \mapsto t^p$. Two pointwise facts do the work.
Fact 1. For $z, w \in \mathbf{F}$, $|z + w| = |z| + |w|$ if and only if $z\overline{w} \in [0, \infty)$: comparing

$$|z+w|^2 = |z|^2 + 2\,\mathrm{Re}(z\overline{w}) + |w|^2, \quad (|z| + |w|)^2 = |z|^2 + 2|z\overline{w}| + |w|^2,$$

equality holds exactly when $\mathrm{Re}(z\overline{w}) = |z\overline{w}|$, that is, when $z\overline{w} \ge 0$.
Fact 2. For $p > 1$, $s \in (0,1)$, and $\alpha, \beta \ge 0$ we have $\big(s\alpha + (1-s)\beta\big)^p \le s\alpha^p + (1-s)\beta^p$, with equality if and only if $\alpha = \beta$; indeed $t \mapsto t^p$ has second derivative $p(p-1)t^{p-2} > 0$ on $(0, \infty)$, hence is strictly convex there, while for $\alpha = 0 < \beta$ the left side is $(1-s)^p\beta^p < (1-s)\beta^p$.
($\Leftarrow$) Suppose $af = bg$ almost everywhere with $a, b \ge 0$ not both 0. If $b = 0$, then $a > 0$, so $f = 0$ almost everywhere and $\|f + g\|_p = \|g\|_p = \|f\|_p + \|g\|_p$. If $b > 0$, then $g = cf$ almost everywhere with $c = a/b \ge 0$, so 7.5(b) gives $\|f + g\|_p = (1 + c)\|f\|_p = \|f\|_p + \|g\|_p$.
($\Rightarrow$) Suppose $\|f + g\|_p = \|f\|_p + \|g\|_p$. If $\|f\|_p = 0$, take $(a, b) = (1, 0)$; if $\|g\|_p = 0$, take $(a, b) = (0, 1)$. Otherwise $A := \|f\|_p$ and $B := \|g\|_p$ lie in $(0, \infty)$; put $u = f/A$, $v = g/B$, and $s = A/(A+B) \in (0,1)$, so that $\|u\|_p = \|v\|_p = 1$ by 7.5(b) and $f + g = (A+B)\big(su + (1-s)v\big)$. The hypothesis becomes $\|su + (1-s)v\|_p = 1$, while pointwise

$$\big|su + (1-s)v\big|^p \le \big(s|u| + (1-s)|v|\big)^p \le s|u|^p + (1-s)|v|^p$$

by the triangle inequality in $\mathbf{F}$ and Fact 2. Integrating turns this chain into $1 \le 1 \le 1$, so each of the two differences is a nonnegative function of integral 0 (all three terms are integrable, being dominated by the right-hand side) and hence vanishes almost everywhere by the second bullet point of 3.43. Thus for almost every $x \in X$:

$$\text{(i)} \quad \big|su(x) + (1-s)v(x)\big| = s|u(x)| + (1-s)|v(x)|,$$

$$\text{(ii)} \quad \big(s|u(x)| + (1-s)|v(x)|\big)^p = s|u(x)|^p + (1-s)|v(x)|^p.$$

where (i) uses the injectivity of $t \mapsto t^p$ on $[0, \infty)$. At such an x, Fact 2 turns (ii) into $|u(x)| = |v(x)|$, and Fact 1 (with $z = su(x)$, $w = (1-s)v(x)$) turns (i) into $u(x)\overline{v(x)} \geq 0$. If $u(x) = 0$, then $v(x) = 0$; otherwise

$$u(x)\overline{v(x)} = |u(x)|\,|v(x)| = |u(x)|^2 = u(x)\overline{u(x)},$$

and dividing by $u(x)$ gives $v(x) = u(x)$. Hence $u = v$ almost everywhere, that is, $Bf = Ag$ almost everywhere; take $a = B$ and $b = A$.

Problem 7A.21. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \in \mathcal{L}^1(\mu)$. Prove that

$$\|f + g\|_1 = \|f\|_1 + \|g\|_1$$

if and only if $f(x)\overline{g(x)} \geq 0$ for almost every $x \in X$.

Solution. The two sides differ by the integral of the nonnegative function

$$\psi = |f| + |g| - |f + g|,$$

so the equality holds if and only if $\psi = 0$ almost everywhere. Indeed $\psi \geq 0$ by the triangle inequality in $\mathbf{F}$, and $|f|$, $|g|$, $|f + g|$ all have finite integral (as $f, g \in \mathcal{L}^1(\mu)$), so additivity of the integral gives

$$\int \psi\, d\mu = \|f\|_1 + \|g\|_1 - \|f + g\|_1,$$

and a nonnegative function has integral 0 exactly when it vanishes almost everywhere, by the second bullet point of 3.43.
Finally $\psi(x) = 0$ says $|f(x) + g(x)| = |f(x)| + |g(x)|$, which holds if and only if $f(x)\overline{g(x)} \geq 0$ by the equality case of the triangle inequality in $\mathbf{F}$ established as Fact 1 in 7A.20.

Problem 7A.22. Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces and $0 < p < \infty$. Prove that if $f \in \mathcal{L}^p(\mu \times \nu)$, then

$$[f]_x \in \mathcal{L}^p(\nu) \quad \text{for almost every } x \in X$$

and

$$[f]^y \in \mathcal{L}^p(\mu) \quad \text{for almost every } y \in Y,$$

where $[f]_x$ and $[f]^y$ are the cross sections of f as defined in 5.7.

Solution. Apply Tonelli's Theorem 5.28 to $|f|^p$. That function is nonnegative and $\mathcal{S} \otimes \mathcal{T}$-measurable, and directly from the definition 5.7 of cross sections,

$$\big[\,|f|^p\,\big]_x = \big|[f]_x\big|^p \qquad \text{and} \qquad \big[\,|f|^p\,\big]^y = \big|[f]^y\big|^p.$$

Both measure spaces are σ-finite, so 5.28 says that

$$h\colon X \to [0, \infty], \qquad h(x) = \int_Y |[f]_x|^p\, d\nu$$

is $\mathcal{S}$-measurable with $\int_X h\, d\mu = \int_{X \times Y} |f|^p\, d(\mu \times \nu) = \|f\|_p^p < \infty$. Writing $E = \{h = \infty\} \in \mathcal{S}$, the bound $c\chi_E \leq h$ gives

$$c\,\mu(E) \leq \int_X h\, d\mu = \|f\|_p^p \qquad \text{for every } c > 0,$$

so $\mu(E) = 0$. Thus for every $x \notin E$ the cross section $[f]_x$ is $\mathcal{T}$-measurable (5.9) with $\int_Y |[f]_x|^p\, d\nu < \infty$, that is, $[f]_x \in \mathcal{L}^p(\nu)$ for almost every $x \in X$. Part (b) of 5.28 gives $[f]^y \in \mathcal{L}^p(\mu)$ for almost every $y \in Y$ by the same argument.

Problem 7A.23. Suppose $1 \le p < \infty$ and $f \in \mathcal{L}^p(\mathbf{R})$.
(a) For $t \in \mathbf{R}$, define $f_t\colon \mathbf{R} \to \mathbf{R}$ by $f_t(x) = f(x-t)$. Prove that the function $t \mapsto \|f - f_t\|_p$ is bounded and uniformly continuous on $\mathbf{R}$.
(b) For $t > 0$, define $f_t\colon \mathbf{R} \to \mathbf{R}$ by $f_t(x) = f(tx)$. Prove that
$$\lim_{t\to 1} \|f - f_t\|_p = 0.$$

Solution. Both parts reduce, via the density result 7A.19, to a continuous function vanishing off a bounded set; the reduction runs on the two invariance identities

$$\int_{-\infty}^{\infty} h(x-t)\,dx = \int_{-\infty}^{\infty} h, \qquad \int_{-\infty}^{\infty} h(tx)\,dx = \frac{1}{t}\int_{-\infty}^{\infty} h$$

for measurable $h \ge 0$, with $t \in \mathbf{R}$ in the first and $t > 0$ in the second. Both hold for $h = \chi_E$ because $|t+E| = |E|$ by 2.7 and $|t^{-1}E| = t^{-1}|E|$ by the definition 2.2 of outer measure applied to dilated covers (translates and dilates of measurable sets are measurable, by 2.70 in the Lebesgue case), hence for simple h by 3.7, hence for all measurable $h \ge 0$ by 2.89 and the Monotone Convergence Theorem 3.11. Taking $h = |f|^p$ gives $f_t \in \mathcal{L}^p(\mathbf{R})$ with

$$\|f_t\|_p = \|f\|_p \quad \text{in (a)}, \qquad \|f_t\|_p = t^{-1/p}\|f\|_p \quad \text{in (b)}.$$

Also, for each $\varepsilon > 0$, 7A.19 supplies a continuous g vanishing outside some $[-M, M]$ with $\|f-g\|_p < \varepsilon$; such a g is uniformly continuous on $\mathbf{R}$, being uniformly continuous on the compact set $[-M-2, M+2]$ and zero on a neighbourhood of the complement. (Check!)
(a) Put $\varphi(t) = \|f - f_t\|_p$. Minkowski 7.14 and translation invariance give $\varphi(t) \le \|f\|_p + \|f_t\|_p = 2\|f\|_p$, so φ is bounded. Since $(f_a)_b = f_{a+b}$, we have $f_t - f_s = (f - f_{s-t})_t$, whence $\|f_t - f_s\|_p = \varphi(s-t)$, and 7.14 applied twice gives
$$|\varphi(t) - \varphi(s)| \le \|f_t - f_s\|_p = \varphi(s-t).$$
So uniform continuity follows once $\varphi(u) \to 0$ as $u \to 0$. Given $\varepsilon > 0$, take g as above with $\|f-g\|_p < \varepsilon/3$; then $\|f_u - g_u\|_p = \|f-g\|_p$, so 7.14 gives $\varphi(u) \le \frac{2\varepsilon}{3} + \|g - g_u\|_p$. For $|u| \le 1$ the function $g - g_u$ vanishes outside $[-M-1, M+1]$, so

$$\|g - g_u\|_p^p \le (2M+2)\Big(\sup_{x\in\mathbf{R}}|g(x) - g(x-u)|\Big)^p,$$

and the supremum tends to 0 as $u \to 0$ by uniform continuity of g. Hence $\varphi(u) < \varepsilon$ for all small $|u|$.
(b) Given $\varepsilon > 0$, take g as above with $\|f-g\|_p < \varepsilon$. For $t \in [\frac{1}{2}, 2]$, dilation invariance gives $\|f_t - g_t\|_p = t^{-1/p}\|f-g\|_p \le 2^{1/p}\varepsilon \le 2\varepsilon$ (here $p \ge 1$), so by 7.14

$$\|f - f_t\|_p \le \|f-g\|_p + \|g - g_t\|_p + \|g_t - f_t\|_p < 3\varepsilon + \|g - g_t\|_p.$$

For such t, if $|x| > 2M$ then $|x| > M$ and $|tx| \ge |x|/2 > M$, so $g - g_t$ vanishes outside $[-2M, 2M]$ and

$$\|g - g_t\|_p^p \le 4M\Big(\sup_{|x|\le 2M}|g(x) - g(tx)|\Big)^p.$$

Since $|tx - x| \le 2M|t-1|$ for $|x| \le 2M$, uniform continuity of g sends that supremum to 0 as $t \to 1$. Hence $\limsup_{t\to1}\|f - f_t\|_p \le 3\varepsilon$ for every $\varepsilon > 0$.

Problem 7A.24. Suppose $1 \le p < \infty$ and $f \in \mathcal{L}^p(\mathbf{R})$. Prove that

$$\lim_{t\downarrow 0}\frac{1}{2t}\int_{b-t}^{b+t}|f - f(b)|^p = 0$$

for almost every $b \in \mathbf{R}$.

Solution. Apply the Lebesgue Differentiation Theorem 4.10 to the countably many localized functions

$$h_{r,n} = |f - r|^p \chi_{(-n,n)}, \qquad r \in D,\ n \in \mathbb{Z}^+,$$

where D is a countable dense subset of $\mathbf{F}$ (take $\mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$), and then let $r \to f(b)$. Convexity of $s \mapsto s^p$ on $[0, \infty)$ (here $p \geq 1$) gives

$$|u + v|^p \leq \big(|u| + |v|\big)^p \leq 2^{p-1}\big(|u|^p + |v|^p\big) \qquad (u, v \in \mathbf{F}),$$

so $\int h_{r,n} \leq 2^{p-1}\big(\|f\|_p^p + 2n|r|^p\big) < \infty$ and $h_{r,n} \in \mathcal{L}^1(\mathbb{R})$. Hence 4.10 supplies null sets $N_{r,n}$ off which the averages of $|h_{r,n} - h_{r,n}(b)|$ over $(b-t, b+t)$ tend to 0; put $N = \bigcup_{r,n} N_{r,n}$, which is null by 2.8. Fix $b \notin N$ and $r \in D$, and choose $n > |b| + 1$. For $0 < t < 1$ we have $(b-t, b+t) \subseteq (-n, n)$, so $h_{r,n} = |f - r|^p$ there and $h_{r,n}(b) = |f(b) - r|^p$, whence

$$\left| \frac{1}{2t} \int_{b-t}^{b+t} |f - r|^p - |f(b) - r|^p \right| \leq \frac{1}{2t} \int_{b-t}^{b+t} \big|h_{r,n} - h_{r,n}(b)\big| \longrightarrow 0.$$

Now let $\varepsilon > 0$ and pick $r \in D$ with $|f(b) - r| < \varepsilon$. The displayed inequality with $u = f(x) - r$ and $v = r - f(b)$, integrated over $(b-t, b+t)$, gives

$$\frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)|^p \leq 2^{p-1} \left(\frac{1}{2t} \int_{b-t}^{b+t} |f - r|^p + |r - f(b)|^p \right),$$

so by the previous paragraph

$$\limsup_{t \downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f - f(b)|^p \leq 2^p |f(b) - r|^p < 2^p \varepsilon^p.$$

As $\varepsilon > 0$ was arbitrary and the averages are nonnegative, the limit is 0 for every $b \notin N$.

7.2 Exercises 7B

Problem 7B.1. Suppose $n > 1$ and $0 < p < 1$. Prove that if $\|\cdot\|$ is defined on $\mathbb{F}^n$ by

$$\|(a_1, \ldots, a_n)\| = \big(|a_1|^p + \cdots + |a_n|^p\big)^{1/p},$$

then $\|\cdot\|$ is not a norm on $\mathbb{F}^n$.

Solution. The triangle inequality fails. Take $a = (1, 0, \ldots, 0)$ and $b = (0, 1, 0, \ldots, 0)$, available because $n > 1$. Then $\|a\| = \|b\| = 1$, while $1/p > 1$ gives

$$\|a + b\| = (1^p + 1^p)^{1/p} = 2^{1/p} > 2 = \|a\| + \|b\|.$$

Problem 7B.2. (a) Suppose $1 \leq p < \infty$. Prove that there is a countable subset of ℓ^p whose closure equals ℓ^p.
(b) Prove that there does not exist a countable subset of ℓ^∞ whose closure equals ℓ^∞.

Solution. (a) Yes: the sequences with finitely many nonzero terms, all of them in $\mathbb{D}$, where $\mathbb{D} = \mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$ is a countable dense subset of $\mathbb{F}$. Call this set D; it is countable, being $\bigcup_{n=1}^\infty D_n$ with D_n (vanishing after coordinate n) in bijection with $\mathbb{D}^n$, and $D \subseteq \ell^p$. Given $a \in \ell^p$ and $\varepsilon > 0$, convergence of $\sum_k |a_k|^p$ gives n with $\sum_{k>n} |a_k|^p < \varepsilon^p/2$; choosing $b_k \in \mathbb{D}$ with $|a_k - b_k|^p < \varepsilon^p/(2n)$ for $k \leq n$ and setting $b = (b_1, \ldots, b_n, 0, 0, \ldots) \in D$ gives

$$\|a - b\|_p^p = \sum_{k=1}^n |a_k - b_k|^p + \sum_{k>n} |a_k|^p < \varepsilon^p.$$

(b) No: the set S of sequences with entries in $\{0,1\}$ lies in ℓ^∞, is uncountable (it is in bijection with the subsets of $\mathbb{Z}^+$), and any two distinct members are at distance exactly 1. If $C \subseteq \ell^\infty$ has closure ℓ^∞, choose for each $a \in S$ some $c(a) \in C$ with $\|a - c(a)\|_\infty < \frac{1}{2}$. This map is injective, since $c(a) = c(b)$ would give

$$\|a - b\|_\infty \leq \|a - c(a)\|_\infty + \|c(b) - b\|_\infty < 1,$$

forcing $a = b$. So C is uncountable.

Problem 7B.3. (a) Suppose $1 \leq p < \infty$. Prove that there is a countable subset of $L^p(\mathbb{R})$ whose closure equals $L^p(\mathbb{R})$.
(b) Prove that there does not exist a countable subset of $L^\infty(\mathbb{R})$ whose closure equals $L^\infty(\mathbb{R})$.

Solution. (a) Yes: with $\mathbb{D} = \mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$ (countable and dense in $\mathbb{F}$), take

$$D = \Big\{ \textstyle\sum_{j=1}^m c_j \chi_{(a_j, b_j)} : m \in \mathbb{Z}^+,\ c_j \in \mathbb{D},\ a_j < b_j \text{ rational} \Big\}.$$

For each m these sums form the image of the countable set $(\mathbb{D} \times \mathbb{Q} \times \mathbb{Q})^m$, so D is countable, and each of its elements is bounded with bounded support, so $D \subseteq \mathcal{L}^p(\mathbb{R})$.
Let $f \in \mathcal{L}^p(\mathbb{R})$ and $\varepsilon > 0$. By 7A.18 there is a step function h with $\|f - h\|_p < \varepsilon/2$, and as in 7A.19 we may write $h = \sum_{j=1}^m a_j \chi_{I_j}$ with the I_j disjoint bounded intervals. Given $\delta \in (0,1)$, enlarge each I_j to an interval J_j with rational endpoints within δ of those of I_j, so that $\lambda(J_j \setminus I_j) < 2\delta$ and hence $\|\chi_{I_j} - \chi_{J_j}\|_p \leq (2\delta)^{1/p}$, and pick $q_j \in \mathbb{D}$ with $|a_j - q_j| < \delta$. Then $u = \sum_{j=1}^m q_j \chi_{J_j} \in D$, and Minkowski 7.14 gives

$$\|h - u\|_p \leq (2\delta)^{1/p} \sum_{j=1}^m |a_j| \ + \ \delta \sum_{j=1}^m \lambda(J_j)^{1/p},$$

which is less than $\varepsilon/2$ once δ is small (the $\lambda(J_j) \leq \lambda(I_j) + 2$ stay bounded). Hence $\|f - u\|_p < \varepsilon$, so the closure of D is $L^p(\mathbb{R})$.
(b) No: for $t \in (0,1)$ put $g_t = \chi_{(0,t)} \in L^\infty(\mathbb{R})$. If $0 < s < t < 1$, then $g_t - g_s = \chi_{[s,t)}$, whose every level set $\{|\cdot| > r\}$ with $r \in (0,1)$ has measure $t - s > 0$, so $\|g_t - g_s\|_\infty = 1$ by 7.1. Thus $\{g_t\}$ is an uncountable family at pairwise distance 1. If $C \subseteq L^\infty(\mathbb{R})$ had closure $L^\infty(\mathbb{R})$, choosing $c_t \in C$ with $\|g_t - c_t\|_\infty < \frac{1}{2}$ would give an injection $t \mapsto c_t$ of $(0,1)$ into C, since $c_s = c_t$ forces

$$\|g_s - g_t\|_\infty \leq \|g_s - c_s\|_\infty + \|c_t - g_t\|_\infty < 1,$$

hence $s = t$. So C is uncountable.

Problem 7B.4. Suppose $(X, \mathcal{S}, \mu)$ is a σ-finite measure space and $1 \leq p \leq \infty$. Prove that if $f : X \to \mathbb{F}$ is an $\mathcal{S}$-measurable function such that $fh \in \mathcal{L}^1(\mu)$ for every $h \in \mathcal{L}^{p'}(\mu)$, then $f \in \mathcal{L}^p(\mu)$.

Solution. If $f \notin \mathcal{L}^p(\mu)$, a single $h \in \mathcal{L}^{p'}(\mu)$ built from a σ-finite exhaustion violates the hypothesis. Fix $X_1 \subseteq X_2 \subseteq \cdots$ in $\mathcal{S}$ with $\mu(X_n) < \infty$ and $\bigcup_n X_n = X$, and recall $\frac{1}{p} + \frac{1}{p'} = 1$ (7.6). We use once and for all that $\int u\chi_N \, d\mu = 0$ for measurable $u \geq 0$ and $\mu(N) = 0$, so that $\int u \, d\mu = \int u\chi_{X \setminus N} \, d\mu$ (Check!); this converts a nonnegative function that is finite almost everywhere into an $\mathbb{F}$-valued one with the same norm.
(i) $p = 1$, so $p' = \infty$. The constant function 1 lies in $\mathcal{L}^\infty(\mu)$, so $f = f \cdot 1 \in \mathcal{L}^1(\mu)$ by hypothesis.
(ii) $p = \infty$, so $p' = 1$. Suppose $\|f\|_\infty = \infty$. By 7.1 each $A_k = \{|f| > 4^k\}$ has $\mu(A_k) > 0$, so continuity from below supplies n_k with $B_k = A_k \cap X_{n_k}$ satisfying $0 < \mu(B_k) < \infty$. Put $h = \sum_{k=1}^\infty 2^{-k} \mu(B_k)^{-1} \chi_{B_k}$; the Monotone Convergence Theorem 3.11 applied to the partial sums gives $\int h \, d\mu = 1$, so $h < \infty$ almost everywhere and $\hat{h} = h\chi_{X \setminus N}$, with $N = \{h = \infty\}$, lies in $\mathcal{L}^1(\mu)$. But for every k,

$$\int |f\hat{h}| \, d\mu = \int |f| h \, d\mu \geq \frac{2^{-k}}{\mu(B_k)} \int_{B_k} |f| \, d\mu \geq 2^{-k} 4^k = 2^k,$$

contradicting $f\hat{h} \in \mathcal{L}^1(\mu)$. Hence $f \in \mathcal{L}^\infty(\mu)$.
(iii) $1 < p < \infty$. Put $E_n = \{x \in X_n : |f(x)| \le n\}$ and $g_n = |f|\chi_{E_n}$. The E_n increase with union X, so 3.11 gives $\|g_n\|_p^p \to \int |f|^p \, d\mu$, while $g_n \le n\chi_{X_n}$ gives $\|g_n\|_p^p \le n^p \mu(X_n) < \infty$. Suppose $\int |f|^p \, d\mu = \infty$. Choose $n_1 < n_2 < \cdots$ with $c_k := \|g_{n_k}\|_p \ge 4^k$, write $G_k = g_{n_k}$, and set $h_k = G_k^{p-1}/c_k^{p-1}$. Since $(p-1)p' = p$,

$$\|h_k\|_{p'}^{p'} = \frac{1}{c_k^p} \int G_k^p \, d\mu = 1 \qquad \text{and} \qquad \int G_k h_k \, d\mu = \frac{c_k^p}{c_k^{p-1}} = c_k \ge 4^k.$$

Put $h = \sum_{k=1}^\infty 2^{-k} h_k$. Minkowski 7.14 bounds the p'-norm of each partial sum by $\sum_k 2^{-k} = 1$, so 3.11 gives $\|h\|_{p'} \le 1$; hence $h < \infty$ almost everywhere and $\hat{h} = h\chi_{X \setminus N}$ lies in $\mathcal{L}^{p'}(\mu)$. But $|f| \ge G_k$ and $\hat{h} \ge 2^{-k} h_k \chi_{X\setminus N}$ give

$$\int |f\hat{h}| \, d\mu \ge 2^{-k} \int G_k h_k \, d\mu \ge 2^{-k} 4^k = 2^k$$

for every k, a contradiction. Hence $\int |f|^p \, d\mu < \infty$.

Problem 7B.5. (a) Prove that if μ is a measure, $1 < p < \infty$, and $f, g \in L^p(\mu)$ are such that

$$\|f\|_p = \|g\|_p = \left\| \frac{f+g}{2} \right\|_p,$$

then $f = g$.
(b) Give an example to show that (a) can fail if $p = 1$.
(c) Give an example to show that (a) can fail if $p = \infty$.

Solution. (a) Everything follows from the strict convexity inequality

$$\left| \frac{a+b}{2} \right|^p \le \left(\frac{|a| + |b|}{2} \right)^p \le \frac{|a|^p + |b|^p}{2} \qquad (a, b \in \mathbb{F}),$$

in which equality forces $a = b$. The first step is the triangle inequality in $\mathbb{F}$ with $t \mapsto t^p$ increasing; the second is convexity of $t \mapsto t^p$ on $[0, \infty)$, strict because $\varphi'(u) = pu^{p-1}$ is strictly increasing there (this is where $p > 1$ enters). So overall equality forces $|a| = |b| =: r$ from the second step and $|a + b| = 2r$ from the first; if $r > 0$, then expanding $|a+b|^2 = 2r^2 + 2\operatorname{Re}(a\bar{b}) = 4r^2$ gives $\operatorname{Re}(a\bar{b}) = r^2 = |a\bar{b}|$, so $a\bar{b} = r^2$ and $a = r^2 b/|b|^2 = b$; and if $r = 0$, then $a = b = 0$.
Now let $c = \|f\|_p = \|g\|_p = \left\|\frac{f+g}{2}\right\|_p < \infty$ and put

$$u = \frac{|f|^p + |g|^p}{2} - \left| \frac{f+g}{2} \right|^p \ge 0.$$

Additivity 3.16 gives $\int u \, d\mu + c^p = c^p$, and $c^p < \infty$ permits the subtraction, so $\int u \, d\mu = 0$. By the second bullet point of 3.43, $u = 0$ almost everywhere, and by the equality case above this says $f = g$ almost everywhere, that is, $f = g$ in $L^p(\mu)$.
(b) Counting measure on $\{1, 2\}$, with $f = (1, 0)$ and $g = (0, 1)$: here $\|f\|_1 = \|g\|_1 = 1$ and $\frac{f+g}{2} = (\frac{1}{2}, \frac{1}{2})$ also has 1-norm 1, yet $f \ne g$.
(c) The same space with $f = (1, 1)$ and $g = (1, 0)$: here $\|f\|_\infty = \|g\|_\infty = 1$ and $\frac{f+g}{2} = (1, \frac{1}{2})$ has ∞-norm 1, yet $f \ne g$.

Problem 7B.6. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < 1$. Show that

$$\|f + g\|_p^p \le \|f\|_p^p + \|g\|_p^p$$

for all $\mathcal{S}$-measurable functions $f, g : X \to \mathbb{F}$.

Solution. Integrate the pointwise inequality $|f+g|^p \le |f|^p + |g|^p$. For that, it suffices that

$$(a+b)^p \le a^p + b^p \qquad (a, b \ge 0),$$

since $|f+g| \le |f| + |g|$ and $t \mapsto t^p$ is increasing. To prove it, assume $a+b > 0$ and put $s = a/(a+b)$, $t = b/(a+b)$; then $s, t \in [0,1]$ with $s+t = 1$, and $u^p \ge u$ for $u \in [0,1]$ (as $u^{p-1} \ge 1$ there, $p < 1$), so

$$\frac{a^p + b^p}{(a+b)^p} = s^p + t^p \ge s + t = 1.$$

Integrating the pointwise bound, 3.8 and additivity 3.16 give

$$\|f+g\|_p^p = \int |f+g|^p \, d\mu \le \int |f|^p \, d\mu + \int |g|^p \, d\mu = \|f\|_p^p + \|g\|_p^p,$$

an inequality in $[0, \infty]$.

Problem 7B.7. Prove that $L^p(\mu)$, with addition and scalar multiplication as defined in 7.16 and norm defined as in 7.17, is a normed vector space. In other words, prove 7.18.

Solution. Everything rests on two facts: $Z(\mu)$ is a subspace of $\mathcal{L}^p(\mu)$, and $\|\cdot\|_p$ is constant on its cosets. Then $L^p(\mu)$ is the quotient $\mathcal{L}^p(\mu)/Z(\mu)$ with the operations of 7.16.
(i) $\mathcal{L}^p(\mu)$ is a vector space: this is 7.5 for $p < \infty$, and for $p = \infty$ it follows from the triangle inequality and homogeneity of $\|\cdot\|_\infty$ proved in 7A.1. Its subset $Z(\mu)$ of functions vanishing almost everywhere is a subspace, the union of two null sets being null, and $Z(\mu) \subseteq \mathcal{L}^p(\mu)$ because such functions have $\|\cdot\|_p = 0$.
(ii) $\widetilde{f} = \widetilde{F}$ if and only if $f - F \in Z(\mu)$, that is, $f = F$ almost everywhere: one direction uses that $Z(\mu)$ is closed under addition and negation, the other that $f \in \widetilde{f}$. (Check!) Consequently, if $\widetilde{f} = \widetilde{F}$ and $\widetilde{g} = \widetilde{G}$, then

$$(f+g) - (F+G) = (f-F) + (g-G) \in Z(\mu), \qquad \alpha f - \alpha F = \alpha(f-F) \in Z(\mu),$$

so the operations of 7.16 do not depend on the representatives. Each vector space axiom for $L^p(\mu)$ then transfers from the corresponding axiom in $\mathcal{L}^p(\mu)$ through $f \mapsto \widetilde{f}$, which by definition carries sums to sums and scalar multiples to scalar multiples; the additive identity is $\widetilde{0} = Z(\mu)$ and $-\widetilde{f} = (-f)^\sim$. (Check!)
(iii) $\|\cdot\|_p$ is well defined on $L^p(\mu)$. Suppose $f = F$ off a set N with $\mu(N) = 0$. If $p < \infty$, then $|f|^p \chi_{X \setminus N} = |F|^p \chi_{X \setminus N}$ and a null set contributes nothing to an integral of a nonnegative function, so $\int |f|^p \, d\mu = \int |F|^p \, d\mu$. If $p = \infty$, then $\{|f| > t\}$ and $\{|F| > t\}$ differ only inside N, so one is null exactly when the other is, and the two infima in 7.1 agree. Since $f \in \mathcal{L}^p(\mu)$, the common value $\|\widetilde{f}\|_p = \|f\|_p$ lies in $[0, \infty)$, as 6.33 requires.
(iv) The three conditions of 6.33. Homogeneity is 7.5(b) for $p < \infty$ and 7A.1 for $p = \infty$, read on representatives; the triangle inequality is Minkowski 7.14 likewise. For positive definiteness, $\|\widetilde{f}\|_p = 0$ forces $f = 0$ almost everywhere: for $p < \infty$ apply the second bullet point of 3.43 to $|f|^p$; for $p = \infty$ each set $\{|f| > \frac{1}{n}\}$ is null by 7.1, and $\{f \ne 0\}$ is their union. Hence $\widetilde{f} = \widetilde{0}$, and the converse is (iii) with $F = 0$.

Problem 7B.8. Prove 7.20 for the case $p = \infty$.

Solution. Take f to be the pointwise limit of $f_1, f_2, \ldots$ off the null set where the essential suprema fail to bound. As in 7A.16(i), $\mu(\{|g| > \|g\|_\infty\}) = 0$ for every $\mathcal{S}$-measurable g, so

$$E = \bigcup_{j,k} \{|f_j - f_k| > \|f_j - f_k\|_\infty\} \;\cup\; \bigcup_k \{|f_k| > \|f_k\|_\infty\}$$

is a countable union of null sets, hence $E \in \mathcal{S}$ with $\mu(E) = 0$. On $X \setminus E$ we have $|f_j - f_k| \le \|f_j - f_k\|_\infty$ for

all j, k, so the Cauchy hypothesis makes $f_1, f_2, \ldots$ uniformly Cauchy there, hence pointwise convergent by completeness of $\mathbf{F}$. Define $f = \lim_{k\to\infty} f_k \chi_{X\setminus E}$, which is $\mathcal{S}$-measurable by 2.48.
Given $\varepsilon > 0$, choose n with $\|f_j - f_k\|_\infty < \varepsilon$ for all $j, k \ge n$. For $k \ge n$ and $x \in X \setminus E$,

$$|f_k(x) - f(x)| = \lim_{j\to\infty} |f_k(x) - f_j(x)| \le \varepsilon,$$

so $\{|f_k - f| > t\} \subseteq E$ is null for every $t > \varepsilon$ and thus $\|f_k - f\|_\infty \le \varepsilon$; hence $\|f_k - f\|_\infty \to 0$. Taking $\varepsilon = 1$ with its n, we get $|f| \le |f_n| + 1 \le \|f_n\|_\infty + 1$ on $X \setminus E$, so $\|f\|_\infty \le \|f_n\|_\infty + 1 < \infty$ and $f \in \mathcal{L}^\infty(\mu)$.

Problem 7B.9. Prove that 7.20 also holds for $p \in (0, 1)$.

Solution. The telescoping proof of 7.20 runs unchanged with $\|\cdot\|_p^p$ in place of the norm, because 7B.6 gives $\|u + v\|_p^p \le \|u\|_p^p + \|v\|_p^p$ and hence, by induction, $\|\sum_{k=1}^m u_k\|_p^p \le \sum_{k=1}^m \|u_k\|_p^p$.
It suffices to find a convergent subsequence: if $\|f_{k_m} - f\|_p \to 0$, then given $\varepsilon > 0$ and n with $\|f_j - f_k\|_p^p < \varepsilon/2$ for $j, k \ge n$, choosing $k_m \ge n$ with $\|f_{k_m} - f\|_p^p < \varepsilon/2$ gives $\|f_k - f\|_p^p < \varepsilon$ for all $k \ge n$. So pick $n_1 < n_2 < \cdots$ with $\|f_j - f_k\|_p^p < 2^{-m}$ for $j, k \ge n_m$, pass to the subsequence (f_{n_m}) without relabelling, and set $f_0 = 0$; then

$$\sum_{k=1}^\infty \|f_k - f_{k-1}\|_p^p < \infty.$$

Let $g = \sum_{k=1}^\infty |f_k - f_{k-1}|$ with partial sums g_m. Finite subadditivity gives $\int g_m^p \, d\mu \le \sum_{k=1}^m \|f_k - f_{k-1}\|_p^p$, and $g_m^p \uparrow g^p$ pointwise, so the Monotone Convergence Theorem 3.11 gives

$$\int g^p \, d\mu \le \sum_{k=1}^\infty \|f_k - f_{k-1}\|_p^p < \infty.$$

Hence $A = \{g < \infty\}$ satisfies $\mu(X \setminus A) = 0$, since $c\chi_{X\setminus A} \le g^p$ for every $c > 0$. On A the series $\sum_k (f_k - f_{k-1})$ converges absolutely with m-th partial sum f_m, so $f := \lim_{m\to\infty} f_m \chi_A$ exists pointwise on X and is $\mathcal{S}$-measurable by 2.48; and $|f| \le g$ on A gives $\|f\|_p^p \le \int g^p \, d\mu < \infty$, so $f \in \mathcal{L}^p(\mu)$.
Finally, let $\varepsilon > 0$ and take n with $\|f_j - f_k\|_p^p < \varepsilon$ for $j, k \ge n$. For $k \ge n$ the functions $|f_k - f_j|^p$ converge almost everywhere to $|f_k - f|^p$, so Fatou's Lemma (3A.17) gives

$$\|f_k - f\|_p^p \le \liminf_{j\to\infty} \|f_k - f_j\|_p^p \le \varepsilon.$$

Problem 7B.10. Prove that 7.23 also holds for $p \in (0, 1)$.

Solution. Choose $k_1 < k_2 < \cdots$ with $\int |f_{k_m} - f|^p \, d\mu < 4^{-m}$, possible because $\|f_k - f\|_p^p \to 0$. Markov's inequality 4.1, applied to $h_m = |f_{k_m} - f|^p \in \mathcal{L}^1(\mu)$ with constant 2^{-m}, gives

$$\mu(E_m) \le 2^m \|h_m\|_1 < 2^{-m}, \qquad E_m = \{h_m \ge 2^{-m}\}.$$

Put $F = \bigcap_{M=1}^\infty \bigcup_{m\ge M} E_m$. Monotonicity 2.57 and countable subadditivity 2.58 give $\mu(F) \le \sum_{m\ge M} 2^{-m} = 2^{-M+1}$ for every M, so $\mu(F) = 0$.
If $x \notin F$, then $x \notin E_m$ for all m beyond some M, so $|f_{k_m}(x) - f(x)| < 2^{-m/p} \to 0$. Hence $f_{k_m} \to f$ pointwise on $X \setminus F$, that is, almost everywhere.

Problem 7B.11. Suppose $1 \le p \le \infty$. Prove that

$$\{(a_1, a_2, \ldots) \in \ell^p : a_k \ne 0 \text{ for every } k \in \mathbf{Z}^+\}$$

is not an open subset of ℓ^p.

Solution. Take $a = (2^{-1}, 2^{-2}, \ldots)$, which lies in the set A in question: every term is nonzero, and $\|a\|_p^p = \sum_k 2^{-kp} < \infty$ for $p < \infty$ while $\|a\|_\infty = \frac{1}{2}$. Given $\varepsilon > 0$, pick n with $2^{-n} < \varepsilon$ and let b agree with a except that $b_n = 0$. Then $b \in \ell^p$, $b \notin A$, and $a - b = 2^{-n} e_n$, so

$$\|a - b\|_p = 2^{-n}\|e_n\|_p = 2^{-n} < \varepsilon$$

for every $p \in [1, \infty]$. So no ball centred at a lies in A; as $a \in A$, the set A is not open.

Problem 7B.12. Show that there exists a sequence $f_1, f_2, \ldots$ of functions in $\mathcal{L}^1([0,1])$ such that $\lim_{k\to\infty} \|f_k\|_1 = 0$ but

$$\sup\{f_k(x) : k \in \mathbf{Z}^+\} = \infty$$

for every $x \in [0,1]$.
[This exercise shows that the conclusion of 7.23 cannot be improved to conclude that $\lim_{k\to\infty} f_k(x) = f(x)$ for almost every $x \in X$.]

Solution. Take the typewriter sequence with growing heights. For $n \in \mathbf{Z}^+$ and $j \in \{0, 1, \ldots, 2^n - 1\}$ put $I_{n,j} = [j2^{-n}, (j+1)2^{-n}]$, and index the pairs by

$$k = 2^n - 1 + j, \qquad f_k = n\,\chi_{I_{n,j}},$$

which is a bijection onto $\mathbf{Z}^+$ because the blocks $\{2^n - 1, \ldots, 2^{n+1} - 2\}$ partition $\mathbf{Z}^+$; write $n(k)$ for the resulting n. Then

$$\|f_k\|_1 = n(k)\,\lambda\big(I_{n(k),j(k)}\big) = \frac{n(k)}{2^{n(k)}},$$

so each $f_k \in \mathcal{L}^1([0,1])$; and $k \geq 2^{n(k)} - 1$ forces $n(k) \to \infty$, whence $\|f_k\|_1 \to 0$ because $n2^{-n} \to 0$.
Fix $x \in [0,1]$ and $n \in \mathbf{Z}^+$. The intervals $I_{n,0}, \ldots, I_{n,2^n-1}$ cover $[0,1]$, so $x \in I_{n,j}$ for some j, and the corresponding $k = 2^n - 1 + j$ gives $f_k(x) = n$. Hence $\sup_k f_k(x) = \infty$.

Problem 7B.13. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \leq p \leq \infty$, $f \in \mathcal{L}^p(\mu)$, and $f_1, f_2, \ldots$ is a sequence in $\mathcal{L}^p(\mu)$ such that $\lim_{k\to\infty} \|f_k - f\|_p = 0$. Show that if $g : X \to \mathbf{F}$ is a function such that $\lim_{k\to\infty} f_k(x) = g(x)$ for almost every $x \in X$, then $f(x) = g(x)$ for almost every $x \in X$.

Solution. By 7.23 some subsequence satisfies $f_{k_m} \to f$ off a set E with $\mu(E) = 0$, while by hypothesis $f_k \to g$ off a set F with $\mu(F) = 0$. For $x \notin E \cup F$, a null set by 2.58, the full sequence $f_k(x)$ converges to $g(x)$, hence so does the subsequence; since limits in $\mathbf{F}$ are unique and $f_{k_m}(x) \to f(x)$, we get $f(x) = g(x)$. Thus $f = g$ almost everywhere.

Problem 7B.14. (a) Give an example of a measure μ such that 7.25 fails for $p = 1$.
(b) Show that if μ is a σ-finite measure, then 7.25 holds for $p = 1$.

Solution. (a) Take $X = \{1\}$, $\mathcal{S} = \{\emptyset, X\}$, and $\mu(X) = \infty$. Then $\int |f|\,d\mu$ is 0 or ∞ according as $f(1) = 0$ or not, so

$$L^1(\mu) = \{0\} \qquad \text{and hence} \qquad \big(L^1(\mu)\big)' = \{0\},$$

whereas $\|f\|_\infty = |f(1)|$ by 7.1 and $\emptyset$ is the only null set, so $L^\infty(\mu)$ is one-dimensional. For h with $h(1) = 1$ we get $\|\varphi_h\| = 0 \neq 1 = \|h\|_\infty$, and $\varphi_h = \varphi_0$ with $h \neq 0$, so $h \mapsto \varphi_h$ is not one-to-one either. (This μ is not semifinite: X has infinite measure but no subset of positive finite measure.)
(b) For every measure, Holder's inequality 7.9 with $p = 1$, $p' = \infty$ gives $\|fh\|_1 \leq \|f\|_1\|h\|_\infty$, so φ_h is a bounded linear functional on $L^1(\mu)$ with $\|\varphi_h\| \leq \|h\|_\infty$, and $h \mapsto \varphi_h$ is linear; only the reverse inequality needs σ-finiteness.
Assume $\|h\|_\infty > 0$ and fix $t \in (0, \|h\|_\infty)$. Then $E = \{|h| > t\}$ has $\mu(E) > 0$, since $\mu(E) = 0$ would give $\|h\|_\infty \leq t$ by 7.1. Writing $X = \bigcup_n X_n$ with $\mu(X_n) < \infty$, countable subadditivity 2.58 supplies n with

$A = E \cap X_n$ satisfying $0 < \mu(A) \le \mu(X_n) < \infty$. Put

$$f = \frac{1}{\mu(A)} \cdot \frac{\overline{h}}{|h|} \chi_A,$$

well defined because $|h| > t > 0$ on A. Then $|f| = \chi_A/\mu(A)$, so $\|f\|_1 = 1$, while $fh = |h|/\mu(A)$ on A and 0 elsewhere gives

$$\varphi_h(f) = \frac{1}{\mu(A)} \int_A |h| \, d\mu \ge t,$$

whence $\|\varphi_h\| \ge t$. Letting $t \uparrow \|h\|_\infty$ gives $\|\varphi_h\| = \|h\|_\infty$. One-to-oneness follows: $\varphi_{h_1} = \varphi_{h_2}$ gives $\|h_1 - h_2\|_\infty = \|\varphi_{h_1 - h_2}\| = 0$.

Problem 7B.15. Let

$$c_0 = \{(a_1, a_2, \ldots) \in \ell^\infty : \lim_{k \to \infty} a_k = 0\}.$$

Give c_0 the norm that it inherits as a subspace of ℓ^∞.
(a) Prove that c_0 is a Banach space.
(b) Prove that the dual space of c_0 can be identified with ℓ^1.

Solution. (a) c_0 is a closed subspace of ℓ^∞, which is a Banach space by 7.24 (it is $L^\infty(\mu)$ for μ counting measure on $\mathbb{Z}^+$, where no nonempty set is null, so the essential supremum is the supremum). It is a subspace because $\lim_{k\to\infty}(a_k + \alpha a'_k) = 0$ whenever $a, a' \in c_0$. For closedness, suppose $a^{(n)} \in c_0$ and $\|a^{(n)} - a\|_\infty \to 0$; given $\varepsilon > 0$, choose n with $\|a^{(n)} - a\|_\infty < \varepsilon/2$ and then K with $|a_k^{(n)}| < \varepsilon/2$ for $k > K$, so that

$$|a_k| \le |a_k - a_k^{(n)}| + |a_k^{(n)}| < \varepsilon \qquad \text{for } k > K,$$

giving $a \in c_0$. A Cauchy sequence in c_0 is Cauchy in ℓ^∞, hence converges there, and the limit lies in the closed set c_0; thus c_0 is a Banach space.
(b) The identification is $b \mapsto \varphi_b$ from ℓ^1 onto $(c_0)'$, where

$$\varphi_b(a) = \sum_{k=1}^\infty a_k b_k \qquad (a \in c_0).$$

This map is linear, and $\sum_{k=1}^\infty |a_k b_k| \le \|a\|_\infty \|b\|_1$ shows the series converges absolutely and $\|\varphi_b\| \le \|b\|_1$. For the reverse, fix n and let $a^{(n)}$ have k^{th} term $\overline{b_k}/|b_k|$ when $k \le n$ and $b_k \ne 0$, and 0 otherwise; then $a^{(n)} \in c_0$ (finitely many nonzero terms) with $\|a^{(n)}\|_\infty \le 1$, and

$$\sum_{k=1}^n |b_k| = \varphi_b\big(a^{(n)}\big) \le \|\varphi_b\| \, \|a^{(n)}\|_\infty \le \|\varphi_b\|.$$

Let $n \to \infty$: $\|\varphi_b\| = \|b\|_1$, so $b \mapsto \varphi_b$ is a linear isometry, in particular one-to-one.
For surjectivity, let $\varphi \in (c_0)'$ and set $b_k = \varphi(e_k)$, where e_k is the k^{th} standard basis sequence. Since $a^{(n)}$ above is a linear combination of $e_1, \ldots, e_n$, the same display with φ in place of φ_b gives $\|b\|_1 \le \|\varphi\| < \infty$, so $b \in \ell^1$. Now φ and φ_b agree at each e_k, hence on the finitely supported sequences, which are dense in c_0 because $\|a - (a_1, \ldots, a_n, 0, 0, \ldots)\|_\infty = \sup_{k>n} |a_k| \to 0$. Both are continuous by 6.48, so $\varphi = \varphi_b$.

Problem 7B.16. Suppose $1 \le p \le 2$.
(a) Prove that if $w, z \in \mathbb{C}$, then

$$\frac{|w+z|^p + |w-z|^p}{2} \le |w|^p + |z|^p \le \frac{|w+z|^p + |w-z|^p}{2^{p-1}}.$$

(b) Prove that if μ is a measure and $f, g \in L^p(\mu)$, then

$$\frac{\|f+g\|_p^p + \|f-g\|_p^p}{2} \le \|f\|_p^p + \|g\|_p^p \le \frac{\|f+g\|_p^p + \|f-g\|_p^p}{2^{p-1}}.$$

Solution. (a) Put $r = p/2 \in (0,1]$, $\alpha = |w+z|^2$, $\beta = |w-z|^2$; the parallelogram equality gives $\alpha+\beta = 2(|w|^2+|z|^2)$. Concavity of $t \mapsto t^r$ on $[0,\infty)$ for $r \le 1$ (second derivative $r(r-1)t^{r-2} \le 0$), applied at α, β with equal weights, together with the subadditivity $(\sigma+\tau)^r \le \sigma^r + \tau^r$ for $\sigma, \tau \ge 0$, yields

$$\frac{|w+z|^p + |w-z|^p}{2} = \frac{\alpha^r+\beta^r}{2} \le \Big(\frac{\alpha+\beta}{2}\Big)^r = \big(|w|^2+|z|^2\big)^r \le |w|^p + |z|^p,$$

which is the left-hand inequality. (Subadditivity: with $t = \sigma/(\sigma+\tau) \in [0,1]$ one has $t^r \ge t$ and $(1-t)^r \ge 1-t$, so $t^r + (1-t)^r \ge 1$; multiply by $(\sigma+\tau)^r$.) Applying this with $w+z, w-z$ in place of w, z, and using $(w+z)+(w-z) = 2w$, $(w+z)-(w-z) = 2z$, gives

$$2^{p-1}\big(|w|^p+|z|^p\big) = \frac{|2w|^p + |2z|^p}{2} \le |w+z|^p + |w-z|^p;$$

divide by 2^{p-1} for the right-hand inequality.
(b) Apply (a) at each x with $w = f(x)$ and $z = g(x)$:

$$|f(x)+g(x)|^p + |f(x)-g(x)|^p \le 2\big(|f(x)|^p + |g(x)|^p\big)$$

and

$$2^{p-1}\big(|f(x)|^p + |g(x)|^p\big) \le |f(x)+g(x)|^p + |f(x)-g(x)|^p.$$

Each of $|f+g|^p, |f-g|^p, |f|^p, |g|^p$ is nonnegative, measurable, and has finite integral, since $L^p(\mu)$ is a vector space by 7.18 and $p < \infty$. Integrating the two displays, using that integration is order preserving (3.8) and additive on nonnegative measurable functions (3.16), gives

$$\|f+g\|_p^p + \|f-g\|_p^p \le 2\big(\|f\|_p^p + \|g\|_p^p\big)$$

and $2^{p-1}(\|f\|_p^p + \|g\|_p^p) \le \|f+g\|_p^p + \|f-g\|_p^p$. Divide the first by 2 and the second by 2^{p-1}.

Problem 7B.17. Suppose $2 \le p < \infty$.
(a) Prove that if $w, z \in \mathbb{C}$, then

$$\frac{|w+z|^p + |w-z|^p}{2^{p-1}} \le |w|^p + |z|^p \le \frac{|w+z|^p + |w-z|^p}{2}.$$

(b) Prove that if μ is a measure and $f, g \in L^p(\mu)$, then

$$\frac{\|f+g\|_p^p + \|f-g\|_p^p}{2^{p-1}} \le \|f\|_p^p + \|g\|_p^p \le \frac{\|f+g\|_p^p + \|f-g\|_p^p}{2}.$$

[The inequalities in the two previous exercises are called Clarkson's inequalities. They were discovered by James Clarkson in 1936.]

Solution. (a) Put $r = p/2 \ge 1$, $\alpha = |w+z|^2$, $\beta = |w-z|^2$; the parallelogram equality gives $\alpha+\beta = 2(|w|^2+|z|^2)$. Convexity of $t \mapsto t^r$ on $[0,\infty)$ for $r \ge 1$ (second derivative $r(r-1)t^{r-2} \ge 0$), applied at α, β with equal weights, together with the superadditivity $\sigma^r + \tau^r \le (\sigma+\tau)^r$ for $\sigma, \tau \ge 0$, yields

$$\frac{|w+z|^p + |w-z|^p}{2} = \frac{\alpha^r+\beta^r}{2} \ge \Big(\frac{\alpha+\beta}{2}\Big)^r = \big(|w|^2+|z|^2\big)^r \ge |w|^p + |z|^p,$$

which is the right-hand inequality. (Superadditivity: with $t = \sigma/(\sigma+\tau) \in [0,1]$ one has $t^r \le t$ and $(1-t)^r \le 1-t$, so $t^r + (1-t)^r \le 1$; multiply by $(\sigma+\tau)^r$.) Applying this with $w+z, w-z$ in place of

w, z, and using $(w+z)+(w-z)=2w$, $(w+z)-(w-z)=2z$, gives

$$|w+z|^p+|w-z|^p \le \frac{|2w|^p+|2z|^p}{2}=2^{p-1}\bigl(|w|^p+|z|^p\bigr);$$

divide by 2^{p-1} for the left-hand inequality.
(b) Apply (a) at each x with $w=f(x)$ and $z=g(x)$:

$$|f(x)+g(x)|^p+|f(x)-g(x)|^p \le 2^{p-1}\bigl(|f(x)|^p+|g(x)|^p\bigr)$$

and

$$2\bigl(|f(x)|^p+|g(x)|^p\bigr) \le |f(x)+g(x)|^p+|f(x)-g(x)|^p.$$

Each of $|f+g|^p, |f-g|^p, |f|^p, |g|^p$ is nonnegative, measurable, and has finite integral, since $L^p(\mu)$ is a vector space by 7.18 and $p<\infty$. Integrating the two displays, using that integration is order preserving (3.8) and additive on nonnegative measurable functions (3.16), gives

$$\|f+g\|_p^p+\|f-g\|_p^p \le 2^{p-1}\bigl(\|f\|_p^p+\|g\|_p^p\bigr)$$

and $2(\|f\|_p^p+\|g\|_p^p) \le \|f+g\|_p^p+\|f-g\|_p^p$. Divide the first by 2^{p-1} and the second by 2.

Problem 7B.18. Suppose $(X,\mathcal{S},\mu)$ is a measure space, $1\le p,q\le\infty$, and $h: X\to\mathbf{F}$ is an $\mathcal{S}$-measurable function such that $hf\in L^q(\mu)$ for every $f\in L^p(\mu)$. Prove that $f\mapsto hf$ is a continuous linear map from $L^p(\mu)$ to $L^q(\mu)$.

Solution. Apply the Closed Graph Theorem 6.85 to $Tf=hf$, which by hypothesis maps $L^p(\mu)$ into $L^q(\mu)$ and is linear because $h(f_1+\alpha f_2)=hf_1+\alpha hf_2$ pointwise; both spaces are Banach by 7.24, so it remains only to show graph(T), a subspace by linearity, is closed.
Suppose $(f_k, hf_k)\to(f,g)$, that is (by 6.84) $\|f_k-f\|_p\to 0$ and $\|hf_k-g\|_q\to 0$. By 7.23 some subsequence has $f_{k_m}\to f$ pointwise almost everywhere, hence

$$h(x)f_{k_m}(x)\longrightarrow h(x)f(x) \quad \text{for a.e. } x.$$

Since $\|hf_{k_m}-g\|_q\to 0$ as well, 7.23 gives a further subsequence with $hf_{k_{m_j}}\to g$ almost everywhere. Off the union of the two exceptional null sets, uniqueness of limits in $\mathbf{F}$ forces $g=hf$; thus $g=Tf$ in $L^q(\mu)$ and the graph is closed.
Hence T is bounded by 6.85, so there is $c<\infty$ with $\|hf\|_q\le c\|f\|_p$ for all $f\in L^p(\mu)$, and T is continuous by 6.48.

Problem 7B.19. Prove that if $1<p<\infty$, then ℓ^p is reflexive.
[A Banach space is called reflexive if the canonical isometry of the Banach space into its double dual space is surjective (see Exercise 20 in Section 6D for the definitions of the double dual space and the canonical isometry).]

Solution. Let $p'=p/(p-1)$, so $1<p'<\infty$ and $(p')'=p$. By 7.26 the map $\Lambda_p b=\varphi_b$, where $\varphi_b(a)=\sum_{k=1}^{\infty}a_kb_k$, is a norm-preserving linear bijection from $\ell^{p'}$ onto $(\ell^p)'$; since $1\le p'<\infty$, 7.26 applies equally with p' in place of p, giving a norm-preserving linear bijection

$$\Lambda_{p'}:\ell^p\to(\ell^{p'})', \qquad (\Lambda_{p'}a)(b)=\sum_{k=1}^{\infty}b_ka_k.$$

The canonical isometry $\Phi:\ell^p\to(\ell^p)''$, $(\Phi a)(\varphi)=\varphi(a)$, is norm preserving by Exercise 20 in Section 6D, so only surjectivity is at issue.
Given $\psi\in(\ell^p)''$, the composition $\psi\circ\Lambda_p$ is linear and bounded, since $|\psi(\Lambda_p b)|\le\|\psi\|\|\Lambda_p b\|=\|\psi\|\|b\|_{p'}$;

hence $\psi \circ \Lambda_p \in (\ell^{p'})'$. Surjectivity of $\Lambda_{p'}$ gives $a \in \ell^p$ with $\Lambda_{p'} a = \psi \circ \Lambda_p$, and evaluating at $b \in \ell^{p'}$,

$$(\Phi a)(\varphi_b) = \varphi_b(a) = \sum_{k=1}^{\infty} b_k a_k = (\Lambda_{p'} a)(b) = \psi(\varphi_b).$$

Every $\varphi \in (\ell^p)'$ is some φ_b because Λ_p is onto, so $\Phi a = \psi$.

Problem 7B.20. Prove that ℓ^1 is not reflexive.

Solution. The element $\psi = \tau \circ \Lambda^{-1}$ of $(\ell^1)''$ constructed below is not in the range of the canonical isometry. Here $\Lambda b = \varphi_b$, with $\varphi_b(a) = \sum_{k=1}^{\infty} a_k b_k$, is the norm-preserving linear bijection of ℓ^∞ onto $(\ell^1)'$ supplied by 7.26 with $p = 1$, so Λ^{-1} is linear with $\|\Lambda^{-1}\varphi\|_\infty = \|\varphi\|$.
Let c_0 be as in Exercise 15 and $\mathbf{1} = (1, 1, 1, \ldots)$. Then c_0 is closed in ℓ^∞ by Exercise 15(a) and $\mathbf{1} \notin c_0$, so 6.73 (with $V = \ell^\infty$, $U = c_0$) gives $\tau \in (\ell^\infty)'$ with $\tau|_{c_0} = 0$ and $\tau(\mathbf{1}) \neq 0$. Then $\psi = \tau \circ \Lambda^{-1}$ is linear with $|\psi(\varphi)| \leq \|\tau\|\|\varphi\|$, so $\psi \in (\ell^1)''$ and $\psi(\varphi_b) = \tau(b)$ for all $b \in \ell^\infty$.
If $\Phi a = \psi$ for some $a \in \ell^1$, where $(\Phi a)(\varphi) = \varphi(a)$, then for every $b \in \ell^\infty$

$$\sum_{k=1}^{\infty} a_k b_k = \varphi_b(a) = (\Phi a)(\varphi_b) = \psi(\varphi_b) = \tau(b).$$

Taking $b = e_j$, the j^{th} standard basis sequence, which lies in c_0, gives $a_j = \tau(e_j) = 0$ for every j; so $a = 0$ and hence $\tau \equiv 0$, contradicting $\tau(\mathbf{1}) \neq 0$. Thus Φ is not surjective.

Problem 7B.21. Show that with the natural identifications, the canonical isometry of c_0 into its double dual space is the inclusion map of c_0 into ℓ^∞ (see Exercise 15 for the definition of c_0 and an identification of its dual space).

Solution. The claim is that $(\Lambda_\infty^{-1} \circ \Gamma)(\Phi a) = a$ for every $a \in c_0$, where the three identifications are as follows. By Exercise 15(b), $\Lambda_1 b = \psi_b$ with $\psi_b(a) = \sum_{k=1}^{\infty} a_k b_k$ is a norm-preserving linear bijection of ℓ^1 onto $(c_0)'$. By 7.26 with $p = 1$, $\Lambda_\infty d = \varphi_d$ with $\varphi_d(b) = \sum_{k=1}^{\infty} b_k d_k$ is a norm-preserving linear bijection of ℓ^∞ onto $(\ell^1)'$. Finally

$$\Gamma : (c_0)'' \to (\ell^1)', \qquad \Gamma\theta = \theta \circ \Lambda_1$$

is linear, is surjective ($\Gamma(\sigma \circ \Lambda_1^{-1}) = \sigma$), and is norm preserving because Λ_1 carries the closed unit ball of ℓ^1 onto that of $(c_0)'$:

$$\|\Gamma\theta\| = \sup_{\|b\|_1 \leq 1} |\theta(\Lambda_1 b)| = \sup_{\|\psi\| \leq 1} |\theta(\psi)| = \|\theta\|.$$

Now fix $a \in c_0 \subseteq \ell^\infty$ and let Φ be the canonical isometry, $(\Phi a)(\psi) = \psi(a)$. For every $b \in \ell^1$,

$$\big(\Gamma(\Phi a)\big)(b) = (\Phi a)(\psi_b) = \psi_b(a) = \sum_{k=1}^{\infty} a_k b_k = \varphi_a(b) = (\Lambda_\infty a)(b).$$

Hence $\Gamma(\Phi a) = \Lambda_\infty a$, so Φ becomes $a \mapsto a$ from c_0 into ℓ^∞: the inclusion map.

Problem 7B.22. Suppose $1 \leq p < \infty$ and V, W are Banach spaces. Show that $V \times W$ is a Banach space if the norm on $V \times W$ is defined by

$$\|(f, g)\| = \big(\|f\|^p + \|g\|^p\big)^{1/p}$$

for $f \in V$ and $g \in W$.

Solution. The norm is squeezed between two constant multiples of the product norm $|(f,g)|_\infty = \max\{\|f\|, \|g\|\}$ of 6.84:

$$|(f,g)|_\infty \le \|(f,g)\| \le 2^{1/p}\,|(f,g)|_\infty,$$

call this $(*)$; it follows from $\max\{a,b\}^p \le a^p + b^p \le 2\max\{a,b\}^p$ and the monotonicity of $t \mapsto t^{1/p}$. Here $V \times W$ carries the componentwise operations of Exercise 10 in Section 6B.

Positivity holds because $\|f\|^p + \|g\|^p = 0$ forces $\|f\| = \|g\| = 0$, and homogeneity because

$$\|\alpha(f,g)\| = \big(|\alpha|^p\|f\|^p + |\alpha|^p\|g\|^p\big)^{1/p} = |\alpha|\,\|(f,g)\|,$$

by homogeneity of the norms on V and W.

For the triangle inequality, let ν be counting measure on $X = \{1,2\}$ with σ-algebra 2^X; integration against ν is summation (Example 3.6), so every $h : X \to \mathbb{R}$ lies in $L^p(\nu)$ with $\|h\|_p = (|h(1)|^p + |h(2)|^p)^{1/p}$. Given $(f_1,g_1),(f_2,g_2)$, define $u,v : X \to \mathbb{R}$ by

$$u(1) = \|f_1\|, \quad u(2) = \|g_1\|, \qquad v(1) = \|f_2\|, \quad v(2) = \|g_2\|.$$

Then $\|f_1 + f_2\| \le (u+v)(1)$ and $\|g_1 + g_2\| \le (u+v)(2)$ by the triangle inequalities in V and W, and $t \mapsto t^p$, $t \mapsto t^{1/p}$ are increasing, so

$$\begin{aligned}
\|(f_1,g_1) + (f_2,g_2)\| &= \big(\|f_1+f_2\|^p + \|g_1+g_2\|^p\big)^{1/p} \\
&\le \Big(\big((u+v)(1)\big)^p + \big((u+v)(2)\big)^p\Big)^{1/p} \\
&= \|u+v\|_p \\
&\le \|u\|_p + \|v\|_p \\
&= \|(f_1,g_1)\| + \|(f_2,g_2)\|,
\end{aligned}$$

the second inequality being Minkowski's inequality 7.14 for ν.

For completeness, let (f_k,g_k) be Cauchy in $V \times W$. The left half of $(*)$ applied to the differences makes (f_k) Cauchy in V and (g_k) Cauchy in W, so $f_k \to f \in V$ and $g_k \to g \in W$; the right half of $(*)$ then gives

$$\|(f_k,g_k) - (f,g)\| \le 2^{1/p}\max\{\|f_k - f\|, \|g_k - g\|\} \longrightarrow 0.$$

8 Hilbert Spaces

Contents

8.1 Exercises 8A

Problem 8A.1. Let V denote the vector space of bounded continuous functions from $\mathbb{R}$ to $\mathbf{F}$. Let $r_1, r_2, \ldots$ be a list of the rational numbers. For $f, g \in V$, define

$$\langle f, g\rangle = \sum_{k=1}^{\infty} \frac{f(r_k)\,\overline{g(r_k)}}{2^k}.$$

Show that $\langle\cdot,\cdot\rangle$ is an inner product on V.

Solution. All four conditions of 8.1 hold. Writing $\|h\|_\infty = \sup_{\mathbb{R}}|h|$, finite since the elements of V are bounded, the estimate

$$\left|\frac{f(r_k)\,\overline{g(r_k)}}{2^k}\right| \le \frac{\|f\|_\infty\|g\|_\infty}{2^k}$$

together with $\sum_k 2^{-k} = 1$ makes every series below absolutely convergent, which legitimizes splitting sums and pulling out scalars.
Positivity. $\langle f, f\rangle = \sum_{k=1}^{\infty}|f(r_k)|^2/2^k \in [0,\infty)$.
Definiteness. If $\langle f, f\rangle = 0$, every term of that nonnegative series vanishes, so $f = 0$ on $\mathbb{Q}$; since $\mathbb{Q}$ is dense and f is continuous, $f = 0$. (Continuity is used only here.)
Linearity in the first slot. For $f, g, h \in V$ and $\alpha \in \mathbf{F}$,

$$\langle f+g, h\rangle = \sum_{k=1}^{\infty}\frac{f(r_k)\overline{h(r_k)}}{2^k} + \sum_{k=1}^{\infty}\frac{g(r_k)\overline{h(r_k)}}{2^k} = \langle f, h\rangle + \langle g, h\rangle,$$
$$\langle \alpha f, h\rangle = \alpha\sum_{k=1}^{\infty}\frac{f(r_k)\overline{h(r_k)}}{2^k} = \alpha\langle f, h\rangle.$$

Conjugate symmetry. Conjugation is continuous and additive, so it passes through the partial sums, and 2^k is real:

$$\overline{\langle g, f\rangle} = \sum_{k=1}^{\infty}\overline{\left(\frac{g(r_k)\overline{f(r_k)}}{2^k}\right)} = \sum_{k=1}^{\infty}\frac{f(r_k)\overline{g(r_k)}}{2^k} = \langle f, g\rangle.$$

Problem 8A.2. Prove that if μ is a measure and $f, g \in L^2(\mu)$, then

$$\|f\|^2\|g\|^2 - |\langle f, g\rangle|^2 = \frac{1}{2}\int\int |f(x)g(y) - g(x)f(y)|^2\,d\mu(y)\,d\mu(x).$$

Solution. Write $A = \|f\|^2 = \int |f|^2 d\mu$, $B = \|g\|^2 = \int |g|^2 d\mu$, and $C = \langle f, g\rangle = \int f\overline{g}\,d\mu$ (8.2, 8.5); then $A, B < \infty$, and $f\overline{g}, g\overline{f} \in L^1(\mu)$ with $\int g\overline{f}\,d\mu = \overline{C}$ by Cauchy–Schwarz 8.14 and conjugate symmetry. Both integrations below use only linearity of the integral, so Tonelli is not needed.
Fix x; expanding $|w|^2 = w\overline{w}$ with $w = f(x)g(y) - g(x)f(y)$ gives, for each y,

$$\begin{aligned}|f(x)g(y) - g(x)f(y)|^2 = {}& |f(x)|^2|g(y)|^2 + |g(x)|^2|f(y)|^2\\ & - f(x)\overline{g(x)}\,g(y)\overline{f(y)} \quad - \overline{f(x)}\,g(x)\,f(y)\overline{g(y)},\end{aligned}$$

a linear combination in y, with finite scalar coefficients, of $|g|^2, |f|^2, g\overline{f}, f\overline{g} \in L^1(\mu)$. Hence by linearity

$$\int |f(x)g(y) - g(x)f(y)|^2\,d\mu(y) = |f(x)|^2 B + |g(x)|^2 A \quad - f(x)\overline{g(x)}\,\overline{C} - \overline{f(x)}g(x)\,C.$$

This is in turn a linear combination in x of the same four $L^1(\mu)$ functions, so integrating again by linearity,

$$\int\int |f(x)g(y) - g(x)f(y)|^2 \, d\mu(y) \, d\mu(x) = AB + BA - \overline{C}\,C - C\,\overline{C}$$
$$= 2AB - 2|C|^2.$$

Halving gives $AB - |C|^2 = \|f\|^2\|g\|^2 - |\langle f, g\rangle|^2$.

Problem 8A.3. Suppose f and g are elements of an inner product space and

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2.$$

(a) Prove that if $\mathbf{F} = \mathbb{R}$, then f and g are orthogonal.
(b) Give an example to show that if $\mathbf{F} = \mathbb{C}$, then f and g can satisfy the equation above without being orthogonal.

Solution. The hypothesis says exactly that $\mathrm{Re}\langle f, g\rangle = 0$, since by 8.4, linearity in the first slot, 8.3(b), and conjugate symmetry,

$$\begin{aligned}\|f + g\|^2 &= \langle f + g, f + g\rangle \\ &= \langle f, f\rangle + \langle g, f\rangle + \langle f, g\rangle + \langle g, g\rangle \\ &= \|f\|^2 + \|g\|^2 + \langle f, g\rangle + \overline{\langle f, g\rangle} \\ &= \|f\|^2 + \|g\|^2 + 2\,\mathrm{Re}\langle f, g\rangle,\end{aligned}$$

(a) If $\mathbf{F} = \mathbb{R}$ then $\langle f, g\rangle$ is real, so $\langle f, g\rangle = \mathrm{Re}\langle f, g\rangle = 0$: orthogonal by 8.7.
(b) Take $V = \mathbb{C}$ with $\langle w, z\rangle = w\overline{z}$ (8.2 with $n = 1$), $f = 1$, $g = i$. Then $\|f\|^2 = \|g\|^2 = 1$ and $\|f + g\|^2 = |1 + i|^2 = 2$, so the displayed equation holds, while $\langle f, g\rangle = \overline{i} = -i \neq 0$.

Problem 8A.4. Find $a, b \in \mathbb{R}^3$ such that a is a scalar multiple of $(1, 6, 3)$, b is orthogonal to $(1, 6, 3)$, and $(5, 4, -2) = a + b$.

Solution. Take

$$a = \left(\tfrac{1}{2},\, 3,\, \tfrac{3}{2}\right), \qquad b = \left(\tfrac{9}{2},\, 1,\, -\tfrac{7}{2}\right).$$

These come from the orthogonal decomposition 8.10 applied to $u = (5, 4, -2)$ and $v = (1, 6, 3) \neq 0$ in $\mathbb{R}^3$ with its standard inner product (8.2): $a = \frac{\langle u,v\rangle}{\|v\|^2} v$ and $b = u - a$, where $\langle u, v\rangle = 5 + 24 - 6 = 23$ and $\|v\|^2 = 1 + 36 + 9 = 46$, so the scalar is $\frac{23}{46} = \frac{1}{2}$. Then $\langle b, v\rangle = \frac{9}{2} + 6 - \frac{21}{2} = 0$ and $a + b = (5, 4, -2)$. (Check!)

Problem 8A.5. Prove that

$$16 \leq (a + b + c + d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right)$$

for all positive numbers a, b, c, d, with equality if and only if $a = b = c = d$.

Solution. Apply Cauchy–Schwarz 8.11 in $\mathbb{R}^4$ with its standard inner product (8.2) to the nonzero vectors

$$u = \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right), \qquad v = \left(\tfrac{1}{\sqrt{a}}, \tfrac{1}{\sqrt{b}}, \tfrac{1}{\sqrt{c}}, \tfrac{1}{\sqrt{d}}\right),$$

which are defined because $a, b, c, d > 0$. Since $\langle u, v\rangle = 4$, $\|u\|^2 = a + b + c + d$, and $\|v\|^2 = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$, squaring $|\langle u, v\rangle| \leq \|u\|\|v\|$ gives

$$16 \leq (a + b + c + d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right).$$

By 8.11 equality holds if and only if one of u, v is a scalar multiple of the other; as both are nonzero this means $v = \lambda u$ with $\lambda \neq 0$, and comparing coordinates gives $\lambda = \frac{1}{a} = \frac{1}{b} = \frac{1}{c} = \frac{1}{d}$, that is, $a = b = c = d$. Conversely $a = b = c = d$ gives $(4a)(4/a) = 16$.

Problem 8A.6. Prove that the square of the average of each finite list of real numbers containing at least two distinct real numbers is less than the average of the squares of the numbers in that list.

Solution. Apply Cauchy–Schwarz 8.11 in $\mathbb{R}^n$ (8.2) to $u = (x_1, \dots, x_n)$ and $v = (1, \dots, 1)$, where $x_1, \dots, x_n$ is the list (so $n \geq 2$). Since $\langle u, v\rangle = x_1 + \cdots + x_n$, $\|u\|^2 = x_1^2 + \cdots + x_n^2$, and $\|v\|^2 = n$, squaring $|\langle u, v\rangle| \leq \|u\|\|v\|$ and dividing by n^2 gives

$$\left(\frac{x_1 + \cdots + x_n}{n}\right)^2 \leq \frac{x_1^2 + \cdots + x_n^2}{n}.$$

Equality would force, by 8.11 and $v \neq 0$, that $u = \lambda v$, that is $x_1 = \cdots = x_n$, contradicting the presence of two distinct entries. Hence the inequality is strict.

Problem 8A.7. Suppose f and g are elements of an inner product space and $\|f\| \leq 1$ and $\|g\| \leq 1$. Prove that

$$\sqrt{1 - \|f\|^2}\,\sqrt{1 - \|g\|^2} \leq 1 - |\langle f, g\rangle|.$$

Solution. Write $s = \|f\|$ and $t = \|g\|$, both in $[0, 1]$ by hypothesis, so the square roots are defined. Cauchy–Schwarz 8.11 gives $|\langle f, g\rangle| \leq st$, so it suffices to prove $\sqrt{1 - s^2}\sqrt{1 - t^2} \leq 1 - st$. Both sides are nonnegative ($1 - st \geq 0$ since $s, t \leq 1$), so this is equivalent to the squared form, and

$$(1 - s^2)(1 - t^2) \leq (1 - st)^2 \iff 2st \leq s^2 + t^2 \iff 0 \leq (s - t)^2.$$

Hence

$$\sqrt{1 - \|f\|^2}\,\sqrt{1 - \|g\|^2} \leq 1 - st \leq 1 - |\langle f, g\rangle|.$$

Problem 8A.8. Suppose a and b are nonzero elements of $\mathbb{R}^2$. Prove that

$$\langle a, b\rangle = \|a\|\,\|b\| \cos\theta,$$

where θ is the angle between a and b (thinking of a as the vector whose initial point is the origin and whose end point is a, and similarly for b).
Hint: Draw the triangle formed by a, b, and $a - b$; then use the law of cosines.

Solution. Compare two evaluations of $\|a - b\|^2$. Algebraically, by additivity in each slot and symmetry of the standard (real) inner product on $\mathbb{R}^2$,

$$\begin{aligned}\|a - b\|^2 &= \langle a - b,\, a - b\rangle \\ &= \langle a, a\rangle - \langle a, b\rangle - \langle b, a\rangle + \langle b, b\rangle \\ &= \|a\|^2 + \|b\|^2 - 2\langle a, b\rangle.\end{aligned}$$

Geometrically, when O, $A = a$, $B = b$ are not collinear the triangle OAB has sides $\|a\|, \|b\|, \|a - b\|$ with angle θ at O, so the law of cosines gives

$$\|a - b\|^2 = \|a\|^2 + \|b\|^2 - 2\|a\|\,\|b\| \cos\theta.$$

Cancelling $\|a\|^2 + \|b\|^2$ and dividing by -2 gives $\langle a, b\rangle = \|a\|\|b\| \cos\theta$.
In the collinear case write $b = ta$ with $t \neq 0$, and use $\|b\| = |t|\|a\|$ (8.6).
(i) $t > 0$: $\theta = 0$ and $\langle a, b\rangle = t\|a\|^2 = \|a\|\,\|b\| = \|a\|\|b\| \cos\theta$.
(ii) $t < 0$: $\theta = \pi$ and $\langle a, b\rangle = t\|a\|^2 = -\|a\|\,\|b\| = \|a\|\|b\| \cos\theta$.

Problem 8A.9. The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbb{R}^2$ or $\mathbb{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbb{R}^n$ for $n > 3$. Thus the angle between two nonzero vectors $a, b \in \mathbb{R}^n$ is defined to be

$$\arccos \frac{\langle a, b\rangle}{\|a\|\,\|b\|},$$

where the motivation for this definition comes from the previous exercise. Explain why the Cauchy–Schwarz inequality is needed to show that this definition makes sense.

Solution. Because arccos, the inverse of the bijection $\cos : [0, \pi] \to [-1, 1]$, has domain $[-1, 1]$. So the definition makes sense only if

$$\frac{\langle a, b\rangle}{\|a\|\,\|b\|} \in [-1, 1],$$

which, since $\|a\|\|b\| > 0$ for nonzero a, b, says exactly that $|\langle a, b\rangle| \le \|a\|\|b\|$ — the Cauchy–Schwarz inequality 8.11 in $\mathbb{R}^n$. Granted it, each pair of nonzero vectors gets a unique angle in $[0, \pi]$.

Problem 8A.10. (a) Suppose f and g are elements of a real inner product space. Prove that f and g have the same norm if and only if $f + g$ is orthogonal to $f - g$.
(b) Use (a) to show that the diagonals of a parallelogram are perpendicular to each other if and only if the parallelogram is a rhombus.

Solution. (a) Both conditions say $\|f\|^2 = \|g\|^2$. Indeed, symmetry of the real inner product (8.1), additivity, and 8.3(b),(c) give

$$\begin{aligned}\langle f + g,\, f - g\rangle &= \langle f, f\rangle - \langle f, g\rangle \quad + \langle g, f\rangle - \langle g, g\rangle \\ &= \|f\|^2 - \langle f, g\rangle + \langle f, g\rangle - \|g\|^2 \\ &= \|f\|^2 - \|g\|^2.\end{aligned}$$

By 8.7, orthogonality of $f+g$ and $f-g$ means this vanishes, that is $\|f\|^2 = \|g\|^2$, equivalently $\|f\| = \|g\|$ since norms are nonnegative.
(b) Translate the parallelogram so its vertices are $0, f, f + g, g$, where $f, g \in \mathbb{R}^2$ are the nonzero non-parallel spanning vectors (orthogonality in $\mathbb{R}^2$ is perpendicularity, by the discussion after 8.8). Its diagonals are the arrows $f + g$ and $g - f$, and by 8.3(c) with $\alpha = -1$,

$$\langle f + g,\, g - f\rangle = -\langle f + g,\, f - g\rangle,$$

so the diagonals are perpendicular exactly when $\langle f + g, f - g\rangle = 0$, which by (a) means $\|f\| = \|g\|$. The four side lengths are $\|f\|, \|g\|, \|-f\| = \|f\|, \|-g\| = \|g\|$ by 8.6, so they are all equal — the parallelogram is a rhombus — under the same condition $\|f\| = \|g\|$.

Problem 8A.11. Suppose f and g are elements of an inner product space. Prove that $\|f\| = \|g\|$ if and only if $\|sf + tg\| = \|tf + sg\|$ for all $s, t \in \mathbb{R}$.

Solution. Everything follows from the identity

$$\|sf + tg\|^2 - \|tf + sg\|^2 = (s^2 - t^2)\big(\|f\|^2 - \|g\|^2\big) \qquad (s, t \in \mathbb{R}).$$

To prove it, expand by additivity in each slot, homogeneity in the first, and 8.3(c) in the second (the real scalars s, t are their own conjugates):

$$\begin{aligned}\|sf + tg\|^2 &= s^2\langle f, f\rangle + st\langle f, g\rangle \quad + ts\langle g, f\rangle + t^2\langle g, g\rangle \\ &= s^2\|f\|^2 + t^2\|g\|^2 + st\big(\langle f, g\rangle + \overline{\langle f, g\rangle}\big) \\ &= s^2\|f\|^2 + t^2\|g\|^2 + 2st\,\mathrm{Re}\langle f, g\rangle,\end{aligned}$$

using $\langle g, f\rangle = \overline{\langle f, g\rangle}$; interchanging s and t leaves the cross term $2st\,\mathrm{Re}\langle f, g\rangle$ fixed, so subtracting gives the identity.
If $\|f\| = \|g\|$, the right side vanishes for all s, t, so $\|sf + tg\| = \|tf + sg\|$. Conversely, taking $s = 1$, $t = 0$ in the hypothesis gives $\|f\| = \|g\|$ at once.

Problem 8A.12. Suppose f and g are elements of an inner product space and $\|f\| = \|g\| = 1$ and $\langle f, g\rangle = 1$. Prove that $f = g$.

Solution. $\|f - g\| = 0$, so $f = g$ by the definiteness requirement in 8.1. Indeed $\langle g, f\rangle = \overline{\langle f, g\rangle} = 1$ by conjugate symmetry, so expanding by additivity in each slot and 8.3(c) with $\alpha = -1$,

$$\begin{aligned}\|f - g\|^2 &= \langle f - g,\ f - g\rangle \\ &= \langle f, f\rangle - \langle f, g\rangle - \langle g, f\rangle + \langle g, g\rangle \\ &= \|f\|^2 - \langle f, g\rangle - \langle g, f\rangle + \|g\|^2 \\ &= 1 - 1 - 1 + 1 \\ &= 0.\end{aligned}$$

Method (2): the hypotheses give $|\langle f, g\rangle| = 1 = \|f\|\|g\|$, equality in Cauchy–Schwarz 8.11, so one of f, g is a scalar multiple of the other; both being nonzero, $f = \alpha g$ for some scalar α, and then $1 = \langle f, g\rangle = \alpha\|g\|^2 = \alpha$.

Problem 8A.13. Suppose f and g are elements of a real inner product space. Prove that

$$\langle f, g\rangle = \frac{\|f + g\|^2 - \|f - g\|^2}{4}.$$

Solution. Expand both squared norms. Conjugate symmetry over $\mathbb{R}$ reads $\langle g, f\rangle = \langle f, g\rangle$, so by additivity in each slot and 8.3(b),

$$\begin{aligned}\|f + g\|^2 &= \langle f + g,\ f + g\rangle \\ &= \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle \\ &= \|f\|^2 + \|g\|^2 + 2\langle f, g\rangle.\end{aligned}$$

Replacing g by $-g$ (and using 8.3(c) with $\alpha = -1$, so that $\langle f, -g\rangle = -\langle f, g\rangle$, together with $\|-g\| = \|g\|$ from 8.6) gives

$$\|f - g\|^2 = \|f\|^2 + \|g\|^2 - 2\langle f, g\rangle.$$

Subtracting, $\|f\|^2 + \|g\|^2$ cancels and $\|f + g\|^2 - \|f - g\|^2 = 4\langle f, g\rangle$; divide by 4.

Problem 8A.14. Suppose f and g are elements of a complex inner product space. Prove that

$$\langle f, g\rangle = \frac{\|f + g\|^2 - \|f - g\|^2 + \|f + ig\|^2\, i - \|f - ig\|^2\, i}{4}.$$

Solution. The four terms recover the real and imaginary parts of $\langle f, g\rangle = x + yi$ separately. Throughout, $\langle g, f\rangle = \overline{\langle f, g\rangle}$ and $\langle f, \alpha g\rangle = \overline{\alpha}\langle f, g\rangle$ by 8.3(c). Expanding as in the proof of 8.15,

$$\begin{aligned}\|f + g\|^2 &= \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle \\ &= \|f\|^2 + \|g\|^2 + 2\,\mathrm{Re}\langle f, g\rangle, \\ \|f - g\|^2 &= \langle f, f\rangle - \langle f, g\rangle - \langle g, f\rangle + \langle g, g\rangle \\ &= \|f\|^2 + \|g\|^2 - 2\,\mathrm{Re}\langle f, g\rangle,\end{aligned}$$

since $\langle f, g\rangle + \langle g, f\rangle = 2\,\mathrm{Re}\langle f, g\rangle$; subtracting gives $\|f + g\|^2 - \|f - g\|^2 = 4x$.

Applying the first identity with ig in place of g, where $\|ig\| = \|g\|$ by 8.6 and $\langle f, ig\rangle = \bar{i}\langle f, g\rangle = -i(x+yi)$ has real part y, gives $\|f+ig\|^2 = \|f\|^2 + \|g\|^2 + 2y$; with $-ig$ in place of g the real part is $-y$, so $\|f-ig\|^2 = \|f\|^2 + \|g\|^2 - 2y$. Subtracting, $\|f+ig\|^2 - \|f-ig\|^2 = 4y$. Hence

$$\|f+g\|^2 - \|f-g\|^2 + \|f+ig\|^2\, i - \|f-ig\|^2\, i = 4x + 4yi = 4\langle f, g\rangle.$$

Problem 8A.15. Suppose f, g, h are elements of an inner product space. Prove that

$$\left\|h - \tfrac{1}{2}(f+g)\right\|^2 = \frac{\|h-f\|^2 + \|h-g\|^2}{2} - \frac{\|f-g\|^2}{4}.$$

Solution. Apply the parallelogram equality 8.20 to $u = h - f$ and $v = h - g$, for which $u + v = 2\big(h - \frac{1}{2}(f+g)\big)$ and $u - v = g - f$:

$$4\left\|h - \tfrac{1}{2}(f+g)\right\|^2 + \|f-g\|^2 = 2\|h-f\|^2 + 2\|h-g\|^2,$$

using $\|2w\| = 2\|w\|$ and $\|g-f\| = \|f-g\|$ (8.6). Solve for the first term and divide by 4.

Problem 8A.16. Prove that a norm satisfying the parallelogram equality comes from an inner product. In other words, show that if V is a normed vector space whose norm $\|\cdot\|$ satisfies the parallelogram equality, then there is an inner product $\langle\cdot,\cdot\rangle$ on V such that $\|f\| = \langle f, f\rangle^{1/2}$ for all $f \in V$.

Solution. Take the polarization formulas of 8A.13 and 8A.14 — the only candidates — and verify the axioms 8.1; here $\|\cdot\|$ satisfies $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$ throughout.
(1) Real case. For $\mathbf{F} = \mathbf{R}$ define

$$\langle f, g\rangle = \frac{\|f+g\|^2 - \|f-g\|^2}{4} \qquad \text{for } f, g \in V.$$

(i) Taking $g = f$ gives $\langle f, f\rangle = \frac{1}{4}\|2f\|^2 = \|f\|^2$, which is nonnegative and vanishes only for $f = 0$, and gives $\|f\| = \langle f, f\rangle^{1/2}$.
(ii) Symmetry: $\|g-f\| = \|f-g\|$, and all values are real.
(iii) Additivity in the first slot. Applying the parallelogram equality to the pairs f_1+g, f_2 and f_1-g, f_2 gives

$$\begin{aligned}\|f_1+f_2+g\|^2 + \|f_1-f_2+g\|^2 &= 2\|f_1+g\|^2 + 2\|f_2\|^2,\\ \|f_1+f_2-g\|^2 + \|f_1-f_2-g\|^2 &= 2\|f_1-g\|^2 + 2\|f_2\|^2.\end{aligned}$$

Subtracting the second equation from the first and dividing by 4 gives

$$\langle f_1+f_2, g\rangle + \langle f_1-f_2, g\rangle = \frac{\|f_1+g\|^2 - \|f_1-g\|^2}{2} = 2\langle f_1, g\rangle. \qquad (*)$$

Since $\langle 0, g\rangle = 0$, taking $f_2 = f_1$ in $(*)$ gives $\langle 2f, g\rangle = 2\langle f, g\rangle$. Now $(*)$ with $f_1 = \frac{u+v}{2}$, $f_2 = \frac{u-v}{2}$, so that $f_1 + f_2 = u$ and $f_1 - f_2 = v$, gives

$$\langle u, g\rangle + \langle v, g\rangle = 2\left\langle \tfrac{u+v}{2}, g\right\rangle = \langle u+v, g\rangle$$

by that doubling identity: additivity.
(iv) Homogeneity in the first slot. From the definition,

$$\langle -f, g\rangle = \frac{\|-f+g\|^2 - \|-f-g\|^2}{4} = \frac{\|f-g\|^2 - \|f+g\|^2}{4} = -\langle f, g\rangle.$$

With additivity and induction this gives $\langle nf, g\rangle = n\langle f, g\rangle$ for $n \in \mathbf{Z}$, whence (applying it to $\frac{1}{m}f$) $\langle tf, g\rangle = t\langle f, g\rangle$ for all rational t. Both $t \mapsto \langle tf, g\rangle$ and $t \mapsto t\langle f, g\rangle$ are continuous on $\mathbf{R}$, the first because $\big|\|tf \pm g\| - \|sf \pm g\|\big| \le |t-s|\|f\|$, so agreeing on the dense set $\mathbf{Q}$ they agree on $\mathbf{R}$.

By (i)-(iv) the requirements of 8.1 hold, so this is an inner product with associated norm $\|\cdot\|$.
(2) Complex case. Restricting scalars to $\mathbf{R}$ leaves the norm and the parallelogram equality intact, so by (1)

$$\langle f,g\rangle_{\mathbf{R}} = \frac{\|f+g\|^2 - \|f-g\|^2}{4}$$

is a real inner product on V as a real vector space: real valued, symmetric, additive and $\mathbf{R}$-homogeneous in each slot, with $\langle f,f\rangle_{\mathbf{R}} = \|f\|^2$. Define

$$\langle f,g\rangle = \langle f,g\rangle_{\mathbf{R}} + i\,\langle f,ig\rangle_{\mathbf{R}}.$$

Since $\|iu\| = \|u\|$, we have $\langle if,ig\rangle_{\mathbf{R}} = \langle f,g\rangle_{\mathbf{R}}$; call this (†). Replacing g by ig there and using $i(ig) = -g$ gives $\langle f,ig\rangle_{\mathbf{R}} = -\langle if,g\rangle_{\mathbf{R}}$, call it (‡).
(i) $\langle f,if\rangle_{\mathbf{R}} = \frac{1}{4}(|1+i|^2 - |1-i|^2)\|f\|^2 = 0$, so $\langle f,f\rangle = \|f\|^2$: nonnegative, zero only at $f=0$, and $\|f\| = \langle f,f\rangle^{1/2}$.
(ii) $f \mapsto \langle f,g\rangle$ is additive and $\mathbf{R}$-homogeneous, since $f \mapsto \langle f,g\rangle_{\mathbf{R}}$ and $f \mapsto \langle f,ig\rangle_{\mathbf{R}}$ are, by (1).
(iii) By the definition, then (†), then (‡),

$$\langle if,g\rangle = \langle if,g\rangle_{\mathbf{R}} + i\langle if,ig\rangle_{\mathbf{R}} = -\langle f,ig\rangle_{\mathbf{R}} + i\langle f,g\rangle_{\mathbf{R}} = i\langle f,g\rangle.$$

Hence for $\alpha = s+it$ with $s,t \in \mathbf{R}$,

$$\langle \alpha f,g\rangle = \langle sf,g\rangle + \langle t(if),g\rangle = s\langle f,g\rangle + t\,i\langle f,g\rangle = \alpha\langle f,g\rangle,$$

so the first slot is $\mathbf{C}$-linear.
(iv) Conjugate symmetry, by symmetry of $\langle\cdot,\cdot\rangle_{\mathbf{R}}$, then (‡), the two real terms being real:

$$\langle g,f\rangle = \langle f,g\rangle_{\mathbf{R}} + i\langle if,g\rangle_{\mathbf{R}} = \langle f,g\rangle_{\mathbf{R}} - i\langle f,ig\rangle_{\mathbf{R}} = \overline{\langle f,g\rangle}.$$

Problem 8A.17. Let λ denote Lebesgue measure on $[1,\infty)$.
(a) Prove that if $f\colon [1,\infty) \to [0,\infty)$ is Borel measurable, then

$$\left(\int_1^\infty f(x)\,d\lambda(x)\right)^2 \le \int_1^\infty x^2\big(f(x)\big)^2\,d\lambda(x).$$

(b) Describe the set of Borel measurable functions $f\colon [1,\infty) \to [0,\infty)$ such that the inequality in (a) is an equality.

Solution. (a) This is Cauchy–Schwarz 8.14 applied to $g(x) = xf(x)$ and $h(x) = 1/x$, for which $gh = f$, $g^2(x) = x^2 f(x)^2$, and

$$\|h\|_2^2 = \int_1^\infty \frac{1}{x^2}\,d\lambda(x) = \lim_{n\to\infty}\left(1 - \frac{1}{n}\right) = 1$$

by the Monotone Convergence Theorem 3.11 applied to $x^{-2}\chi_{[1,n]}$, the integral over $[1,n]$ being the Riemann integral there (3.34). We may assume $g \in L^2(\lambda)$, as otherwise the right side is ∞. Since $g,h \ge 0$, $\langle g,h\rangle = \int_1^\infty f\,d\lambda \ge 0$, so

$$\int_1^\infty f\,d\lambda = \big|\langle g,h\rangle\big| \le \|g\|_2\|h\|_2 = \left(\int_1^\infty x^2 f(x)^2\,d\lambda(x)\right)^{1/2};$$

square both sides.
(b) Equality holds exactly when $\int_1^\infty f\,d\lambda = \infty$, or $f(x) = c/x^2$ for almost every x and some $c \in [0,\infty)$.
(i) If $\int_1^\infty f\,d\lambda = \infty$, then (a) forces both sides to be ∞. If $f = c/x^2$ a.e., both sides equal c^2, by the computation of $\int_1^\infty x^{-2} d\lambda = 1$ above.
(ii) Conversely, suppose equality holds with $\int_1^\infty f\,d\lambda < \infty$; then $\|g\|_2^2 = (\int_1^\infty f d\lambda)^2 < \infty$, so $g \in L^2(\lambda)$ and the equality reads $|\langle g,h\rangle| = \|g\|_2\|h\|_2$. By the equality condition in 8.11, and since $h \ne 0$, we get

$g = \alpha h$ in $L^2(\lambda)$, where

$$\alpha = \alpha\|h\|_2^2 = \langle \alpha h, h\rangle = \langle g, h\rangle = \int_1^\infty f \, d\lambda \geq 0.$$

Thus $xf(x) = \alpha/x$, that is $f(x) = \alpha/x^2$, for almost every x.

Problem 8A.18. Suppose μ is a measure. For $f, g \in L^2(\mu)$, define $\langle f, g\rangle$ by

$$\langle f, g\rangle = \int f\overline{g} \, d\mu.$$

(a) Using the inequality

$$\bigl|f(x)\,\overline{g(x)}\bigr| \leq \tfrac{1}{2}\bigl(|f(x)|^2 + |g(x)|^2\bigr),$$

verify that the integral above makes sense and the map sending f, g to $\langle f, g\rangle$ defines an inner product on $L^2(\mu)$ (without using Hölder's inequality).
(b) Show that the Cauchy–Schwarz inequality implies that

$$\|fg\|_1 \leq \|f\|_2 \, \|g\|_2$$

for all $f, g \in L^2(\mu)$ (again, without using Hölder's inequality).

Solution. (a) The integral converges absolutely, because $0 \leq (|f| - |g|)^2$ gives $|f\overline{g}| = |f|\,|g| \leq \frac{1}{2}(|f|^2 + |g|^2)$ pointwise, whence by 3.8

$$\int |f\overline{g}| \, d\mu \leq \tfrac{1}{2}\bigl(\|f\|_2^2 + \|g\|_2^2\bigr) < \infty$$

for $f, g \in L^2(\mu)$; thus $\langle f, g\rangle \in \mathbf{F}$ (6.21). Now check the four requirements of 8.1. Linearity in the first slot is homogeneity and additivity of the integral (3.20, 3.21), all three integrands being absolutely integrable by the display:

$$\begin{aligned}\langle \alpha f_1 + f_2, g\rangle &= \alpha \int f_1\overline{g} \, d\mu + \int f_2\overline{g} \, d\mu \\ &= \alpha\langle f_1, g\rangle + \langle f_2, g\rangle.\end{aligned}$$

Conjugate symmetry follows from $\overline{f\overline{g}} = g\overline{f}$ and 6.24, and positivity from $\langle f, f\rangle = \int |f|^2 d\mu = \|f\|_2^2 \geq 0$. For definiteness, suppose $\int |f|^2 d\mu = 0$ and put $E_n = \{|f|^2 > 1/n\}$; then $\frac{1}{n}\chi_{E_n} \leq |f|^2$ forces $\mu(E_n) = 0$ by 3.8, so $\{f \neq 0\} = \bigcup_n E_n$ is null by 2.58 and f is the 0 element of $L^2(\mu)$. Hence $\langle \cdot, \cdot\rangle$ is an inner product, with associated norm (8.4) equal to $\|\cdot\|_2$; Hölder's inequality was nowhere used.
(b) Apply Cauchy–Schwarz (8.11) to $|f|, |g| \in L^2(\mu)$, which are real valued, so $\bigl\langle |f|, |g|\bigr\rangle = \int |f|\,|g| \, d\mu$:

$$\|fg\|_1 = \int |f|\,|g| \, d\mu \leq \bigl\||f|\bigr\|_2 \, \bigl\||g|\bigr\|_2 = \|f\|_2 \, \|g\|_2,$$

the first equality because $|f||g| = |fg|$.

Problem 8A.19. Suppose $V_1, \ldots, V_m$ are inner product spaces. Show that the equation

$$\bigl\langle (f_1, \ldots, f_m), (g_1, \ldots, g_m)\bigr\rangle = \langle f_1, g_1\rangle + \cdots + \langle f_m, g_m\rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.
[Each of the inner product spaces $V_1, \ldots, V_m$ may have a different inner product, even though the same inner product notation is used on all these spaces.]

Solution. Each of the four requirements of 8.1 for $V = V_1 \times \cdots \times V_m$ (a vector space over the common field $\mathbf{F}$ under coordinatewise operations) reduces coordinatewise to the same requirement in V_k. Write $f = (f_1, \ldots, f_m)$, $g = (g_1, \ldots, g_m)$, $h = (h_1, \ldots, h_m)$.
Positivity and definiteness: $\langle f, f\rangle = \sum_{k=1}^m \langle f_k, f_k\rangle = \sum_{k=1}^m \|f_k\|^2 \geq 0$, and this sum of nonnegative

terms vanishes exactly when every $\|f_k\| = 0$, which by definiteness in V_k says $f_k = 0$ for each k, that is, $f = 0$.
Linearity in the first slot, from linearity of $u \mapsto \langle u, g_k\rangle$ on each V_k and $\alpha f + h = (\alpha f_1 + h_1, \dots, \alpha f_m + h_m)$:

$$\langle \alpha f + h, g\rangle = \sum_{k=1}^{m} \bigl(\alpha\langle f_k, g_k\rangle + \langle h_k, g_k\rangle\bigr) = \alpha\langle f, g\rangle + \langle h, g\rangle.$$

Conjugate symmetry, from conjugate symmetry in each V_k and additivity of conjugation (6.24):

$$\langle g, f\rangle = \sum_{k=1}^{m} \overline{\langle f_k, g_k\rangle} = \overline{\langle f, g\rangle}.$$

Hence $\langle\cdot,\cdot\rangle$ is an inner product on V, with associated norm $\|f\| = \bigl(\|f_1\|^2 + \cdots + \|f_m\|^2\bigr)^{1/2}$.

Problem 8A.20. Suppose V is an inner product space. Make $V \times V$ an inner product space as in the exercise above. Prove that the function that takes an ordered pair $(f, g) \in V \times V$ to the inner product $\langle f, g\rangle \in \mathbf{F}$ is a continuous function from $V \times V$ to $\mathbf{F}$.

Solution. Fix $(f, g) \in V \times V$ and $\varepsilon > 0$; the required δ is

$$\delta = \min\left\{1,\ \frac{\varepsilon}{\|f\| + \|g\| + 1}\right\} > 0.$$

By Exercise 19 the norm on $V \times V$ is $\|(f_1, f_2)\| = (\|f_1\|^2 + \|f_2\|^2)^{1/2}$, so $\|(u, v) - (f, g)\| < \delta$ forces both $\|u - f\| < \delta$ and $\|v - g\| < \delta$. Linearity in the first slot together with 8.3(b) gives $\langle u, v\rangle - \langle f, g\rangle = \langle u - f, v\rangle + \langle f, v - g\rangle$, so Cauchy–Schwarz (8.11) and $\|v\| \le \|g\| + \|v - g\| < \|g\| + 1$ (by 8.15 and $\delta \le 1$) yield

$$\bigl|\langle u, v\rangle - \langle f, g\rangle\bigr| \le \|u - f\|\,\|v\| + \|f\|\,\|v - g\| < \delta\bigl(\|f\| + \|g\| + 1\bigr) \le \varepsilon.$$

Thus $(f, g) \mapsto \langle f, g\rangle$ is continuous at each point of $V \times V$.

Problem 8A.21. Suppose $1 \le p \le \infty$.
(a) Show the norm on ℓ^p comes from an inner product if and only if $p = 2$.
(b) Show the norm on $L^p(\mathbf{R})$ comes from an inner product if and only if $p = 2$.

Solution. In both parts the answer is $p = 2$: for $p = 2$ the norm is the one associated with the standard inner product (8.2 and 8.5), and for $p \ne 2$ two disjointly supported unit vectors violate the parallelogram equality 8.20,

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2,$$

which every norm coming from an inner product satisfies.
(a) Take $u = (1, 0, 0, \dots)$ and $v = (0, 1, 0, \dots)$ in ℓ^p, so $\|u\|_p = \|v\|_p = 1$.
(i) $p < \infty$: here $\|u + v\|_p = \|u - v\|_p = 2^{1/p}$, so 8.20 reads $2 \cdot 2^{2/p} = 4$, that is $2^{1+2/p} = 2^2$; injectivity of $t \mapsto 2^t$ forces $p = 2$.
(ii) $p = \infty$: here $\|u + v\|_\infty = \|u - v\|_\infty = 1$, so the left side of 8.20 is 2 and the right side is 4.
(b) Take $u = \chi_{(0,1)}$ and $v = \chi_{(1,2)}$, which lie in $L^p(\mathbf{R})$ for every $p \in [1, \infty]$ and satisfy $\|u\|_p = \|v\|_p = 1$. Since $\lambda(\{1\}) = 0$,

$$u + v = \chi_{(0,2)} \quad \text{and} \quad |u - v| = \chi_{(0,2)} \qquad \text{almost everywhere,}$$

so $\|u + v\|_p = \|u - v\|_p = 2^{1/p}$ when $p < \infty$ and $= 1$ when $p = \infty$. Cases (i) and (ii) of part (a) now apply verbatim.

Problem 8A.22. Use inner products to prove Apollonius's identity:
In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$

Solution. The identity is the parallelogram equality 8.20 applied to the two sides emanating from the vertex opposite the side of length c. Work in $\mathbf{R}^2$ with the standard inner product, whose associated norm is Euclidean distance (8.5), and label the vertices P, Q, R so that

$$c = \|P - Q\|, \qquad b = \|P - R\|, \qquad a = \|Q - R\|.$$

Setting $f = P - R$ and $g = Q - R$ gives $\|f\| = b$, $\|g\| = a$, and $\|f - g\| = \|P - Q\| = c$; moreover the midpoint $M = \frac{1}{2}(P + Q)$ satisfies $M - R = \frac{1}{2}(f + g)$, so homogeneity of the norm (8.6) gives $\|f + g\| = 2d$. Hence 8.20 reads

$$(2d)^2 + c^2 = 2b^2 + 2a^2,$$

and dividing by 2 gives $a^2 + b^2 = \frac{1}{2}c^2 + 2d^2$.

8.2 Exercises 8B

Problem 8B.1. Show that each of the inner product spaces in Example 8.23 is not a Hilbert space.

Solution. Each of the two spaces in 8.23 carries a Cauchy sequence with no limit in it, so by 8.21 neither is a Hilbert space.
(i) ℓ^1 with $\|a\| = \big(\sum_k |a_k|^2\big)^{1/2}$. Take $a^{(n)} = (1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, 0, \ldots) \in \ell^1$. For $m < n$,

$$\|a^{(n)} - a^{(m)}\|^2 = \sum_{k=m+1}^{n} \frac{1}{k^2} \leq \sum_{k=m+1}^{\infty} \frac{1}{k^2} \to 0$$

as $m \to \infty$, so the sequence is Cauchy. If $b \in \ell^1$ had $\|a^{(n)} - b\| \to 0$, then $|a_j^{(n)} - b_j| \leq \|a^{(n)} - b\|$ forces $b_j = 1/j$ for every j, contradicting $\sum_j 1/j = \infty$.
(ii) $C([0,1])$ with $\|f\| = \big(\int_0^1 |f|^2\big)^{1/2}$. Take the continuous ramps

$$f_n(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \frac{1}{2}, \\ n\left(x - \frac{1}{2}\right) & \text{if } \frac{1}{2} \leq x \leq \frac{1}{2} + \frac{1}{n}, \\ 1 & \text{if } \frac{1}{2} + \frac{1}{n} \leq x \leq 1, \end{cases}$$

for $n \geq 3$. Since $f_n = f_m$ off $\left(\frac{1}{2}, \frac{1}{2} + \frac{1}{m}\right)$ and $|f_n - f_m| \leq 1$, we get $\|f_n - f_m\|^2 \leq \frac{1}{m}$ for $3 \leq m \leq n$, so the sequence is Cauchy. Suppose $f \in C([0,1])$ and $\|f_n - f\| \to 0$. Because $f_n = 0$ on $[0, \frac{1}{2}]$ and $f_n = 1$ on $[a, 1]$ once $\frac{1}{2} + \frac{1}{n} < a$,

$$\int_0^{1/2} |f|^2 \leq \|f_n - f\|^2 \to 0, \qquad \int_a^1 |1 - f|^2 \leq \|f_n - f\|^2 \to 0,$$

so both integrals vanish; continuity of the nonnegative integrands then gives $f = 0$ on $[0, \frac{1}{2}]$ and $f = 1$ on $[a, 1]$ for every $a \in (\frac{1}{2}, 1)$. Continuity of f at $\frac{1}{2}$ forces $f(\frac{1}{2}) = 1$ as well as $f(\frac{1}{2}) = 0$, a contradiction.

Problem 8B.2. Prove or disprove: The inner product space in Exercise 1 in Section 8A is a Hilbert space.

Solution. Disproof: the ramps

$$\varphi_n(x) = \begin{cases} 0 & \text{if } x \le 0, \\ nx & \text{if } 0 \le x \le \frac{1}{n}, \\ 1 & \text{if } x \ge \frac{1}{n} \end{cases}$$

form a Cauchy sequence with no limit in the space V of bounded continuous functions on $\mathbf{R}$ with $\|f\|^2 = \sum_k 2^{-k}|f(r_k)|^2$, so $\|\cdot\|$ is not complete and 8.21 rules out a Hilbert space. Each φ_n is continuous with $0 \le \varphi_n \le 1$, hence lies in V.

Cauchy: for $m \le n$ the functions φ_n and φ_m agree outside $(0, \frac{1}{m})$ and differ by at most 1, so with $A_m = \{k : 0 < r_k < \frac{1}{m}\}$,

$$\|\varphi_n - \varphi_m\|^2 \le \sum_{k \in A_m} 2^{-k} \longrightarrow 0 \quad \text{as } m \to \infty,$$

since the A_m decrease with $\bigcap_m A_m = \emptyset$ and $\sum_k 2^{-k} < \infty$.

No limit: if $f \in V$ had $\|\varphi_n - f\| \to 0$, then $2^{-j/2}|\varphi_n(r_j) - f(r_j)| \le \|\varphi_n - f\|$ forces $f(r_j) = \lim_n \varphi_n(r_j)$, which is 0 for $r_j \le 0$ and 1 for $r_j > 0$. Taking rationals $-1/n \uparrow 0$ and $1/n \downarrow 0$, continuity of f at 0 gives both $f(0) = 0$ and $f(0) = 1$.

Problem 8B.3. Suppose $V_1, V_2, \dots$ are Hilbert spaces. Let

$$V = \Big\{(f_1, f_2, \dots) \in V_1 \times V_2 \times \cdots : \sum_{k=1}^{\infty} \|f_k\|^2 < \infty\Big\}.$$

Show that the equation

$$\langle (f_1, f_2, \dots), (g_1, g_2, \dots) \rangle = \sum_{k=1}^{\infty} \langle f_k, g_k \rangle$$

defines an inner product on V that makes V a Hilbert space.

[Each of the Hilbert spaces $V_1, V_2, \dots$ may have a different inner product, even though the same notation is used for the norm and inner product on all these Hilbert spaces.]

Solution. First, V is a subspace of $V_1 \times V_2 \times \cdots$: it is closed under scalar multiplication because $\sum_k \|\alpha f_k\|^2 = |\alpha|^2 \sum_k \|f_k\|^2$, and under addition because 8.15 gives $\|f_k + g_k\|^2 \le (\|f_k\| + \|g_k\|)^2 \le 2\|f_k\|^2 + 2\|g_k\|^2$.

Second, the defining series converges absolutely for $f, g \in V$, by Cauchy–Schwarz (8.11) in V_k and $st \le \frac{1}{2}(s^2 + t^2)$:

$$\sum_{k=1}^{\infty} \big|\langle f_k, g_k \rangle\big| \le \sum_{k=1}^{\infty} \|f_k\|\,\|g_k\| \le \tfrac{1}{2} \sum_{k=1}^{\infty} \big(\|f_k\|^2 + \|g_k\|^2\big) < \infty.$$

The four requirements of 8.1 now hold termwise exactly as in Exercise 19 of Section 8A, absolutely convergent series being added and conjugated termwise; in particular $\langle f, f \rangle = \sum_k \|f_k\|^2$, which vanishes only when every $f_k = 0$, and the associated norm is $\|f\| = \big(\sum_k \|f_k\|^2\big)^{1/2}$.

Completeness. Let $f^{(1)}, f^{(2)}, \dots$ be Cauchy in V. Since $\|f_k^{(n)} - f_k^{(m)}\| \le \|f^{(n)} - f^{(m)}\|$ for each k, completeness of V_k supplies $f_k \in V_k$ with $f_k^{(n)} \to f_k$. Given $\varepsilon > 0$, choose N with $\|f^{(n)} - f^{(m)}\| \le \varepsilon$ for $m, n \ge N$; for $n \ge N$ and each K, letting $m \to \infty$ in the finite sum $\sum_{k=1}^{K} \|f_k^{(n)} - f_k^{(m)}\|^2 \le \varepsilon^2$ (legitimate since $\big|\|u\| - \|v\|\big| \le \|u - v\|$) and then $K \to \infty$ gives

$$\sum_{k=1}^{\infty} \|f_k^{(n)} - f_k\|^2 \le \varepsilon^2 \qquad \text{for all } n \ge N.$$

Thus $f^{(N)} - f \in V$, so $f = f^{(N)} - (f^{(N)} - f) \in V$ and the display reads $\|f^{(n)} - f\| \le \varepsilon$ for $n \ge N$. Hence V is complete, so by 8.21 it is a Hilbert space.

Problem 8B.4. Suppose V is a real Hilbert space. The complexification of V is the complex vector space $V_{\mathbf{C}}$ defined by $V_{\mathbf{C}} = V \times V$, but we write a typical element of $V_{\mathbf{C}}$ as $f + ig$ instead of (f, g). Addition and scalar multiplication are defined on $V_{\mathbf{C}}$ by

$$(f_1 + ig_1) + (f_2 + ig_2) = (f_1 + f_2) + i(g_1 + g_2)$$

and

$$(\alpha + i\beta)(f + ig) = (\alpha f - \beta g) + i(\alpha g + \beta f)$$

for $f_1, f_2, f, g_1, g_2, g \in V$ and $\alpha, \beta \in \mathbf{R}$. Show that

$$\langle f_1 + ig_1, f_2 + ig_2 \rangle = \langle f_1, f_2 \rangle + \langle g_1, g_2 \rangle + i\big(\langle g_1, f_2 \rangle - \langle f_1, g_2 \rangle\big)$$

defines an inner product on $V_{\mathbf{C}}$ that makes $V_{\mathbf{C}}$ into a complex Hilbert space.

Solution. Write $[\cdot,\cdot]$ for the proposed form and $A = \langle f_1, f_2 \rangle + \langle g_1, g_2 \rangle$, $B = \langle g_1, f_2 \rangle - \langle f_1, g_2 \rangle$, so $[f_1 + ig_1, f_2 + ig_2] = A + iB$ with $A, B \in \mathbf{R}$; here $\langle \cdot, \cdot \rangle$ is the inner product of V, symmetric and $\mathbf{R}$-bilinear.

Positivity and definiteness: taking $f_1 = f_2 = f$, $g_1 = g_2 = g$ makes $B = 0$ by symmetry, so

$$[f + ig, f + ig] = \|f\|^2 + \|g\|^2,$$

which is nonnegative and vanishes exactly when $f = g = 0$.

Additivity in the first slot is additivity of each $\langle \cdot, f_2 \rangle$ and $\langle \cdot, g_2 \rangle$ in its first argument, since $(f_1 + ig_1) + (f_1' + ig_1') = (f_1 + f_1') + i(g_1 + g_1')$. Complex homogeneity: for $\lambda = \alpha + i\beta$ we have $\lambda(f_1 + ig_1) = (\alpha f_1 - \beta g_1) + i(\alpha g_1 + \beta f_1)$, so $\mathbf{R}$-bilinearity gives

$$\begin{aligned}\big[\lambda(f_1 + ig_1),\, f_2 + ig_2\big] &= \langle \alpha f_1 - \beta g_1, f_2 \rangle + \langle \alpha g_1 + \beta f_1, g_2 \rangle \\ &\quad + i\big(\langle \alpha g_1 + \beta f_1, f_2 \rangle - \langle \alpha f_1 - \beta g_1, g_2 \rangle\big) \\ &= (\alpha A - \beta B) + i(\alpha B + \beta A) = \lambda(A + iB).\end{aligned}$$

Conjugate symmetry: swapping the two arguments fixes A and negates B, by symmetry of $\langle \cdot, \cdot \rangle$, so $[f_2 + ig_2, f_1 + ig_1] = A - iB = \overline{[f_1 + ig_1, f_2 + ig_2]}$. Hence $[\cdot,\cdot]$ is an inner product on $V_{\mathbf{C}}$, with associated norm $\|f + ig\|_{\mathbf{C}} = (\|f\|^2 + \|g\|^2)^{1/2}$.

Completeness: since $(f_1 + ig_1) - (f_2 + ig_2) = (f_1 - f_2) + i(g_1 - g_2)$, a Cauchy sequence $f_n + ig_n$ has $\|f_n - f_m\|, \|g_n - g_m\| \le \|(f_n + ig_n) - (f_m + ig_m)\|_{\mathbf{C}}$, so $f_n \to f$ and $g_n \to g$ for some $f, g \in V$ by completeness of V, and then

$$\|(f_n + ig_n) - (f + ig)\|_{\mathbf{C}}^2 = \|f_n - f\|^2 + \|g_n - g\|^2 \to 0.$$

By 8.21, $V_{\mathbf{C}}$ is a complex Hilbert space.

Problem 8B.5. Prove that if V is a normed vector space, $f \in V$, and $r > 0$, then the open ball $B(f, r)$ centered at f with radius r is convex.

Solution. For $g, h \in B(f, r)$ and $t \in [0, 1]$, the identity $\big((1 - t)g + th\big) - f = (1 - t)(g - f) + t(h - f)$ together with the triangle inequality and homogeneity of the norm gives

$$\big\|\big((1 - t)g + th\big) - f\big\| \le (1 - t)\|g - f\| + t\,\|h - f\| < (1 - t)r + tr = r,$$

the last inequality strict because $\|g - f\| < r$, $\|h - f\| < r$, and t cannot be both 0 and 1, so at least one of the two terms is bounded strictly. Hence $(1 - t)g + th \in B(f, r)$, which is convexity (8.25).

Problem 8B.6. (a) Suppose V is an inner product space and B is the open unit ball in V (thus $B = \{f \in V : \|f\| < 1\}$). Prove that if U is a subset of V such that $B \subseteq U \subseteq \overline{B}$, then U is convex.
(b) Give an example to show that (a) can fail if the phrase inner product space is replaced by Banach space.

Solution. (a) Let $f, g \in U$, let $t \in (0,1)$ (the values $t = 0, 1$ being trivial), and set $h = (1-t)f + tg$. Since $\overline{B} = \{f : \|f\| \le 1\}$ (the norm is continuous, and $(1 - \frac{1}{n})f \to f$ from inside B), we have $\|f\|, \|g\| \le 1$, so by 8.15

$$\|h\| \le (1-t)\|f\| + t\|g\| \le (1-t) + t = 1.$$

(i) $\|h\| < 1$: then $h \in B \subseteq U$.
(ii) $\|h\| = 1$: then equality holds throughout, which forces $\|f\| = \|g\| = 1$ (since $1 - t, t > 0$). Put $u = (1-t)f$ and $v = tg$, so $u, v \neq 0$ and $\|u + v\| = 1 = \|u\| + \|v\|$. Equality in the triangle inequality (8.18) makes one of u, v a nonnegative multiple of the other, so $v = cu$ with $c > 0$; taking norms gives $c = t/(1-t)$, whence $tg = tf$ and $g = f$, so $h = f \in U$.
(b) Take $V = \mathbf{R}^2$ with $\|(x_1, x_2)\| = \max\{|x_1|, |x_2|\}$, a Banach space because convergence in this norm is coordinatewise convergence. Then B and $\overline{B}$ are the open and closed unit squares, and

$$U = B \cup \{(1,1), (1,-1)\}$$

satisfies $B \subseteq U \subseteq \overline{B}$ but is not convex, since $\frac{1}{2}(1,1) + \frac{1}{2}(1,-1) = (1,0) \notin U$.

Problem 8B.7. Suppose V is a normed vector space and U is a closed subset of V. Prove that U is convex if and only if

$$\frac{f+g}{2} \in U \quad \text{for all } f, g \in U.$$

Solution. If U is convex, take $t = \frac{1}{2}$ in 8.25. Conversely, suppose $\frac{f+g}{2} \in U$ whenever $f, g \in U$, fix $f, g \in U$, and set

$$D = \bigl\{t \in [0,1] : (1-t)f + tg \in U\bigr\};$$

we show $D = [0,1]$, which is convexity.
D is closed under averaging: for $s, t \in D$ the hypothesis applied to $(1-s)f + sg$ and $(1-t)f + tg$ puts their average, namely $\bigl(1 - \frac{s+t}{2}\bigr)f + \frac{s+t}{2}g$, in U, so $\frac{s+t}{2} \in D$. Since $0, 1 \in D$, induction on n now gives $k/2^n \in D$ for all $0 \le k \le 2^n$: an even k is covered by the previous stage, and an odd $k = 2j+1$ satisfies $\frac{k}{2^{n+1}} = \frac{1}{2}\bigl(\frac{j}{2^n} + \frac{j+1}{2^n}\bigr)$. (Check!)
D is closed in $[0,1]$: the map $\gamma(t) = (1-t)f + tg$ satisfies $\|\gamma(t) - \gamma(s)\| = |t - s|\,\|g - f\|$, hence is continuous, and U is closed, so $D = \gamma^{-1}(U)$ is closed. Being closed and containing the dense set of dyadic rationals, $D = [0,1]$.

Problem 8B.8. Prove that if U is a convex subset of a normed vector space, then $\overline{U}$ is also convex.

Solution. Given $f, g \in \overline{U}$ and $t \in [0,1]$, pick sequences $f_k, g_k \in U$ with $f_k \to f$ and $g_k \to g$ (each ball $B(f, 1/k)$ meets U, else f would lie in the open set $V \setminus \overline{U}$). Convexity of U puts $h_k = (1-t)f_k + tg_k$ in U, and the triangle inequality with homogeneity gives

$$\bigl\|\bigl((1-t)f + tg\bigr) - h_k\bigr\| \le (1-t)\|f - f_k\| + t\|g - g_k\| \longrightarrow 0.$$

Hence $(1-t)f + tg \in \overline{U}$, so $\overline{U}$ is convex (8.25).

Problem 8B.9. Prove that if U is a convex subset of a normed vector space, then the interior of U is also convex.
[The interior of U is the set $\{f \in U : B(f, r) \subseteq U \text{ for some } r > 0\}$.]

Solution. Given $f, g \in \operatorname{int} U$ with $B(f,r) \subseteq U$ and $B(g,s) \subseteq U$, and $t \in [0,1]$, the ball $B(h,\rho)$ around $h = (1-t)f + tg$ with $\rho = \min\{r,s\}$ lies in U, so $h \in \operatorname{int} U$. Indeed, for $k \in B(h,\rho)$ put $u = k - h$; then $\|u\| < \rho$ places $f+u$ in $B(f,r) \subseteq U$ and $g+u$ in $B(g,s) \subseteq U$, so convexity of U gives

$$(1-t)(f+u) + t(g+u) = h + \big((1-t)+t\big)u = k \in U.$$

Hence $\operatorname{int} U$ is convex (vacuously so if it is empty).

Problem 8B.10. Suppose V is a Hilbert space, U is a nonempty closed convex subset of V, and $g \in U$ is the unique element of U with smallest norm (obtained by taking $f = 0$ in 8.28). Prove that

$$\operatorname{Re}\langle g, h\rangle \geq \|g\|^2$$

for all $h \in U$.

Solution. Perturb g toward h inside U. For $h \in U$ and $t \in (0,1]$, convexity (8.25) puts $(1-t)g + th = g + t(h-g)$ in U, so minimality of $\|g\|$ and expansion of the norm give

$$\|g\|^2 \leq \|g + t(h-g)\|^2 = \|g\|^2 + 2t\operatorname{Re}\langle g, h-g\rangle + t^2\|h-g\|^2,$$

the cross terms combining because $\langle h-g, g\rangle = \overline{\langle g, h-g\rangle}$ and $z + \bar{z} = 2\operatorname{Re} z$. Subtracting $\|g\|^2$ and dividing by $2t > 0$ gives $0 \leq \operatorname{Re}\langle g, h-g\rangle + \frac{t}{2}\|h-g\|^2$ for every $t \in (0,1]$; letting $t \to 0^+$ yields $\operatorname{Re}\langle g, h-g\rangle \geq 0$. Since $\operatorname{Re}\langle g, h-g\rangle = \operatorname{Re}\langle g, h\rangle - \|g\|^2$ by additivity in the second slot, the inequality $\operatorname{Re}\langle g, h\rangle \geq \|g\|^2$ follows.

Problem 8B.11. Suppose V is a Hilbert space. A *closed half-space* of V is a set of the form

$$\{g \in V : \operatorname{Re}\langle g, h\rangle \geq c\}$$

for some $h \in V$ and some $c \in \mathbf{R}$. Prove that every closed convex subset of V is the intersection of all the closed half-spaces that contain it.

Solution. Let C be closed and convex and let I be the intersection of all closed half-spaces containing C. Then $C \subseteq I$ trivially, so only $I \subseteq C$ needs proof: given $f \notin C$, we exhibit a closed half-space containing C but not f.
(i) $C = \varnothing$: taking $h = 0$ and $c = 1$ makes $\{g : \operatorname{Re}\langle g, 0\rangle \geq 1\} = \varnothing$ itself a closed half-space containing C, so $I = \varnothing = C$.
(ii) $C \neq \varnothing$: by 8.28 (V a Hilbert space, C nonempty closed convex) there is $g \in C$ with $\|f - g\| = \operatorname{distance}(f, C)$, and $u = f - g \neq 0$. For $k \in C$ and $t \in (0,1]$, convexity puts $g + t(k-g)$ in C, so minimality gives

$$\|u\|^2 \leq \|u - t(k-g)\|^2 = \|u\|^2 - 2t\operatorname{Re}\langle u, k-g\rangle + t^2\|k-g\|^2,$$

whence $\operatorname{Re}\langle u, k-g\rangle \leq \frac{t}{2}\|k-g\|^2$ for all such t, and letting $t \to 0^+$ gives the variational inequality $\operatorname{Re}\langle u, k-g\rangle \leq 0$ for every $k \in C$. Now set

$$H = \big\{k \in V : \operatorname{Re}\langle k, -u\rangle \geq -\operatorname{Re}\langle g, u\rangle\big\},$$

a closed half-space. It contains C, since for $k \in C$ the variational inequality and $\operatorname{Re}\langle k, u\rangle = \operatorname{Re}\langle u, k\rangle$ give $\operatorname{Re}\langle k, u\rangle \leq \operatorname{Re}\langle g, u\rangle$. It omits f, since $f = g + u$ gives

$$\operatorname{Re}\langle f, -u\rangle = -\operatorname{Re}\langle g, u\rangle - \|u\|^2 < -\operatorname{Re}\langle g, u\rangle.$$

Hence $f \notin I$, so $I \subseteq C$ and $C = I$.

Problem 8B.12. Give an example of a nonempty closed subset U of the Hilbert space ℓ^2 and $a \in \ell^2$ such that there does not exist $b \in U$ with $\|a - b\| = \text{distance}(a, U)$.
[By 8.28, U cannot be a convex subset of ℓ^2.]

Solution. Take $a = 0$ and $U = \{u_k : k \in \mathbf{Z}^+\}$, where $u_k = \bigl(1 + \frac{1}{k}\bigr)e_k$ and $e_k \in \ell^2$ is the k-th standard basis sequence.
U is closed: distinct u_j, u_k are supported on disjoint coordinates, so

$$\|u_j - u_k\|^2 = \Bigl(1 + \tfrac{1}{j}\Bigr)^2 + \Bigl(1 + \tfrac{1}{k}\Bigr)^2 > 2,$$

and a convergent sequence in U, being Cauchy, must therefore be eventually constant, with its limit that constant value in U.
The distance is not attained: $\text{distance}(0, U) = \inf_k \bigl(1 + \frac{1}{k}\bigr) = 1$, while $\|0 - u_k\| = 1 + \frac{1}{k} > 1$ for every k.

Problem 8B.13. In the real Banach space $\mathbf{R}^2$ with norm defined by $\|(x, y)\|_\infty = \max\{|x|, |y|\}$, give an example of a closed convex set $U \subseteq \mathbf{R}^2$ and $z \in \mathbf{R}^2$ such that there exist infinitely many choices of $w \in U$ with $\|z - w\|_\infty = \text{distance}(z, U)$.

Solution. Take $U = \{(x, y) : x \geq 1\}$ and $z = (0, 0)$; every $w = (1, y)$ with $|y| \leq 1$ is then a closest point.
U is closed as the inverse image of $[1, \infty)$ under the $\|\cdot\|_\infty$-continuous map $(x, y) \mapsto x$, and convex because $(1 - t)x + tx' \geq 1$ whenever $x, x' \geq 1$ and $t \in [0, 1]$. For $w = (x, y) \in U$ we have $\|z - w\|_\infty = \max\{|x|, |y|\} \geq x \geq 1$, with $\|z - (1, 0)\|_\infty = 1$, so $\text{distance}(z, U) = 1$. Finally

$$\|z - (1, y)\|_\infty = \max\{1, |y|\} = 1 \qquad \text{for all } |y| \leq 1,$$

giving infinitely many minimizers.

Problem 8B.14. Suppose f and g are elements of an inner product space. Prove that $\langle f, g\rangle = 0$ if and only if

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

for all $\alpha \in \mathbf{F}$.

Solution. Both directions run off the expansion

$$\|f + \alpha g\|^2 = \|f\|^2 + 2\,\text{Re}\bigl(\overline{\alpha}\langle f, g\rangle\bigr) + |\alpha|^2\|g\|^2,$$

which holds for all $\alpha \in \mathbf{F}$ by additivity in each slot, $\langle f, \alpha g\rangle = \overline{\alpha}\langle f, g\rangle$, and $\langle \alpha g, f\rangle = \overline{\overline{\alpha}\langle f, g\rangle}$.
If $\langle f, g\rangle = 0$, the middle term vanishes and $\|f + \alpha g\|^2 = \|f\|^2 + |\alpha|^2\|g\|^2 \geq \|f\|^2$.
Conversely, suppose $\|f\| \leq \|f + \alpha g\|$ for all α. If $g = 0$ then $\langle f, g\rangle = 0$; otherwise take $\alpha = -\langle f, g\rangle/\|g\|^2$, for which $\overline{\alpha}\langle f, g\rangle = -|\langle f, g\rangle|^2/\|g\|^2$ is real and $|\alpha|^2\|g\|^2 = |\langle f, g\rangle|^2/\|g\|^2$, so the expansion becomes

$$\|f\|^2 \leq \|f + \alpha g\|^2 = \|f\|^2 - \frac{|\langle f, g\rangle|^2}{\|g\|^2}.$$

As $\|g\|^2 > 0$, this forces $\langle f, g\rangle = 0$.

Problem 8B.15. Suppose U is a closed subspace of a Hilbert space V and $f \in V$. Prove that

$$\|P_U f\| \leq \|f\|,$$

with equality if and only if $f \in U$.
[This exercise asks you to prove 8.37(d).]

Solution. Since $f - P_U f$ is orthogonal to every element of U (8.37(a)) and $P_U f \in U$, the Pythagorean Theorem (8.9) applied to $f = (f - P_U f) + P_U f$ gives

$$\|f\|^2 = \|f - P_U f\|^2 + \|P_U f\|^2 \geq \|P_U f\|^2.$$

Equality holds exactly when $\|f - P_U f\| = 0$, that is, when $f = P_U f$, which happens if and only if $f \in U$: one direction because $P_U f \in U$, the other because $f \in U$ makes f itself the unique closest point of U to f (8.28).

Problem 8B.16. Suppose V is a Hilbert space and $P : V \to V$ is a linear map such that $P^2 = P$ and $\|Pf\| \leq \|f\|$ for every $f \in V$. Prove that there exists a closed subspace U of V such that $P = P_U$.

Solution. Take $U = \operatorname{range} P$.
U is a closed subspace: $P^2 = P$ gives $\operatorname{range} P = \{f : Pf = f\} = \operatorname{null}(I - P)$, and $\|Pf\| \leq \|f\|$ makes P, hence $I - P$, continuous, so this null space is the inverse image of the closed set $\{0\}$.
Every $h \in \operatorname{null} P$ is orthogonal to every $g \in U$: linearity and $Pg = g$ give $P(g + \alpha h) = g$ for all $\alpha \in \mathbf{F}$, so the hypothesis applied to $g + \alpha h$ yields

$$\|g\| = \|P(g + \alpha h)\| \leq \|g + \alpha h\| \qquad \text{for all } \alpha \in \mathbf{F},$$

whence $\langle g, h\rangle = 0$ by Exercise 14 in this section.
Now let $f \in V$. Then $Pf \in U$, and $P(f - Pf) = Pf - P^2 f = 0$ puts $f - Pf$ in $\operatorname{null} P$, so $\langle f - Pf, g\rangle = 0$ for every $g \in U$ (conjugating the previous paragraph). Hence $Pf = P_U f$ by 8.37(b).

Problem 8B.17. Suppose U is a subspace of a Hilbert space V. Suppose also that W is a Banach space and $S : U \to W$ is a bounded linear map. Prove that there exists a bounded linear map $T : V \to W$ such that $T|_U = S$ and $\|T\| = \|S\|$.
[If $W = \mathbf{F}$, then this result is just the Hahn-Banach Theorem (6.69) for Hilbert spaces. The result here is stronger because it allows W to be an arbitrary Banach space instead of requiring W to be $\mathbf{F}$. Also, the proof in this Hilbert space context does not require use of Zorn's Lemma or the Axiom of Choice.]

Solution. Take

$$T = \widetilde{S} \circ P_{\overline{U}} : V \to W,$$

where $\overline{U}$ is the closure of U, a closed subspace of V (Exercise 12 in Section 6C), and $\widetilde{S} : \overline{U} \to W$ is the extension of S with $\widetilde{S}|_U = S$ and $\|\widetilde{S}\| = \|S\|$ supplied by Exercise 14 in Section 6C (W is a Banach space, so the limits defining it exist).
T is linear as a composition of linear maps, $P_{\overline{U}}$ being linear by 8.37(c). For $f \in U$ we have $P_{\overline{U}} f = f$, so $Tf = \widetilde{S} f = Sf$ and $T|_U = S$. Finally 8.37(d) gives $\|P_{\overline{U}} f\| \leq \|f\|$, so

$$\|Tf\| \leq \|\widetilde{S}\|\,\|P_{\overline{U}} f\| \leq \|S\|\,\|f\|,$$

whence $\|T\| \leq \|S\|$; and $\|T\| \geq \|S\|$ because T extends S. Thus $\|T\| = \|S\|$.

Problem 8B.18. Suppose U and W are subspaces of a Hilbert space V. Prove that $\overline{U} = \overline{W}$ if and only if $U^\perp = W^\perp$.

Solution. Each direction is one application of a numbered result. If $\overline{U} = \overline{W}$, then 8.40(d) gives

$$U^\perp = \left(\overline{U}\right)^\perp = \left(\overline{W}\right)^\perp = W^\perp.$$

If $U^\perp = W^\perp$, then 8.41 (applicable since U and W are subspaces of the Hilbert space V) gives

$$\overline{U} = \big(U^\perp\big)^\perp = \big(W^\perp\big)^\perp = \overline{W}.$$

Problem 8B.19. Suppose U and W are closed subspaces of a Hilbert space. Prove that $P_U P_W = 0$ if and only if $\langle f, g\rangle = 0$ for all $f \in U$ and all $g \in W$.

Solution. Both directions are the statement $W \subseteq U^\perp$ read through 8.45(a), which gives $\operatorname{null} P_U = U^\perp$ and $\operatorname{range} P_W = W$.
If $\langle f, g\rangle = 0$ for all $f \in U$ and $g \in W$, then $W \subseteq U^\perp$ by 8.38, so for every h in the Hilbert space, $P_W h \in W \subseteq \operatorname{null} P_U$ and hence $P_U P_W h = 0$.
Conversely, if $P_U P_W = 0$ and $g \in W$, then $P_W g = g$ (as W is a closed subspace containing g), so

$$P_U g = P_U P_W g = 0,$$

putting g in $\operatorname{null} P_U = U^\perp$, that is, $\langle f, g\rangle = 0$ for every $f \in U$.

Problem 8B.20. Verify the assertions in Example 8.46. That example states: suppose U is the closed subspace of $L^2(\mathbb{R})$ defined by

$$U = \{f \in L^2(\mathbb{R}) : f(x) = 0 \text{ for almost every } x < 0\}.$$

Then

$$U^\perp = \{g \in L^2(\mathbb{R}) : g(x) = 0 \text{ for almost every } x \geq 0\},$$

and if $h \in L^2(\mathbb{R})$, then

$$P_U h = h\chi_{[0,\infty)} \quad \text{and} \quad P_{U^\perp} h = h\chi_{(-\infty,0)}.$$

Thus $P_{U^\perp} h = h(1 - \chi_{[0,\infty)}) = (I - P_U)h$ and hence $P_{U^\perp} = I - P_U$, as asserted in 8.45(c).

Solution. Write $W = \{g \in L^2(\mathbf{R}) : g = 0 \text{ almost everywhere on } [0,\infty)\}$; the whole example follows from the splitting $h = h\chi_{[0,\infty)} + h\chi_{(-\infty,0)}$ into a member of U and a member of W.
U is a closed subspace: it is closed under linear combinations (a union of two null sets is null), and if $f_k \in U$ with $f_k \to f$ in L^2, then

$$\int_{(-\infty,0)} |f|^2 \, d\lambda = \int_{(-\infty,0)} |f - f_k|^2 \, d\lambda \leq \|f - f_k\|^2 \longrightarrow 0,$$

forcing $f = 0$ almost everywhere on $(-\infty, 0)$, that is, $f \in U$.
$U^\perp = W$. If $g \in W$ and $f \in U$, then $f\overline{g} = 0$ almost everywhere (each factor vanishes on one half-line), so $\langle f, g\rangle = 0$ and $W \subseteq U^\perp$. Conversely, if $g \in U^\perp$, then $f = g\chi_{[0,\infty)}$ lies in U (and in L^2, since $|f| \leq |g|$), so

$$0 = \langle f, g\rangle = \int_{[0,\infty)} |g|^2 \, d\lambda,$$

giving $g = 0$ almost everywhere on $[0, \infty)$, that is, $g \in W$.
The projections. For $h \in L^2(\mathbf{R})$, the element $h\chi_{[0,\infty)}$ lies in U and $h - h\chi_{[0,\infty)} = h\chi_{(-\infty,0)}$ lies in $W = U^\perp$, hence is orthogonal to every element of U; so $P_U h = h\chi_{[0,\infty)}$ by 8.37(b). Symmetrically $h\chi_{(-\infty,0)} \in U^\perp$ (a closed subspace by 8.40(a)) and $h - h\chi_{(-\infty,0)} = h\chi_{[0,\infty)} \in U$ is orthogonal to every element of $U^\perp$, so 8.37(b) applied to $U^\perp$ gives $P_{U^\perp} h = h\chi_{(-\infty,0)}$. Consequently

$$P_{U^\perp} h = h\big(1 - \chi_{[0,\infty)}\big) = h - P_U h = (I - P_U)h,$$

confirming 8.45(c) here.

Problem 8B.21. Show that every inner product space is a subspace of some Hilbert space.
Hint: See Exercise 13 in Section 6C.

Solution. Take $B = \overline{V}$, the closure of V in a Banach space C that contains the normed space V isometrically (Exercise 13 in Section 6C). Then B is a subspace of C (Exercise 12 in Section 6C) and is closed, hence complete; so B is a Banach space in which V is dense.

Extend the inner product of V to B by continuity. Given $f, g \in B$, choose $f_k \to f$ and $g_k \to g$ with $f_k, g_k \in V$; convergent sequences are bounded, say by M, so the Cauchy-Schwarz inequality (8.11) in V gives

$$\begin{aligned}\big|\langle f_k, g_k\rangle_V - \langle f_j, g_j\rangle_V\big| &= \big|\langle f_k - f_j, g_k\rangle_V + \langle f_j, g_k - g_j\rangle_V\big|\\ &\le M\big(\|f_k - f_j\| + \|g_k - g_j\|\big).\end{aligned}$$

The right side tends to 0, so $\langle f_k, g_k\rangle_V$ converges; the same estimate with f_k', g_k' in place of f_j, g_j shows the limit is independent of the approximating sequences. Hence

$$\langle f, g\rangle = \lim_{k\to\infty} \langle f_k, g_k\rangle_V \qquad (f, g \in B)$$

is well defined, and constant sequences give $\langle f, g\rangle = \langle f, g\rangle_V$ for $f, g \in V$.

Linearity in the first slot and conjugate symmetry pass to the limit because addition, scalar multiplication, and conjugation are continuous (Check!). For positivity, continuity of the norm gives

$$\langle f, f\rangle = \lim_{k\to\infty} \|f_k\|^2 = \|f\|^2,$$

so $\langle f, f\rangle \ge 0$ with equality only when $f = 0$, and the norm induced by $\langle \cdot, \cdot\rangle$ is the norm of B. Being complete in that norm, B is a Hilbert space (8.21) containing V as a subspace with its original inner product.

Problem 8B.22. Prove that if V is a Hilbert space and $T : V \to V$ is a bounded linear map such that the dimension of range T is 1, then there exist $g, h \in V$ such that

$$Tf = \langle f, g\rangle h$$

for all $f \in V$.

Solution. Take $h \ne 0$ spanning the one-dimensional subspace range T, and take g to be the Riesz vector of the functional φ below.

For each $f \in V$ the equation $Tf = \varphi(f)h$ determines the scalar $\varphi(f)$ uniquely, because $h \ne 0$. That uniqueness converts the linearity of T into linearity of φ: applying it to

$$\varphi(f_1 + \alpha f_2)h = T(f_1 + \alpha f_2) = \big(\varphi(f_1) + \alpha\varphi(f_2)\big)h$$

gives $\varphi(f_1 + \alpha f_2) = \varphi(f_1) + \alpha\varphi(f_2)$. Furthermore

$$|\varphi(f)|\,\|h\| = \|Tf\| \le \|T\|\,\|f\|,$$

so φ is bounded, with $\|\varphi\| \le \|T\|/\|h\|$. Because V is a Hilbert space, the Riesz Representation Theorem (8.47) supplies $g \in V$ with $\varphi(f) = \langle f, g\rangle$ for all $f \in V$. Hence

$$Tf = \varphi(f)h = \langle f, g\rangle h \qquad \text{for all } f \in V.$$

Problem 8B.23. (a) Give an example of a Banach space V and a bounded linear functional φ on V such that $|\varphi(f)| < \|\varphi\|\,\|f\|$ for all $f \in V \setminus \{0\}$.
(b) Show there does not exist an example in part (a) where V is a Hilbert space.

Solution. (a) Take $V = \ell^1$, a Banach space by 7.24, and

$$\varphi(a) = \sum_{k=1}^{\infty}\Big(1 - \frac{1}{k}\Big)a_k \qquad \text{for } a = (a_1, a_2, \dots) \in \ell^1,$$

the series converging absolutely since $|(1 - 1/k)a_k| \le |a_k|$.
Here $\|\varphi\| = 1$: the same bound gives $|\varphi(a)| \le \|a\|_1$, while $\varphi(e_k) = 1 - 1/k$ and $\|e_k\|_1 = 1$ for the kth standard basis vector e_k, so $\|\varphi\| \ge 1 - 1/k$ for every k. If $a \ne 0$, choose m with $a_m \ne 0$; then

$$|\varphi(a)| \le \|a\|_1 - \sum_{k=1}^{\infty} \frac{|a_k|}{k} \le \|a\|_1 - \frac{|a_m|}{m} < \|a\|_1 = \|\varphi\|\,\|a\|_1,$$

all quantities being finite because $a \in \ell^1$.
(b) On a Hilbert space every bounded linear functional attains its norm, so the strict inequality always fails somewhere. Indeed, let $V \ne \{0\}$ be a Hilbert space and φ a bounded linear functional on V; the Riesz Representation Theorem (8.47) gives $h \in V$ with $\varphi(f) = \langle f, h\rangle$ for all f and $\|\varphi\| = \|h\|$.
(i) If $h \ne 0$, then $f = h$ is nonzero and

$$|\varphi(h)| = \|h\|^2 = \|\varphi\|\,\|h\|.$$

(ii) If $h = 0$, then $\varphi = 0$ and $\|\varphi\| = 0$, so $|\varphi(f)| = 0 = \|\varphi\|\,\|f\|$ for any nonzero $f \in V$.
(For $V = \{0\}$ the condition in (a) is vacuous.)

Problem 8B.24. (a) Suppose φ and ψ are bounded linear functionals on a Hilbert space V such that $\|\varphi + \psi\| = \|\varphi\| + \|\psi\|$. Prove that one of φ, ψ is a scalar multiple of the other.
(b) Give an example to show that (a) can fail if the hypothesis that V is a Hilbert space is replaced by the hypothesis that V is a Banach space.

Solution. (a) The hypothesis forces equality in Cauchy-Schwarz for the two Riesz vectors. By 8.47 there are $g, h \in V$ with $\varphi = \langle\cdot, g\rangle$, $\psi = \langle\cdot, h\rangle$, $\|\varphi\| = \|g\|$, and $\|\psi\| = \|h\|$; since $\varphi + \psi = \langle\cdot, g + h\rangle$, the uniqueness in 8.47 gives $\|\varphi + \psi\| = \|g + h\|$. So $\|g + h\| = \|g\| + \|h\|$, and squaring both sides yields

$$\operatorname{Re}\langle g, h\rangle = \|g\|\,\|h\|.$$

As $\operatorname{Re}\langle g, h\rangle \le |\langle g, h\rangle| \le \|g\|\,\|h\|$ by the Cauchy–Schwarz inequality (8.11), both inequalities are equalities, and the equality case of 8.11 makes one of g, h a scalar multiple of the other. If $g = \alpha h$, then

$$\varphi(f) = \langle f, \alpha h\rangle = \overline{\alpha}\,\psi(f) \qquad \text{for all } f \in V,$$

so $\varphi = \overline{\alpha}\psi$; symmetrically, $h = \alpha g$ gives $\psi = \overline{\alpha}\varphi$.
(b) Take $V = \mathbf{F}^2$ with $\|(a_1, a_2)\| = \max\{|a_1|, |a_2|\}$, a Banach space because convergence in this norm is coordinatewise convergence in the complete field $\mathbf{F}$, and take

$$\varphi(a_1, a_2) = a_1, \qquad \psi(a_1, a_2) = a_2.$$

Then $\|\varphi\| = \|\psi\| = 1$, attained at $(1, 0)$ and $(0, 1)$, while $|a_1 + a_2| \le 2\|(a_1, a_2)\|$ with equality at $(1, 1)$ gives

$$\|\varphi + \psi\| = 2 = \|\varphi\| + \|\psi\|.$$

Yet $\varphi = c\psi$ fails at $(1, 0)$ and $\psi = c\varphi$ fails at $(0, 1)$.

Problem 8B.25. (a) Suppose that μ is a finite measure, $1 \le p \le 2$, and φ is a bounded linear functional on $L^p(\mu)$. Prove that there exists $h \in L^{p'}(\mu)$ such that $\varphi(f) = \int f h\, d\mu$ for every $f \in L^p(\mu)$.
(b) Same as (a), but with the hypothesis that μ is a finite measure replaced by the hypothesis that

μ is a measure, and assume that $1 < p \le 2$.
[See 7.25, which along with this exercise shows that we can identify the dual of $L^p(\mu)$ with $L^{p'}(\mu)$ for $1 < p \le 2$. See 9.42 for an extension to all $p \in (1, \infty)$.]

Solution. Take $h = \overline{g}$, where $g \in L^2(\mu)$ is the Riesz vector of the restriction of φ to $L^2(\mu)$. Throughout, $(X, \mathcal{S}, \mu)$ is the measure space, $\frac{1}{p} + \frac{1}{p'} = 1$, and $\sigma = \overline{h}/|h|$ on $\{h \neq 0\}$ and $\sigma = 0$ elsewhere, so $|\sigma| \le 1$ and $\sigma h = |h|$.
(a) Assume $\mu(X) > 0$ (otherwise $h = 0$ works), and note $2 \le p' \le \infty$. Hölder's inequality (7.9) applied to $|f|^p$ and 1 with the conjugate exponents $\frac{2}{p}, \frac{2}{2-p}$ gives

$$\|f\|_p \le c\,\|f\|_2, \qquad c = \mu(X)^{\frac{1}{p}-\frac{1}{2}} \in (0, \infty),$$

so $L^2(\mu) \subseteq L^p(\mu)$ and $|\varphi(f)| \le c\|\varphi\|\|f\|_2$ there. As $L^2(\mu)$ is a Hilbert space (8.22), the Riesz Representation Theorem (8.47) supplies $g \in L^2(\mu)$ with $\varphi(f) = \langle f, g\rangle = \int f\overline{g}\,d\mu$ for $f \in L^2(\mu)$; put $h = \overline{g}$, so $\varphi(f) = \int fh\,d\mu$ on $L^2(\mu)$.
Next, $h \in L^{p'}(\mu)$ with $\|h\|_{p'} \le \|\varphi\|$. Because μ is finite, every bounded measurable function lies in $L^2(\mu) \cap L^p(\mu)$, and $|h| < \infty$ almost everywhere since $h \in L^2(\mu)$.
(i) $1 < p \le 2$, so $2 \le p' < \infty$. With $E_n = \{|h| \le n\}$ and $f_n = \chi_{E_n}|h|^{p'-1}\sigma$, the function f_n is bounded by $n^{p'-1}$ and $f_n h = \chi_{E_n}|h|^{p'}$; since $(p'-1)p = p'$, the number $I_n = \int_{E_n}|h|^{p'}\,d\mu \le n^{p'}\mu(X)$ satisfies $\|f_n\|_p^p = I_n$ and hence

$$I_n = |\varphi(f_n)| \le \|\varphi\|\,\|f_n\|_p = \|\varphi\|\,I_n^{1/p}.$$

So $I_n^{1/p'} \le \|\varphi\|$ (trivially if $I_n = 0$), and since $\chi_{E_n}|h|^{p'}$ increases pointwise to $|h|^{p'}$, the Monotone Convergence Theorem (3.11) gives $\int |h|^{p'}\,d\mu \le \|\varphi\|^{p'}$.
(ii) $p = 1$, so $p' = \infty$. Given $\varepsilon > 0$, put $E = \{|h| > \|\varphi\| + \varepsilon\}$; the bounded function $f = \chi_E\sigma$ has $fh = \chi_E|h|$ and $\|f\|_1 = \mu(E)$, whence

$$(\|\varphi\| + \varepsilon)\mu(E) \le \int_E |h|\,d\mu = \varphi(f) \le \|\varphi\|\,\mu(E).$$

As $\mu(E) \le \mu(X) < \infty$, this forces $\mu(E) = 0$; so $\|h\|_\infty \le \|\varphi\|$.
Finally $\psi(f) = \int fh\,d\mu$ is a bounded linear functional on $L^p(\mu)$ by Hölder (7.9), and $\varphi = \psi$ on $L^2(\mu)$. For $f \in L^p(\mu)$ the truncations $f\chi_{\{|f|\le n\}}$ are bounded, hence in $L^2(\mu)$, and

$$\|f - f\chi_{\{|f|\le n\}}\|_p^p = \int_{\{|f|>n\}} |f|^p\,d\mu \longrightarrow 0$$

by the Dominated Convergence Theorem (3.31) (dominant $|f|^p \in L^1(\mu)$; here $p < \infty$). So $L^2(\mu)$ is dense in $L^p(\mu)$ and the continuous functional $\varphi - \psi$ vanishes identically.
(b) Now μ is arbitrary and $1 < p \le 2$, so $2 \le p' < \infty$; that finiteness is essential below.
Local representations. For $E \in \mathcal{S}$ with $\mu(E) < \infty$ let μ_E be the restriction of μ to $\{A \in \mathcal{S} : A \subseteq E\}$, a finite measure, and write $\widetilde{f}$ for extension by 0, a linear isometry of $L^p(\mu_E)$ into $L^p(\mu)$. Then $f \mapsto \varphi(\widetilde{f})$ is a bounded linear functional on $L^p(\mu_E)$ of norm at most $\|\varphi\|$, so part (a) gives $h_E \in L^{p'}(\mu_E)$ with

$$\varphi(\widetilde{f}) = \int_E fh_E\,d\mu \qquad \text{for all } f \in L^p(\mu_E).$$

Since $1 < p \le 2$, the map of 7.25 from $L^{p'}(\mu_E)$ to $L^p(\mu_E)'$ is one-to-one and norm-preserving, so h_E is unique and $\|h_E\|_{p'} \le \|\varphi\|$; uniqueness applied to $E \subseteq F$ with $\mu(F) < \infty$ gives the consistency relation $h_F|_E = h_E$ almost everywhere on E.
A maximizing σ-finite set. Let

$$M = \sup\{\|h_E\|_{p'} : E \in \mathcal{S},\ \mu(E) < \infty\} \le \|\varphi\| < \infty,$$

choose F_n of finite measure with $\|h_{F_n}\|_{p'}^{p'} > M^{p'} - \frac{1}{n}$, and set $E_n = F_1 \cup \cdots \cup F_n$. Consistency gives $\|h_{E_n}\|_{p'}^{p'} \ge \int_{F_n} |h_{F_n}|^{p'}\,d\mu > M^{p'} - \frac{1}{n}$, while $\|h_{E_n}\|_{p'} \le M$ is nondecreasing in n; hence $\|h_{E_n}\|_{p'} \to M$.

Put $E = \bigcup_n E_n$. Discarding the null set $\bigcup_{n<m}\{x \in E_n : h_{E_m}(x) \neq h_{E_n}(x)\}$ (a countable union of null sets, by consistency) makes the representatives agree, so

$$h = h_{E_n} \text{ on } E_n, \qquad h = 0 \text{ off } E$$

is unambiguous and measurable, being the pointwise limit of $h_{E_n}\chi_{E_n}$. Since $\chi_{E_n}|h|^{p'}$ increases to $|h|^{p'}$, the Monotone Convergence Theorem (3.11) gives $\|h\|_{p'}^{p'} = \lim_n \|h_{E_n}\|_{p'}^{p'} = M^{p'}$, so $h \in L^{p'}(\mu)$.
Nothing lives outside E. If $\mu(A) < \infty$ and $A \cap E = \varnothing$, then $h_{E_n \cup A}$ agrees with h_{E_n} on E_n and with h_A on A, so disjointness gives

$$M^{p'} \geq \|h_{E_n}\|_{p'}^{p'} + \|h_A\|_{p'}^{p'} \longrightarrow M^{p'} + \|h_A\|_{p'}^{p'}.$$

Because $M < \infty$ and $p' < \infty$ (this additivity of $\|\cdot\|_{p'}^{p'}$ over disjoint pieces is exactly what fails for $p' = \infty$), we get $h_A = 0$, hence $\varphi(\widetilde{f}) = 0$ for every $f \in L^p(\mu_A)$.
Conclusion. Fix $f \in L^p(\mu)$. Then $\{f \neq 0\}$ is σ-finite, since

$$\tfrac{1}{k^p}\,\mu\big(\{|f| > \tfrac{1}{k}\}\big) \leq \|f\|_p^p < \infty.$$

Write $A = \{f \neq 0\} \setminus E = \bigcup_k A_k$ with A_k increasing of finite measure, each disjoint from E; then $\varphi(f\chi_{A_k}) = 0$, and $f\chi_{A_k} \to f\chi_A$ in $L^p(\mu)$ by the Dominated Convergence Theorem (3.31), so continuity of φ gives

$$\varphi(f\chi_{X\setminus E}) = \varphi(f\chi_A) = 0 = \int f\chi_{X\setminus E}\, h\, d\mu,$$

the last equality because $h = 0$ off E. Also

$$\varphi(f\chi_{E_n}) = \int_{E_n} f h_{E_n}\, d\mu = \int f\chi_{E_n} h\, d\mu,$$

and $\|f\chi_E - f\chi_{E_n}\|_p \to 0$ (3.31) sends the left side to $\varphi(f\chi_E)$ by continuity and the right side to $\int f\chi_E h\, d\mu$ by Hölder (7.9). Adding the two lines,

$$\varphi(f) = \varphi(f\chi_E) + \varphi(f\chi_{X\setminus E}) = \int fh\, d\mu.$$

Problem 8B.26. Prove that if V is an infinite-dimensional Hilbert space, then the Banach space $\mathcal{B}(V,V)$ is nonseparable.

Solution. The orthogonal projections $T_S = P_{U_S}$ onto the closed subspaces $U_S = \overline{\operatorname{span}}\{e_k : k \in S\}$, indexed by the uncountably many subsets $S \subseteq \mathbf{Z}^+$, are pairwise at distance at least 1, and no metric space containing such a family is separable.
Here $e_1, e_2, \ldots$ is an orthonormal sequence in V, obtained by choosing $f_n \notin \operatorname{span}(f_1, \ldots, f_{n-1})$ inductively (possible since V is infinite-dimensional) and applying Gram–Schmidt:

$$v_n = f_n - \sum_{k=1}^{n-1}\langle f_n, e_k\rangle e_k \neq 0, \qquad e_n = \frac{v_n}{\|v_n\|}$$

(Check! the e_n are orthonormal and span the same subspaces as the f_n).
Each T_S exists and lies in $\mathcal{B}(V,V)$ with $\|T_S\| \leq 1$, by 8.37(c) and 8.37(d), since V is a Hilbert space and U_S is closed. If $k \in S$, then $e_k \in U_S$, so $T_S e_k = e_k$ by 8.37(b). If $k \notin S$, then e_k is orthogonal to every finite combination of $\{e_j : j \in S\}$, so $e_k \in U_S^\perp$ by 8.40(d), and null $P_{U_S} = U_S^\perp$ by 8.45(a) gives $T_S e_k = 0$. Hence for distinct S, S', picking k in their symmetric difference (say $k \in S \setminus S'$),

$$\|T_S - T_{S'}\| \geq \|(T_S - T_{S'})e_k\| = \|e_k\| = 1.$$

Now if D were a countable dense subset of $\mathcal{B}(V,V)$, choosing $d_S \in D$ with $\|T_S - d_S\| < \frac{1}{2}$ for each S

would give an injection of the uncountable collection of subsets of $\mathbf{Z}^+$ into D: if $d_S = d_{S'}$ with $S \neq S'$, the triangle inequality would give $\|T_S - T_{S'}\| < 1$.

8.3 Exercises 8C

Problem 8C.1. Verify that the family $\{e_k\}_{k\in\mathbb{Z}}$ as defined in the third bullet point of Example 8.51 is an orthonormal family in $L^2\big((-\pi,\pi]\big)$. The following formulas should help:

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

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

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

Solution. Everything reduces to three product integrals, since all the e_k are real valued (so $\langle f,g\rangle = \int_{-\pi}^{\pi} fg\,d\lambda$) and, cos being even, the family is the constant $e_0 = \frac{1}{\sqrt{2\pi}}$ together with

$$e_m(t) = \tfrac{1}{\sqrt{\pi}}\sin(mt), \qquad e_{-m}(t) = \tfrac{1}{\sqrt{\pi}}\cos(mt) \qquad (m\in\mathbb{Z}^+),$$

each continuous and bounded on $(-\pi,\pi]$, hence in $L^2\big((-\pi,\pi]\big)$.
Antidifferentiating gives $\int_{-\pi}^{\pi}\sin(nt)\,dt = 0$ for every $n\in\mathbb{Z}$, and $\int_{-\pi}^{\pi}\cos(nt)\,dt = 0$ for $n\neq 0$ while it is 2π for $n=0$ (Check!). Hence for $m,n\in\mathbb{Z}^+$ the three displayed product formulas give

$$\begin{aligned}
\int_{-\pi}^{\pi}\sin(mt)\sin(nt)\,dt &= \tfrac12\int_{-\pi}^{\pi}\cos\big((m-n)t\big)dt - \tfrac12\int_{-\pi}^{\pi}\cos\big((m+n)t\big)dt,\\
\int_{-\pi}^{\pi}\cos(mt)\cos(nt)\,dt &= \tfrac12\int_{-\pi}^{\pi}\cos\big((m-n)t\big)dt + \tfrac12\int_{-\pi}^{\pi}\cos\big((m+n)t\big)dt,\\
\int_{-\pi}^{\pi}\sin(mt)\cos(nt)\,dt &= \tfrac12\int_{-\pi}^{\pi}\sin\big((m-n)t\big)dt + \tfrac12\int_{-\pi}^{\pi}\sin\big((m+n)t\big)dt.
\end{aligned}$$

Since $m+n\geq 2$, the first two equal π when $m=n$ and 0 when $m\neq n$, and the third is 0. Dividing by π: $\|e_{\pm m}\| = 1$, while $\langle e_j, e_k\rangle = 0$ for distinct nonzero j,k.
Finally $\|e_0\|^2 = \int_{-\pi}^{\pi}\frac{1}{2\pi}dt = 1$, and $\langle e_m, e_0\rangle$ and $\langle e_{-m}, e_0\rangle$ are multiples of $\int_{-\pi}^{\pi}\sin(mt)\,dt$ and $\int_{-\pi}^{\pi}\cos(mt)\,dt$, both 0. Every pair $j,k\in\mathbb{Z}$ is now covered, so the family is orthonormal (8.50).

Problem 8C.2. Suppose $\{a_k\}_{k\in\Gamma}$ is a family in $\mathbb{R}$ and $a_k\geq 0$ for each $k\in\Gamma$. Prove the unordered sum $\sum_{k\in\Gamma} a_k$ converges if and only if

$$\sup\Big\{\sum_{j\in\Omega} a_j : \Omega \text{ is a finite subset of } \Gamma\Big\} < \infty.$$

Furthermore, prove that if $\sum_{k\in\Gamma} a_k$ converges then it equals the supremum above.

Solution. Both directions run off monotonicity of finite subsums: if $\Omega\subseteq\Omega'$ are finite subsets of Γ, then $\sum_{j\in\Omega'} a_j - \sum_{j\in\Omega} a_j = \sum_{j\in\Omega'\setminus\Omega} a_j \geq 0$. Write $S\in[0,\infty]$ for the supremum in question.
(i) Suppose $S<\infty$ and let $\varepsilon>0$. Choose a finite $\Omega\subseteq\Gamma$ with $\sum_{j\in\Omega} a_j > S-\varepsilon$. Then for every finite Ω' with $\Omega\subseteq\Omega'\subseteq\Gamma$, monotonicity and the definition of S give

$$S-\varepsilon < \sum_{j\in\Omega} a_j \leq \sum_{j\in\Omega'} a_j \leq S,$$

so $\big|S - \sum_{j\in\Omega'} a_j\big| < \varepsilon$; by 8.53 the unordered sum converges, with value S.

(ii) Suppose $\sum_{k\in\Gamma} a_k = g \in \mathbb{R}$, and given $\varepsilon > 0$ take the finite set Ω_ε provided by 8.53. For an arbitrary finite $\Omega \subseteq \Gamma$, the set $\Omega \cup \Omega_\varepsilon$ contains Ω_ε, so

$$\sum_{j\in\Omega} a_j \le \sum_{j\in\Omega\cup\Omega_\varepsilon} a_j < g + \varepsilon.$$

Hence $S \le g + \varepsilon$: with $\varepsilon = 1$ this makes S finite, and letting $\varepsilon \to 0$ gives $S \le g$. In the other direction $S \ge \sum_{j\in\Omega_\varepsilon} a_j > g - \varepsilon$ for every $\varepsilon > 0$, so $S = g$.

Problem 8C.3. Suppose $\{e_k\}_{k\in\Gamma}$ is an orthonormal family in an inner product space V. Prove that if $f \in V$, then $\{k \in \Gamma : \langle f, e_k\rangle \neq 0\}$ is a countable set.

Solution. The set in question is $\bigcup_{n=1}^{\infty} \Gamma_n$, where

$$\Gamma_n = \left\{k \in \Gamma : |\langle f, e_k\rangle|^2 > \tfrac{1}{n}\right\},$$

since $\langle f, e_k\rangle \neq 0$ forces $|\langle f, e_k\rangle|^2 > \frac{1}{n}$ for some n by the Archimedean property, and conversely each Γ_n consists of such indices.

Each Γ_n is finite. Bessel's inequality (8.57), combined with Exercise 2 of this section identifying a nonnegative unordered sum with the supremum of its finite subsums, gives

$$\sum_{j\in\Omega} |\langle f, e_j\rangle|^2 \le \|f\|^2 \qquad \text{for every finite } \Omega \subseteq \Gamma.$$

So if $\Omega \subseteq \Gamma_n$ is finite with $m \ge 1$ elements, then $\frac{m}{n} < \|f\|^2$, that is, $m < n\|f\|^2$; every finite subset of Γ_n is thus bounded in size by $n\|f\|^2$, which forces Γ_n itself to be finite.

Hence the set is a countable union of finite sets, so it is countable.

Problem 8C.4. Suppose $\{f_k\}_{k\in\Gamma}$ and $\{g_k\}_{k\in\Gamma}$ are families in a normed vector space such that $\sum_{k\in\Gamma} f_k$ and $\sum_{k\in\Gamma} g_k$ converge. Prove that $\sum_{k\in\Gamma}(f_k + g_k)$ converges and

$$\sum_{k\in\Gamma}(f_k + g_k) = \sum_{k\in\Gamma} f_k + \sum_{k\in\Gamma} g_k.$$

Solution. Write $f = \sum_{k\in\Gamma} f_k$ and $g = \sum_{k\in\Gamma} g_k$. Given $\varepsilon > 0$, take $\Omega = \Omega_1 \cup \Omega_2$, where Ω_1 and Ω_2 are the finite sets that 8.53 supplies for these two sums at tolerance $\varepsilon/2$.

If Ω' is finite with $\Omega \subseteq \Omega' \subseteq \Gamma$, then both estimates apply to Ω', and finite sums split, so the triangle inequality gives

$$\begin{aligned}\Big\|(f+g) - \sum_{j\in\Omega'}(f_j + g_j)\Big\| &\le \Big\|f - \sum_{j\in\Omega'} f_j\Big\| + \Big\|g - \sum_{j\in\Omega'} g_j\Big\| \\ &< \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

By 8.53 this says exactly that $\sum_{k\in\Gamma}(f_k + g_k)$ converges to $f + g$.

Problem 8C.5. Suppose $\{f_k\}_{k\in\Gamma}$ is a family in a normed vector space such that $\sum_{k\in\Gamma} f_k$ converges. Prove that if $c \in \mathbf{F}$, then $\sum_{k\in\Gamma}(cf_k)$ converges and

$$\sum_{k\in\Gamma}(cf_k) = c\sum_{k\in\Gamma} f_k.$$

Solution. Write $f = \sum_{k\in\Gamma} f_k$. Given $\varepsilon > 0$, take the finite set Ω that 8.53 supplies for this sum at tolerance $\varepsilon/(1 + |c|)$.

If Ω' is finite with $\Omega \subseteq \Omega' \subseteq \Gamma$, then scalar multiplication distributes over the finite sum, so homogeneity of the norm gives

$$\Big\|cf - \sum_{j\in\Omega'}(cf_j)\Big\| = |c|\,\Big\|f - \sum_{j\in\Omega'} f_j\Big\| \le \frac{|c|}{1+|c|}\,\varepsilon < \varepsilon.$$

By 8.53 this says exactly that $\sum_{k\in\Gamma}(cf_k)$ converges to cf.

Problem 8C.6. Suppose $\{a_k\}_{k\in\Gamma}$ is a family in $\mathbb{R}$. Prove that the unordered sum $\sum_{k\in\Gamma} a_k$ converges if and only if $\sum_{k\in\Gamma}|a_k| < \infty$.

Solution. By Exercise 2 of this section, the condition $\sum_{k\in\Gamma}|a_k| < \infty$ says precisely that

$$S = \sup\Big\{\sum_{j\in\Omega}|a_j| : \Omega \text{ is a finite subset of } \Gamma\Big\} < \infty.$$

(i) Suppose $S<\infty$, and put $a_k^+ = \max\{a_k, 0\}$ and $a_k^- = \max\{-a_k, 0\}$, so $a_k = a_k^+ - a_k^-$ with $a_k^\pm \le |a_k|$. Every finite subsum of the nonnegative families $\{a_k^+\}$ and $\{a_k^-\}$ is then at most S, so Exercise 2 makes $\sum_{k\in\Gamma} a_k^+$ and $\sum_{k\in\Gamma} a_k^-$ converge; Exercise 5 (with $c=-1$) and Exercise 4 then give convergence of $\sum_{k\in\Gamma} a_k = \sum_{k\in\Gamma}\big(a_k^+ + (-a_k^-)\big)$.
(ii) Suppose $\sum_{k\in\Gamma} a_k = g$, and apply 8.53 with $\varepsilon = 1$ to get a finite $\Omega_0\subseteq\Gamma$ with $\big|g - \sum_{j\in\Omega'} a_j\big| < 1$ for all finite Ω' with $\Omega_0\subseteq\Omega'\subseteq\Gamma$. For finite $\Omega\subseteq\Gamma\setminus\Omega_0$, disjointness gives

$$\sum_{j\in\Omega} a_j = \Big(\sum_{j\in\Omega_0\cup\Omega} a_j - g\Big) + \Big(g - \sum_{j\in\Omega_0} a_j\Big),$$

so $\big|\sum_{j\in\Omega} a_j\big| < 2$. Splitting Ω into $\Omega^+ = \{j\in\Omega : a_j\ge 0\}$ and $\Omega^- = \Omega\setminus\Omega^+$, both again finite subsets of $\Gamma\setminus\Omega_0$,

$$\sum_{j\in\Omega}|a_j| = \Big|\sum_{j\in\Omega^+} a_j\Big| + \Big|\sum_{j\in\Omega^-} a_j\Big| < 4.$$

Hence any finite $\Omega\subseteq\Gamma$ satisfies $\sum_{j\in\Omega}|a_j| \le \sum_{j\in\Omega_0}|a_j| + 4$, a fixed finite bound, so $S<\infty$.

Problem 8C.7. Suppose $\{f_k\}_{k\in\mathbb{Z}^+}$ is a family in a normed vector space V and $f\in V$. Prove that the unordered sum $\sum_{k\in\mathbb{Z}^+} f_k$ equals f if and only if the usual ordered sum $\sum_{k=1}^\infty f_{p(k)}$ equals f for every injective and surjective function $p:\mathbb{Z}^+\to\mathbb{Z}^+$.

Solution. For injective p, the ordered partial sum $\sum_{k=1}^n f_{p(k)}$ is the finite unordered sum over $\Omega = \{p(1),\dots,p(n)\}$, and each direction exploits this.
(i) Suppose $\sum_{k\in\mathbb{Z}^+} f_k = f$ and p is a bijection. Given $\varepsilon>0$, take the finite set Ω from 8.53 and put $N = \max\{p^{-1}(j) : j\in\Omega\}$ (with $N=1$ if $\Omega=\varnothing$). For $n\ge N$ we have $\Omega\subseteq\{p(1),\dots,p(n)\}$, so 8.53 applied to that set gives

$$\Big\|f - \sum_{k=1}^n f_{p(k)}\Big\| < \varepsilon \qquad (n\ge N).$$

Hence $\sum_{k=1}^\infty f_{p(k)} = f$ in the sense of 6.40.
(ii) Conversely, suppose $\sum_{k\in\mathbb{Z}^+} f_k = f$ fails. Negating 8.53 produces $\varepsilon>0$ such that every finite $A\subseteq\mathbb{Z}^+$ is contained in some finite $\Omega\subseteq\mathbb{Z}^+$ with $\big\|f - \sum_{j\in\Omega} f_j\big\| \ge \varepsilon$. Starting from $\Omega_0 = \varnothing$ and applying this to $A_m = \Omega_{m-1}\cup\{1,\dots,m\}$ gives finite sets $\Omega_1\subseteq\Omega_2\subseteq\cdots$ with

$$\Big\|f - \sum_{j\in\Omega_m} f_j\Big\| \ge \varepsilon, \qquad \{1,\dots,m\}\subseteq\Omega_m,$$

so $\bigcup_m \Omega_m = \mathbb{Z}^+$ and $n_m = |\Omega_m| \ge m$ tends to ∞. Let p list Ω_1 in increasing order and then, for each $m\ge 2$, the elements of $\Omega_m\setminus\Omega_{m-1}$ in increasing order; the blocks are disjoint with union $\mathbb{Z}^+$, so p is a bijection with $\{p(1), \dots, p(n_m)\} = \Omega_m$. Then

$$\Big\| f - \sum_{k=1}^{n_m} f_{p(k)} \Big\| = \Big\| f - \sum_{j \in \Omega_m} f_j \Big\| \geq \varepsilon$$

for every m, and $n_m \to \infty$, so $\sum_{k=1}^{\infty} f_{p(k)}$ does not equal f.

Problem 8C.8. Explain why 8.58 implies that if Γ is a finite set and $\{e_k\}_{k \in \Gamma}$ is an orthonormal family in a Hilbert space V, then $\operatorname{span}\{e_k\}_{k \in \Gamma}$ is a closed subspace of V.

Solution. For finite Γ the set on the right side of 8.58(a), namely

$$\Big\{ \sum_{k \in \Gamma} \alpha_k e_k : \{\alpha_k\}_{k \in \Gamma} \text{ is a family in } \mathbf{F} \text{ and } \sum_{k \in \Gamma} |\alpha_k|^2 < \infty \Big\},$$

is exactly $\operatorname{span}\{e_k\}_{k \in \Gamma}$, so 8.58(a) reads $\overline{\operatorname{span}\{e_k\}_{k \in \Gamma}} = \operatorname{span}\{e_k\}_{k \in \Gamma}$, and a set equal to its own closure is closed (6.7).
Three consequences of finiteness give that identification. The constraint $\sum_{k \in \Gamma} |\alpha_k|^2 < \infty$ is vacuous, since the sum has finitely many finite terms. The unordered sum $\sum_{k \in \Gamma} \alpha_k e_k$ of 8.53 converges to the ordinary finite sum, because taking $\Omega = \Gamma$ in 8.53 leaves $\Omega' = \Gamma$ as the only option. And extending a family on a finite $\Omega \subseteq \Gamma$ by zeros on $\Gamma \setminus \Omega$ changes no sum, so those finite sums are precisely the elements of the span (6.54).

Problem 8C.9. Suppose V is an infinite-dimensional Hilbert space. Prove that there does not exist a basis of V that is an orthonormal family.

Solution. The witness is $f = \sum_{n=1}^{\infty} \frac{1}{n} e_{k_n}$, which lies in V but in no finite span of the e_k.
Suppose $\{e_k\}_{k \in \Gamma}$ is an orthonormal family that is a basis of V, so $\operatorname{span}\{e_k\}_{k \in \Gamma} = V$ by 6.54. Then Γ is infinite, since otherwise V would be the span of a finite list and hence finite-dimensional. Choose distinct $k_1, k_2, \dots \in \Gamma$ and set $\alpha_{k_n} = \frac{1}{n}$, with $\alpha_k = 0$ for all other k. As $\sum_{k \in \Gamma} |\alpha_k|^2 = \sum_n \frac{1}{n^2} < \infty$ and V is complete, 8.54(a) makes $f = \sum_{k \in \Gamma} \alpha_k e_k$ converge in V.
Moreover $\langle f, e_j \rangle = \alpha_j$ for every $j \in \Gamma$: given $\varepsilon > 0$, take the finite Ω from 8.53 and put $\Omega' = \Omega \cup \{j\}$, so that orthonormality gives $\langle \sum_{i \in \Omega'} \alpha_i e_i, e_j \rangle = \alpha_j$ and hence, by Cauchy–Schwarz (8.11) with $\|e_j\| = 1$,

$$\big| \langle f, e_j \rangle - \alpha_j \big| = \Big| \Big\langle f - \sum_{i \in \Omega'} \alpha_i e_i,\, e_j \Big\rangle \Big| \leq \Big\| f - \sum_{i \in \Omega'} \alpha_i e_i \Big\| < \varepsilon.$$

In particular $\langle f, e_{k_n} \rangle = \frac{1}{n} \neq 0$ for every n.
But $f \in V = \operatorname{span}\{e_k\}_{k \in \Gamma}$, so $f = \sum_{j \in \Omega_0} \beta_j e_j$ for some finite $\Omega_0 \subseteq \Gamma$; choosing n with $k_n \notin \Omega_0$ (possible since Ω_0 is finite and the k_n are distinct), orthonormality gives $\langle f, e_{k_n} \rangle = 0$, a contradiction.

Problem 8C.10. (a) Show that the orthonormal family given in the first bullet point of Example 8.51 is an orthonormal basis of ℓ^2.
(b) Show that the orthonormal family given in the second bullet point of Example 8.51 is an orthonormal basis of $\ell^2(\Gamma)$.
(c) Show that the orthonormal family given in the fourth bullet point of Example 8.51 is not an orthonormal basis of $L^2\big([0,1)\big)$.
(d) Show that the orthonormal family given in the fifth bullet point of Example 8.51 is not an orthonormal basis of $L^2(\mathbf{R})$.

Solution. Lemma. An orthonormal family $\{e_k\}_{k \in \Gamma}$ in a Hilbert space V is an orthonormal basis of V if and only if the only $f \in V$ with $\langle f, e_k \rangle = 0$ for all $k \in \Gamma$ is $f = 0$. Indeed, with $U = \operatorname{span}\{e_k\}_{k \in \Gamma}$, additivity and conjugate homogeneity of the second slot give

$$U^{\perp} = \{f \in V : \langle f, e_k \rangle = 0 \text{ for all } k \in \Gamma\}$$

(for the reverse inclusion take $u = e_k$), while 8.42 says $\overline{U} = V$ if and only if $U^{\perp} = \{0\}$, and $\overline{U} = V$ is Definition 8.61.
(a) For $f = (a_1, a_2, \ldots) \in \ell^2$ the standard inner product gives $\langle f, e_k \rangle = a_k$, so orthogonality to every e_k forces $f = 0$; apply the Lemma.
(b) Here $e_k = \chi_{\{k\}}$ and μ is counting measure on Γ, so $f\overline{e_k}$ is the simple function $f(k)\chi_{\{k\}}$ and

$$\langle f, e_k \rangle = \int_{\Gamma} f\overline{e_k}\, d\mu = f(k)\mu(\{k\}) = f(k).$$

Hence orthogonality to every e_k forces $f = 0$ (with counting measure the only null set is $\varnothing$); apply the Lemma. Part (a) is the case $\Gamma = \mathbf{Z}^+$.
(c) Take $f = e_1 e_2$, which on the four quarter-intervals of $[0,1)$ has values $(1, -1, -1, 1)$, because there $e_1 = (1, 1, -1, -1)$ and $e_2 = (1, -1, 1, -1)$. Since $|f| = 1$ we have $\|f\|_2 = 1$, so $f \neq 0$, yet $\langle f, e_k \rangle = \int_0^1 f e_k = 0$ for every $k \geq 0$ (all functions here are real valued):
(i) $k \in \{0, 1, 2\}$. The products $fe_0 = (1, -1, -1, 1)$, $fe_1 = (1, -1, 1, -1)$, and $fe_2 = (1, 1, -1, -1)$ are constant on the four quarter-intervals, of equal length $\frac{1}{4}$, with values summing to 0.
(ii) $k \geq 3$. Writing $f = c_j$ on $I_j = [j2^{-2}, (j+1)2^{-2})$, the Fact below with $m = 2 < k$ gives $\int_0^1 fe_k = \sum_{j=0}^{3} c_j \int_{I_j} e_k = 0$.
Fact. For integers $k > m \geq 0$ and $j \in \{0, \ldots, 2^m - 1\}$,

$$\int_{[\,j2^{-m},\,(j+1)2^{-m})} e_k = 0,$$

since that interval is the disjoint union of the intervals $\left[\frac{n-1}{2^k}, \frac{n}{2^k}\right)$ for 2^{k-m} consecutive integers n, an even number of them, so e_k takes the value 1 on exactly half of these equal-length pieces and -1 on the other half.
Thus $f \neq 0$ is orthogonal to every e_k, and the Lemma applies.
(d) Take the extension by 0 of the f of (c), so $|f| = \chi_{[0,1)}$ and $\|f\|_2 = 1$. If $m \neq 0$, then $[m, m+1) \cap [0,1) = \varnothing$, so $f\overline{e_{k,m}} \equiv 0$; if $m = 0$, then $e_{k,0}$ agrees with e_k on $[0,1)$ and

$$\langle f, e_{k,0} \rangle = \int_0^1 fe_k = 0$$

by (c). So $f \neq 0$ is orthogonal to the whole family, and the Lemma applies.

Problem 8C.11. Suppose μ is a σ-finite measure on $(X, \mathcal{S})$ and ν is a σ-finite measure on $(Y, \mathcal{T})$. Suppose also that $\{e_j\}_{j \in \Omega}$ is an orthonormal basis of $L^2(\mu)$ and $\{f_k\}_{k \in \Gamma}$ is an orthonormal basis of $L^2(\nu)$ for some countable set Γ. For $j \in \Omega$ and $k \in \Gamma$, define $g_{j,k} : X \times Y \to \mathbf{F}$ by

$$g_{j,k}(x, y) = e_j(x) f_k(y).$$

Prove that $\{g_{j,k}\}_{j \in \Omega,\, k \in \Gamma}$ is an orthonormal basis of $L^2(\mu \times \nu)$.

Solution. By the Lemma of Exercise 10 it suffices to prove the family orthonormal and to show that no nonzero $h \in L^2(\mu \times \nu)$ is orthogonal to every $g_{j,k}$. The σ-finiteness of μ and ν is what makes $\mu \times \nu$ a measure and licenses each use below of Tonelli (5.28) and Fubini (5.32); for $\mathbf{F} = \mathbf{C}$ apply those to real and imaginary parts (and, for Tonelli, to their positive and negative parts).
Orthonormal. Each $g_{j,k}$ is $\mathcal{S} \otimes \mathcal{T}$-measurable, being the product of $(x, y) \mapsto e_j(x)$ and $(x, y) \mapsto f_k(y)$, whose inverse images are rectangles. Tonelli (5.28) applied to $|g_{j,k}|^2 = |e_j(x)|^2 |f_k(y)|^2 \geq 0$ gives $\|g_{j,k}\|^2 = \|e_j\|^2 \|f_k\|^2 = 1$, so $g_{j,k} \in L^2(\mu \times \nu)$; then $g_{j,k}\overline{g_{j',k'}} \in L^1(\mu \times \nu)$ by Cauchy–Schwarz (8.11),

and Fubini (5.32) gives

$$\langle g_{j,k}, g_{j',k'}\rangle = \langle e_j, e_{j'}\rangle\,\langle f_k, f_{k'}\rangle,$$

which vanishes unless $(j,k) = (j',k')$, since then $j \neq j'$ or $k \neq k'$.
Nothing nonzero is orthogonal. Suppose $\langle h, g_{j,k}\rangle = 0$ for all $j \in \Omega$, $k \in \Gamma$. Tonelli (5.28) makes

$$\varphi(x) = \int_Y |h(x,y)|^2\,d\nu(y)$$

an $\mathcal{S}$-measurable function with $\int_X \varphi\,d\mu = \|h\|^2 < \infty$, so $\varphi < \infty$ on some $E \in \mathcal{S}$ with $\mu(X \setminus E) = 0$; for $x \in E$ the cross section h_x, which is $\mathcal{T}$-measurable by 5.9, lies in $L^2(\nu)$ with $\|h_x\|^2 = \varphi(x)$. Define

$$H_k(x) = \langle h_x, f_k\rangle \text{ for } x \in E, \qquad H_k(x) = 0 \text{ for } x \notin E.$$

Cauchy–Schwarz (8.11) in $L^2(\nu)$ gives $|H_k| \leq \varphi^{1/2}$ on X, so $H_k \in L^2(\mu)$; and H_k is $\mathcal{S}$-measurable by 5.28 applied to the positive and negative parts of the real and imaginary parts of $(x,y) \mapsto h(x,y)\overline{f_k(y)}$, each finite on E by that bound.
Now use each basis once. Parseval's identity 8.63(c) for $\{f_k\}_{k\in\Gamma}$, applied to h_x, gives

$$\varphi(x) = \sum_{k\in\Gamma} |H_k(x)|^2 \qquad (x \in E),$$

while $h\overline{g_{j,k}} \in L^1(\mu \times \nu)$ as above, so Fubini (5.32) gives

$$0 = \langle h, g_{j,k}\rangle = \int_X H_k \overline{e_j}\,d\mu = \langle H_k, e_j\rangle \qquad (j \in \Omega),$$

the inner integral being $H_k(x)$ for μ-almost every x. Parseval for $\{e_j\}_{j\in\Omega}$ then forces $\|H_k\|^2 = 0$, so $H_k = 0$ almost everywhere.
Countability of Γ enters here: $N = \bigcup_{k\in\Gamma}\{H_k \neq 0\}$ has $\mu(N) = 0$ by countable subadditivity, so $\varphi = 0$ on $E \setminus N$ and hence μ-almost everywhere, giving

$$\|h\|^2 = \int_X \varphi\,d\mu = 0.$$

Problem 8C.12. Prove the converse of Parseval's identity. More specifically, prove that if $\{e_k\}_{k\in\Gamma}$ is an orthonormal family in a Hilbert space V and

$$\|f\|^2 = \sum_{k\in\Gamma} |\langle f, e_k\rangle|^2$$

for every $f \in V$, then $\{e_k\}_{k\in\Gamma}$ is an orthonormal basis of V.

Solution. Put $U = \operatorname{span}\{e_k\}_{k\in\Gamma}$. Then $U^\perp = \{0\}$, since every $f \in U^\perp$ satisfies

$$\|f\|^2 = \sum_{k\in\Gamma} |\langle f, e_k\rangle|^2 = 0$$

by the hypothesis.
Here $U^\perp$ is exactly $\{f \in V : \langle f, e_k\rangle = 0 \text{ for all } k \in \Gamma\}$, because the inner product is additive and conjugate homogeneous in its second slot, so orthogonality to each e_k gives orthogonality to every finite combination of them (and conversely, take $u = e_k$). As V is a Hilbert space and U is a subspace, 8.42 turns $U^\perp = \{0\}$ into $\overline{U} = V$; with the family orthonormal by hypothesis, that is Definition 8.61.

Problem 8C.13. (a) Show that the Hilbert space $L^2([0,1])$ is separable.
(b) Show that the Hilbert space $L^2(\mathbf{R})$ is separable.
(c) Show that the Banach space ℓ^∞ is not separable.

Solution. For (a) and (b) the countable dense set is

$$D_J = \Big\{ \sum_{j=1}^n q_j\, \chi_{(c_j,d_j)}\big|_J \ :\ n \in \mathbf{Z}^+,\ q_j \in \mathbf{Q_F},\ c_j, d_j \in \mathbf{Q},\ c_j < d_j \Big\},$$

where $\mathbf{Q_F} = \mathbf{Q}$ if $\mathbf{F} = \mathbf{R}$ and $\mathbf{Q_F} = \{r + is : r, s \in \mathbf{Q}\}$ if $\mathbf{F} = \mathbf{C}$. Each element of D_J is bounded and supported in a bounded set, hence in $L^2(J)$, and D_J is countable, being the image of the countable set of finite sequences from $\mathbf{Q_F} \times \mathbf{Q} \times \mathbf{Q}$.
Claim 1. Simple measurable functions are dense in $L^2(J)$. Write $f = u + iv$ with u, v real, so $u, v \in L^2(J)$; by 2.89 there are simple measurable u_k with $|u_k| \le |u|$ and $u_k \to u$ pointwise, whence $|u - u_k|^2 \le 4u^2 \in L^1(J)$ and the Dominated Convergence Theorem (3.31) gives $\|u - u_k\|_2 \to 0$; likewise for v.
Claim 2. If $|A| < \infty$ and $\delta > 0$, some finite union W of pairwise disjoint bounded open intervals with rational endpoints has $|A \triangle W| < \delta$. By 2.71(e) choose open $G \supseteq A$ with $|G \setminus A| < \delta/4$, so $|G| < \infty$; write $G = \bigcup_j I_j$ as a countable disjoint union of open intervals, each bounded, and choose N with $|G \setminus U| = \sum_{j>N} |I_j| < \delta/4$, where $U = \bigcup_{j \le N} I_j$. Shrink each $I_j = (\alpha_j, \beta_j)$ with $j \le N$ to a rational interval $(c_j, d_j) \subseteq I_j$ with $|I_j| - (d_j - c_j) < \delta/(4N)$, and set $W = \bigcup_{j \le N} (c_j, d_j)$, still pairwise disjoint. Then

$$|A \setminus W| \le |A \setminus G| + |G \setminus U| + |U \setminus W| < \tfrac{\delta}{2}, \qquad |W \setminus A| \le |G \setminus A| < \tfrac{\delta}{4}.$$

Claim 3. D_J is dense in $L^2(J)$ when $|J| < \infty$. Given f and $\varepsilon > 0$, Claim 1 supplies a simple $s = \sum_{k=1}^n a_k \chi_{A_k}$ with the a_k nonzero and the $A_k \subseteq J$ disjoint, such that $\|f - s\|_2 < \varepsilon/3$. Choosing $q_k \in \mathbf{Q_F}$ with $|a_k - q_k| < \varepsilon / \big(3n(1 + |J|)^{1/2}\big)$ gives, since $\|\chi_{A_k}\|_2 = |A_k|^{1/2} \le |J|^{1/2}$,

$$\Big\| s - \sum_{k=1}^n q_k \chi_{A_k} \Big\|_2 \le \sum_{k=1}^n |a_k - q_k|\, |A_k|^{1/2} < \frac{\varepsilon}{3}.$$

Now take δ with $\delta^{1/2} \sum_k |q_k| < \varepsilon/3$ and, by Claim 2, sets W_k with $|A_k \triangle W_k| < \delta$; as $|\chi_{A_k} - \chi_{W_k \cap J}| \le \chi_{A_k \triangle W_k}$ on J,

$$\Big\| \sum_{k=1}^n q_k \chi_{A_k} - \sum_{k=1}^n q_k \chi_{W_k}\big|_J \Big\|_2 \le \sum_{k=1}^n |q_k|\, \delta^{1/2} < \frac{\varepsilon}{3}.$$

The intervals composing each W_k are disjoint, so χ_{W_k} is the sum of their characteristic functions and $\sum_k q_k \chi_{W_k}|_J \in D_J$. Adding the three estimates finishes the claim.
(a) $|[0,1]| < \infty$, so Claim 3 makes $D_{[0,1]}$ a countable dense subset of $L^2([0,1])$ (8.64).
(b) Given $f \in L^2(\mathbf{R})$ and $\varepsilon > 0$, the Dominated Convergence Theorem (3.31) applied to $|f - f\chi_{[-m,m]}|^2 \le |f|^2$ gives m with $\|f - f\chi_{[-m,m]}\|_2 < \varepsilon/2$, and Claim 3 on $[-m,m]$ gives $g \in D_{[-m,m]}$ within $\varepsilon/2$ of $f\chi_{[-m,m]}$ there. Intersecting each of the finitely many intervals of g with $(-m,m)$ yields $\widetilde{g} \in D_{\mathbf{R}}$ vanishing off $[-m,m]$, so $\|f\chi_{[-m,m]} - \widetilde{g}\|_{L^2(\mathbf{R})} < \varepsilon/2$ and $\|f - \widetilde{g}\|_{L^2(\mathbf{R})} < \varepsilon$.
(c) For $S \subseteq \mathbf{Z}^+$ let $x_S \in \ell^\infty$ have kth term 1 if $k \in S$ and 0 otherwise. Distinct S, S' differ in some coordinate while all coordinates of the difference are at most 1 in absolute value, so $\|x_S - x_{S'}\|_\infty = 1$ and the uncountably many balls $B(x_S, \frac{1}{2})$ are pairwise disjoint by the triangle inequality. A countable dense $C \subseteq \ell^\infty$ would meet each of them, injecting an uncountable set into C.

Problem 8C.14. Prove that every subspace of a separable normed vector space is separable.

Solution. Take the countable dense set $\{f_n\}_{n \in I}$ of V given by 8.64, indexed by some $I \subseteq \mathbf{Z}^+$, and put

$$C = \{u_{n,k} : (n,k) \in P\}, \qquad P = \Big\{(n,k) \in I \times \mathbf{Z}^+ : B\big(f_n, \tfrac{1}{k}\big) \cap U \neq \varnothing \Big\},$$

where $u_{n,k}$ is a chosen element of $B(f_n, \frac{1}{k}) \cap U$. Then $C \subseteq U$ is countable, being the image of $P \subseteq \mathbf{Z}^+ \times \mathbf{Z}^+$.
C is dense in U: given $u \in U$ and $\varepsilon > 0$, choose k with $\frac{2}{k} < \varepsilon$ and then $n \in I$ with $\|u - f_n\| < \frac{1}{k}$, which puts u in $B(f_n, \frac{1}{k}) \cap U$, so $(n, k) \in P$ and

$$\|u - u_{n,k}\| \leq \|u - f_n\| + \|f_n - u_{n,k}\| < \tfrac{2}{k} < \varepsilon.$$

Since the metric of U is the restriction of the metric of V, this says the closure of C in U is U (6.7).

Problem 8C.15. Suppose V is an infinite-dimensional Hilbert space. Prove that there does not exist a translation invariant measure on the Borel subsets of V that assigns positive but finite measure to each open ball in V.
[A subset of V is called a **Borel set** if it is in the smallest σ-algebra containing all the open subsets of V. A measure μ on the Borel subsets of V is called **translation invariant** if $\mu(f + E) = \mu(E)$ for every $f \in V$ and every Borel set E of V.]

Solution. Infinitely many disjoint balls of one fixed radius fit inside a single ball, which no such μ can tolerate. Write $B(f, r) = \{g \in V : \|g - f\| < r\}$, a Borel set, and suppose μ is translation invariant with $0 < \mu(B) < \infty$ for every open ball B.
By 8.75 the space V has an orthonormal basis $\{e_k\}_{k \in \Gamma}$, and Γ is infinite: a finite Γ would make Parseval's identity 8.63(a) write every $f \in V$ as $\sum_{k \in \Gamma} \langle f, e_k \rangle e_k$, so that V would be finite-dimensional.
Choose distinct $k_1, k_2, \ldots \in \Gamma$ and set $u_n = e_{k_n}$. For $m \neq n$ the Pythagorean theorem (8.9) gives

$$\|u_m - u_n\|^2 = \|u_m\|^2 + \|u_n\|^2 = 2,$$

so with $r = \frac{\sqrt{2}}{2}$ the balls $B(u_n, r)$ are pairwise disjoint (a common point h would give $\sqrt{2} \leq \|u_m - h\| + \|h - u_n\| < \sqrt{2}$), and each lies in $B(0, 1 + r)$ because $\|h\| \leq \|h - u_n\| + 1 < r + 1$ there.
Since $B(u_n, r) = u_n + B(0, r)$, translation invariance gives $\mu(B(u_n, r)) = \mu(B(0, r)) = c > 0$ for every n, so monotonicity and countable additivity give

$$\mu\big(B(0, 1 + r)\big) \;\geq\; \sum_{n=1}^{\infty} \mu\big(B(u_n, r)\big) \;=\; \sum_{n=1}^{\infty} c \;=\; \infty,$$

contradicting $\mu(B(0, 1 + r)) < \infty$.

Problem 8C.16. Find the polynomial g of degree at most 4 that minimizes

$$\int_0^1 \big|x^5 - g(x)\big|^2 \, dx.$$

Solution. The minimizer is

$$g(x) = \frac{5}{2}x^4 - \frac{20}{9}x^3 + \frac{5}{6}x^2 - \frac{5}{42}x + \frac{1}{252}.$$

Work in the real Hilbert space $L^2([0, 1])$ with $\langle p, q \rangle = \int_0^1 pq$, let $f(x) = x^5$, and let U be the space of polynomials of degree at most 4, a 5-dimensional and hence closed subspace (Exercise 9 in Section 6D). The minimizer therefore exists, is unique, and equals $P_U f$ (8.34); by 8.37(a) and 8.37(b) it is the unique $g \in U$ with $f - g \in U^{\perp}$, that is,

$$\int_0^1 \big(x^5 - g(x)\big) x^j \, dx = 0 \qquad \text{for } j = 0, 1, 2, 3, 4.$$

So $x^5 - g$ is the monic polynomial of degree 5 orthogonal to U. (Complex coefficients change nothing: $g = g_1 + ig_2$ gives $\int_0^1 |x^5 - g|^2 = \int_0^1 (x^5 - g_1)^2 + \int_0^1 g_2^2$, minimized only at $g_2 = 0$.)

That polynomial is $q = \frac{5!}{10!}u^{(5)}$, where $u(x) = (x^2-x)^5 = x^5(x-1)^5$. It is monic of degree 5, since u has leading term x^{10}. It is orthogonal to each x^j with $j \leq 4$ because 0 and 1 are zeros of u of multiplicity 5, so integrating by parts $j+1$ times leaves

$$\int_0^1 u^{(5)}(x)\, x^j\, dx = \Big[\sum_{i=0}^{j}(-1)^i u^{(4-i)}(x)\,(x^j)^{(i)}\Big]_0^1 + (-1)^{j+1}\int_0^1 u^{(4-j)}(x)\,(x^j)^{(j+1)}\, dx = 0,$$

every boundary term involving some $u^{(i)}$ with $i \leq 4$, and $(x^j)^{(j+1)} = 0$.
Expanding $u(x) = x^{10} - 5x^9 + 10x^8 - 10x^7 + 5x^6 - x^5$ gives

$$u^{(5)}(x) = 30240x^5 - 75600x^4 + 67200x^3 - 25200x^2 + 3600x - 120,$$

and $\frac{5!}{10!} = \frac{1}{30240}$, so

$$q(x) = x^5 - \frac{5}{2}x^4 + \frac{20}{9}x^3 - \frac{5}{6}x^2 + \frac{5}{42}x - \frac{1}{252}.$$

Hence $g = f - q$ is the polynomial displayed at the outset.

Problem 8C.17. Prove that each orthonormal family in a Hilbert space can be extended to an orthonormal basis of the Hilbert space. Specifically, suppose $\{e_j\}_{j\in\Omega}$ is an orthonormal family in a Hilbert space V. Prove that there exists a set Γ containing Ω and an orthonormal basis $\{f_k\}_{k\in\Gamma}$ of V such that $f_j = e_j$ for every $j \in \Omega$.

Solution. Apply Zorn's Lemma to the orthonormal subsets of V containing $E = \{e_j : j \in \Omega\}$, then index the maximal one.
The map $j \mapsto e_j$ is injective, since $\langle e_j, e_j\rangle = 1$ while $\langle e_j, e_{j'}\rangle = 0$, so E is an orthonormal subset of V in bijection with Ω. Let

$$\mathcal{A} = \{\varnothing\} \cup \{S \subseteq V : S \text{ is orthonormal and } E \subseteq S\},$$

the empty set included only so the empty chain's union lies in $\mathcal{A}$. If $\mathcal{C} \subseteq \mathcal{A}$ is a chain with union $L \neq \varnothing$, then some member of $\mathcal{C}$ contains E, so $E \subseteq L$; each $f \in L$ has $\|f\| = 1$, and distinct $f, g \in L$ lie in a common member of the chain, hence $\langle f, g\rangle = 0$. So $L \in \mathcal{A}$, and Zorn's Lemma (6.60) gives a maximal $F \in \mathcal{A}$.
Then $E \subseteq F$ (were $F = \varnothing$ with $E \neq \varnothing$, the element $E \in \mathcal{A}$ would contradict maximality), and F is maximal among all orthonormal subsets of V: any orthonormal $F' \supsetneq F$ contains E, hence lies in $\mathcal{A}$. So F is an orthonormal basis of V by 8.74.
For the indexing put $\Lambda = (F \setminus E) \times \{\Omega\}$, which is disjoint from Ω (else $(h, \Omega) \in \Omega$ would give the membership cycle $\Omega \in \{h, \Omega\} \in (h, \Omega) \in \Omega$), and let $\Gamma = \Omega \cup \Lambda \supseteq \Omega$ with

$$f_j = e_j \text{ for } j \in \Omega, \qquad f_{(h,\Omega)} = h \text{ for } (h, \Omega) \in \Lambda.$$

Disjointness makes this well defined, and $k \mapsto f_k$ is a bijection of Γ onto $E \cup (F \setminus E) = F$. Hence $\{f_k\}_{k\in\Gamma}$ is an orthonormal family whose set of values is F; since the condition in Definition 8.61 depends only on that set of values, $\{f_k\}_{k\in\Gamma}$ is an orthonormal basis of V with $f_j = e_j$ on Ω.

Problem 8C.18. Prove that every vector space has a basis.

Solution. Apply Zorn's Lemma to the collection $\mathcal{A}$ of linearly independent subsets of V; a maximal element is a basis.
The hypothesis of 6.60 holds. Let $\mathcal{C} \subseteq \mathcal{A}$ be a chain with union L, and suppose

$$a_1 g_1 + \cdots + a_n g_n = 0$$

with $g_1, \ldots, g_n$ distinct elements of L. Each g_m lies in some $\Gamma_m \in \mathcal{C}$, and one of the finitely many Γ_m contains all the others (induction on n, since $\mathcal{C}$ is a chain); linear independence of that one forces

$a_1 = \cdots = a_n = 0$. (For $\mathcal{C} = \varnothing$ the union $\varnothing$ is vacuously independent.) So $\mathcal{A}$ has a maximal element Γ.
Then $\operatorname{span}\Gamma = V$, by the argument of 6.57: otherwise choose $f \notin \operatorname{span}\Gamma$, so $f \notin \Gamma$, and $\Gamma \cup \{f\}$ is still linearly independent, contradicting maximality. Indeed, in a relation $af + a_1 g_1 + \cdots + a_n g_n = 0$ with $g_1, \dots, g_n$ distinct in Γ, the case $a \neq 0$ would give

$$f = -\frac{a_1}{a} g_1 - \cdots - \frac{a_n}{a} g_n \in \operatorname{span}\Gamma,$$

so $a = 0$, and then independence of Γ kills the remaining coefficients. Being linearly independent and spanning, Γ is a basis of V.

Problem 8C.19. Find the polynomial g of degree at most 4 such that

$$f\left(\tfrac{1}{2}\right) = \int_0^1 fg$$

for every polynomial f of degree at most 4.

Solution. The answer is

$$g(x) = \frac{15}{8}\left(126x^4 - 252x^3 + 154x^2 - 28x + 1\right) = \frac{945}{4}x^4 - \frac{945}{2}x^3 + \frac{1155}{4}x^2 - \frac{105}{2}x + \frac{15}{8}.$$

Let $U \subseteq L^2([0,1])$ be the space of polynomials of degree at most 4, with $\langle p, q\rangle = \int_0^1 pq$; being finite-dimensional it is closed, hence a Hilbert space [8.22], and $\varphi(f) = f\left(\tfrac{1}{2}\right)$ is a linear functional on U, bounded because U is finite-dimensional (Exercise 7 in Section 6D). The Riesz Representation Theorem (8.47) thus gives a unique $g \in U$ with $f\left(\tfrac{1}{2}\right) = \int_0^1 fg$ for every such f, and 8.77 computes it from any orthonormal basis of U (the scalars here being real, so conjugation is invisible):

$$g = \sum_{k=0}^{4} \varphi(e_k)\, e_k = \sum_{k=0}^{4} e_k\left(\tfrac{1}{2}\right) e_k.$$

An orthonormal basis of U. Take the shifted Legendre polynomials

$$p_k(x) = \frac{1}{k!}\,\frac{d^k}{dx^k}\left[(x^2 - x)^k\right] \qquad (k = 0, \dots, 4),$$

of degree k with leading coefficient $\frac{(2k)!}{(k!)^2}$, so that they form a basis of U. Writing $u(x) = (x^2 - x)^k$, both 0 and 1 are zeros of u of multiplicity k, so for $0 \le j < k$ integrating by parts $j+1$ times kills every boundary term:

$$\int_0^1 u^{(k)}(x)\, x^j\, dx = (-1)^{j+1} \int_0^1 u^{(k-j-1)}(x)\, (x^j)^{(j+1)}\, dx = 0,$$

whence $\langle p_k, p_j\rangle = 0$ for $j < k$. Integrating by parts k times instead, and using the beta integral $\int_0^1 x^k (1-x)^k dx = \frac{(k!)^2}{(2k+1)!}$,

$$\int_0^1 u^{(k)}(x)\, x^k\, dx = (-1)^k k! \int_0^1 (x^2 - x)^k\, dx = k! \cdot \frac{(k!)^2}{(2k+1)!},$$

so $\langle p_k, x^k\rangle = \frac{(k!)^2}{(2k+1)!}$ and, by orthogonality to lower degrees,

$$\|p_k\|^2 = \left\langle p_k, \frac{(2k)!}{(k!)^2} x^k \right\rangle = \frac{1}{2k+1}.$$

Hence $e_k = \sqrt{2k+1}\, p_k$ for $k = 0, \dots, 4$ is an orthonormal basis of U.

Evaluating at $\frac{1}{2}$. As $(x^2-x)^k$ is unchanged by $x \mapsto 1-x$, differentiating k times gives $p_k(1-x) = (-1)^k p_k(x)$, so $p_k(\frac{1}{2}) = 0$ for odd k, while

$$p_0 = 1, \quad p_2 = 6x^2 - 6x + 1, \quad p_4 = 70x^4 - 140x^3 + 90x^2 - 20x + 1$$

give $p_0(\frac{1}{2}) = 1$, $p_2(\frac{1}{2}) = -\frac{1}{2}$, and $p_4(\frac{1}{2}) = \frac{3}{8}$. Since $e_k(\frac{1}{2})e_k = (2k+1)p_k(\frac{1}{2})p_k$,

$$g = p_0 - \tfrac{5}{2}\, p_2 + \tfrac{27}{8}\, p_4,$$

which expands to the polynomial displayed at the outset.

Problem 8C.20. Suppose G is a nonempty open subset of $\mathbf{C}$. The **Bergman space** $L^2_a(G)$ is defined to be the set of analytic functions $f : G \to \mathbf{C}$ such that

$$\int_G |f|^2 \, d\lambda_2 < \infty,$$

where λ_2 is the usual Lebesgue measure on $\mathbf{R}^2$, which is identified with $\mathbf{C}$. For $f, h \in L^2_a(G)$, define $\langle f, h\rangle$ to be $\int_G f\overline{h}\, d\lambda_2$.
(a) Show that $L^2_a(G)$ is a Hilbert space.
(b) Show that if $w \in G$, then $f \mapsto f(w)$ is a bounded linear functional on $L^2_a(G)$.

Solution. Everything follows from one estimate: writing $D(w,r) = \{z : |z-w| < r\}$ and $\|f\| = (\int_G |f|^2 d\lambda_2)^{1/2}$, if f is analytic on G and $\overline{D(w,r)} \subseteq G$, then

$$|f(w)| \le \frac{1}{\sqrt{\pi}\, r}\, \|f\|.$$

Call this the *pointwise bound*. Indeed, Cauchy's integral formula on the circle $|z-w| = s$ (contained in G with its inside, for $0 < s \le r$) gives $f(w) = \frac{1}{2\pi}\int_0^{2\pi} f(w + se^{i\theta})\, d\theta$; multiplying by $2\pi s$, integrating over $s \in (0,r)$ and using polar coordinates (legitimate since f is bounded on the compact set $\overline{D(w,r)}$) yields the area mean value property

$$\pi r^2 f(w) = \int_{D(w,r)} f \, d\lambda_2,$$

and then Cauchy-Schwarz (8.11) against the constant 1 on $D(w,r)$ gives $|f(w)| \le \frac{1}{\pi r^2}\|f\|(\pi r^2)^{1/2}$, which is the pointwise bound.
(a) $L^2_a(G)$ is a subspace of $L^2(\lambda_2|_G)$, since sums and scalar multiples of analytic functions are analytic and $|f+h|^2 \le 2|f|^2 + 2|h|^2$; $\langle f, h\rangle$ is the L^2 inner product, and it is definite on $L^2_a(G)$ as a genuine function space because $f(w) \ne 0$ forces $|f| > 0$ on a disk of positive measure. By 8.21 it remains to prove completeness. Let $f_1, f_2, \ldots$ be Cauchy in $L^2_a(G)$. For compact $K \subseteq G$ put $r_K = \min\{1, \frac{1}{2}\operatorname{dist}(K, \mathbf{C} \setminus G)\} > 0$ (positive as K is compact and disjoint from the closed set $\mathbf{C} \setminus G$), so $\overline{D(w, r_K)} \subseteq G$ for $w \in K$ and the pointwise bound applied to $f_n - f_m$ gives

$$\sup_{w \in K} |f_n(w) - f_m(w)| \le \frac{1}{\sqrt{\pi}\, r_K}\|f_n - f_m\|.$$

Thus $f_n \to f$ uniformly on compact subsets of G, and f is analytic by the Weierstrass convergence theorem. Given $\varepsilon > 0$ choose N with $\|f_n - f_m\| < \varepsilon$ for $n, m \ge N$; for $n \ge N$ and compact $K \subseteq G$, uniform convergence on K and $\lambda_2(K) < \infty$ give

$$\int_K |f_n - f|^2 \, d\lambda_2 = \lim_{m \to \infty} \int_K |f_n - f_m|^2 \, d\lambda_2 \le \varepsilon^2.$$

Taking $K_j = \{z : |z| \le j,\ \operatorname{dist}(z, \mathbf{C} \setminus G) \ge \frac{1}{j}\}$, which are compact subsets of G increasing to G, the Monotone Convergence Theorem (3.11) upgrades this to $\|f_n - f\|^2 \le \varepsilon^2$. Hence $f = f_n - (f_n - f) \in L^2_a(G)$ and $f_n \to f$ in norm, so $L^2_a(G)$ is a Hilbert space.

(b) Fix $w \in G$ and pick $r > 0$ with $\overline{D(w,r)} \subseteq G$. Evaluation $\varphi(f) = f(w)$ is linear (operations in $L^2_a(G)$ are pointwise) and the pointwise bound gives $|\varphi(f)| \le \frac{1}{\sqrt{\pi}\, r}\|f\|$, so φ is bounded with $\|\varphi\| \le \frac{1}{\sqrt{\pi}\, r}$.

Problem 8C.21. Let $\mathbf{D}$ denote the open unit disk in $\mathbf{C}$; thus

$$\mathbf{D} = \{z \in \mathbf{C} : |z| < 1\}.$$

(a) Find an orthonormal basis of $L^2_a(\mathbf{D})$.
(b) Suppose $f \in L^2_a(\mathbf{D})$ has Taylor series

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

for $z \in \mathbf{D}$. Find a formula for $\|f\|$ in terms of $a_0, a_1, a_2, \ldots$.
(c) Suppose $w \in \mathbf{D}$. By the previous exercise and the Riesz Representation Theorem (8.47 and 8.76), there exists $\Gamma_w \in L^2_a(\mathbf{D})$ such that

$$f(w) = \langle f, \Gamma_w \rangle \quad \text{for all } f \in L^2_a(\mathbf{D}).$$

Find an explicit formula for Γ_w.

Solution. (a) $\{\sqrt{(k+1)/\pi}\, z^k\}_{k\ge 0}$; (b) $\|f\| = \sqrt{\pi}\big(\sum_k \frac{|a_k|^2}{k+1}\big)^{1/2}$; (c) $\Gamma_w(z) = \frac{1}{\pi(1-\overline{w}z)^2}$.
For $j, k \ge 0$ and $0 < s \le 1$, polar coordinates give

$$\int_{D(0,s)} z^j \overline{z^k}\, d\lambda_2 = \int_0^s \int_0^{2\pi} t^{j+k} e^{i(j-k)\theta}\, t\, d\theta\, dt = \begin{cases} 0 & \text{if } j \ne k, \\ \dfrac{\pi s^{2k+2}}{k+1} & \text{if } j = k. \end{cases}$$

Now let $f \in L^2_a(\mathbf{D})$ have Taylor series $f(z) = \sum_{n\ge 0} a_n z^n$, convergent on $\mathbf{D}$ and uniformly on each $\overline{D(0,s)}$ with $s < 1$. Fix $k \ge 0$ and $0 < s < 1$; uniform convergence on the finite-measure set $\overline{D(0,s)}$ permits term-by-term integration, so

$$\int_{D(0,s)} f(z)\overline{z^k}\, d\lambda_2 = \sum_{n=0}^{\infty} a_n \int_{D(0,s)} z^n \overline{z^k}\, d\lambda_2 = \frac{\pi s^{2k+2}}{k+1} a_k.$$

Letting $s \uparrow 1$, the integrands $\chi_{D(0,s)} f \overline{z^k}$ are dominated by $|f|$, integrable on $\mathbf{D}$ by Cauchy-Schwarz (8.11), so the Dominated Convergence Theorem (3.31) gives

$$\langle f, z^k \rangle = \int_{\mathbf{D}} f \overline{z^k}\, d\lambda_2 = \frac{\pi}{k+1} a_k \qquad (k = 0, 1, 2, \ldots)$$

(a) Put $e_k(z) = \sqrt{\frac{k+1}{\pi}}\, z^k$. The first display with $s = 1$ gives $\|z^k\|^2 = \frac{\pi}{k+1}$ and $\langle z^j, z^k \rangle = 0$ for $j \ne k$, so $\{e_k\}_{k\ge 0}$ is an orthonormal subset of $L^2_a(\mathbf{D})$. It is maximal: if $g \in L^2_a(\mathbf{D})$ has $\langle g, e_k \rangle = 0$ for all k, then the coefficient formula forces every Taylor coefficient of g to vanish, so $g = 0$. Hence by 8.74 it is an orthonormal basis.
(b) With $f(z) = \sum_k a_k z^k$, the coefficient formula gives

$$\langle f, e_k \rangle = \sqrt{\frac{k+1}{\pi}}\, \langle f, z^k \rangle = \sqrt{\frac{k+1}{\pi}} \cdot \frac{\pi a_k}{k+1} = a_k \sqrt{\frac{\pi}{k+1}}.$$

Hence Parseval's identity [8.63(c)] yields

$$\|f\|^2 = \sum_{k=0}^{\infty} |\langle f, e_k \rangle|^2 = \pi \sum_{k=0}^{\infty} \frac{|a_k|^2}{k+1},$$

that is, $\|f\| = \sqrt{\pi}\big(\sum_{k\geq 0} \frac{|a_k|^2}{k+1}\big)^{1/2}$.
(c) Let $\varphi(f) = f(w)$, bounded on $L_a^2(\mathbf{D})$ by 8C.20(b). By 8.76 and 8.77 the representing vector is

$$\Gamma_w = \sum_{k=0}^{\infty} \overline{\varphi(e_k)}\, e_k = \sum_{k=0}^{\infty} \sqrt{\frac{k+1}{\pi}}\, \overline{w}^{\,k}\, e_k,$$

the series converging in norm, with n-th partial sum $S_n(z) = \sum_{k=0}^{n} \frac{k+1}{\pi} \overline{w}^{\,k} z^k$. Since $|\overline{w}z| \leq |w| < 1$ on $\mathbf{D}$, the identity $\sum_{k\geq 0}(k+1)u^k = (1-u)^{-2}$ shows $S_n(z) \to \frac{1}{\pi(1-\overline{w}z)^2}$ pointwise on $\mathbf{D}$. But $\|S_n - \Gamma_w\| \to 0$, and norm convergence forces pointwise convergence by the pointwise bound of 8C.20 (take $r = \frac{1-|z|}{2}$). The two limits agree, so

$$\Gamma_w(z) = \frac{1}{\pi(1-\overline{w}z)^2} \qquad (z \in \mathbf{D}).$$

Problem 8C.22. Suppose G is the annulus defined by

$$G = \{z \in \mathbb{C} : 1 < |z| < 2\}.$$

(a) Find an orthonormal basis of $L_a^2(G)$.
(b) Suppose $f \in L_a^2(G)$ has Laurent series

$$f(z) = \sum_{k=-\infty}^{\infty} a_k z^k$$

for $z \in G$. Find a formula for $\|f\|$ in terms of $\dots, a_{-1}, a_0, a_1, \dots$.

Solution. (a) The normalized powers $e_k = u_k/\|u_k\|$, where $u_k(z) = z^k$ for $k \in \mathbb{Z}$ and

$$\|u_k\|^2 = 2\pi \int_1^2 \rho^{2k+1}\, d\rho = \begin{cases} \dfrac{\pi\,(4^{\,k+1}-1)}{k+1} & \text{if } k \neq -1, \\ 2\pi \ln 2 & \text{if } k = -1, \end{cases}$$

a finite positive number in each case. Throughout, polar coordinates turn an integral over an annulus $\{r_1 < |z| < r_2\}$ into $\int_{r_1}^{r_2}\int_0^{2\pi} h(\rho e^{i\theta})\rho\, d\theta\, d\rho$, valid for $h \geq 0$ by Tonelli (5.28) and for $h \in L^1$ by Fubini (5.32).
Orthogonality is immediate from $\int_0^{2\pi} e^{i(j-k)\theta} d\theta = 0$ for $j \neq k$, the product $u_j\overline{u_k}$ being bounded on G:

$$\langle u_j, u_k\rangle = \int_1^2\int_0^{2\pi} \rho^{\,j+k+1} e^{i(j-k)\theta}\, d\theta\, d\rho = 0 \qquad (j \neq k).$$

So $\{e_k\}_{k\in\mathbb{Z}}$ is orthonormal in $L_a^2(G)$, a Hilbert space by 8C.20.
For maximality, let $f \in L_a^2(G)$ have Laurent series $f(z) = \sum_m a_m z^m$ on G. The Laurent series converges uniformly on each circle $|z| = \rho$, and the circle has finite measure, so term-by-term integration against $e^{-ik\theta}$ gives

$$\int_0^{2\pi} f(\rho e^{i\theta})\, e^{-ik\theta}\, d\theta = 2\pi a_k \rho^{\,k} \qquad (1 < \rho < 2).$$

Since $f\overline{u_k} \in L^1(\lambda_2|_G)$ by Cauchy-Schwarz (8.11), the polar formula and this identity give

$$\langle f, u_k\rangle = \int_1^2 \rho^{\,k+1} \cdot 2\pi a_k \rho^{\,k}\, d\rho = a_k\|u_k\|^2.$$

Hence $\langle f, e_k\rangle = 0$ for all k forces every $a_k = 0$, i.e. $f = 0$; by 8.74 the family is an orthonormal basis.
(b) The last display gives $\langle f, e_k\rangle = a_k\|u_k\|$, so Parseval [8.63(c)] yields

$$\|f\| = \left(2\pi(\ln 2)\,|a_{-1}|^2 \;+\; \pi \sum_{k\neq -1} \frac{4^{\,k+1}-1}{k+1}\,|a_k|^2\right)^{1/2}.$$

Problem 8C.23. Prove that if $f \in L^2_a(\mathbb{D} \setminus \{0\})$, then f has a removable singularity at 0 (meaning that f can be extended to a function that is analytic on $\mathbb{D}$).

Solution. Square integrability kills the principal part. Let $f \in L^2_a(\mathbb{D} \setminus \{0\})$ have Laurent series $f(z) = \sum_{k\in\mathbb{Z}} a_k z^k$ on the punctured disk, converging uniformly on each circle $|z| = \rho$ with $0 < \rho < 1$. As in 8C.22, uniform convergence on the finite-measure circle and orthonormality of $\{e^{ik\theta}/\sqrt{2\pi}\}_{k\in\mathbb{Z}}$ in $L^2((0, 2\pi])$ give, via 8.54,

$$\int_0^{2\pi} |f(\rho e^{i\theta})|^2\, d\theta = 2\pi \sum_{k=-\infty}^{\infty} |a_k|^2 \rho^{2k}.$$

(The coefficients are square summable for each fixed ρ because $\sum_k |a_k|\rho^k < \infty$ makes $|a_k|\rho^k$ bounded.) Polar coordinates and Tonelli (5.28), applied twice to nonnegative terms, then give the identity in $[0, \infty]$

$$\int_{\mathbb{D}\setminus\{0\}} |f|^2\, d\lambda_2 = 2\pi \sum_{k=-\infty}^{\infty} |a_k|^2 \int_0^1 \rho^{2k+1}\, d\rho.$$

The left side is finite by hypothesis. For $k \le -1$ we have $2k + 1 \le -1$, so $\int_0^1 \rho^{2k+1} d\rho = \infty$ and the k-th term is ∞ unless $a_k = 0$. Hence $a_k = 0$ for all $k \le -1$, so

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \qquad (0 < |z| < 1),$$

a power series converging on all of $\mathbb{D}$. Its sum is analytic on $\mathbb{D}$ and agrees with f off 0, so f has a removable singularity at 0.

Problem 8C.24. The Dirichlet space $\mathcal{D}$ is defined to be the set of analytic functions $f\colon \mathbb{D} \to \mathbb{C}$ such that

$$\int_{\mathbb{D}} |f'|^2\, d\lambda_2 < \infty.$$

For $f, g \in \mathcal{D}$, define $\langle f, g\rangle$ to be $f(0)\overline{g(0)} + \int_{\mathbb{D}} f'\, \overline{g'}\, d\lambda_2$.
(a) Show that $\mathcal{D}$ is a Hilbert space.
(b) Show that if $w \in \mathbb{D}$, then $f \mapsto f(w)$ is a bounded linear functional on $\mathcal{D}$.
(c) Find an orthonormal basis of $\mathcal{D}$.
(d) Suppose $f \in \mathcal{D}$ has Taylor series

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

for $z \in \mathbb{D}$. Find a formula for $\|f\|$ in terms of $a_0, a_1, a_2, \ldots$.
(e) Suppose $w \in \mathbb{D}$. Find an explicit formula for $\Gamma_w \in \mathcal{D}$ such that

$$f(w) = \langle f, \Gamma_w\rangle \quad \text{for all } f \in \mathcal{D}.$$

Solution. Everything comes from one coefficient identity: for f analytic on $\mathbb{D}$ with $f(z) = \sum_{k\ge0} a_k z^k$,

$$\int_{\mathbb{D}} |f'|^2\, d\lambda_2 = \pi \sum_{k=1}^{\infty} k\, |a_k|^2 \qquad \text{in } [0, \infty],$$

so that $f \in \mathcal{D}$ exactly when $\sum_{k\ge1} k|a_k|^2 < \infty$, and then

$$\|f\|^2 = |a_0|^2 + \pi \sum_{k=1}^{\infty} k\, |a_k|^2,$$

which is the answer to (d). For the identity, fix $\rho \in (0, 1)$: the series $\sum_{k\ge1} k a_k \rho^{k-1} e^{i(k-1)\theta}$ converges uniformly in θ, hence in $L^2((0, 2\pi])$, to $\theta \mapsto f'(\rho e^{i\theta})$, and its coefficients against the orthonormal family

$\{e^{ij\theta}/\sqrt{2\pi}\}_{j\in\mathbb{Z}}$ are square summable (absolute convergence at ρ makes $k|a_k|\rho^{k-1}$ a bounded summable family), so 8.54 gives

$$\int_0^{2\pi} |f'(\rho e^{i\theta})|^2\, d\theta = 2\pi \sum_{k=1}^{\infty} k^2 |a_k|^2 \rho^{2k-2};$$

polar coordinates and Tonelli (5.28) on the nonnegative terms then give

$$\int_{\mathbb{D}} |f'|^2\, d\lambda_2 = 2\pi \sum_{k=1}^{\infty} k^2 |a_k|^2 \cdot \frac{1}{2k} = \pi \sum_{k=1}^{\infty} k|a_k|^2.$$

(a) $\mathcal{D}$ is a vector space (differentiation is linear and $L^2(\lambda_2|_{\mathbb{D}})$ is closed under sums), and $\langle\cdot,\cdot\rangle$ is an inner product: finiteness is Cauchy-Schwarz (8.11), and $\langle f, f\rangle = 0$ forces $f' \equiv 0$ by continuity, hence f constant on the connected set $\mathbb{D}$, hence $f = 0$ since $f(0) = 0$. For completeness, let $Tf = (a_0, \sqrt{\pi}\, a_1, \sqrt{2\pi}\, a_2, \ldots)$, read off the Taylor coefficients of f. The norm formula says $\|Tf\|_{\ell^2} = \|f\|$, so T is a linear isometry of $\mathcal{D}$ into ℓ^2. It is onto: given $c \in \ell^2$, set $a_0 = c_0$ and $a_k = c_k/\sqrt{\pi k}$; boundedness of (c_k) gives $\limsup |a_k|^{1/k} \le 1$, so $f(z) = \sum a_k z^k$ is analytic on $\mathbb{D}$, and the identity gives $\int_{\mathbb{D}} |f'|^2 = \sum_{k\ge1} |c_k|^2 < \infty$, so $f \in \mathcal{D}$ with $Tf = c$. A Cauchy sequence in $\mathcal{D}$ is carried to a Cauchy sequence in ℓ^2, which converges (8.22); pulling the limit back by T^{-1} shows $\mathcal{D}$ is complete, hence a Hilbert space.

(b) Fix $w \in \mathbb{D}$. For $f \in \mathcal{D}$, splitting off a_0 and applying Cauchy-Schwarz (8.11) to $\sqrt{\pi k}\,|a_k|$ against $|w|^k/\sqrt{\pi k}$,

$$|f(w)| \le |a_0| + \Big(\pi \sum_{k\ge1} k|a_k|^2\Big)^{1/2} \Big(\frac{1}{\pi} \sum_{k\ge1} \frac{|w|^{2k}}{k}\Big)^{1/2}.$$

Since $\sum_{k\ge1} t^k/k = \ln\frac{1}{1-t}$, the last factor is $t_w = \big(\frac{1}{\pi}\ln\frac{1}{1-|w|^2}\big)^{1/2} < \infty$. Writing $A = |a_0|$ and $B = (\pi\sum_{k\ge1} k|a_k|^2)^{1/2}$, so $\|f\|^2 = A^2 + B^2$, Cauchy-Schwarz in $\mathbb{R}^2$ gives

$$|f(w)| \le A + Bt_w \le \|f\| \left(1 + \frac{1}{\pi} \ln\frac{1}{1-|w|^2}\right)^{1/2},$$

so evaluation at w is a bounded linear functional.

(c) $e_0(z) = 1$ and $e_k(z) = z^k/\sqrt{\pi k}$ for $k \ge 1$. These are polynomials, and T carries them to the standard basis of ℓ^2, an orthonormal basis by 8C.10(a); a linear isometry onto preserves inner products, so $\{e_k\}_{k\ge0}$ is an orthonormal basis of $\mathcal{D}$.

(e) By 8.76 and 8.77 the representing vector for $\varphi(f) = f(w)$ is $\Gamma_w = \sum_k \overline{\varphi(e_k)}\, e_k$, and $\varphi(e_0) = 1$, $\varphi(e_k) = w^k/\sqrt{\pi k}$, so

$$\Gamma_w(z) = 1 + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{(\overline{w}z)^k}{k} = 1 + \frac{1}{\pi} \operatorname{Log}\frac{1}{1-\overline{w}z},$$

the principal branch being legitimate because $1 - \overline{w}z$ lies in the open right half-plane for $z \in \mathbb{D}$. Its coefficients $b_k = \overline{w}^k/(\pi k)$ satisfy $\pi\sum_{k\ge1} k|b_k|^2 = \frac{1}{\pi}\ln\frac{1}{1-|w|^2} < \infty$, so indeed $\Gamma_w \in \mathcal{D}$.

Problem 8C.25. (a) Prove that the Dirichlet space $\mathcal{D}$ is contained in the Bergman space $L^2_a(\mathbb{D})$.
(b) Prove that there exists a function $f \in L^2_a(\mathbb{D})$ such that f is uniformly continuous on $\mathbb{D}$ and $f \notin \mathcal{D}$.

Solution. Both parts read off the two coefficient criteria: for f analytic on $\mathbb{D}$ with $f(z) = \sum_{k\ge0} a_k z^k$,

$$f \in L^2_a(\mathbb{D}) \iff \sum_{k=0}^{\infty} \frac{|a_k|^2}{k+1} < \infty, \qquad f \in \mathcal{D} \iff \sum_{k=1}^{\infty} k\,|a_k|^2 < \infty,$$

the first from 8C.21(b) and the second from 8C.24, both being identities in $[0,\infty]$ ($\int_{\mathbb{D}} |f|^2 = \pi\sum_k \frac{|a_k|^2}{k+1}$ and $\int_{\mathbb{D}} |f'|^2 = \pi\sum_{k\ge1} k|a_k|^2$).

(a) For $k \geq 1$ we have $\frac{1}{k+1} \leq 1 \leq k$, so

$$\sum_{k=0}^{\infty} \frac{|a_k|^2}{k+1} \leq |a_0|^2 + \sum_{k=1}^{\infty} k\,|a_k|^2 < \infty$$

whenever $f \in \mathcal{D}$. Hence $\mathcal{D} \subseteq L_a^2(\mathbb{D})$.
(b) Take the lacunary series

$$f(z) = \sum_{k=1}^{\infty} \frac{z^{2^k}}{2^{k/2}} \qquad (z \in \mathbb{D}),$$

so $a_n = 2^{-k/2}$ when $n = 2^k$ with $k \geq 1$, and $a_n = 0$ otherwise. Since $\sum_{k\geq 1} 2^{-k/2} < \infty$, the Weierstrass M-test makes the series converge uniformly on the compact set $\overline{\mathbb{D}}$; so f is analytic on $\mathbb{D}$, extends continuously to $\overline{\mathbb{D}}$, and is therefore uniformly continuous on $\mathbb{D}$ and bounded, whence $f \in L_a^2(\mathbb{D})$ because $\lambda_2(\mathbb{D}) = \pi < \infty$. But

$$\int_{\mathbb{D}} |f'|^2\, d\lambda_2 = \pi \sum_{n=1}^{\infty} n\,|a_n|^2 = \pi \sum_{k=1}^{\infty} 2^k \cdot 2^{-k} = \infty,$$

so $f \notin \mathcal{D}$.

9 Real and Complex Measures

Contents

9.1 Exercises 9A

Problem 9A.1. Prove or give a counterexample: If ν is a real measure on a measurable space $(X, \mathcal{S})$ and $A, B \in \mathcal{S}$ are such that $\nu(A) \geq 0$ and $\nu(B) \geq 0$, then $\nu(A \cup B) \geq 0$.

Solution. False. Take $X = \{1, 2, 3\}$ with $\mathcal{S} = 2^X$ and let ν be the weighted counting measure with weights $3, -2, -2$ at $1, 2, 3$:

$$\nu(E) = 3\,\chi_E(1) - 2\,\chi_E(2) - 2\,\chi_E(3).$$

This is a real measure (countable additivity is finite additivity here, since a disjoint sequence in 2^X has at most three nonempty terms). With $A = \{1, 2\}$ and $B = \{1, 3\}$,

$$\nu(A) = \nu(B) = 1 \geq 0, \qquad \nu(A \cup B) = \nu(X) = -1 < 0.$$

Problem 9A.2. Suppose ν is a real measure on $(X, \mathcal{S})$. Define $\mu : \mathcal{S} \to [0, \infty)$ by

$$\mu(E) = |\nu(E)|.$$

Prove that μ is a (positive) measure on $(X, \mathcal{S})$ if and only if the range of ν is contained in $[0, \infty)$ or the range of ν is contained in $(-\infty, 0]$.

Solution. If the range of ν lies in $[0, \infty)$ then $\mu = \nu$, and if it lies in $(-\infty, 0]$ then $\mu = -\nu$; either way μ takes values in $[0, \infty)$, satisfies $\mu(\varnothing) = 0$ by 9.3(a), and is countably additive, so μ is a positive measure.
Conversely, suppose μ is a positive measure. For disjoint $E, F \in \mathcal{S}$, additivity of μ and of ν gives $|\nu(E) + \nu(F)| = |\nu(E)| + |\nu(F)|$; squaring both sides (all quantities real) leaves

$$\nu(E)\,\nu(F) = |\nu(E)|\,|\nu(F)| \geq 0,$$

so ν never takes a strictly positive and a strictly negative value on two disjoint sets. Suppose now, for a contradiction, that $\nu(A) > 0$ and $\nu(B) < 0$ for some $A, B \in \mathcal{S}$. The three sets $A \cap B$, $A \setminus B$, $B \setminus A$ are pairwise disjoint, so by the previous sentence their ν-values are all ≥ 0 or all ≤ 0. By finite additivity (9.3(a) plus countable additivity),

$$\nu(A) = \nu(A \cap B) + \nu(A \setminus B), \qquad \nu(B) = \nu(A \cap B) + \nu(B \setminus A).$$

(i) All three ≥ 0: then $\nu(B) \geq 0$, contradiction. (ii) All three ≤ 0: then $\nu(A) \leq 0$, contradiction. Hence the range of ν lies in $[0, \infty)$ or in $(-\infty, 0]$.

Problem 9A.3. Suppose ν is a complex measure on a measurable space $(X, \mathcal{S})$. Prove that $|\nu|(X) = \nu(X)$ if and only if ν is a (positive) measure.

Solution. Suppose first that ν is a positive measure, so $\nu(E) \in [0, \infty)$ for every $E \in \mathcal{S}$. Taking $n = 1$, $E_1 = X$ in the definition 9.8 of $|\nu|$ gives $|\nu|(X) \geq \nu(X)$; conversely, for disjoint $E_1, \dots, E_n \in \mathcal{S}$ contained in X, positivity and monotonicity give

$$\sum_{k=1}^{n} |\nu(E_k)| = \nu(E_1 \cup \cdots \cup E_n) \leq \nu(X),$$

and taking the supremum yields $|\nu|(X) \leq \nu(X)$. Hence $|\nu|(X) = \nu(X)$.
Conversely, suppose $|\nu|(X) = \nu(X)$; in particular $\nu(X) \in [0, \infty)$. For $E \in \mathcal{S}$, the pair E, $X \setminus E$ is admissible in 9.8, so additivity of ν and $\operatorname{Re} z \leq |z|$ give

$$|\nu(E)| + |\nu(X \setminus E)| \leq \nu(X) = \operatorname{Re}\nu(E) + \operatorname{Re}\nu(X \setminus E) \leq |\nu(E)| + |\nu(X \setminus E)|,$$

where taking real parts is legitimate because the outer terms are real. Equality throughout forces it termwise, so $\operatorname{Re}\nu(E) = |\nu(E)|$, which for a complex number means $\nu(E) \in [0, \infty)$. As E was arbitrary and ν is countably additive with $\nu(\varnothing) = 0$ by 9.3(a), ν is a (finite) positive measure.

Problem 9A.4. Suppose ν is a complex measure on a measurable space $(X, \mathcal{S})$. Prove that if $E \in \mathcal{S}$ then

$$|\nu|(E) = \sup\Big\{\sum_{k=1}^{\infty} |\nu(E_k)| : E_1, E_2, \ldots \text{ is a disjoint sequence in } \mathcal{S} \text{ such that } E = \bigcup_{k=1}^{\infty} E_k\Big\}.$$

Solution. Write s for the supremum on the right; each sum in it converges absolutely by 9.3(b), and the sequence $E, \varnothing, \varnothing, \ldots$ shows the set is nonempty, so $s \in [0, \infty]$.

$s \leq |\nu|(E)$: if $E_1, E_2, \ldots$ is a disjoint sequence in $\mathcal{S}$ with union E, then for each n the sets $E_1, \ldots, E_n$ are disjoint with union contained in E, so $\sum_{k=1}^{n} |\nu(E_k)| \leq |\nu|(E)$ by 9.8; let $n \to \infty$ and take the supremum.

$|\nu|(E) \leq s$: given disjoint $E_1, \ldots, E_n \in \mathcal{S}$ with $E_1 \cup \cdots \cup E_n \subseteq E$, pad them into a disjoint sequence with union E by setting $F_k = E_k$ for $k \leq n$, $F_{n+1} = E \setminus (E_1 \cup \cdots \cup E_n)$, and $F_k = \varnothing$ for $k > n + 1$. Then

$$\sum_{k=1}^{n} |\nu(E_k)| \leq \sum_{k=1}^{\infty} |\nu(F_k)| \leq s,$$

since $\nu(\varnothing) = 0$ by 9.3(a). Taking the supremum over such choices gives $|\nu|(E) \leq s$ by 9.8.

Problem 9A.5. Suppose μ is a (positive) measure on a measurable space $(X, \mathcal{S})$ and h is a nonnegative function in $\mathcal{L}^1(\mu)$. Let ν be the (positive) measure on $(X, \mathcal{S})$ defined by $d\nu = h\,d\mu$. Prove that

$$\int f\,d\nu = \int fh\,d\mu$$

for all $\mathcal{S}$-measurable functions $f : X \to [0, \infty]$.

Solution. Both sides agree on characteristic functions and then propagate up the standard approximation ladder.

(i) $f = \chi_A$ with $A \in \mathcal{S}$: by 3.15 and the definition of ν,

$$\int \chi_A\,d\nu = \nu(A) = \int_A h\,d\mu = \int \chi_A h\,d\mu.$$

(ii) $g = \sum_{k=1}^{n} c_k \chi_{A_k}$ simple and nonnegative, with c_k the distinct values of g and $A_k = g^{-1}(\{c_k\})$ disjoint: 3.15 with respect to ν, and additivity (3.16) plus homogeneity (3.20) of the μ-integral, give

$$\int g\,d\nu = \sum_{k=1}^{n} c_k\,\nu(A_k) = \sum_{k=1}^{n} c_k \int \chi_{A_k} h\,d\mu = \int gh\,d\mu.$$

(iii) General f: by 2.89 choose simple $\mathcal{S}$-measurable g_k with $|g_k| \leq |g_{k+1}| \leq |f| = f$ and $g_k \to f$ pointwise, and set $f_k = |g_k|$, so $0 \leq f_1 \leq f_2 \leq \cdots$ increases pointwise to f. Since $h \geq 0$, the products $f_k h$ increase pointwise to fh (with $0 \cdot \infty = 0$ when $h(x) = 0$). The Monotone Convergence Theorem (3.11), applied once for ν and once for μ, together with (ii), gives

$$\int f\,d\nu = \lim_{k\to\infty} \int f_k\,d\nu = \lim_{k\to\infty} \int f_k h\,d\mu = \int fh\,d\mu.$$

Problem 9A.6. Suppose $(X, \mathcal{S}, \mu)$ is a (positive) measure space. Prove that

$$\{h\, d\mu : h \in \mathcal{L}^1(\mu)\}$$

is a closed subspace of $M_{\mathbf{F}}(\mathcal{S})$.

Solution. Write $M = \{h\, d\mu : h \in \mathcal{L}^1(\mu)\}$, where $(h\, d\mu)(E) = \int_E h\, d\mu$ is in $M_{\mathbf{F}}(\mathcal{S})$ by 9.4. The whole exercise turns on the isometry

$$\|h\, d\mu - g\, d\mu\| = \|(h-g)\, d\mu\| = \|h-g\|_1 \qquad (h, g \in \mathcal{L}^1(\mu)),$$

which is 9.10 with $E = X$ (equivalently the second bullet of 9.16) applied to $h - g$.
M is a subspace: $\mathcal{L}^1(\mu)$ is a vector space (7.5 with $p = 1$), and linearity of integration (3.20, 3.21) with the definition 9.13 of the vector operations on measures gives $h\, d\mu + \alpha(g\, d\mu) = (h + \alpha g)\, d\mu \in M$.
M is closed: suppose $\nu_n = h_n\, d\mu \to \nu$ in $M_{\mathbf{F}}(\mathcal{S})$. A convergent sequence is Cauchy, so the isometry makes $\|h_j - h_k\|_1 \to 0$, i.e. the corresponding classes $\tilde{h}_n$ form a Cauchy sequence in $L^1(\mu)$ (7.17 identifies the two norms). Since $L^1(\mu)$ is a Banach space (7.24), $\tilde{h}_n \to \tilde{h}$ for some $h \in \mathcal{L}^1(\mu)$, whence $\|h_n - h\|_1 \to 0$ and, by the isometry again, $\|\nu_n - h\, d\mu\| \to 0$. Limits in a normed space are unique, so $\nu = h\, d\mu \in M$.

Problem 9A.7. (a) Suppose $\mathcal{B}$ is the collection of Borel subsets of $\mathbb{R}$. Show that the Banach space $M_{\mathbf{F}}(\mathcal{B})$ is not separable.
(b) Give an example of a measurable space $(X, \mathcal{S})$ such that the Banach space $M_{\mathbf{F}}(\mathcal{S})$ is infinite-dimensional and separable.

Solution. (a) The Dirac measures $\{\delta_t\}_{t \in \mathbb{R}}$ form an uncountable 2-separated set. Each δ_t (value 1 if $t \in E$, else 0) is countably additive, since t lies in at most one term of a disjoint sequence, so $\delta_t \in M_{\mathbf{F}}(\mathcal{B})$; and for $s \neq t$, taking $E_1 = \{s\}$, $E_2 = \{t\}$ in 9.8 gives

$$\|\delta_s - \delta_t\| \geq |(\delta_s - \delta_t)(\{s\})| + |(\delta_s - \delta_t)(\{t\})| = 2.$$

If D were a countable dense set, choosing $\nu_t \in D$ with $\|\nu_t - \delta_t\| < 1$ would give an injection $t \mapsto \nu_t$ of $\mathbb{R}$ into D (two equal choices would force $2 \leq \|\delta_s - \delta_t\| < 2$), contradicting countability. Hence $M_{\mathbf{F}}(\mathcal{B})$ is not separable.
(b) Take $X = \mathbb{Z}^+$ with $\mathcal{S} = 2^X$; then $T\nu = (\nu(\{1\}), \nu(\{2\}), \ldots)$ is a linear isometry of $M_{\mathbf{F}}(\mathcal{S})$ onto ℓ^1, which is infinite-dimensional and separable.
T maps into ℓ^1 by 9.3(b), and is linear by 9.13. Countable additivity gives $\nu(E) = \sum_{k \in E} \nu(\{k\})$ for every E, so T is injective; conversely each $a \in \ell^1$ defines $\nu(E) = \sum_{k \in E} a_k$, countably additive because an absolutely convergent series may be grouped and rearranged freely, so T is onto. For the isometry, disjointness of $E_1, \ldots, E_n$ gives

$$\sum_{j=1}^{n} |\nu(E_j)| \leq \sum_{j=1}^{n} \sum_{k \in E_j} |a_k| \leq \|a\|_1,$$

so $\|\nu\| \leq \|a\|_1$ by 9.8, while the choice $E_j = \{j\}$ for $j \leq n$ gives $\sum_{j \leq n} |a_j| \leq \|\nu\|$, hence $\|a\|_1 \leq \|\nu\|$.
Finally ℓ^1 is infinite-dimensional (the standard basis vectors are independent) and separable, a countable dense set being the finitely supported sequences with coordinates in $\mathbb{Q}$ (or $\mathbb{Q} + i\mathbb{Q}$): truncate a so the tail has ℓ^1-norm below $\varepsilon/2$, then approximate each of the remaining n coordinates within $\varepsilon/(2n)$. A surjective isometry transports both properties to $M_{\mathbf{F}}(\mathcal{S})$.

Problem 9A.8. Suppose $t > 0$ and λ is Lebesgue measure on the σ-algebra of Borel subsets of $[0, t]$.

Suppose $h : [0,t] \to \mathbb{C}$ is the function defined by

$$h(x) = \cos x + i \sin x.$$

Let ν be the complex measure defined by $d\nu = h\,d\lambda$.
(a) Show that $\|\nu\| = t$.
(b) Show that if $E_1, E_2, \ldots$ is a sequence of disjoint Borel subsets of $[0,t]$, then

$$\sum_{k=1}^{\infty} |\nu(E_k)| < t.$$

[This exercise shows that the supremum in the definition of $|\nu|([0,t])$ is not attained, even if countably many disjoint sets are allowed.]

Solution. (a) Here $h(x) = e^{ix}$ has $|h| = 1$, so 9.10 gives $|\nu|(E) = \int_E |h|\,d\lambda = \lambda(E)$ for every Borel $E \subseteq [0,t]$; in particular $\|\nu\| = \lambda([0,t]) = t$.
(b) Everything follows from the strict inequality: $|\nu(E)| < \lambda(E)$ whenever $\lambda(E) > 0$. To see it, pick α with $|\alpha| = 1$ and $\alpha\,\nu(E) = |\nu(E)|$; since the left side is real,

$$\lambda(E) - |\nu(E)| = \int_E \bigl(1 - \mathrm{Re}(\alpha e^{ix})\bigr)\,d\lambda(x),$$

whose integrand is nonnegative and vanishes only where $\alpha e^{ix} = 1$, i.e. on a countable (hence λ-null) set. An integral of a nonnegative function that vanishes is 0 a.e. by Markov's inequality (4.1) applied at each level $1/n$, so the integral above is strictly positive.
Also $|\nu(E)| \le \int_E |h|\,d\lambda = \lambda(E)$ for every Borel E. Now let $E_1, E_2, \ldots$ be disjoint Borel subsets of $[0,t]$, so $\sum_k \lambda(E_k) \le \lambda([0,t]) = t$.
(i) If $\lambda(E_k) = 0$ for all k, then every $\nu(E_k) = 0$ and $\sum_k |\nu(E_k)| = 0 < t$ (using $t > 0$).
(ii) Otherwise $\lambda(E_j) > 0$ for some j; putting $\varepsilon = \lambda(E_j) - |\nu(E_j)| > 0$,

$$\sum_{k=1}^{\infty} |\nu(E_k)| \le \sum_{k=1}^{\infty} \lambda(E_k) - \varepsilon \le t - \varepsilon < t,$$

the sums being finite.

Problem 9A.9. Give an example to show that 9.9 can fail if the hypothesis that ν is a real measure is replaced by the hypothesis that ν is a complex measure.

Solution. Take $X = \{1,2,3\}$ with $\mathcal{S} = 2^X$, μ counting measure, $\omega = e^{2\pi i/3}$, and $\nu = h\,d\mu$ where $h(1) = 1$, $h(2) = \omega$, $h(3) = \omega^2$; then 9.9 fails at $E = X$. Since X is finite, $h \in L^1(\mu)$ and ν is a complex measure by 9.4, with $\nu(E) = \sum_{k \in E} h(k)$.
By 9.10, $|\nu|(X) = \int_X |h|\,d\mu = 3$. But $1 + \omega + \omega^2 = 0$, so

$$\nu(\{1,2\}) = -\omega^2, \quad \nu(\{1,3\}) = -\omega, \quad \nu(\{2,3\}) = -1, \quad \nu(X) = 0,$$

and as $|\omega| = |\omega^2| = 1$ this gives $|\nu(E)| = 1$ for every E other than $\varnothing$ and X, where $|\nu(E)| = 0$. Hence

$$\sup\{|\nu(A)| + |\nu(B)| : A, B \in \mathcal{S} \text{ disjoint, } A \cup B \subseteq X\} = 2 < 3 = |\nu|(X),$$

the value 2 being attained at $A = \{1\}$, $B = \{2\}$.

Problem 9A.10. Suppose $(X, \mathcal{S})$ is a measurable space with $\mathcal{S} \neq \{\emptyset, X\}$. Prove that the total variation norm on $M_{\mathbb{F}}(\mathcal{S})$ does not come from an inner product. In other words, show that there does not exist an inner product $\langle\cdot,\cdot\rangle$ on $M_{\mathbb{F}}(\mathcal{S})$ such that $\|\nu\| = \langle\nu,\nu\rangle^{1/2}$ for all $\nu \in M_{\mathbb{F}}(\mathcal{S})$, where $\|\cdot\|$ is the usual total variation norm on $M_{\mathbb{F}}(\mathcal{S})$.

Solution. Two Dirac measures at points separated by $\mathcal{S}$ violate the parallelogram equality. Since $\mathcal{S} \neq \{\varnothing, X\}$ there is $E \in \mathcal{S}$ with $\varnothing \neq E \neq X$; pick $b \in E$ and $c \in X \setminus E$, and let $\nu = \delta_b$, $\eta = \delta_c$ be the corresponding Dirac measures, which are finite positive measures and so lie in $M_{\mathbb{F}}(\mathcal{S})$.
For a finite positive measure μ, the first bullet of 9.16 gives $\|\mu\| = \mu(X)$. Hence $\|\nu\| = \|\eta\| = 1$ and $\|\nu + \eta\| = (\delta_b + \delta_c)(X) = 2$. For the difference, any disjoint $E_1, \ldots, E_n \in \mathcal{S}$ satisfy

$$\sum_{k=1}^{n} |(\delta_b - \delta_c)(E_k)| \leq \delta_b(X) + \delta_c(X) = 2$$

by finite additivity and monotonicity, so $\|\nu - \eta\| \leq 2$ by 9.8; and E, $X \setminus E$ give values 1 and -1, so $\|\nu - \eta\| \geq 2$. Thus $\|\nu - \eta\| = 2$ and

$$\|\nu + \eta\|^2 + \|\nu - \eta\|^2 = 8 \neq 4 = 2\|\nu\|^2 + 2\|\eta\|^2.$$

A norm coming from an inner product satisfies the parallelogram equality (expand $\langle u \pm v, u \pm v\rangle$ and add), so the total variation norm does not.

Problem 9A.11. For $(X, \mathcal{S})$ a measurable space and $b \in X$, define a finite (positive) measure δ_b on $(X, \mathcal{S})$ by

$$\delta_b(E) = \begin{cases} 1 & \text{if } b \in E, \\ 0 & \text{if } b \notin E \end{cases}$$

for $E \in \mathcal{S}$.
(a) Show that if $b, c \in X$, then $\|\delta_b + \delta_c\| = 2$.
(b) Give an example of a measurable space $(X, \mathcal{S})$ and $b, c \in X$ with $b \neq c$ such that $\|\delta_b - \delta_c\| \neq 2$.

Solution. (a) $\delta_b + \delta_c$ is a finite positive measure, so the first bullet of 9.16 gives

$$\|\delta_b + \delta_c\| = (\delta_b + \delta_c)(X) = \delta_b(X) + \delta_c(X) = 2.$$

(b) Take $X = \{1, 2\}$ with the trivial σ-algebra $\mathcal{S} = \{\varnothing, X\}$, and $b = 1$, $c = 2$. Since $b, c \in X$, we have $\delta_b = \delta_c$ on $\mathcal{S}$, so $\delta_b - \delta_c$ is the zero measure and $\|\delta_b - \delta_c\| = 0 \neq 2$.

9.2 Exercises 9B

Problem 9B.1. Suppose ν is a real measure on a measurable space $(X, \mathcal{S})$. Prove that the Hahn decomposition of ν is almost unique, in the sense that if A, B and A', B' are pairs satisfying the Hahn Decomposition Theorem (9.23), then

$$|\nu|(A \setminus A') = |\nu|(A' \setminus A) = |\nu|(B \setminus B') = |\nu|(B' \setminus B) = 0.$$

Solution. The four sets are really two: $A \cup B = X$ and $A \cap B = \varnothing$ give $B = X \setminus A$, likewise $B' = X \setminus A'$, so

$$A \setminus A' = A \cap B' = B' \setminus B, \qquad A' \setminus A = A' \cap B = B \setminus B'.$$

If $E \in \mathcal{S}$ with $E \subseteq A \setminus A'$, then $E \subseteq A$ gives $\nu(E) \geq 0$ by 9.23(b) and $E \subseteq B'$ gives $\nu(E) \leq 0$ by 9.23(c), so $\nu(E) = 0$. Every finite disjoint family inside $A \setminus A'$ therefore contributes 0 to the supremum in 9.8, whence $|\nu|(A \setminus A') = 0$. Interchanging the roles of the two decompositions gives $|\nu|(A' \setminus A) = 0$, and the set identities above supply the remaining two equalities.

Problem 9B.2. Suppose μ is a (positive) measure and $g, h \in L^1(\mu)$. Prove that $g\,d\mu \perp h\,d\mu$ if and only if $g(x)h(x) = 0$ for almost every $x \in X$.

Solution. Write $\nu_g = g\,d\mu$ and $\nu_h = h\,d\mu$, and read almost every with respect to μ. The bridge in both directions is 9.10: $|\nu_g|(F) = \int_F |g|\,d\mu$, and this vanishes exactly when $g = 0$ a.e. on F (Markov's inequality (4.1) at each level $1/n$, then countable subadditivity).
Suppose $\nu_g \perp \nu_h$, and take A, B as in 9.28: disjoint with union X, $\nu_g(E) = \nu_g(E \cap A)$, $\nu_h(E) = \nu_h(E \cap B)$. Every measurable $F \subseteq B$ then has $\nu_g(F) = \nu_g(\varnothing) = 0$, so $|\nu_g|(B) = 0$ by the third bullet following 9.8, whence $\int_B |g|\,d\mu = 0$ and $g = 0$ a.e. on B. Symmetrically $h = 0$ a.e. on A. Off the union of these two null sets, $x \in A$ gives $h(x) = 0$ and $x \in B$ gives $g(x) = 0$; either way $g(x)h(x) = 0$.
Conversely, suppose $gh = 0$ a.e., and set $A = \{g \neq 0\}$, $B = \{g = 0\}$, disjoint sets in $\mathcal{S}$ with union X. On A minus the null set $\{gh \neq 0\}$ we have $h = 0$, so $\int_{E\cap A} h\,d\mu = 0$ and therefore

$$\nu_h(E) = \nu_h(E \cap A) + \nu_h(E \cap B) = \nu_h(E \cap B),$$

while $g = 0$ on B gives $\nu_g(E) = \nu_g(E \cap A)$, for every $E \in \mathcal{S}$. By 9.28, $\nu_g \perp \nu_h$.

Problem 9B.3. Suppose ν and μ are complex measures on a measurable space $(X, \mathcal{S})$. Show that the following are equivalent.
(a) $\nu \perp \mu$.
(b) $|\nu| \perp |\mu|$.
(c) $\operatorname{Re}\nu \perp \mu$ and $\operatorname{Im}\nu \perp \mu$.

Solution. The engine is a restatement of 9.28: for complex measures σ and a partition $X = A \cup B$ into disjoint measurable sets,

$$\sigma(E) = \sigma(E \cap A) \text{ for all } E \in \mathcal{S} \iff |\sigma|(B) = 0.$$

Indeed, the left side makes $\sigma(F) = \sigma(\varnothing) = 0$ for every measurable $F \subseteq B$, which is $|\sigma|(B) = 0$ by the third bullet following 9.8; conversely $|\sigma|(B) = 0$ forces $|\sigma(E \cap B)| \leq |\sigma|(B) = 0$ (first bullet following 9.8, with $|\sigma|$ a positive measure by 9.11), so $\sigma(E) = \sigma(E \cap A)$. The same equivalence holds for a finite positive measure σ, with $|\sigma|$ replaced by σ. Hence the *singularity criterion*: $\nu \perp \mu$ iff there are disjoint $A, B \in \mathcal{S}$ with $A \cup B = X$, $|\nu|(B) = 0$, $|\mu|(A) = 0$.
(a) $\iff$ (b). Since $|\nu|, |\mu|$ are finite positive measures (9.11, 9.17), the positive form of the equivalence says $|\nu| \perp |\mu|$ iff there are disjoint A, B with union X, $|\nu|(B) = 0$ and $|\mu|(A) = 0$ – verbatim the singularity criterion for $\nu \perp \mu$.
(a) $\implies$ (c). Taking real and imaginary parts of $\nu(E) = \nu(E \cap A)$ shows the same pair A, B from 9.28 witnesses $\operatorname{Re}\nu \perp \mu$ and $\operatorname{Im}\nu \perp \mu$.
(c) $\implies$ (a). Take pairs A_1, B_1 and A_2, B_2 from the criterion with $|\operatorname{Re}\nu|(B_1) = |\operatorname{Im}\nu|(B_2) = |\mu|(A_1) = |\mu|(A_2) = 0$, and set $A = A_1 \cup A_2$, $B = X \setminus A = B_1 \cap B_2$. Then $|\mu|(A) \leq |\mu|(A_1) + |\mu|(A_2) = 0$. For $|\nu|(B)$, any disjoint $E_1, \ldots, E_n \subseteq B$ satisfy $|z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|$ termwise, so by 9.8

$$\sum_{k=1}^{n} |\nu(E_k)| \leq |\operatorname{Re}\nu|(B) + |\operatorname{Im}\nu|(B) \leq |\operatorname{Re}\nu|(B_1) + |\operatorname{Im}\nu|(B_2) = 0,$$

using $B \subseteq B_1 \cap B_2$ and monotonicity of total variation measures. Hence $|\nu|(B) = 0$, and the criterion gives $\nu \perp \mu$.

Problem 9B.4. Suppose ν and μ are complex measures on a measurable space $(X, \mathcal{S})$. Prove that if $\nu \perp \mu$, then $|\nu + \mu| = |\nu| + |\mu|$ and $\|\nu + \mu\| = \|\nu\| + \|\mu\|$.

Solution. Take A, B as in 9.28: disjoint with union X, $\nu(E) = \nu(E \cap A)$ and $\mu(E) = \mu(E \cap B)$ for all $E \in \mathcal{S}$. Then every measurable $F \subseteq B$ has $\nu(F) = \nu(\varnothing) = 0$, so $|\nu|(B) = 0$ by the third bullet

following 9.8, and symmetrically $|\mu|(A) = 0$.
Consequently $\mu(F) = 0$ for every measurable $F \subseteq A$ (since $|\mu(F)| \leq |\mu|(A) = 0$), so $\nu + \mu$ and ν agree on all measurable subsets of A; as the supremum in 9.8 defining $|\cdot|(E \cap A)$ ranges only over disjoint families inside A, this gives $|\nu + \mu|(E \cap A) = |\nu|(E \cap A)$. Symmetrically $|\nu + \mu|(E \cap B) = |\mu|(E \cap B)$. Also $|\nu|(E \cap B) \leq |\nu|(B) = 0$ and $|\mu|(E \cap A) \leq |\mu|(A) = 0$. Hence, by additivity of the positive measures $|\nu|, |\mu|, |\nu + \mu|$ (9.11),

$$\begin{aligned} |\nu + \mu|(E) &= |\nu + \mu|(E \cap A) + |\nu + \mu|(E \cap B) \\ &= |\nu|(E \cap A) + |\mu|(E \cap B) = |\nu|(E) + |\mu|(E). \end{aligned}$$

Taking $E = X$ and using the definition 9.15 of the norm gives $\|\nu + \mu\| = \|\nu\| + \|\mu\|$.

Problem 9B.5. Suppose ν and μ are finite (positive) measures on a measurable space $(X, \mathcal{S})$. Prove that $\nu \perp \mu$ if and only if $\|\nu - \mu\| = \|\nu\| + \|\mu\|$.

Solution. Throughout, $\|\sigma\| = \sigma(X)$ for a finite positive measure σ (first bullet of 9.16), and $\sigma = \nu - \mu$ is a real measure.
Suppose $\nu \perp \mu$, with A, B as in 9.28. Taking $E = B$ and $E = A$ in the two identities of 9.28 gives $\nu(B) = 0$ and $\mu(A) = 0$, so $\nu(A) = \nu(X)$ and $\mu(B) = \mu(X)$ by finiteness. The disjoint pair A, B in 9.8 then gives

$$\|\sigma\| \geq |\sigma(A)| + |\sigma(B)| = \nu(X) + \mu(X) = \|\nu\| + \|\mu\|,$$

while for any disjoint $E_1, \ldots, E_n \in \mathcal{S}$, two applications of 9.8 give $\sum_k |\sigma(E_k)| \leq |\nu|(X) + |\mu|(X) = \|\nu\| + \|\mu\|$, so $\|\sigma\| \leq \|\nu\| + \|\mu\|$. Hence $\|\nu - \mu\| = \|\nu\| + \|\mu\|$.
Conversely, suppose $\|\nu - \mu\| = \|\nu\| + \|\mu\|$, and take a Hahn decomposition A, B of σ (9.23). With $\sigma^+(E) = \sigma(E \cap A)$ and $\sigma^-(E) = -\sigma(E \cap B)$ as in 9.30, we have $|\sigma| = \sigma^+ + \sigma^-$ by 9.31, so

$$\|\sigma\| = \sigma(A) - \sigma(B) = \nu(A) - \nu(B) - \mu(A) + \mu(B),$$

whereas $\|\nu\| + \|\mu\| = \nu(A) + \nu(B) + \mu(A) + \mu(B)$. Equating and cancelling the finite terms $\nu(A), \mu(B)$ gives $2(\nu(B) + \mu(A)) = 0$, and positivity forces

$$\nu(B) = 0 \qquad \text{and} \qquad \mu(A) = 0.$$

Monotonicity then gives $\nu(E) = \nu(E \cap A)$ and $\mu(E) = \mu(E \cap B)$ for every $E \in \mathcal{S}$, so $\nu \perp \mu$ by 9.28.

Problem 9B.6. Suppose μ is a complex or positive measure on a measurable space $(X, \mathcal{S})$. Prove that

$$\{\nu \in \mathcal{M}_{\mathbf{F}}(\mathcal{S}) : \nu \perp \mu\}$$

is a closed subspace of $\mathcal{M}_{\mathbf{F}}(\mathcal{S})$.

Solution. Write $V = \{\nu \in \mathcal{M}_{\mathbf{F}}(\mathcal{S}) : \nu \perp \mu\}$, and call a partition $X = A \cup B$ into disjoint measurable sets a *witness* for ν when $|\nu|(B) = 0$ and $\mu(F) = 0$ for every measurable $F \subseteq A$. Unwinding 9.28 by additivity (and the third bullet following 9.8 for the $|\nu|$ half), $\nu \perp \mu$ iff ν has a witness; this phrasing avoids $|\mu|$, which need not be finite when μ is a positive measure.
Subspace. The pair $\varnothing, X$ is a witness for 0. A witness for ν is a witness for $\alpha\nu$, since $|\alpha\nu| = |\alpha|\,|\nu|$ directly from 9.8. Given witnesses A_1, B_1 and A_2, B_2 for ν_1, ν_2, set $A = A_1 \cup A_2$ and $B = X \setminus A = B_1 \cap B_2$: a measurable $F \subseteq A$ splits as $(F \cap A_1) \cup (F \setminus A_1)$ with the pieces inside A_1 and A_2, so $\mu(F) = 0$; and $B \subseteq B_1 \cap B_2$ with 9.8 gives, for disjoint $E_1, \ldots, E_n$ inside B,

$$\sum_{k=1}^{n} |(\nu_1 + \nu_2)(E_k)| \leq |\nu_1|(B) + |\nu_2|(B) = 0,$$

so $|\nu_1 + \nu_2|(B) = 0$ and A, B is a witness for $\nu_1 + \nu_2$.

Closed. Suppose $\nu_k \in V$ with witnesses A_k, B_k and $\|\nu - \nu_k\| \to 0$. Put $A = \bigcup_k A_k$ and $B = X \setminus A = \bigcap_k B_k$. A measurable $F \subseteq A$ is the disjoint union of $F_1 = F \cap A_1$ and $F_k = (F \cap A_k) \setminus (A_1 \cup \cdots \cup A_{k-1})$, each inside A_k, so countable additivity gives $\mu(F) = \sum_k \mu(F_k) = 0$. And for each k, the subadditivity estimate above applied to $\nu = (\nu - \nu_k) + \nu_k$ gives

$$0 \leq |\nu|(B) \leq |\nu - \nu_k|(B) + |\nu_k|(B) \leq \|\nu - \nu_k\| \to 0,$$

using $B \subseteq B_k$. So A, B is a witness for ν and $\nu \in V$.

Problem 9B.7. Use the Cantor set to prove that there exists a (positive) measure ν on $(\mathbf{R}, \mathcal{B})$ such that $\nu \perp \lambda$ and $\nu(\mathbf{R}) \neq 0$ but $\nu(\{x\}) = 0$ for every $x \in \mathbf{R}$; here λ denotes Lebesgue measure on the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathbf{R}$.
[The second bullet point in Example 9.29 does not provide an example of the desired behavior because in that example, $\nu(\{r_k\}) \neq 0$ for all $k \in \mathbf{Z}^+$ with $w_k \neq 0$.]

Solution. Take $\nu(E) = \lambda(f^{-1}(E))$, the pushforward of Lebesgue measure on $[0,1]$ under the right inverse

$$f(y) = \inf\{x \in [0,1] : \Lambda(x) \geq y\} \qquad (y \in [0,1])$$

of the Cantor function Λ (2.77), where C is the Cantor set (2.74). The infimum is over a nonempty set since $\Lambda(1) = 1$.
(i) $\Lambda(f(y)) = y$. Continuity of Λ (2.79) makes $S_y = \Lambda^{-1}([y,\infty))$ closed, so $f(y) \in S_y$ and $\Lambda(f(y)) \geq y$. If $f(y) = 0$ then $y \leq \Lambda(0) = 0$, so both sides are 0; if $f(y) > 0$ then $\Lambda(x) < y$ for all $x < f(y)$, and letting $x \uparrow f(y)$ gives $\Lambda(f(y)) \leq y$.
(ii) f is increasing, hence Borel measurable by 2.43, and injective: $y \leq y'$ gives $S_{y'} \subseteq S_y$ hence $f(y) \leq f(y')$, and $f(y) = f(y')$ gives $y = y'$ by (i).
(iii) $f(y) \in C$. Both 0 and 1 lie in C, since every interval removed in 2.74 is an open middle third of some $[c,d]$, hence inside (c,d). So if $f(y) \notin C$ then $0 < f(y) < 1$, and (as noted in the proof of 2.79) Λ is constant, say $= c$, on an open interval $(a,b) \subseteq (0,1)$ containing $f(y)$. Then $c = \Lambda(f(y)) \geq y$ by (i), while any $x \in (a, f(y))$ lies outside S_y, giving $c = \Lambda(x) < y$ – a contradiction.
Countable additivity of ν follows from that of λ applied to the disjoint sets $f^{-1}(E_k)$, so ν is a positive measure on $(\mathbf{R}, \mathcal{B})$, and $f^{-1}(\mathbf{R}) = [0,1]$ gives $\nu(\mathbf{R}) = 1 \neq 0$. By (iii), $\nu(\mathbf{R} \setminus C) = \lambda(\varnothing) = 0$, while $\lambda(C) = 0$ by 2.76(b); so the disjoint Borel pair $A = C$, $B = \mathbf{R} \setminus C$ satisfies $\nu(E) = \nu(E \cap A)$ and $\lambda(E) = \lambda(E \cap B)$ for every $E \in \mathcal{B}$, i.e. $\nu \perp \lambda$ by 9.28. Finally $f^{-1}(\{x\})$ has at most one point by (ii), so $\nu(\{x\}) = 0$ for every $x \in \mathbf{R}$.

Problem 9B.8. Suppose ν is a real measure on a measurable space $(X, \mathcal{S})$. Prove that

$$\nu^+(E) = \sup\{\nu(D) : D \in \mathcal{S} \text{ and } D \subseteq E\}$$

and

$$\nu^-(E) = -\inf\{\nu(D) : D \in \mathcal{S} \text{ and } D \subseteq E\}$$

for all $E \in \mathcal{S}$.

Solution. Both suprema are attained, at $D = E \cap A$ and $D = E \cap B$ respectively, where A, B is a Hahn decomposition of ν (9.23). The proof of 9.30 gives $\nu^+(E) = \nu(E \cap A)$ and $\nu^-(E) = -\nu(E \cap B)$, and the uniqueness in 9.30 makes these independent of the decomposition chosen.
Fix $E \in \mathcal{S}$ and let s and i be the supremum and infimum of $\{\nu(D) : D \in \mathcal{S},\ D \subseteq E\}$, a set containing $\nu(\varnothing) = 0$ by 9.3(a).
$\nu^+(E) = s$: the admissible set $D = E \cap A$ gives $s \geq \nu(E \cap A) = \nu^+(E)$, while any admissible D satisfies, using $\nu = \nu^+ - \nu^-$ (9.31) and monotonicity of the positive measure ν^+,

$$\nu(D) = \nu^+(D) - \nu^-(D) \leq \nu^+(D) \leq \nu^+(E),$$

so $s \leq \nu^+(E)$.

$\nu^-(E) = -i$: the admissible set $D = E \cap B$ gives $i \le \nu(E \cap B) = -\nu^-(E)$, while any admissible D satisfies $\nu(D) \ge -\nu^-(D) \ge -\nu^-(E)$, so $i \ge -\nu^-(E)$.

Problem 9B.9. Suppose μ is a (positive) finite measure on a measurable space $(X, \mathcal{S})$ and h is a nonnegative function in $L^1(\mu)$. Thus $h\,d\mu \ll d\mu$. Find a reasonable condition on h that is equivalent to the condition $d\mu \ll h\,d\mu$.

Solution. The condition is $\mu(Z) = 0$, where $Z = \{x \in X : h(x) = 0\}$; that is, $h > 0$ almost everywhere. Write $d\nu = h\,d\mu$, so $\nu(E) = \int_E h\,d\mu$.
Suppose $\mu(Z) = 0$, and suppose $E \in \mathcal{S}$ has $\nu(E) = 0$. With $E_n = \{x \in E : h(x) > \frac{1}{n}\}$ we have $\frac{1}{n}\chi_{E_n} \le h\chi_E$, so

$$\tfrac{1}{n}\,\mu(E_n) \le \int_E h\,d\mu = 0.$$

Since $h \ge 0$, the sets E_n increase to $E \setminus Z$, so 2.59 gives $\mu(E \setminus Z) = \lim_{n\to\infty} \mu(E_n) = 0$ and hence

$$\mu(E) = \mu(E \setminus Z) + \mu(E \cap Z) \le 0 + \mu(Z) = 0.$$

Thus $\mu \ll \nu$.
Conversely, if $\mu \ll \nu$ then $\nu(Z) = \int_Z h\,d\mu = 0$ forces $\mu(Z) = 0$.

Problem 9B.10. Suppose μ is a (positive) measure on a measurable space $(X, \mathcal{S})$ and ν is a complex measure on $(X, \mathcal{S})$. Show that the following are equivalent.
(a) $\nu \ll \mu$.
(b) $|\nu| \ll \mu$.
(c) $\operatorname{Re}\nu \ll \mu$ and $\operatorname{Im}\nu \ll \mu$.

Solution. Everything follows from the definition 9.8 of $|\nu|$ and the bound $|\nu(E)| \le |\nu|(E)$ recorded just after it. Fix $E \in \mathcal{S}$ with $\mu(E) = 0$.
(i) (a) $\Rightarrow$ (b). If $E_1, \dots, E_n \in \mathcal{S}$ are disjoint subsets of E, then $\mu(E_k) \le \mu(E) = 0$, so (a) gives $\nu(E_k) = 0$ and hence

$$|\nu(E_1)| + \cdots + |\nu(E_n)| = 0.$$

The supremum of the left side over all such families is $|\nu|(E)$, so $|\nu|(E) = 0$.
(ii) (b) $\Rightarrow$ (a). Now $|\nu(E)| \le |\nu|(E) = 0$, so $\nu(E) = 0$.
(iii) (a) $\Leftrightarrow$ (c). Because $(\operatorname{Re}\nu)(E) = \operatorname{Re}\big(\nu(E)\big)$ and $(\operatorname{Im}\nu)(E) = \operatorname{Im}\big(\nu(E)\big)$, the equation $\nu(E) = 0$ holds if and only if $(\operatorname{Re}\nu)(E) = (\operatorname{Im}\nu)(E) = 0$.

Problem 9B.11. Suppose μ is a (positive) measure on a measurable space $(X, \mathcal{S})$ and ν is a real measure on $(X, \mathcal{S})$. Show that $\nu \ll \mu$ if and only if $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.

Solution. Both directions read off the Jordan decomposition 9.30, which supplies finite (positive) measures with

$$\nu = \nu^+ - \nu^- \qquad \text{and} \qquad |\nu| = \nu^+ + \nu^-.$$

Fix $E \in \mathcal{S}$ with $\mu(E) = 0$.
If $\nu^+ \ll \mu$ and $\nu^- \ll \mu$, then $\nu(E) = \nu^+(E) - \nu^-(E) = 0$; thus $\nu \ll \mu$.
Conversely, if $\nu \ll \mu$ then $|\nu| \ll \mu$ by Exercise 9B.10, so

$$\nu^+(E) + \nu^-(E) = |\nu|(E) = 0.$$

Both summands lie in $[0, \infty)$, so both vanish; thus $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.

Problem 9B.12. Suppose μ is a (positive) measure on a measurable space $(X, \mathcal{S})$. Prove that

$$\{\nu \in \mathcal{M}_{\mathbf{F}}(\mathcal{S}) : \nu \ll \mu\}$$

is a closed subspace of $\mathcal{M}_{\mathbf{F}}(\mathcal{S})$.

Solution. Write V for the set in question and $\mathcal{N} = \{E \in \mathcal{S} : \mu(E) = 0\}$; by 9.32,

$$V = \bigcap_{E \in \mathcal{N}} \{\nu \in \mathcal{M}_{\mathbf{F}}(\mathcal{S}) : \nu(E) = 0\},$$

so it suffices to show that each evaluation $\nu \mapsto \nu(E)$ is linear and continuous.
Linearity is the definition 9.13 of the operations on $\mathcal{M}_{\mathbf{F}}(\mathcal{S})$:

$$(\nu_1 + \nu_2)(E) = \nu_1(E) + \nu_2(E), \qquad (\alpha\nu)(E) = \alpha\, \nu(E).$$

Hence each set in the intersection is a subspace, and so is V.
For continuity, let $\lambda \in \mathcal{M}_{\mathbf{F}}(\mathcal{S})$. The bound $|\lambda(E)| \le |\lambda|(E)$ noted after 9.8, together with monotonicity of the (positive) measure $|\lambda|$ (9.11), gives

$$|\lambda(E)| \le |\lambda|(E) \le |\lambda|(X) = \|\lambda\|.$$

So if $\nu_k \in V$ and $\|\nu - \nu_k\| \to 0$, then for $E \in \mathcal{N}$,

$$|\nu(E)| = \big|(\nu - \nu_k)(E)\big| \le \|\nu - \nu_k\| \to 0,$$

whence $\nu(E) = 0$ and $\nu \in V$. Thus each such set, and therefore V, is closed.

Problem 9B.13. Give an example to show that the Radon–Nikodym Theorem (9.36) can fail if the σ-finite hypothesis is eliminated.

Solution. Take $X = [0, 1]$ with $\mathcal{S} = \mathcal{B}_{[0,1]}$, let μ be counting measure, and let $\nu = \lambda|_{\mathcal{S}}$ be Lebesgue measure, a finite (hence complex) measure.
Here μ is not σ-finite: a set of finite counting measure is finite, and the uncountable set $[0, 1]$ is not a countable union of finite sets.
Also $\nu \ll \mu$, because $\mu(E) = 0$ forces $E = \varnothing$.
But no $h \in L^1(\mu)$ satisfies $d\nu = h\, d\mu$. Indeed, for each $x \in [0, 1]$ the function $h\chi_{\{x\}}$ equals the simple function $h(x)\chi_{\{x\}}$, so

$$h(x) = h(x)\, \mu(\{x\}) = \int_{\{x\}} h\, d\mu = \nu(\{x\}) = 0.$$

Thus h is identically 0, contradicting

$$1 = \nu([0, 1]) = \int_{[0,1]} h\, d\mu = 0.$$

Every hypothesis of 9.36 except σ-finiteness of μ holds, so that hypothesis cannot be eliminated.

Problem 9B.14. Suppose μ is a (positive) σ-finite measure on a measurable space $(X, \mathcal{S})$ and ν is a complex measure on $(X, \mathcal{S})$. Show that the following are equivalent.
(a) $\nu \ll \mu$.
(b) For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|\nu(E)| < \varepsilon$ for every set $E \in \mathcal{S}$ with $\mu(E) < \delta$.
(c) For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|\nu|(E) < \varepsilon$ for every set $E \in \mathcal{S}$ with $\mu(E) < \delta$.

Solution. We run the cycle (c) $\Rightarrow$ (b) $\Rightarrow$ (a) $\Rightarrow$ (c).
(i) (c) $\Rightarrow$ (b). The same δ works, since $|\nu(E)| \le |\nu|(E)$ by the bound after 9.8.
(ii) (b) $\Rightarrow$ (a). If $\mu(E) = 0$ then $\mu(E) < \delta$ for every δ, so $|\nu(E)| < \varepsilon$ for every $\varepsilon > 0$ and hence $\nu(E) = 0$.
(iii) (a) $\Rightarrow$ (c). Here $|\nu| \ll \mu$ by Exercise 9B.10, and $|\nu|$ is a finite (positive) measure by 9.11 and 9.17. If (c) fails, then for some $\varepsilon > 0$ and every $k \in \mathbf{Z}^+$ there is $E_k \in \mathcal{S}$ with

$$\mu(E_k) < 2^{-k} \qquad \text{and} \qquad |\nu|(E_k) \ge \varepsilon.$$

Put $F_n = \bigcup_{k=n}^{\infty} E_k$, a decreasing sequence, and $F = \bigcap_{n=1}^{\infty} F_n$. Countable subadditivity (2.58) and monotonicity give

$$\mu(F) \le \mu(F_n) \le \sum_{k=n}^{\infty} 2^{-k} = 2^{-(n-1)}$$

for every n, so $\mu(F) = 0$ and hence $|\nu|(F) = 0$. But $|\nu|(F_n) \ge |\nu|(E_n) \ge \varepsilon$, and $|\nu|$ is a finite (positive) measure, so 9.7(d) applied to $F_1 \supseteq F_2 \supseteq \cdots$ gives

$$|\nu|(F) = \lim_{n\to\infty} |\nu|(F_n) \ge \varepsilon > 0,$$

a contradiction.
Method (2) for (iii), using the σ-finiteness hypothesis. The Radon–Nikodym Theorem (9.36) gives $h \in L^1(\mu)$ with $d\nu = h\, d\mu$, and then $|\nu|(E) = \int_E |h|\, d\mu$ by 9.10. Given $\varepsilon > 0$, set $h_n = \min\{|h|, n\}$; since $h_n \uparrow |h|$ with $\int_X |h|\, d\mu < \infty$, the Monotone Convergence Theorem (3.11) supplies n with $\int_X (|h| - h_n)\, d\mu < \varepsilon/2$. Take $\delta = \varepsilon/(2n)$; then $\mu(E) < \delta$ forces

$$\begin{aligned}|\nu|(E) &\le \int_X \big(|h| - h_n\big) d\mu + n\,\mu(E)\\ &< \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Problem 9B.15. Prove 9.42 [with the extra hypothesis that μ is a σ-finite (positive) measure] in the case where $p = 1$.

Solution. For $h \in L^\infty(\mu)$ put $\varphi_h(f) = \int fh\, d\mu$; we show $h \mapsto \varphi_h$ is a one-to-one linear map of $L^\infty(\mu)$ onto $\big(L^1(\mu)\big)'$ with $\|\varphi_h\| = \|h\|_\infty$.
Boundedness is 7.9 with the exponent pair $1, \infty$: $|\varphi_h(f)| \le \|h\|_\infty \|f\|_1$, so $\|\varphi_h\| \le \|h\|_\infty$; linearity in h is linearity of the integral.
For the reverse inequality, let $0 \le c < \|h\|_\infty$. Then $A = \{|h| > c\}$ has $\mu(A) > 0$, so σ-finiteness and countable subadditivity give $B \subseteq A$ with $0 < \mu(B) < \infty$. Setting $f = \chi_B \overline{h}/|h|$ (well defined, as $|h| > c \ge 0$ on B) gives $|f| = \chi_B$, $\|f\|_1 = \mu(B)$, and $fh = \chi_B |h|$, so by 3.8

$$\varphi_h(f) = \int_B |h|\, d\mu \ge c\,\mu(B) = c\,\|f\|_1.$$

Hence $\|\varphi_h\| \ge c$, and letting $c \uparrow \|h\|_\infty$ gives $\|\varphi_h\| = \|h\|_\infty$. Injectivity follows: $\varphi_{h_1} = \varphi_{h_2}$ forces $\|h_1 - h_2\|_\infty = \|\varphi_{h_1 - h_2}\| = 0$.
Surjectivity, first for $\mu(X) < \infty$. Given $\varphi \in \big(L^1(\mu)\big)'$, set $\nu(E) = \varphi(\chi_E)$, legitimate since $\chi_E \in L^1(\mu)$. For disjoint $E_1, E_2, \ldots$ with union E,

$$\Big\|\chi_E - \sum_{k=1}^{n} \chi_{E_k}\Big\|_1 = \sum_{k=n+1}^{\infty} \mu(E_k) \longrightarrow 0$$

because $\sum_k \mu(E_k) = \mu(E) < \infty$; continuity of φ then gives $\nu(E) = \sum_{k=1}^{\infty} \nu(E_k)$, so ν is a complex measure. If $\mu(E) = 0$ then $\chi_E = 0$ in $L^1(\mu)$, so $\nu(E) = 0$; thus $\nu \ll \mu$, and since μ is finite the Radon–Nikodym Theorem (9.36) supplies $h \in L^1(\mu)$ with $d\nu = h\, d\mu$. Hence $\varphi(f) = \int fh\, d\mu$ for simple f, by linearity.

This extends to $f \in L^\infty(\mu)$: taking simple $f_k \to f$ uniformly (2.89, applied to the real and imaginary parts of a bounded representative of f),

$$\|f - f_k\|_1 \leq \mu(X) \sup_X |f - f_k| \to 0, \qquad \Big|\int (f - f_k) h \, d\mu\Big| \leq \|h\|_1 \sup_X |f - f_k| \to 0.$$

Next, $\|h\|_\infty \leq \|\varphi\|$. Otherwise $A_n = \{|h| \geq \|\varphi\| + \frac{1}{n}\}$ has $0 < \mu(A_n) \leq \mu(X) < \infty$ for some n, and $f = \chi_{A_n} \overline{h}/|h| \in L^\infty(\mu)$ with $\|f\|_1 = \mu(A_n)$ gives

$$\begin{aligned} \|\varphi\| \, \mu(A_n) \geq |\varphi(f)| &= \int_{A_n} |h| \, d\mu \\ &\geq \big(\|\varphi\| + \tfrac{1}{n}\big)\mu(A_n), \end{aligned}$$

impossible. So $h \in L^\infty(\mu)$, and φ and φ_h are continuous functionals on $L^1(\mu)$ agreeing on each $f_k = f\chi_{\{|f| \leq k\}}$, where $\|f - f_k\|_1 \to 0$ by dominated convergence (3.31). Thus $\varphi = \varphi_h$.
Now let μ be σ-finite, say $X_1 \subseteq X_2 \subseteq \cdots$ with union X and $\mu(X_k) < \infty$. Identifying $L^1(\mu_{X_k})$ with the functions in $L^1(\mu)$ vanishing off X_k, the finite case applied to $\varphi|_{L^1(\mu_{X_k})}$ (of norm at most $\|\varphi\|$) yields h_k with $\|h_k\|_\infty \leq \|\varphi\|$ and $\varphi(f) = \int_{X_k} f h_k \, d\mu$ there. For $j < k$ both h_j and $h_k|_{X_j}$ represent the same functional, so the injectivity above (on X_j) gives $h_j = h_k$ almost everywhere on X_j; patching, there is h with $h = h_k$ almost everywhere on X_k for each k, and $|h| \leq \|\varphi\|$ almost everywhere, so $h \in L^\infty(\mu)$. For $f \in L^1(\mu)$, dominated convergence (3.31) gives $\|f - f\chi_{X_k}\|_1 \to 0$ and hence

$$\varphi(f) = \lim_{k \to \infty} \int_{X_k} f h_k \, d\mu = \lim_{k \to \infty} \int f \chi_{X_k} h \, d\mu = \int f h \, d\mu,$$

the last step dominated by $\|h\|_\infty |f| \in L^1(\mu)$. Thus $\varphi = \varphi_h$.

Problem 9B.16. Explain where the proof of 9.42 fails if $p = \infty$.

Solution. The proof fails at its first move: for $p = \infty$ the set function $\nu(E) = \varphi(\chi_E)$ is only finitely additive.
Countable additivity of ν is deduced from continuity of φ applied to $\sum_{k=1}^\infty \chi_{E_k} = \chi_E$, which requires convergence of the partial sums in norm. For $p < \infty$ that holds, since

$$\Big\|\chi_E - \sum_{k=1}^n \chi_{E_k}\Big\|_p = \mu\Big(\bigcup_{k=n+1}^\infty E_k\Big)^{1/p} \longrightarrow 0,$$

but the $L^\infty(\mu)$-norm of a characteristic function is 1 whenever its set has positive measure, so that quantity equals 1 for every n in the typical case. Hence ν need not be a complex measure, the Radon–Nikodym Theorem (9.36) is unavailable, and with it $d\nu = h \, d\mu$ and every later step (9.43, 9.44).
A second breakdown occurs in the reduction to sets of finite measure, where one truncates a near-maximizer f_k to $D_k = \{|f_k| > 1/n_k\}$ and needs $\mu(D_k) < \infty$. The truncation itself survives $p = \infty$ (the discarded part has sup-norm at most $1/n_k$), but the bound does not: for $p < \infty$ it is Markov's inequality (4.1) applied to $|f_k|^p$, giving $\mu(D_k) \leq n_k^p \|f_k\|_p^p$, whereas membership in $L^\infty(\mu)$ bounds no measure at all, as $\|\chi_X\|_\infty = 1$ even when $\mu(X) = \infty$.
That defect also kills the closing approximation of an arbitrary element of $L^p(\mu)$ by functions living on σ-finite sets: for counting measure on an uncountable X, such a g vanishes off a countable set, so $\|\chi_X - g\|_\infty \geq 1$.
No repair is possible, since the conclusion itself is false for $p = \infty$: a Hahn–Banach extension (6.69) to $L^\infty(\lambda)$ of $f \mapsto f(0)$ on $C([0,1])$ has $\varphi(g_n) = 1$ for $g_n(x) = \max\{0, 1 - nx\}$, while $\int g_n h \, d\lambda \to 0$ for every $h \in L^1(\lambda)$ by dominated convergence (3.31).

Problem 9B.17. Prove that if μ is a (positive) measure and $1 < p < \infty$, then $L^p(\mu)$ is reflexive. [See the definition before Exercise 19 in Section 7B for the meaning of reflexive.]

Solution. Surjectivity of the canonical isometry Φ, where $(\Phi g)(\varphi) = \varphi(g)$ (Exercise 6D.20), comes from applying 9.42 to both p and its dual exponent p'.
Since $\frac{1}{p'} = 1 - \frac{1}{p} \in (0,1)$, we have $1 < p' < \infty$ with dual exponent p; so 9.42 applies to each exponent (its σ-finiteness hypothesis is needed only for exponent 1) and gives linear bijections

$$\Lambda : L^{p'}(\mu) \to \big(L^p(\mu)\big)', \qquad (\Lambda h)(f) = \int fh \, d\mu,$$

$$\Gamma : L^{p}(\mu) \to \big(L^{p'}(\mu)\big)', \qquad (\Gamma g)(h) = \int hg \, d\mu,$$

with $\|\Lambda h\| = \|h\|_{p'}$.
Let $\Psi \in L^p(\mu)''$. Then $\Psi \circ \Lambda$ is linear with

$$|(\Psi \circ \Lambda)(h)| \le \|\Psi\| \, \|\Lambda h\| = \|\Psi\| \, \|h\|_{p'},$$

so $\Psi \circ \Lambda \in \big(L^{p'}(\mu)\big)'$, and surjectivity of Γ yields $g \in L^p(\mu)$ with $\Psi(\Lambda h) = \int hg \, d\mu$ for all $h \in L^{p'}(\mu)$.
Given $\varphi \in \big(L^p(\mu)\big)'$, surjectivity of Λ provides h with $\varphi = \Lambda h$, and then

$$(\Phi g)(\varphi) = \varphi(g) = \int gh \, d\mu = \Psi(\Lambda h) = \Psi(\varphi).$$

Thus $\Phi g = \Psi$, so Φ is surjective and $L^p(\mu)$ is reflexive.

Problem 9B.18. Prove that $L^1(\mathbb{R})$ is not reflexive.

Solution. The element of $L^1(\mathbb{R})''$ outside the range of the canonical isometry Φ is $\Psi = \psi \circ \Lambda^{-1}$, where ψ extends evaluation at 0.
Lebesgue measure λ is σ-finite, so Exercise 9B.15 (that is, 9.42 with $p = 1$) makes

$$\Lambda : L^{\infty}(\mathbb{R}) \to \big(L^1(\mathbb{R})\big)', \qquad (\Lambda h)(f) = \int fh \, d\lambda,$$

a linear bijection with $\|\Lambda h\| = \|h\|_\infty$; hence Λ^{-1} is a norm-preserving linear bijection as well.
On the subspace $C_b(\mathbb{R})$ of $L^\infty(\mathbb{R})$ the essential supremum is the supremum, since $|h(x_0)| > c$ for continuous h forces $|h| > c$ on an open interval, a set of positive measure. So continuous functions agreeing almost everywhere agree everywhere, and $\psi_0(h) = h(0)$ is a well-defined linear functional on $C_b(\mathbb{R})$ with $|\psi_0(h)| \le \|h\|_\infty$. The Hahn–Banach Theorem (6.69) extends it to $\psi \in \big(L^\infty(\mathbb{R})\big)'$.
Then $|\Psi(\varphi)| \le \|\psi\| \, \|\Lambda^{-1}\varphi\|_\infty = \|\psi\| \, \|\varphi\|$, so $\Psi \in L^1(\mathbb{R})''$. Suppose $\Phi f = \Psi$ for some $f \in L^1(\mathbb{R})$. Taking $\varphi = \Lambda h$ gives

$$\int fh \, d\lambda = (\Phi f)(\Lambda h) = \Psi(\Lambda h) = \psi(h) \qquad \text{for every } h \in L^\infty(\mathbb{R}).$$

Apply this to $h_n(x) = \max\{0, 1 - n|x|\} \in C_b(\mathbb{R})$, so that $\psi(h_n) = h_n(0) = 1$. But $|fh_n| \le |f| \in L^1(\mathbb{R})$ and $h_n \to 0$ off the origin, so dominated convergence (3.31) gives

$$1 = \lim_{n\to\infty} \int fh_n \, d\lambda = 0,$$

a contradiction. Thus Φ is not surjective and $L^1(\mathbb{R})$ is not reflexive.

10 Linear Maps on Hilbert Spaces

Contents

10.1 Exercises 10A

Problem 10A.1. Define $T : \ell^2 \to \ell^2$ by $T(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$. Find a formula for T^*.

Solution. The answer is the left shift:

$$T^*(b_1, b_2, b_3, \ldots) = (b_2, b_3, b_4, \ldots).$$

Write S for this left shift, a bounded linear map since $\|Sb\| \le \|b\|$. Because the n^{th} coordinate of Ta is 0 for $n = 1$ and a_{n-1} for $n \ge 2$,

$$\begin{aligned}\langle Ta, b\rangle &= \sum_{n=2}^{\infty} a_{n-1}\overline{b_n} = \sum_{k=1}^{\infty} a_k\overline{b_{k+1}} \\ &= \sum_{k=1}^{\infty} a_k\overline{(Sb)_k} = \langle a, Sb\rangle\end{aligned}$$

for all $a, b \in \ell^2$, the reindexing being legitimate because the series converge absolutely by the Cauchy–Schwarz inequality (8.11). Since T^*b is the unique element of ℓ^2 with $\langle Ta, b\rangle = \langle a, T^*b\rangle$ for all a (10.1), we get $T^* = S$.

Problem 10A.2. Suppose V is a Hilbert space, U is a closed subspace of V, and $T : U \to V$ is defined by $Tf = f$. Describe the linear operator $T^* : V \to U$.

Solution. T^* is the orthogonal projection of V onto U: $T^*g = P_U g$ for every $g \in V$.
By 8.43 we may write $g = P_U g + (g - P_U g)$ with $P_U g \in U$ and $g - P_U g \in U^{\perp}$. Hence for $f \in U$,

$$\langle Tf, g\rangle = \langle f, g\rangle = \langle f, P_U g\rangle + \langle f, g - P_U g\rangle = \langle f, P_U g\rangle,$$

the last term vanishing by 8.37(a). Since $P_U g \in U$ and T^*g is the unique element of U with $\langle Tf, g\rangle = \langle f, T^*g\rangle$ for all $f \in U$ (10.1), we get $T^*g = P_U g$.

Problem 10A.3. Suppose V and W are Hilbert spaces and $g \in V$, $h \in W$. Define $T \in \mathcal{B}(V, W)$ by $Tf = \langle f, g\rangle h$. Find a formula for T^*.

Solution. The answer is $T^*u = \langle u, h\rangle g$ for $u \in W$.
Indeed, for $f \in V$ and $u \in W$, using homogeneity in the first slot, conjugate symmetry, and conjugate homogeneity in the second slot,

$$\begin{aligned}\langle Tf, u\rangle &= \langle f, g\rangle\langle h, u\rangle = \overline{\langle u, h\rangle}\,\langle f, g\rangle \\ &= \big\langle f, \langle u, h\rangle g\big\rangle.\end{aligned}$$

Since T^*u is the unique element of V with $\langle Tf, u\rangle = \langle f, T^*u\rangle$ for all $f \in V$ (10.1), the formula follows. As a sanity check with 10.13: if $g \ne 0$ and $h \ne 0$, then $\operatorname{range} T = \{\alpha h : \alpha \in \mathbf{F}\}$, and indeed $\operatorname{null} T^* = \{u \in W : \langle u, h\rangle = 0\} = \{h\}^{\perp} = (\operatorname{range} T)^{\perp}$, as 10.13(a) requires. Also $\|T\| = \|T^*\| = \|g\|\,\|h\|$, in agreement with 10.11 (take $f = g$ to see that the bound $\|g\|\,\|h\|$ above is attained).

Problem 10A.4. Suppose V and W are Hilbert spaces and $T \in \mathcal{B}(V, W)$ has finite-dimensional range. Prove that T^* also has finite-dimensional range.

Solution. With $n = \dim \operatorname{range} T$ (the case $n = 0$ being trivial), the range of T^* lies in the span of $T^*g_1, \ldots, T^*g_n$, where $g_1, \ldots, g_n$ is an orthonormal basis of $\operatorname{range} T$ obtained by Gram–Schmidt.
Indeed, let $u \in W$ and $f \in V$. Expanding $Tf = \sum_{j=1}^{n}\langle Tf, g_j\rangle g_j$ and using the defining property 10.2

of the adjoint together with $\langle g_j, u\rangle = \overline{\langle u, g_j\rangle}$,

$$\begin{aligned}\langle f, T^*u\rangle = \langle Tf, u\rangle &= \sum_{j=1}^{n} \langle Tf, g_j\rangle \langle g_j, u\rangle \\ &= \sum_{j=1}^{n} \langle f, T^*g_j\rangle\, \overline{\langle u, g_j\rangle} = \Big\langle f, \sum_{j=1}^{n} \langle u, g_j\rangle T^*g_j \Big\rangle.\end{aligned}$$

Taking f to be the difference of the two second slots shows that difference has norm 0, so

$$T^*u = \sum_{j=1}^{n} \langle u, g_j\rangle T^*g_j.$$

Hence $\dim \operatorname{range} T^* \le n < \infty$.

Problem 10A.5. Prove or give a counterexample: If V is a Hilbert space and $T : V \to V$ is a bounded linear map such that $\dim \operatorname{null} T < \infty$, then $\dim \operatorname{null} T^* < \infty$.

Solution. False. Take $V = \ell^2$ and

$$T(a_1, a_2, a_3, \ldots) = (0, a_1, 0, a_2, 0, a_3, \ldots),$$

so that $(Ta)_{2k} = a_k$ and $(Ta)_{2k-1} = 0$. Then $\|Ta\| = \|a\|$, so $T \in \mathcal{B}(\ell^2)$ is an isometry and $\operatorname{null} T = \{0\}$. Its adjoint is $T^*b = (b_2, b_4, b_6, \ldots)$: writing Sb for that (bounded, as $\|Sb\| \le \|b\|$),

$$\langle Ta, b\rangle = \sum_{k=1}^{\infty} a_k \overline{b_{2k}} = \sum_{k=1}^{\infty} a_k \overline{(Sb)_k} = \langle a, Sb\rangle$$

for all $a, b \in \ell^2$, the odd-indexed terms dropping out and the series converging absolutely by the Cauchy–Schwarz inequality (8.11); uniqueness of the adjoint (10.1) gives $T^* = S$.
Hence $\operatorname{null} T^* = \{b \in \ell^2 : b_{2k} = 0 \text{ for all } k\}$, which contains the orthonormal list $e_1, e_3, e_5, \ldots$, so $\dim \operatorname{null} T^* = \infty$.

Problem 10A.6. Suppose T is a bounded linear map from a Hilbert space V to a Hilbert space W. Prove that $\|T^*T\| = \|T\|^2$.
[This formula for $\|T^*T\|$ leads to the important subject of C^*-algebras.]

Solution. One inequality is 10.20 (the norm of a composition) combined with $\|T^*\| = \|T\|$ from 10.11:

$$\|T^*T\| \le \|T^*\|\,\|T\| = \|T\|^2.$$

For the other, let $f \in V$ with $\|f\| \le 1$. Then

$$\begin{aligned}\|Tf\|^2 &= \langle Tf, Tf\rangle = \langle f, (T^*T)f\rangle = |\langle f, (T^*T)f\rangle| \\ &\le \|f\|\,\|(T^*T)f\| \le \|T^*T\|\,\|f\|^2 \le \|T^*T\|,\end{aligned}$$

where the second equality is the defining property 10.2 of the adjoint, the third holds because that number equals the nonnegative real $\|Tf\|^2$, and the first inequality is Cauchy–Schwarz (8.11). Taking the supremum over such f gives $\|T\|^2 \le \|T^*T\|$.

Problem 10A.7. Suppose V is a Hilbert space and $\operatorname{Inv}(V)$ is the set of invertible bounded operators on V. Think of $\operatorname{Inv}(V)$ as a metric space with the metric it inherits as a subset of $\mathcal{B}(V)$. Show that

$T \mapsto T^{-1}$ is a continuous function from $\operatorname{Inv}(V)$ to $\operatorname{Inv}(V)$.

Solution. Everything comes from the identity $S^{-1} - T^{-1} = S^{-1}(T-S)T^{-1}$ plus a bound on $\|S^{-1}\|$ that is uniform for S near T. (Each T^{-1} is indeed in $\mathcal{B}(V)$: V is complete, so the Bounded Inverse Theorem 6.83 applies.)
Fix $T \in \operatorname{Inv}(V)$; we may assume $V \neq \{0\}$, so $\|T^{-1}\| > 0$. Expanding the identity and applying 10.20 twice,

$$\|S^{-1} - T^{-1}\| \leq \|S^{-1}\|\,\|T-S\|\,\|T^{-1}\|.$$

Suppose $\|T-S\| < \frac{1}{2\|T^{-1}\|}$. Then $A = I - T^{-1}S = T^{-1}(T-S)$ has $\|A\| < \frac{1}{2}$, so $T^{-1}S = I - A$ is invertible with $(T^{-1}S)^{-1} = \sum_{k=0}^{\infty} A^k$ by 10.22, whence

$$\|(T^{-1}S)^{-1}\| \leq \sum_{k=0}^{\infty} \|A\|^k = \frac{1}{1-\|A\|} \leq 2.$$

Since $S^{-1} = (T^{-1}S)^{-1}T^{-1}$, this gives $\|S^{-1}\| \leq 2\|T^{-1}\|$.
So given $\varepsilon > 0$, the choice $\delta = \min\{\frac{1}{2\|T^{-1}\|}, \frac{\varepsilon}{2\|T^{-1}\|^2}\}$ yields, for $S \in \operatorname{Inv}(V)$ with $\|S-T\| < \delta$,

$$\|S^{-1} - T^{-1}\| \leq 2\|T^{-1}\|^2\,\|T-S\| < \varepsilon.$$

Hence inversion is continuous at each $T \in \operatorname{Inv}(V)$.

Problem 10A.8. Suppose T is a bounded operator on a Hilbert space.
(a) Prove that T is left invertible if and only if T^* is right invertible.
(b) Prove that T is invertible if and only if T is both left and right invertible.

Solution. (a) Each direction is one application of the adjoint rules 10.12(c), 10.12(d) and $(T^*)^* = T$ (10.11), the adjoints being bounded by 10.11. If $ST = I$, then

$$T^*S^* = (ST)^* = I^* = I,$$

so T^* is right invertible. If $T^*R = I$, then

$$R^*T = R^*(T^*)^* = (T^*R)^* = I,$$

so T is left invertible.
(b) If T is invertible, then T^{-1} is bounded by the Bounded Inverse Theorem (6.83), and it is both a left and a right inverse. Conversely, if $ST = I$ and $TR = I$ with $S, R \in \mathcal{B}(V)$, then

$$S = S(TR) = (ST)R = R,$$

so $ST = TS = I$. Hence T is injective ($Tf = 0$ gives $f = S(Tf) = 0$) and surjective ($g = T(Sg)$), that is, invertible in the sense of 10.18.

Problem 10A.9. Suppose $b_1, b_2, \ldots$ is a bounded sequence in $\mathbf{F}$. Define a bounded linear map $T : \ell^2 \to \ell^2$ by

$$T(a_1, a_2, \ldots) = (a_1b_1, a_2b_2, \ldots).$$

(a) Find a formula for T^*.
(b) Show that T is injective if and only if $b_k \neq 0$ for every $k \in \mathbf{Z}^+$.
(c) Show that T has dense range if and only if $b_k \neq 0$ for every $k \in \mathbf{Z}^+$.
(d) Show that T has closed range if and only if

$$\inf\{|b_k| : k \in \mathbf{Z}^+ \text{ and } b_k \neq 0\} > 0.$$

(e) Show that T is invertible if and only if $\inf\{|b_k| : k \in \mathbf{Z}^+\} > 0$.

Solution. Write e_k for the k^{th} standard basis vector, so $Te_k = b_k e_k$, and put $N = \{k : b_k \neq 0\}$ with complement Z.
(a) $T^*(c_1, c_2, \ldots) = (\overline{b_1}c_1, \overline{b_2}c_2, \ldots)$, since

$$\langle Ta, c\rangle = \sum_{k=1}^{\infty} a_k b_k \overline{c_k} = \sum_{k=1}^{\infty} a_k \overline{\overline{b_k}\, c_k} = \big\langle a, (\overline{b_k}c_k)_k\big\rangle,$$

the series converging absolutely by Cauchy–Schwarz because $(b_k c_k) \in \ell^2$.
(b) If $b_k = 0$ then $Te_k = 0$ with $e_k \neq 0$; conversely if every $b_k \neq 0$ then $Ta = 0$ forces $a_k b_k = 0$, hence $a = 0$.
(c) By 10.14, T has dense range if and only if T^* is injective; by (a), T^* is the multiplier operator of $(\overline{b_k})$, so (b) applies to it and gives the criterion $b_k \neq 0$ for every k.
(d) Set $c = \inf\{|b_k| : k \in N\}$, with $\inf \varnothing = \infty$ (if $N = \varnothing$ then $T = 0$ has closed range and the equivalence holds).
(i) $c > 0$. Then $\operatorname{range} T = W := \{a \in \ell^2 : a_k = 0 \text{ for } k \in Z\}$, which is closed because $|a_k^{(n)} - a_k| \leq \|a^{(n)} - a\|$ makes each coordinate map continuous. The inclusion $\operatorname{range} T \subseteq W$ is clear; conversely, given $a \in W$, the sequence d with $d_k = a_k/b_k$ on N and $d_k = 0$ on Z satisfies $\|d\| \leq \|a\|/c < \infty$ and $Td = a$.
(ii) $c = 0$. Choose $k_1 < k_2 < \cdots$ in N with $|b_{k_j}| < 1/j$: having picked k_{j-1}, the finitely many $k \in N$ with $k \leq k_{j-1}$ have $|b_k| \geq \delta$ for some $\delta > 0$, while $c = 0$ supplies $k \in N$ with $|b_k| < \min\{1/j, \delta\}$, forcing $k > k_{j-1}$. Let $a_{k_j} = 1/j$ and $a_k = 0$ otherwise, so $a \in \ell^2$. Its truncations to slots $k_1, \ldots, k_J$ lie in $\operatorname{range} T$ (finitely supported preimages) and converge to a, so $a \in \overline{\operatorname{range} T}$. But $Td = a$ would give $d_{k_j} b_{k_j} = 1/j$, hence

$$|d_{k_j}| = \frac{1}{j\,|b_{k_j}|} > 1 \qquad \text{for every } j,$$

contradicting $d \in \ell^2$. So $\operatorname{range} T$ is not closed.
(e) If $c = \inf_k |b_k| > 0$, then $S(a_1, a_2, \ldots) = (a_1/b_1, a_2/b_2, \ldots)$ is bounded with $\|S\| \leq 1/c$ and $ST = TS = I$. Conversely, if T is invertible then every $b_k \neq 0$ by (b), and $T^{-1}e_k = e_k/b_k$ gives

$$\frac{1}{|b_k|} = \|T^{-1}e_k\| \leq \|T^{-1}\|,$$

so $\inf_k |b_k| \geq 1/\|T^{-1}\| > 0$.

Problem 10A.10. Suppose $h \in L^\infty(\mathbf{R})$ and $M_h : L^2(\mathbf{R}) \to L^2(\mathbf{R})$ is the bounded operator defined by $M_h f = fh$.
(a) Show that M_h is injective if and only if $|\{x \in \mathbf{R} : h(x) = 0\}| = 0$.
(b) Find a necessary and sufficient condition (in terms of h) for M_h to have dense range.
(c) Find a necessary and sufficient condition (in terms of h) for M_h to have closed range.
(d) Find a necessary and sufficient condition (in terms of h) for M_h to be invertible.

Solution. Write $Z = \{x \in \mathbf{R} : h(x) = 0\}$ and $|A|$ for Lebesgue measure. The adjoint is $(M_h)^* = M_{\overline{h}}$, since

$$\langle M_h f, g\rangle = \int fh\overline{g}\, d\lambda = \int f\overline{\overline{h}g}\, d\lambda = \langle f, M_{\overline{h}}g\rangle.$$

(a) If $|Z| > 0$, then $Z = \bigcup_n (Z \cap [-n, n])$ and 2.59 give $E = Z \cap [-n, n]$ with $0 < |E| < \infty$; then $\chi_E \neq 0$ in L^2 while $M_h \chi_E = 0$. If $|Z| = 0$ and $M_h f = 0$, then $fh = 0$ almost everywhere forces $f = 0$ almost everywhere off Z, hence $f = 0$ in $L^2(\mathbf{R})$.
(b) The condition is again $|Z| = 0$: by 10.14 the range is dense if and only if $(M_h)^* = M_{\overline{h}}$ is injective, and $\{\overline{h} = 0\} = Z$, so (a) applies.

(c) The condition is that $\big|\{0 < |h| < \delta\}\big| = 0$ for some $\delta > 0$.
(i) Such a δ exists. Then $\operatorname{range} M_h = W := \{f : f = 0 \text{ a.e. on } Z\}$, and $W = \operatorname{null} M_{\chi_Z}$ is closed. The inclusion $\subseteq$ is clear; conversely, for $f \in W$ put $g = f/h$ off Z and $g = 0$ on Z, so $|g| \le |f|/\delta$ almost everywhere and $M_h g = f$.
(ii) No such δ. With $B_n = \{\frac{1}{n+1} \le |h| < \frac{1}{n}\}$, infinitely many B_n have positive measure: otherwise countable additivity would give $|\{0 < |h| < 1/M\}| = 0$ for large M. Pick $n_1 < n_2 < \cdots$ with $|B_{n_j}| > 0$ (so $n_j \ge j$) and, as in (a), disjoint $E_j \subseteq B_{n_j}$ with $0 < |E_j| < \infty$. Set

$$f = \sum_{j=1}^{\infty} \frac{1}{j\,|E_j|^{1/2}}\,\chi_{E_j}, \qquad \|f\|_2^2 = \sum_{j=1}^{\infty} \frac{1}{j^2} < \infty.$$

Each partial sum $f^{(J)}$ lies in $\operatorname{range} M_h$ (divide by h, which is at least $\min_{j \le J} \frac{1}{n_j+1} > 0$ on $\bigcup_{j \le J} E_j$), and $\|f - f^{(J)}\|_2 \to 0$. But $M_h g = f$ would force, almost everywhere on E_j, where $|h| < 1/n_j \le 1/j$,

$$|g| = \frac{|f|}{|h|} > \frac{1}{|E_j|^{1/2}},$$

so $\int_{E_j} |g|^2\, d\lambda \ge 1$ for every j and $\|g\|_2 = \infty$. Hence the range is not closed.
(d) The condition is that $\big|\{|h| < \delta\}\big| = 0$ for some $\delta > 0$, that is, $1/h \in L^\infty(\mathbf{R})$.
If it holds, then $\|1/h\|_\infty \le 1/\delta$, so $M_{1/h}$ is bounded and $M_{1/h}M_h = M_h M_{1/h} = I$. Conversely, if M_h is invertible, then $(M_h)^{-1}$ is bounded by the Bounded Inverse Theorem (6.83); with $\alpha = \|(M_h)^{-1}\| > 0$ we have $\|f\|_2 \le \alpha \|M_h f\|_2$ for all f. Were $\{|h| < \delta\}$ of positive measure for $\delta = 1/(2\alpha)$, choosing E inside it with $0 < |E| < \infty$ and $f = \chi_E$ would give

$$|E|^{1/2} \le \alpha \|M_h \chi_E\|_2 \le \alpha\delta |E|^{1/2} = \tfrac{1}{2}|E|^{1/2},$$

impossible.

Problem 10A.11. (a) Prove or give a counterexample: If T is a bounded operator on a Hilbert space such that T and T^* are both injective, then T is invertible.
(b) Prove or give a counterexample: If T is a bounded operator on a Hilbert space such that T and T^* are both surjective, then T is invertible.

Solution. (a) False. Take the operator of Example 10.28,

$$T(a_1, a_2, a_3, \ldots) = \Big(a_1, \frac{a_2}{2}, \frac{a_3}{3}, \ldots\Big),$$

the multiplier operator of the real sequence $b_k = 1/k$, so $T^* = T$ by Exercise 10A.9(a). It is injective, since $Ta = 0$ forces every $a_k = 0$. But $\inf_k |b_k| = 0$, so T is not invertible by Exercise 10A.9(e).
(b) True. Because T^* is surjective, 10.13(b) gives

$$V = \overline{\operatorname{range} T^*} = (\operatorname{null} T)^\perp,$$

and since $\operatorname{null} T$ is a closed subspace, 8.41 yields

$$\operatorname{null} T = \big((\operatorname{null} T)^\perp\big)^\perp = V^\perp = \{0\}.$$

Thus T is injective as well as surjective, hence invertible by 10.18.

Problem 10A.12. Define $T : \ell^2 \to \ell^2$ by $T(a_1, a_2, a_3, \ldots) = (a_2, a_3, a_4, \ldots)$. Suppose $\alpha \in \mathbf{F}$.
(a) Prove that $T - \alpha I$ is injective if and only if $|\alpha| \ge 1$.
(b) Prove that $T - \alpha I$ is invertible if and only if $|\alpha| > 1$.

(c) Prove that $T - \alpha I$ is surjective if and only if $|\alpha| \neq 1$.
(d) Prove that $T - \alpha I$ is left invertible if and only if $|\alpha| > 1$.

Solution. Let S be the right shift, $S(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$, so that $\|Ta\| \leq \|a\|$ and $\|Sa\| = \|a\|$. Then $T^* = S$ and $S^* = T$, since

$$\langle Sa, c\rangle = \sum_{j=1}^{\infty} a_j \overline{c_{j+1}} = \langle a, Tc\rangle$$

and $(S^*)^* = S$ by 10.11.
Also, for $|\alpha| = 1$ the operator $T - \alpha I$ is not bounded below: with $a^{(n)} = (1, \alpha, \ldots, \alpha^{n-1}, 0, \ldots)$, every coordinate of $(T - \alpha I)a^{(n)}$ vanishes except the n^{th}, which is $-\alpha^n$, so

$$\frac{\|(T - \alpha I)a^{(n)}\|}{\|a^{(n)}\|} = \frac{1}{\sqrt{n}} \longrightarrow 0.$$

(a) The equation $(T - \alpha I)a = 0$ says $a_{k+1} = \alpha a_k$, that is, $a_k = \alpha^{k-1} a_1$.
(i) $|\alpha| < 1$: then $a = (1, \alpha, \alpha^2, \ldots)$ is a nonzero element of ℓ^2 killed by $T - \alpha I$.
(ii) $|\alpha| \geq 1$: if $a_1 \neq 0$ then $|a_k| \geq |a_1| > 0$ for every k, contradicting $a \in \ell^2$; so $a_1 = 0$ and $a = 0$.
(b) If $|\alpha| > 1$, then $\|T/\alpha\| \leq 1/|\alpha| < 1$, so $I - T/\alpha$ is invertible by 10.22 (ℓ^2 being a Banach space), and hence so is $T - \alpha I = -\alpha(I - T/\alpha)$. Conversely, invertibility gives injectivity, so $|\alpha| \geq 1$ by (a); and $|\alpha| = 1$ is impossible, since a bounded inverse (6.83) would give $\|(T - \alpha I)a\| \geq c\|a\|$ with $c = 1/\|(T - \alpha I)^{-1}\| > 0$, contradicting the display above.
(c) (i) $|\alpha| > 1$: surjective, being invertible by (b).
(ii) $|\alpha| < 1$: here $(T - \alpha I)^* = S - \overline{\alpha} I$ by 10.12, and

$$\|(S - \overline{\alpha} I)b\| \geq \|Sb\| - |\alpha|\,\|b\| = (1 - |\alpha|)\|b\|,$$

so 10.29(b) makes $(T - \alpha I)^*$ left invertible; then Exercise 10A.8(a) applied to it, with $\big((T - \alpha I)^*\big)^* = T - \alpha I$ (10.11), makes $T - \alpha I$ right invertible and hence surjective by 10.31.
(iii) $|\alpha| = 1$: injective by (a), so surjectivity would force invertibility by 10.18, contradicting (b).
(d) (i) $|\alpha| > 1$: invertible by (b), and its inverse is bounded (6.83), hence a left inverse.
(ii) $|\alpha| < 1$: not injective by (a), so 10.29(c) fails.
(iii) $|\alpha| = 1$: not bounded below, by the display above, so 10.29(b) fails.

Problem 10A.13. Suppose V is a Hilbert space.
(a) Show that $\{T \in \mathcal{B}(V) : T \text{ is left invertible}\}$ is an open subset of $\mathcal{B}(V)$.
(b) Show that $\{T \in \mathcal{B}(V) : T \text{ is right invertible}\}$ is an open subset of $\mathcal{B}(V)$.

Solution. (a) If T is left invertible, 10.29 supplies $\alpha \in (0, \infty)$ with $\|f\| \leq \alpha\|Tf\|$ for all $f \in V$. Then the ball of radius $1/\alpha$ about T consists of left invertible operators: if $\|T - R\| < 1/\alpha$ and $\varepsilon = 1/\alpha - \|T - R\| > 0$, then

$$\|Rf\| \geq \|Tf\| - \|(T - R)f\| \geq \frac{\|f\|}{\alpha} - \|T - R\|\,\|f\| = \varepsilon\|f\|$$

for every f, so 10.29(b) holds for R with constant $1/\varepsilon$.
(b) Adjunction is an isometry of $\mathcal{B}(V)$: $\|T^* - R^*\| = \|(T - R)^*\| = \|T - R\|$ by 10.12(a), 10.12(b), and 10.11. So if T is right invertible, then T^* is left invertible by Exercise 10A.8(a) (using $(T^*)^* = T$), and (a) gives $r > 0$ such that every A with $\|T^* - A\| < r$ is left invertible. Any R with $\|T - R\| < r$ then has R^* left invertible, hence R right invertible by Exercise 10A.8(a) again.

Problem 10A.14. Suppose T is a bounded operator on a Hilbert space V.
(a) Prove that T is invertible if and only if T has a unique left inverse. In other words, prove that T is invertible if and only if there exists a unique $S \in \mathcal{B}(V)$ such that $ST = I$.
(b) Prove that T is invertible if and only if T has a unique right inverse. In other words, prove that T is invertible if and only if there exists a unique $S \in \mathcal{B}(V)$ such that $TS = I$.

Solution. (a) If T is invertible, then $T^{-1} \in \mathcal{B}(V)$ by the Bounded Inverse Theorem (6.83) is a left inverse, and it is the only one: $ST = I$ gives

$$S = S(TT^{-1}) = (ST)T^{-1} = T^{-1}.$$

Conversely, if S is the unique left inverse, then $S' = S + (I - TS) \in \mathcal{B}(V)$ satisfies

$$S'T = ST + (I - TS)T = I + T - T(ST) = I,$$

so $S' = S$, forcing $TS = I$; now Exercise 10A.8(b) makes T invertible.
(b) Likewise with the factors reversed. If T is invertible and $TS = I$, then $S = (T^{-1}T)S = T^{-1}$. Conversely, if S is the unique right inverse, then $S' = S + (I - ST) \in \mathcal{B}(V)$ satisfies

$$TS' = TS + T(I - ST) = I + T - (TS)T = I,$$

so $S' = S$, forcing $ST = I$, and Exercise 10A.8(b) applies.

10.2 Exercises 10B

Problem 10B.1. Verify all the assertions in Example 10.33.
Example 10.33 (eigenvalues and spectrum) asserts the following.

- Suppose $b_1, b_2, \ldots$ is a bounded sequence in $\mathbf{F}$. Define a bounded linear map $T : \ell^2 \to \ell^2$ by
$$T(a_1, a_2, \ldots) = (a_1 b_1, a_2 b_2, \ldots).$$

Then the set of eigenvalues of T equals $\{b_k : k \in \mathbf{Z}^+\}$ and the spectrum of T equals the closure of $\{b_k : k \in \mathbf{Z}^+\}$.

- Suppose $h \in L^\infty(\mathbf{R})$. Define a bounded linear map $M_h : L^2(\mathbf{R}) \to L^2(\mathbf{R})$ by $M_h f = fh$. Then $\alpha \in \mathbf{F}$ is an eigenvalue of M_h if and only if $|\{t \in \mathbf{R} : h(t) = \alpha\}| > 0$. Also, $\alpha \in \operatorname{sp}(M_h)$ if and only if $|\{t \in \mathbf{R} : |h(t) - \alpha| < \varepsilon\}| > 0$ for all $\varepsilon > 0$.
- Define the right shift $T : \ell^2 \to \ell^2$ and the left shift $S : \ell^2 \to \ell^2$ by
$$T(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots) \quad \text{and} \quad S(a_1, a_2, a_3, \ldots) = (a_2, a_3, a_4, \ldots).$$

Then T has no eigenvalues, and $\operatorname{sp}(T) = \{\alpha \in \mathbf{F} : |\alpha| \le 1\}$. Also, the set of eigenvalues of S is the open set $\{\alpha \in \mathbf{F} : |\alpha| < 1\}$, and the spectrum of S is the closed set $\{\alpha \in \mathbf{F} : |\alpha| \le 1\}$.

Solution. Write e_k for the k^{th} standard basis vector of ℓ^2. Each non-invertibility below is certified the same way: an invertible R satisfies $\|f\| \le \|R^{-1}\| \, \|Rf\|$, so unit vectors f_n with $\|Rf_n\| \to 0$ rule invertibility out.
Multiplier operator on ℓ^2: here $(T - \alpha I)a = \big(a_k(b_k - \alpha)\big)_k$.
Eigenvalues: if $\alpha = b_j$ then $(T - \alpha I)e_j = 0$, while if $\alpha \ne b_k$ for every k, then $a_k(b_k - \alpha) = 0$ forces $a = 0$. So the eigenvalues are exactly $\{b_k : k \in \mathbf{Z}^+\}$.
Spectrum: if $c = \inf_k |b_k - \alpha| > 0$, the multiplier operator of $\big((b_k - \alpha)^{-1}\big)_k$ is bounded by $1/c$ and is a two-sided inverse of $T - \alpha I$. If $c = 0$, choose k_n with $|b_{k_n} - \alpha| \to 0$; then $\|e_{k_n}\| = 1$ and $\|(T - \alpha I)e_{k_n}\| = |b_{k_n} - \alpha| \to 0$. Since $c = 0$ says exactly that α lies in the closure of $\{b_k\}$, the spectrum is that closure.
Multiplication operator on $L^2(\mathbf{R})$: put $g = h - \alpha$, so $M_h - \alpha I = M_g$.
Eigenvalues: by Exercise 10A.10(a) applied to g, the operator M_g fails to be injective precisely when $|\{g = 0\}| > 0$, that is, $|\{h = \alpha\}| > 0$.
Spectrum: if $|\{|g| < \varepsilon\}| = 0$ for some $\varepsilon > 0$, then $\|1/g\|_\infty \le 1/\varepsilon$ and $M_{1/g}$ inverts M_g. If instead

$|\{|g| < \varepsilon\}| > 0$ for every $\varepsilon > 0$, take (as in Exercise 10A.10(a)) sets $F_n \subseteq \{|g| < 1/n\}$ with $0 < |F_n| < \infty$ and put $f_n = \chi_{F_n}/|F_n|^{1/2}$, so $\|f_n\|_2 = 1$ and

$$\|M_g f_n\|_2^2 = \frac{1}{|F_n|} \int_{F_n} |g|^2 \, d\lambda \le \frac{1}{n^2}.$$

Shifts on ℓ^2: here $\|Ta\| = \|a\|$, while $\|Sa\| \le \|a\|$ with $\|Se_2\| = 1$, so $\|T\| = \|S\| = 1$.
T has no eigenvalue: $Ta = \alpha a$ gives $0 = \alpha a_1$ and $a_k = \alpha a_{k+1}$. If $\alpha = 0$ the latter forces every $a_k = 0$; if $\alpha \ne 0$ the former gives $a_1 = 0$ and then $a_{k+1} = a_k/\alpha = 0$ inductively.
$\operatorname{sp}(T) = \{|\alpha| \le 1\}$: the inclusion $\subseteq$ is 10.34(a) with $\|T\| = 1$. For $|\alpha| < 1$, the nonzero vector $v = (1, \overline{\alpha}, \overline{\alpha}^2, \ldots) \in \ell^2$ satisfies, for every $a \in \ell^2$,

$$\begin{aligned} \langle (T - \alpha I)a, v \rangle &= -\alpha a_1 + \sum_{k=1}^{\infty} (a_k - \alpha a_{k+1}) \alpha^k \\ &= -\alpha a_1 + a_1 \alpha = 0, \end{aligned}$$

the series converging absolutely by Cauchy–Schwarz (8.11). So $\operatorname{range}(T - \alpha I) \subseteq \{v\}^\perp \ne \ell^2$, and $T - \alpha I$ is not surjective; since $\operatorname{sp}(T)$ is closed (10.36), the whole closed disk lies in it.
Eigenvalues of S: $Sa = \alpha a$ says $a_{k+1} = \alpha a_k$, that is, $a_k = \alpha^{k-1} a_1$. For $|\alpha| < 1$ this gives the eigenvector $w = (1, \alpha, \alpha^2, \ldots) \in \ell^2$; conversely $a \ne 0$ forces $a_1 \ne 0$, and then $\sum_k |a_1|^2 |\alpha|^{2(k-1)} < \infty$ forces $|\alpha| < 1$.
$\operatorname{sp}(S) = \{|\alpha| \le 1\}$: the eigenvalues lie in $\operatorname{sp}(S)$, which is closed (10.36), giving $\supseteq$; and 10.34(a) with $\|S\| = 1$ gives $\subseteq$.

Problem 10B.2. Suppose T is a bounded operator on a Hilbert space V.
- (a) Prove that $\operatorname{sp}(S^{-1}TS) = \operatorname{sp}(T)$ for all bounded invertible operators S on V.
- (b) Prove that $\operatorname{sp}(T^*) = \{\overline{\alpha} : \alpha \in \operatorname{sp}(T)\}$.
- (c) Prove that if T is invertible, then $\operatorname{sp}(T^{-1}) = \left\{\frac{1}{\alpha} : \alpha \in \operatorname{sp}(T)\right\}$.

Solution. Each part rests on one factorization of $T - \alpha I$ together with the fact that ABC is invertible if and only if B is (for A, C invertible): if R inverts ABC then CRA inverts B. (Check!)
(a) For $\alpha \in \mathbf{F}$,

$$S^{-1}TS - \alpha I = S^{-1}(T - \alpha I)S,$$

so $S^{-1}TS - \alpha I$ is invertible exactly when $T - \alpha I$ is. Hence $\operatorname{sp}(S^{-1}TS) = \operatorname{sp}(T)$.
(b) Taking adjoints in $RR^{-1} = R^{-1}R = I$ and using 10.11 and 10.12(c), (d) shows R is invertible if and only if R^* is, with $(R^*)^{-1} = (R^{-1})^*$. By 10.12(a), (b), (c),

$$T^* - \beta I = T^* - \overline{\overline{\beta}}\, I^* = \bigl(T - \overline{\beta} I\bigr)^*,$$

so $\beta \in \operatorname{sp}(T^*)$ if and only if $\overline{\beta} \in \operatorname{sp}(T)$. Because conjugation is an involution of $\mathbf{F}$, this says $\operatorname{sp}(T^*) = \{\overline{\alpha} : \alpha \in \operatorname{sp}(T)\}$.
(c) Invertibility of T and of T^{-1} puts 0 outside both $\operatorname{sp}(T)$ and $\operatorname{sp}(T^{-1})$, so all reciprocals below are defined. For $\alpha \ne 0$,

$$\Bigl(-\tfrac{1}{\alpha} T^{-1}\Bigr)(T - \alpha I) = -\tfrac{1}{\alpha}\bigl(I - \alpha T^{-1}\bigr) = T^{-1} - \tfrac{1}{\alpha} I,$$

and $-\frac{1}{\alpha} T^{-1}$ is invertible, so

$$\tfrac{1}{\alpha} \in \operatorname{sp}(T^{-1}) \iff \alpha \in \operatorname{sp}(T).$$

Reading this forward gives $\{1/\alpha : \alpha \in \operatorname{sp}(T)\} \subseteq \operatorname{sp}(T^{-1})$; reading it with $\alpha = 1/\beta$ for $\beta \in \operatorname{sp}(T^{-1})$ (legitimate as $\beta \ne 0$) gives the reverse inclusion. Thus $\operatorname{sp}(T^{-1}) = \{1/\alpha : \alpha \in \operatorname{sp}(T)\}$.

Problem 10B.3. Suppose E is a bounded subset of $\mathbf{F}$. Show that there exists a Hilbert space V and $T \in \mathcal{B}(V)$ such that the set of eigenvalues of T equals E.

Solution. Take $V = \ell^2(E)$, the space $L^2(\mu)$ for μ counting measure on E (second bullet of Example 8.51), and let T be multiplication by the coordinate:

$$(Tf)(\alpha) = \alpha f(\alpha) \qquad (\alpha \in E).$$

(If $E = \varnothing$ this is $V = \{0\}$, $T = 0$, which has no eigenvalues since eigenvectors are nonzero.)
T is bounded: choosing M with $|\alpha| \leq M$ on E, which exists as E is bounded,

$$\|Tf\|^2 = \sum_{\alpha \in E} |\alpha|^2 |f(\alpha)|^2 \leq M^2 \|f\|^2,$$

so $T \in \mathcal{B}(V)$ with $\|T\| \leq M$.
Each $\beta \in E$ is an eigenvalue: the indicator e_β of $\{\beta\}$ has $\|e_\beta\| = 1$ and $Te_\beta = \beta e_\beta$, both sides vanishing off β. Conversely, if $Tf = \beta f$ with $f \neq 0$, then $(\alpha - \beta) f(\alpha) = 0$ for all $\alpha \in E$, and picking α_0 with $f(\alpha_0) \neq 0$ gives $\beta = \alpha_0 \in E$. Thus the eigenvalues of T are exactly E.

Problem 10B.4. Suppose E is a nonempty closed bounded subset of $\mathbf{F}$. Show that there exists $T \in \mathcal{B}(\ell^2)$ such that $\operatorname{sp}(T) = E$.

Solution. Take $b_1, b_2, \ldots$ an enumeration of a countable dense subset A of E (repeating terms if A is finite) and let T be the diagonal operator

$$T(a_1, a_2, \ldots) = (a_1 b_1, a_2 b_2, \ldots).$$

The sequence (b_k) is bounded because E is, so with $M = \sup_k |b_k|$ we get $\|Ta\|^2 = \sum_k |b_k|^2 |a_k|^2 \leq M^2 \|a\|^2$ and $T \in \mathcal{B}(\ell^2)$. By the first bullet of Example 10.33 (Exercise 1 of this section),

$$\operatorname{sp}(T) = \overline{\{b_k : k \in \mathbf{Z}^+\}} = \overline{A} = E,$$

the last equality because $A \subseteq E$ is dense in E and E is closed.
It remains to produce A. Let D be $\mathbf{Q}$ or $\mathbf{Q} + i\mathbf{Q}$ according as $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$, and for each $(d, n) \in D \times \mathbf{Z}^+$ with $\{x \in E : |x - d| < 1/n\} \neq \varnothing$ choose a point $x_{d,n}$ of that set; let A be the set of chosen points. Then $A \subseteq E$ is countable, and nonempty since $E \neq \varnothing$. Given $x \in E$ and $\varepsilon > 0$, pick n with $2/n < \varepsilon$ and $d \in D$ with $|x - d| < 1/n$; then $x_{d,n}$ exists and

$$|x_{d,n} - x| \leq |x_{d,n} - d| + |d - x| < \tfrac{2}{n} < \varepsilon.$$

Problem 10B.5. Give an example of a bounded operator T on a normed vector space such that for every $\alpha \in \mathbf{F}$, the operator $T - \alpha I$ is not invertible.

Solution. Take V to be the (incomplete) subspace of ℓ^2 of sequences with only finitely many nonzero coordinates, with the ℓ^2-norm, and T the right shift

$$T(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots).$$

Shifting preserves finite support, so T maps V into V, and $\|Ta\| = \|a\|$, so $T \in \mathcal{B}(V)$ with $\|T\| = 1$. (By 10.34(b) no example on a Banach space exists, so incompleteness is forced.)
Fix $\alpha \in \mathbf{F}$ and let $e_1 = (1, 0, 0, \ldots)$. If $(T - \alpha I)a = e_1$ for some $a \in V$, then comparing coordinates in

$$(T - \alpha I)a = \big(-\alpha a_1,\ a_1 - \alpha a_2,\ a_2 - \alpha a_3,\ \ldots\big)$$

gives $-\alpha a_1 = 1$ (so $\alpha \neq 0$ and $a_1 = -1/\alpha$) and $a_{k+1} = a_k/\alpha$ for $k \geq 1$, whence $a_k = -1/\alpha^k \neq 0$ for all k, contradicting $a \in V$. So $T - \alpha I$ is not surjective, hence not invertible, for every $\alpha \in \mathbf{F}$.

Problem 10B.6. Suppose T is a bounded operator on a complex nonzero Banach space V.

- (a) Prove that the function
$$\alpha \mapsto \varphi\big((T-\alpha I)^{-1}f\big)$$
is analytic on $\mathbf{C} \setminus \mathrm{sp}(T)$ for every $f \in V$ and every $\varphi \in V'$.
- (b) Prove that $\mathrm{sp}(T) \neq \varnothing$.

Solution. (a) The resolvent expands in a norm-convergent power series around each $\beta \notin \mathrm{sp}(T)$ (the complement is open by 10.36). Put $R = (T-\beta I)^{-1}$, which is nonzero, and take $|\alpha-\beta| < 1/\|R\|$. Then $\|(\alpha-\beta)R\| < 1$, so $I-(\alpha-\beta)R$ is invertible by 10.22 with inverse $\sum_{k\geq 0}(\alpha-\beta)^k R^k$. Since

$$(T-\beta I)\big(I-(\alpha-\beta)R\big) = (T-\beta I)-(\alpha-\beta)I = T-\alpha I,$$

the operator $T-\alpha I$ is invertible and

$$(T-\alpha I)^{-1} = \big(I-(\alpha-\beta)R\big)^{-1}R = \sum_{k=0}^{\infty}(\alpha-\beta)^k R^{k+1}.$$

The functional $S \mapsto \varphi(Sf)$ on $\mathcal{B}(V)$ is bounded, with $|\varphi(Sf)| \leq \|\varphi\|\|f\|\|S\|$, so it passes through this norm-convergent sum:

$$\varphi\big((T-\alpha I)^{-1}f\big) = \sum_{k=0}^{\infty}\varphi\big(R^{k+1}f\big)(\alpha-\beta)^k.$$

The coefficient bound $|\varphi(R^{k+1}f)| \leq \|\varphi\|\|f\|\|R\|^{k+1}$ gives radius of convergence at least $1/\|R\|$. Hence the function is analytic at each $\beta \in \mathbf{C} \setminus \mathrm{sp}(T)$.
(b) Suppose $\mathrm{sp}(T) = \varnothing$, so $T-\alpha I$ is invertible for all α, in particular T is. Pick $f \neq 0$ (possible as $V \neq \{0\}$) and $\varphi \in V'$, and set $g(\alpha) = \varphi((T-\alpha I)^{-1}f)$, entire by (a). Then

$$|g(\alpha)| \leq \|\varphi\|\,\|f\|\,\big\|(T-\alpha I)^{-1}\big\| \longrightarrow 0$$

as $|\alpha| \to \infty$, by 10.34(c); being continuous, g is bounded on the complementary closed disc, so g is a bounded entire function. By Liouville's theorem g is constant, and the limit forces $g \equiv 0$; in particular $\varphi(T^{-1}f) = g(0) = 0$. As $\varphi \in V'$ was arbitrary and 6.72 supplies φ with $\varphi(h) = \|h\|$ for $h = T^{-1}f \neq 0$, we get $h = 0$, so $f = Th = 0$, a contradiction. Thus $\mathrm{sp}(T) \neq \varnothing$.

Problem 10B.7. Prove that if T is an operator on a Hilbert space V such that $\langle Tf, g\rangle = \langle f, Tg\rangle$ for all $f, g \in V$, then T is a bounded operator.

Solution. Apply the Closed Graph Theorem 6.85, whose hypotheses hold because a Hilbert space is a Banach space; so it suffices to show that

$$\mathrm{graph}(T) = \{(f, Tf) : f \in V\}$$

is closed in $V \times V$ with the norm $\|(f,g)\| = \max\{\|f\|, \|g\|\}$ of 6.84. It is a subspace, by linearity of T. (Check!)
Suppose $(f_k, Tf_k) \to (f, g)$, that is $f_k \to f$ and $Tf_k \to g$ by 6.84, and let $h \in V$. Cauchy–Schwarz (8.11) makes both $\langle Tf_k, h\rangle \to \langle g, h\rangle$ and $\langle f_k, Th\rangle \to \langle f, Th\rangle$, so the hypothesis applied to the pairs (f_k, h) and then (f, h) gives

$$\langle g, h\rangle = \lim_{k\to\infty}\langle Tf_k, h\rangle = \lim_{k\to\infty}\langle f_k, Th\rangle = \langle f, Th\rangle = \langle Tf, h\rangle.$$

Taking $h = g - Tf$ yields $\|g-Tf\|^2 = 0$, so $(f,g) = (f, Tf) \in \mathrm{graph}(T)$. Hence the graph is closed and T is bounded.

Problem 10B.8. Suppose P is a bounded operator on a Hilbert space V such that $P^2 = P$. Prove that P is self-adjoint if and only if there exists a closed subspace U of V such that $P = P_U$.

Solution. If $P = P_U$ for a closed subspace U, then P is self-adjoint by the last bullet of 10.45. Conversely, suppose P is self-adjoint and take $U = \operatorname{range} P$. Then $U = \operatorname{null}(I - P)$: for $f = Pg \in U$ we have $Pf = P^2 g = Pg = f$, and conversely $(I - P)f = 0$ puts $f = Pf \in \operatorname{range} P$. Hence U is the null space of the bounded operator $I - P$, so U is closed and P_U is defined.
Let $f \in V$. Then $Pf \in U$, and for $g \in U$ (so $Pg = g$),

$$\langle f - Pf, g\rangle = \langle f, g\rangle - \langle f, Pg\rangle = 0,$$

using self-adjointness in the first step. So $f - Pf \perp U$ with $Pf \in U$, and 8.37(b) with $h = Pf$ gives $Pf = P_U f$. Thus $P = P_U$.

Problem 10B.9. Suppose V is a real Hilbert space and $T \in \mathcal{B}(V)$. The *complexification* of T is the function $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ defined by

$$T_{\mathbb{C}}(f + ig) = Tf + iTg$$

for $f, g \in V$ (see Exercise 4 in Section 8B for the definition of $V_{\mathbb{C}}$).
(a) Show that $T_{\mathbb{C}}$ is a bounded operator on the complex Hilbert space $V_{\mathbb{C}}$ and $\|T_{\mathbb{C}}\| = \|T\|$.
(b) Show that $T_{\mathbb{C}}$ is invertible if and only if T is invertible.
(c) Show that $(T_{\mathbb{C}})^* = (T^*)_{\mathbb{C}}$.
(d) Show that T is self-adjoint if and only if $T_{\mathbb{C}}$ is self-adjoint.
(e) Use the previous parts of this exercise and 10.49 and 10.38 to show that if T is self-adjoint and $V \neq \{0\}$, then $\operatorname{sp}(T) \neq \emptyset$.

Solution. Throughout we use the norm identity on $V_{\mathbb{C}}$ from Exercise 4 in Section 8B, immediate from its inner product formula and the symmetry of the real inner product on V:

$$\|f + ig\|^2 = \|f\|^2 + \|g\|^2.$$

(a) Additivity of $T_{\mathbb{C}}$ is inherited from T, and for $\alpha, \beta \in \mathbb{R}$ the $\mathbb{R}$-linearity of T gives

$$\begin{aligned} T_{\mathbb{C}}\big((\alpha + i\beta)(f + ig)\big) &= T(\alpha f - \beta g) + iT(\alpha g + \beta f) \\ &= (\alpha + i\beta)(Tf + iTg) = (\alpha + i\beta)\, T_{\mathbb{C}}(f + ig), \end{aligned}$$

so $T_{\mathbb{C}}$ is $\mathbb{C}$-linear. For the norm,

$$\|T_{\mathbb{C}}(f + ig)\|^2 = \|Tf\|^2 + \|Tg\|^2 \le \|T\|^2\big(\|f\|^2 + \|g\|^2\big) = \|T\|^2\, \|f + ig\|^2.$$

Hence $\|T_{\mathbb{C}}\| \le \|T\|$; and $\|T_{\mathbb{C}}(f + i0)\| = \|Tf\|$ with $\|f + i0\| = \|f\|$ gives the reverse. Thus $\|T_{\mathbb{C}}\| = \|T\|$.
(b) If T is invertible, then $(T^{-1})_{\mathbb{C}}$, bounded by (a), inverts $T_{\mathbb{C}}$:

$$(T^{-1})_{\mathbb{C}} T_{\mathbb{C}}(f + ig) = T^{-1}Tf + iT^{-1}Tg = f + ig,$$

and symmetrically in the other order.
Conversely, let $S = (T_{\mathbb{C}})^{-1}$. Then T is injective, since $Tf = 0$ forces $T_{\mathbb{C}}(f + i0) = 0$ and so $f = 0$. It is surjective: writing $S(f + i0) = u + iv$ and applying $T_{\mathbb{C}}$ gives $Tu = f$ and $Tv = 0$, so $v = 0$ by injectivity and $u = T^{-1}f$. The inverse of a linear bijection is linear, and it is bounded because

$$\|T^{-1}f\| = \|u + i0\| = \|S(f + i0)\| \le \|S\|\, \|f\|.$$

(c) Expand both sides in the inner product formula of $V_{\mathbb{C}}$:

$$\begin{aligned} \big\langle f_1 + ig_1, (T^*)_{\mathbb{C}}(f_2 + ig_2)\big\rangle &= \langle f_1, T^*f_2\rangle + \langle g_1, T^*g_2\rangle \\ &\qquad + i\big(\langle g_1, T^*f_2\rangle - \langle f_1, T^*g_2\rangle\big) \\ &= \langle Tf_1, f_2\rangle + \langle Tg_1, g_2\rangle \\ &\qquad + i\big(\langle Tg_1, f_2\rangle - \langle Tf_1, g_2\rangle\big) \\ &= \big\langle T_{\mathbb{C}}(f_1 + ig_1), f_2 + ig_2\big\rangle, \end{aligned}$$

the middle equality by $\langle u, T^*w\rangle = \langle Tu, w\rangle$ on V. Uniqueness of the adjoint now gives $(T_{\mathbb{C}})^* = (T^*)_{\mathbb{C}}$.
(d) The map $S \mapsto S_{\mathbb{C}}$ is injective, since $S_{\mathbb{C}}(f + i0) = Sf + i0$. So by (c),

$$T = T^* \iff T_{\mathbb{C}} = (T^*)_{\mathbb{C}} = (T_{\mathbb{C}})^*.$$

(e) $V_{\mathbb{C}}$ is a nonzero complex Hilbert space (Exercise 4 in Section 8B; nonzero because $\|f + i0\| = \|f\|$) and $T_{\mathbb{C}}$ is a self-adjoint bounded operator on it by (a) and (d). So 10.38 gives some $\alpha \in \operatorname{sp}(T_{\mathbb{C}})$, and 10.49 gives $\alpha \in \mathbb{R}$. Reality of α makes

$$(T - \alpha I)_{\mathbb{C}}(f + ig) = (Tf - \alpha f) + i(Tg - \alpha g) = (T_{\mathbb{C}} - \alpha I)(f + ig),$$

which is not invertible, so $T - \alpha I$ is not invertible by (b). Hence $\alpha \in \operatorname{sp}(T) \neq \emptyset$.

Problem 10B.10. Suppose T is a bounded operator on a Hilbert space V such that $\langle Tf, f\rangle \geq 0$ for all $f \in V$. Prove that $\operatorname{sp}(T) \subseteq [0, \infty)$.

Solution. The spectrum is real, and $T + cI$ is invertible for every $c > 0$; these two facts give $\operatorname{sp}(T) \subseteq [0, \infty)$.
Reality: if $\mathbb{F} = \mathbb{R}$ this is the definition 10.32; if $\mathbb{F} = \mathbb{C}$, the hypothesis makes $\langle Tf, f\rangle$ real for all f, so T is self-adjoint by 10.48 and $\operatorname{sp}(T) \subseteq \mathbb{R}$ by 10.49.
Invertibility of $T + cI$: for $f \in V$, Cauchy–Schwarz (8.11) gives

$$\|(T + cI)f\|\,\|f\| \geq \langle (T + cI)f, f\rangle \geq c\|f\|^2,$$

since $\langle Tf, f\rangle \geq 0$; dividing by $\|f\|$ when $f \neq 0$ yields $\|f\| \leq \frac{1}{c}\|(T + cI)f\|$ for all f. So $T + cI$ is left invertible by 10.29, say $R(T + cI) = I$. The adjoint obeys the same hypothesis, because

$$\langle T^*f, f\rangle = \overline{\langle f, T^*f\rangle} = \overline{\langle Tf, f\rangle} = \langle Tf, f\rangle \geq 0,$$

so the same argument gives $R'(T + cI)^* = R'(T^* + cI) = I$ (here c real), and taking adjoints gives $(T + cI)(R')^* = I$. Setting $S = (R')^*$,

$$R = R(T + cI)S = S,$$

so R is a two-sided inverse of $T + cI$.
Now if $\alpha \in \operatorname{sp}(T)$ then α is real, and $\alpha < 0$ would make $T - \alpha I = T + (-\alpha)I$ invertible. Hence $\alpha \geq 0$.

Problem 10B.11. Suppose P is a bounded operator on a Hilbert space V such that $P^2 = P$. Prove that P is self-adjoint if and only if P is normal.

Solution. Self-adjoint implies normal, since $P^* = P$ gives $P^*P = P^2 = PP^*$.
Conversely, suppose P is normal. Set $U = \operatorname{range} P$; as in Exercise 8, $P^2 = P$ makes $U = \operatorname{null}(I - P)$, a closed subspace. Its orthogonal complement is

$$U^{\perp} = \operatorname{null} P^*,$$

because $\langle Pf, g\rangle = \langle f, P^*g\rangle$ vanishes for all f exactly when $P^*g = 0$ (take $f = P^*g$). Normality gives $\|Ph\| = \|P^*h\|$ by 10.53, so

$$\operatorname{null} P = \operatorname{null} P^* = U^{\perp}.$$

Now for $f \in V$ we have $Pf \in U$ and $P(f - Pf) = Pf - P^2f = 0$, so $f - Pf \in \operatorname{null} P = U^{\perp}$. Hence 8.37(b) with $h = Pf$ gives $Pf = P_U f$. Thus $P = P_U$ is self-adjoint by the last bullet of 10.45.

Problem 10B.12. Prove that a normal operator on a separable Hilbert space has at most countably many eigenvalues.

Solution. The set E of eigenvalues injects into a countable dense subset of V. For each $\alpha \in E$ pick a unit eigenvector f_α. Eigenvectors of a normal operator for distinct eigenvalues are orthogonal by 10.57, so for $\alpha \neq \beta$ the Pythagorean Theorem (8.9) gives

$$\|f_\alpha - f_\beta\|^2 = \|f_\alpha\|^2 + \|f_\beta\|^2 = 2.$$

Separability supplies a countable dense $D \subseteq V$ (8.64); choose $d_\alpha \in D$ with $\|f_\alpha - d_\alpha\| < \sqrt{2}/2$. If $d_\alpha = d_\beta$ then

$$\|f_\alpha - f_\beta\| \leq \|f_\alpha - d_\alpha\| + \|d_\beta - f_\beta\| < \sqrt{2},$$

forcing $\alpha = \beta$. Hence $\alpha \mapsto d_\alpha$ is injective and E is at most countable.

Problem 10B.13. Prove or give a counterexample: If T is a normal operator on a Hilbert space and $T = A + iB$, where A and B are self-adjoint, then $\|T\| = \sqrt{\|A\|^2 + \|B\|^2}$.

Solution. False. On $V = \mathbb{C}^2$ take

$$T(z_1, z_2) = (z_1, iz_2), \quad A(z_1, z_2) = (z_1, 0), \quad B(z_1, z_2) = (0, z_2).$$

Then $T^*(z_1, z_2) = (z_1, -iz_2)$ by 10.45, so $T^*T = TT^* = I$ and T is normal (indeed unitary). Both A and B are orthogonal projections onto coordinate axes, hence self-adjoint by the last bullet of 10.45, and

$$(A + iB)(z_1, z_2) = (z_1, iz_2) = T(z_1, z_2).$$

Since $T^*T = I$, T is an isometry by 10.60, so $\|T\| = 1$; and $\|A\| = \|B\| = 1$, attained at $(1, 0)$ and $(0, 1)$. Thus

$$\sqrt{\|A\|^2 + \|B\|^2} = \sqrt{2} \neq 1 = \|T\|.$$

Problem 10B.14. A number $\alpha \in \mathbb{F}$ is called an *approximate eigenvalue* of a bounded operator T on a Hilbert space V if

$$\inf\bigl\{\|(T - \alpha I)f\| : f \in V \text{ and } \|f\| = 1\bigr\} = 0.$$

Suppose T is a normal operator on a Hilbert space and $\alpha \in \mathbb{F}$. Prove that $\alpha \in \operatorname{sp}(T)$ if and only if α is an approximate eigenvalue of T.

Solution. We prove both directions in contrapositive form, on the Hilbert space V on which T acts. (If $V = \{0\}$ both conditions fail for every α: $\operatorname{sp}(T) = \emptyset$ and the infimum over the empty set is ∞.) Note $T - \alpha I$ is normal, since expanding $(T - \alpha I)^* = T^* - \overline{\alpha} I$ gives

$$(T - \alpha I)^*(T - \alpha I) = T^*T - \alpha T^* - \overline{\alpha} T + |\alpha|^2 I,$$

which by $T^*T = TT^*$ equals the same product in the other order.

(i) If α is not an approximate eigenvalue, the infimum c is positive, and scaling any $f \neq 0$ to a unit vector gives $\|f\| \leq \frac{1}{c}\|(T - \alpha I)f\|$ for all $f \in V$. So $T - \alpha I$ is left invertible by 10.29, hence invertible by 10.55 as it is normal; thus $\alpha \notin \operatorname{sp}(T)$.

(ii) If $\alpha \notin \operatorname{sp}(T)$, put $S = (T - \alpha I)^{-1} \neq 0$. For $\|f\| = 1$,

$$1 = \bigl\|S(T - \alpha I)f\bigr\| \leq \|S\| \, \bigl\|(T - \alpha I)f\bigr\|,$$

so the infimum is at least $1/\|S\| > 0$ and α is not an approximate eigenvalue.

Problem 10B.15. Suppose T is a normal operator on a Hilbert space.
(a) Prove that if α is an eigenvalue of T, then $|\alpha|^2$ is an eigenvalue of T^*T.
(b) Prove that if $\alpha \in \operatorname{sp}(T)$, then $|\alpha|^2 \in \operatorname{sp}(T^*T)$.

Solution. (a) If $Tf = \alpha f$ with $f \neq 0$, then $T^*f = \overline{\alpha} f$ by 10.56 (T normal), so

$$T^*Tf = \alpha T^*f = |\alpha|^2 f.$$

(b) Exercise 14 of this section identifies the spectrum of a normal operator with its set of approximate eigenvalues, and shows in general that approximate eigenvalues lie in the spectrum. So it suffices to produce unit vectors f_n with $\|(T^*T - |\alpha|^2 I)f_n\| \to 0$.
By Exercise 14 applied to the normal T, take unit vectors f_n with $g_n := (T - \alpha I)f_n \to 0$. The operator $T - \alpha I$ is normal (Exercise 14), so 10.53 gives

$$\|h_n\| = \|(T^* - \overline{\alpha} I)f_n\| = \|(T - \alpha I)f_n\| \longrightarrow 0,$$

where $h_n := (T^* - \overline{\alpha} I)f_n$. Substituting $Tf_n = \alpha f_n + g_n$ and $T^*f_n = \overline{\alpha} f_n + h_n$,

$$T^*Tf_n = \alpha T^*f_n + T^*g_n = |\alpha|^2 f_n + \alpha h_n + T^*g_n,$$

whence

$$\|(T^*T - |\alpha|^2 I)f_n\| \le |\alpha|\,\|h_n\| + \|T^*\|\,\|g_n\| \longrightarrow 0.$$

Thus $|\alpha|^2$ is an approximate eigenvalue of T^*T, so $|\alpha|^2 \in \operatorname{sp}(T^*T)$.

Problem 10B.16. Suppose $\{e_k\}_{k \in \mathbf{Z}^+}$ is an orthonormal basis of a Hilbert space V. Suppose also that T is a normal operator on V and e_k is an eigenvector of T for every $k \ge 2$. Prove that e_1 is an eigenvector of T.

Solution. The vector Te_1 is orthogonal to every e_k with $k \ge 2$, so its expansion collapses to a multiple of e_1.
Write $Te_k = \alpha_k e_k$ for $k \ge 2$. Normality gives $T^*e_k = \overline{\alpha_k} e_k$ by 10.56, so for $k \ge 2$,

$$\langle Te_1, e_k \rangle = \langle e_1, T^*e_k \rangle = \alpha_k \langle e_1, e_k \rangle = 0,$$

by orthonormality. Hence the basis expansion 8.63(a) gives

$$Te_1 = \sum_{k=1}^{\infty} \langle Te_1, e_k \rangle e_k = \langle Te_1, e_1 \rangle e_1,$$

and $\|e_1\| = 1 \neq 0$, so e_1 is an eigenvector with eigenvalue $\langle Te_1, e_1 \rangle$.

Problem 10B.17. Prove that if T is a self-adjoint operator on a Hilbert space, then $\|T^n\| = \|T\|^n$ for every $n \in \mathbf{Z}^+$.

Solution. Get the identity first along powers of 2, then interpolate. Assume $T \neq 0$ (else both sides vanish). Each T^n is self-adjoint, since $(T^n)^* = (T^*)^n = T^n$ by 10.12(d); so Exercise 6 in Section 10A gives $\|S^2\| = \|S^*S\| = \|S\|^2$ for $S = T^{2^m}$, and induction on m yields

$$\|T^{2^{m+1}}\| = \|T^{2^m}\|^2 = \left(\|T\|^{2^m}\right)^2 = \|T\|^{2^{m+1}},$$

starting from the trivial case $m = 0$. Thus $\|T^{2^m}\| = \|T\|^{2^m}$ for all $m \ge 0$.
Now fix n; submultiplicativity (10.20) gives $\|T^n\| \le \|T\|^n$. For the converse choose m with $j := 2^m - n \ge 0$; if $j = 0$ we are done, and otherwise two applications of 10.20 give

$$\|T\|^{2^m} = \|T^n T^j\| \le \|T^n\|\,\|T^j\| \le \|T^n\|\,\|T\|^{2^m - n}.$$

Dividing by $\|T\|^{2^m-n} > 0$ gives $\|T\|^n \le \|T^n\|$.

Problem 10B.18. Prove that if T is a normal operator on a Hilbert space, then $\|T^n\| = \|T\|^n$ for every $n \in \mathbf{Z}^+$.

Solution. Let V be the Hilbert space and let $T \in \mathcal{B}(V)$ be normal, so $T^*T = TT^*$. Fix $n \in \mathbf{Z}^+$. Apply Exercise 17 of this section to the self-adjoint operator T^*T, after rewriting $(T^n)^*T^n$ as a power of T^*T.
Normality makes T^* commute with every power of T: from $T^*T = TT^*$, induction gives

$$T^*T^{k+1} = T^kT^*T = T^kTT^* = T^{k+1}T^*.$$

Hence, by a second induction using this commutation,

$$(T^*T)^{n+1} = (T^*)^nT^nT^*T = (T^*)^nT^*T^nT = (T^*)^{n+1}T^{n+1},$$

so $(T^*)^nT^n = (T^*T)^n$ for all n. Also $(T^n)^* = (T^*)^n$ and $(T^*T)^* = T^*T$, both by 10.12(d). Thus by Exercise 6 in Section 10A (used twice) and Exercise 17,

$$\|T^n\|^2 = \|(T^*)^nT^n\| = \|(T^*T)^n\| = \|T^*T\|^n = \|T\|^{2n},$$

and taking square roots gives $\|T^n\| = \|T\|^n$.

Problem 10B.19. Suppose T is an invertible operator on a Hilbert space. Prove that T is unitary if and only if $\|T\| = \|T^{-1}\| = 1$.

Solution. Both directions turn on 10.61(b): unitary is the same as surjective isometry. Assume $V \neq \{0\}$, as the norms are asserted to be 1.
(i) If T is unitary, then T is an isometry by 10.61, so $\|Tf\| = \|f\|$ gives $\|T\| = 1$ (a unit vector exists as $V \neq \{0\}$). Also $T^{-1} = T^*$ is unitary by 10.61, so the same argument gives $\|T^{-1}\| = 1$.
(ii) If $\|T\| = \|T^{-1}\| = 1$, then for $f \in V$,

$$\|Tf\| \le \|f\| = \|T^{-1}(Tf)\| \le \|Tf\|,$$

so T is an isometry, and it is surjective since invertible. Hence T is unitary by 10.61.

Problem 10B.20. Suppose T is a bounded operator on a complex Hilbert space, with $T = A + iB$, where A and B are self-adjoint (see 10.54). Prove that T is unitary if and only if T is normal and $A^2 + B^2 = I$.
[If $z = x + yi$, where $x, y \in \mathbf{R}$, then $|z| = 1$ if and only if $x^2 + y^2 = 1$. Thus this exercise strengthens the analogy between the unit circle in the complex plane and the unitary operators.]

Solution. Everything follows from expanding T^*T and TT^*. Self-adjointness of A and B with 10.12(a), (b) gives $T^* = A^* + \bar{i}B^* = A - iB$, so

$$T^*T = (A - iB)(A + iB) = (A^2 + B^2) + i(AB - BA),$$
$$TT^* = (A + iB)(A - iB) = (A^2 + B^2) - i(AB - BA).$$

(i) If T is unitary, then $T^*T = TT^* = I$, so T is normal by 10.50, whence $AB = BA$ by 10.54(b) and the first display reads $I = A^2 + B^2$.
(ii) If T is normal and $A^2 + B^2 = I$, then $AB = BA$ by 10.54(b), so both displays give $T^*T = TT^* = A^2 + B^2 = I$; by 10.58, T is unitary.

Problem 10B.21. Suppose T is a unitary operator on a complex Hilbert space such that $T - I$ is invertible. Prove that

$$i(T+I)(T-I)^{-1}$$

is a self-adjoint operator.
[The function $z \mapsto i(z+1)(z-1)^{-1}$ maps $\{z \in \mathbf{C} : |z| = 1\} \setminus \{1\}$ to $\mathbf{R}$. Thus this exercise provides another useful illustration of the analogies showing unitary $\approx \{z \in \mathbf{C} : |z| = 1\}$ and self-adjoint $\approx \mathbf{R}$.]

Solution. Set $S = i(T+I)(T-I)^{-1}$, a bounded operator since $T - I$ is invertible; we compute S^* and get S back, using $T^* = T^{-1}$ (10.61).
By 10.12(a), (b), (d), and by 10.19 (which makes $(T-I)^*$ invertible with inverse $((T-I)^{-1})^*$),

$$S^* = -i\left((T-I)^{-1}\right)^*(T+I)^* = -i\,(T^*-I)^{-1}(T^*+I).$$

Substituting $T^* = T^{-1}$ and factoring,

$$T^* - I = -T^{-1}(T-I), \qquad T^* + I = T^{-1}(T+I),$$

so inverting the first product of invertibles gives $(T^*-I)^{-1} = -(T-I)^{-1}T$ and hence

$$S^* = i\,(T-I)^{-1}\,T\,T^{-1}\,(T+I) = i\,(T-I)^{-1}(T+I).$$

Finally $T + I$ commutes with $(T-I)^{-1}$, because $(T+I)(T-I) = T^2 - I = (T-I)(T+I)$ and one may multiply by $(T-I)^{-1}$ on both sides. Therefore $S^* = i(T+I)(T-I)^{-1} = S$.

Problem 10B.22. Suppose T is a self-adjoint operator on a complex Hilbert space. Prove that

$$(T+iI)(T-iI)^{-1}$$

is a unitary operator.
[The function $z \mapsto (z+i)(z-i)^{-1}$ maps $\mathbb{R}$ to $\{z \in \mathbb{C} : |z| = 1\} \setminus \{1\}$. Thus this exercise provides another useful illustration of the analogies showing (a) unitary $\Longleftrightarrow \{z \in \mathbb{C} : |z| = 1\}$; (b) self-adjoint $\Longleftrightarrow \mathbb{R}$.]

Solution. Write $A = T + iI$, $B = T - iI$, so $U = AB^{-1}$; we show $U^*U = UU^* = I$. Self-adjointness gives $\operatorname{sp}(T) \subseteq \mathbb{R}$ by 10.49, so $\pm i \notin \operatorname{sp}(T)$ and both A and B are invertible; thus $U \in \mathcal{B}(V)$.
By 10.12(a), (b), (c) and $T^* = T$, we have $A^* = B$ and $B^* = A$; and $(S^{-1})^* = (S^*)^{-1}$ for invertible S, by taking adjoints in $SS^{-1} = S^{-1}S = I$. Hence

$$U^* = (B^{-1})^*A^* = (B^*)^{-1}B = A^{-1}B.$$

Also $AB = T^2 + I = BA$, and multiplying $AB = BA$ by B^{-1} on both sides gives $B^{-1}A = AB^{-1}$; the same trick with A^{-1} commutes all four operators. Therefore

$$U^*U = A^{-1}B\,A\,B^{-1} = I \quad \text{and} \quad UU^* = AB^{-1}\,A^{-1}B = I,$$

so U is unitary by 10.58.

Problem 10B.23. For T a bounded operator on a Banach space, define e^T by

$$e^T = \sum_{k=0}^{\infty} \frac{T^k}{k!}.$$

(a) Prove that if T is a bounded operator on a Banach space V, then the infinite sum above converges

in $\mathcal{B}(V)$ and $\|e^T\| \le e^{\|T\|}$.
(b) Prove that if S, T are bounded operators on a Banach space V such that $ST = TS$, then $e^S e^T = e^{S+T}$.
(c) Prove that if T is a self-adjoint operator on a complex Hilbert space, then e^{iT} is unitary.

Solution. Throughout, $T^0 = I$ and $\|T^k\| \le \|T\|^k$ for $k \ge 0$, by submultiplicativity of the operator norm.
(a) The series converges absolutely in the Banach space $\mathcal{B}(V)$ (6.47), since

$$\sum_{k=0}^{\infty} \left\| \frac{T^k}{k!} \right\| \le \sum_{k=0}^{\infty} \frac{\|T\|^k}{k!} = e^{\|T\|} < \infty,$$

so it converges by 6.41. The same bound applies to each partial sum $A_n = \sum_{k \le n} T^k/k!$, and continuity of the norm gives $\|e^T\| = \lim_n \|A_n\| \le e^{\|T\|}$.
(b) Commutation gives the binomial theorem

$$(S+T)^m = \sum_{k=0}^{m} \binom{m}{k} S^k T^{m-k},$$

proved by induction on m exactly as for numbers, using $T^j S = S T^j$ and Pascal's identity. (Check!) Set

$$A_n = \sum_{j=0}^{n} \frac{S^j}{j!}, \qquad B_n = \sum_{k=0}^{n} \frac{T^k}{k!}, \qquad C_n = \sum_{m=0}^{n} \frac{(S+T)^m}{m!},$$

which converge to e^S, e^T, e^{S+T} by (a). Multiplying out and using the binomial theorem,

$$A_n B_n = \sum_{j=0}^{n} \sum_{k=0}^{n} \frac{S^j T^k}{j!\,k!}, \qquad C_n = \sum_{m=0}^{n} \frac{1}{m!} \sum_{k=0}^{m} \binom{m}{k} S^k T^{m-k} = \sum_{\substack{j,k \ge 0 \\ j+k \le n}} \frac{S^j T^k}{j!\,k!}.$$

Every index pair in C_n also occurs in $A_n B_n$, so

$$A_n B_n - C_n = \sum_{(j,k) \in D_n} \frac{S^j T^k}{j!\,k!}, \qquad \text{where } D_n = \{(j,k) : 0 \le j,k \le n \text{ and } j+k > n\}.$$

Writing $a = \|S\|$ and $b = \|T\|$, the triangle inequality and $\|S^j T^k\| \le a^j b^k$ give

$$\|A_n B_n - C_n\| \le \sum_{(j,k) \in D_n} \frac{a^j b^k}{j!\,k!} = \left(\sum_{j=0}^{n} \frac{a^j}{j!} \right) \left(\sum_{k=0}^{n} \frac{b^k}{k!} \right) - \sum_{\substack{j,k \ge 0 \\ j+k \le n}} \frac{a^j b^k}{j!\,k!},$$

the same index bookkeeping, now for nonnegative reals. The first product tends to $e^a e^b$ and the subtracted sum equals $\sum_{m \le n} (a+b)^m/m! \to e^{a+b} = e^a e^b$, so $\|A_n B_n - C_n\| \to 0$. Finally $A_n B_n \to e^S e^T$, since

$$\|A_n B_n - e^S e^T\| \le \|A_n - e^S\|\,\|B_n\| + \|e^S\|\,\|B_n - e^T\| \to 0,$$

using $\|B_n\| \le e^b$ from (a). Hence $e^S e^T = \lim_n A_n B_n = \lim_n C_n = e^{S+T}$.
(c) First, $(e^A)^* = e^{A^*}$ for every $A \in \mathcal{B}(V)$: taking adjoints term by term in a partial sum (legitimate by 10.12 and $(A^k)^* = (A^*)^k$, with $1/k!$ real) gives

$$\left(\sum_{k=0}^{n} \frac{A^k}{k!} \right)^* = \sum_{k=0}^{n} \frac{(A^*)^k}{k!},$$

and the adjoint is continuous because $\|A^* - B^*\| = \|A - B\|$ by 10.11. With $A = iT$, self-adjointness gives $(iT)^* = \bar{i} T^* = -iT$, so $(e^{iT})^* = e^{-iT}$. Since iT and $-iT$ commute, part (b) gives

$$(e^{iT})^* e^{iT} = e^{-iT} e^{iT} = e^0 = I = e^{iT} e^{-iT} = e^{iT} (e^{iT})^*,$$

so e^{iT} is unitary by 10.58.

Problem 10B.24. A bounded operator T on a Hilbert space is called a partial isometry if

$$\|Tf\| = \|f\| \ \text{ for all } f \in (\operatorname{null} T)^{\perp}.$$

Suppose $(X, \mathcal{S}, \mu)$ is a σ-finite measure space and $h \in L^{\infty}(\mu)$. As usual, let $M_h \in \mathcal{B}\big(L^2(\mu)\big)$ denote the multiplication operator defined by $M_h f = fh$. Prove that M_h is a partial isometry if and only if there exists a set $E \in \mathcal{S}$ such that $|h| = \chi_E$.

Solution. The condition is that $|h| = 1$ almost everywhere on $W = \{h \neq 0\}$, and then $E = W$ works. Fix an $\mathcal{S}$-measurable representative of h and write $Z = X \setminus W = \{h = 0\}$.
First, $M_h f = 0$ if and only if $f = 0$ a.e. on W, so

$$(\operatorname{null} M_h)^{\perp} = \{f \in L^2(\mu) : f = 0 \text{ a.e. on } Z\} :$$

$\supseteq$ holds because such f and any $g \in \operatorname{null} M_h$ have $f\overline{g} = 0$ a.e.; $\subseteq$ holds because $g = \chi_Z f \in \operatorname{null} M_h$ forces $0 = \langle f, g\rangle = \int_Z |f|^2 \, d\mu$. Since $\|M_h f\|^2 = \int_X |f|^2 |h|^2 \, d\mu$, the operator M_h is a partial isometry if and only if

$$(*) \qquad \int_W |f|^2 \big(|h|^2 - 1\big) \, d\mu = 0 \quad \text{whenever } f = 0 \text{ a.e. on } Z.$$

(i) If $|h| = \chi_E$, then $W = E$ up to a null set and $|h|^2 - 1 = 0$ a.e. on W, so $(*)$ holds.
(ii) Conversely, suppose M_h is a partial isometry but $\mu(\{x \in W : |h(x)| \neq 1\}) > 0$. Since

$$\{x \in W : |h(x)| \neq 1\} = \bigcup_{n=1}^{\infty} \Big(\big\{x \in W : |h(x)| \geq 1 + \tfrac{1}{n}\big\} \cup \big\{x \in W : 0 < |h(x)| \leq 1 - \tfrac{1}{n}\big\}\Big),$$

countable subadditivity gives n and $F \in \mathcal{S}$ with $F \subseteq W$, $\mu(F) > 0$, and either $|h| \geq 1 + \frac{1}{n}$ throughout F or $|h| \leq 1 - \frac{1}{n}$ throughout F. By σ-finiteness, intersecting F with a piece of finite measure produces $A \subseteq W$ with $0 < \mu(A) < \infty$. Then $f = \chi_A$ lies in $(\operatorname{null} M_h)^{\perp}$ with $\|f\|^2 = \mu(A)$, and in the first case

$$\|M_h f\|^2 = \int_A |h|^2 \, d\mu \geq \Big(1 + \tfrac{1}{n}\Big)^2 \mu(A) > \mu(A) = \|f\|^2,$$

and in the second case,

$$\|M_h f\|^2 = \int_A |h|^2 \, d\mu \leq \Big(1 - \tfrac{1}{n}\Big)^2 \mu(A) < \mu(A) = \|f\|^2,$$

the strict inequalities using $0 < \mu(A) < \infty$. Either way $\|M_h f\| \neq \|f\|$, a contradiction. Hence $|h| = 1$ a.e. on W and $|h| = 0$ on Z, so $|h| = \chi_W$ and $E = W$ is the desired set.

Problem 10B.25. Suppose T is an isometry on a Hilbert space. Prove that T^* is a partial isometry.

Solution. The point is that $(\operatorname{null} T^*)^{\perp} = \operatorname{range} T$, on which T^* undoes T.
The range of T is closed: if $Tf_n \to g$, then $\|f_n - f_m\| = \|Tf_n - Tf_m\| \to 0$, so $f_n \to f$ for some f by completeness, and continuity gives $g = Tf$. Hence 10.13(d) gives

$$(\operatorname{null} T^*)^{\perp} = \overline{\operatorname{range} T} = \operatorname{range} T.$$

Now let $g = Tf$ be in this subspace. Since $T^*T = I$ by 10.60,

$$\|T^* g\| = \|T^*Tf\| = \|f\| = \|Tf\| = \|g\|,$$

so T^* is a partial isometry.

Problem 10B.26. Suppose T is a bounded operator on a Hilbert space V. Prove that T is a partial isometry if and only if $T^*T = P_U$ for some closed subspace U of V.

Solution. The subspace is $U = (\operatorname{null} T)^\perp$. Two identities drive both directions: $\|Sf\|^2 = \langle f, S^*Sf\rangle$ for $S \in \mathcal{B}(V)$, and, by orthogonality of $P_W f$ and $f - P_W f$ (8.43),

$$\langle f, P_W f\rangle = \|P_W f\|^2.$$

(i) If $T^*T = P_U$, then $\|Tf\|^2 = \langle f, P_U f\rangle = \|P_U f\|^2$ for all f, so $\|Tf\| = \|P_U f\|$. In particular $\operatorname{null} T = \operatorname{null} P_U = U^\perp$ by 8.45(a), and taking complements of the closed U via 8.41 gives $(\operatorname{null} T)^\perp = U$. For $f \in U$ we have $P_U f = f$, so $\|Tf\| = \|f\|$.
(ii) If T is a partial isometry, put $U = (\operatorname{null} T)^\perp$, a closed subspace with $U^\perp = \operatorname{null} T$ (by 8.41, $\operatorname{null} T$ being closed). Decomposing $f = P_U f + (f - P_U f)$ with $f - P_U f \in \operatorname{null} T$ gives $Tf = TP_U f$, and the partial-isometry hypothesis at $P_U f \in U$ gives

$$\langle f, T^*Tf\rangle = \|Tf\|^2 = \|P_U f\|^2 = \langle f, P_U f\rangle.$$

So $S = T^*T - P_U$ satisfies $\langle f, Sf\rangle = 0$ for all f; it is self-adjoint, since $(T^*T)^* = T^*T$ and $P_U^* = P_U$ (last bullet of 10.45), so $S = 0$ by 10.46.

10.3 Exercises 10C

Problem 10C.1. Prove that if T is a compact operator on a Hilbert space V and $e_1, e_2, \ldots$ is an orthonormal sequence in V, then $\lim_{n\to\infty} Te_n = 0$.

Solution. An orthonormal sequence converges weakly to 0, so Exercise 3 of this section applies. Indeed, Bessel's inequality (8.57), applied to the orthonormal family $\{e_n\}$ and any $h \in V$, gives

$$\sum_{n=1}^{\infty} |\langle h, e_n\rangle|^2 \le \|h\|^2 < \infty.$$

The terms of a convergent series tend to 0, so $\langle h, e_n\rangle \to 0$, and hence

$$\lim_{n\to\infty} \langle e_n, h\rangle = \lim_{n\to\infty} \overline{\langle h, e_n\rangle} = 0 \qquad \text{for every } h \in V.$$

Thus Exercise 3 (whose hypothesis is exactly this weak convergence, with T compact as assumed) yields $\lim_{n\to\infty} \|Te_n\| = 0$.

Problem 10C.2. Prove that if T is a compact operator on $L^2([0,1])$, then

$$\lim_{n\to\infty} \sqrt{n}\, \|T(x^n)\|_2 = 0,$$

where x^n means the element of $L^2([0,1])$ defined by $x \mapsto x^n$.

Solution. Normalize: $\|x^n\|_2^2 = \int_0^1 x^{2n}\,dx = 1/(2n+1)$, so

$$g_n = \sqrt{2n+1}\, x^n$$

has $\|g_n\|_2 = 1$, and $\sqrt{n}\,\|T(x^n)\|_2 \le 2^{-1/2}\|Tg_n\|_2$. It therefore suffices to show $g_n \to 0$ weakly and quote Exercise 3 of this section.
Fix $h \in L^2([0,1])$ and $\varepsilon > 0$. Since $|h|^2\chi_{[1-1/m,1]} \to 0$ pointwise on $[0,1)$ and is dominated by the integrable $|h|^2$, the Dominated Convergence Theorem (3.31) gives $\int_{1-\delta}^1 |h|^2 \to 0$ as $\delta \downarrow 0$; choose

$\delta \in (0,1)$ with $\left(\int_{1-\delta}^1 |h|^2\right)^{1/2} < \varepsilon/2$. Splitting at $1-\delta$ and using Cauchy–Schwarz on each piece,

$$\begin{aligned}
|\langle g_n, h\rangle| &\le \sqrt{2n+1}\int_0^{1-\delta} x^n|h|\,dx \;+\; \sqrt{2n+1}\int_{1-\delta}^1 x^n|h|\,dx \\
&\le \sqrt{2n+1}\,(1-\delta)^n\|h\|_2 \;+\; \sqrt{2n+1}\Big(\int_{1-\delta}^1 x^{2n}\Big)^{1/2}\Big(\int_{1-\delta}^1 |h|^2\Big)^{1/2} \\
&\le \sqrt{2n+1}\,(1-\delta)^n\|h\|_2 + \tfrac{\varepsilon}{2},
\end{aligned}$$

where the first term used $x^n \le (1-\delta)^n$ on $[0,1-\delta]$ together with $\int_0^1 |h| \le \|h\|_2$, and the last step used $\int_{1-\delta}^1 x^{2n} \le 1/(2n+1)$. Because $0 < 1-\delta < 1$, exponential decay beats $\sqrt{2n+1}$, so the first term is $< \varepsilon/2$ for all large n. Hence $\langle g_n, h\rangle \to 0$ for every h.
Thus $g_1, g_2, \ldots$ is a sequence converging weakly to 0, so Exercise 3 gives $\|Tg_n\|_2 \to 0$ and therefore

$$\lim_{n\to\infty} \sqrt{n}\,\|T(x^n)\|_2 = \lim_{n\to\infty} \sqrt{\tfrac{n}{2n+1}}\,\|Tg_n\|_2 = 0.$$

Problem 10C.3. Suppose T is a compact operator on a Hilbert space V and $f_1, f_2, \ldots$ is a sequence in V such that $\lim_{n\to\infty}\langle f_n, g\rangle = 0$ for every $g \in V$. Prove that $\lim_{n\to\infty}\|Tf_n\| = 0$.

Solution. The sequence $f_1, f_2, \ldots$ is bounded, and then compactness of T leaves no room for $\|Tf_n\|$ to stay away from 0.
Boundedness: define $\varphi_n \in \mathcal{B}(V, \mathbf{F})$ by $\varphi_n(g) = \langle g, f_n\rangle$, which is linear in g and satisfies $\|\varphi_n\| = \|f_n\|$ (Cauchy–Schwarz gives $\le$; the choice $g = f_n$ gives $\ge$). For each fixed g,

$$\varphi_n(g) = \overline{\langle f_n, g\rangle} \longrightarrow 0,$$

so $\sup_n |\varphi_n(g)| < \infty$ for every $g \in V$. A Hilbert space is a Banach space, so the Principle of Uniform Boundedness (6.86) applies and yields $\sup_n \|f_n\| = \sup_n \|\varphi_n\| < \infty$.
Now suppose $\|Tf_n\| \not\to 0$, so that $\|Tf_{n_k}\| \ge \varepsilon > 0$ for some subsequence. That subsequence is bounded, so compactness of T (10.66) gives a further subsequence with $Tf_{n_{k_j}} \to u$ for some $u \in V$. Since T is bounded (10.68), the adjoint T^* exists, and continuity of the inner product in its first slot gives

$$\|u\|^2 = \lim_{j\to\infty}\langle Tf_{n_{k_j}}, u\rangle = \lim_{j\to\infty}\langle f_{n_{k_j}}, T^*u\rangle = 0,$$

the last equality by the hypothesis with $g = T^*u$. Hence $u = 0$ and $\|Tf_{n_{k_j}}\| \to 0$, contradicting $\|Tf_{n_{k_j}}\| \ge \varepsilon$. Therefore $\lim_{n\to\infty}\|Tf_n\| = 0$.

Problem 10C.4. Suppose $h \in L^\infty(\mathbb{R})$. Define $M_h \in \mathcal{B}\big(L^2(\mathbb{R})\big)$ by $M_h f = fh$. Prove that if $\|h\|_\infty > 0$, then M_h is not compact.

Solution. The witnesses are normalized indicators of infinitely many disjoint sets on which $|h|$ is bounded below; their images under M_h are mutually far apart. Write λ for Lebesgue measure and put $c = \frac{1}{2}\|h\|_\infty > 0$.
The set $E = \{x : |h(x)| > c\}$ has $\lambda(E) > 0$, since $\lambda(E) = 0$ would force $\|h\|_\infty \le c < \|h\|_\infty$. Because E is the countable union of the sets $E \cap [n, n+1)$, countable additivity gives an $n_0 \in \mathbb{Z}$ with

$$E_0 = E \cap [n_0, n_0+1), \qquad 0 < \lambda(E_0) \le 1.$$

Split E_0 into infinitely many disjoint pieces: the function $\varphi(t) = \lambda\big(E_0 \cap (-\infty, t]\big)$ on $[n_0, n_0+1]$ is nondecreasing and Lipschitz, hence continuous, with $\varphi(n_0) = 0$ and $\varphi(n_0+1) = \lambda(E_0)$. By the Intermediate Value Theorem pick t_k with $\varphi(t_k) = \lambda(E_0)(1 - 2^{-k})$; these values increase strictly, so $t_1 < t_2 < \cdots$. Setting $A_1 = E_0 \cap (-\infty, t_1]$ and $A_k = E_0 \cap (t_{k-1}, t_k]$ for $k \ge 2$ gives pairwise disjoint

Borel sets with

$$\lambda(A_1) = \tfrac{1}{2}\lambda(E_0), \qquad \lambda(A_k) = 2^{-k}\lambda(E_0) \in (0,\infty).$$

Now let $f_k = \chi_{A_k}/\sqrt{\lambda(A_k)}$, so $\|f_k\|_2 = 1$ and $f_1, f_2, \ldots$ is bounded. Since $|h| > c$ on $A_k \subseteq E$,

$$\|M_h f_k\|_2^2 = \frac{1}{\lambda(A_k)} \int_{A_k} |h|^2 \, d\lambda \geq c^2,$$

and $M_h f_j, M_h f_k$ have disjoint supports for $j \neq k$, so

$$\|M_h f_j - M_h f_k\|_2^2 = \|M_h f_j\|_2^2 + \|M_h f_k\|_2^2 \geq 2c^2.$$

Hence no subsequence of $M_h f_1, M_h f_2, \ldots$ is Cauchy, so the definition of compactness (10.66) fails and M_h is not compact.

Problem 10C.5. Suppose $(b_1, b_2, \ldots) \in \ell^\infty$. Define $T\colon \ell^2 \to \ell^2$ by

$$T(a_1, a_2, \ldots) = (a_1 b_1, a_2 b_2, \ldots).$$

Prove that T is compact if and only if $\lim_{n\to\infty} b_n = 0$.

Solution. Both directions run through the standard orthonormal basis $e_1, e_2, \ldots$ of ℓ^2, for which $Te_k = b_k e_k$ and hence $\|Te_k\| = |b_k|$. (That $T \in \mathcal{B}(\ell^2)$ with $\|T\| \leq \|b\|_\infty$ follows from $\sum_k |a_k b_k|^2 \leq \|b\|_\infty^2 \|a\|^2$.)
(i) Suppose $b_n \to 0$. Let $T_n(a_1, a_2, \ldots) = (a_1 b_1, \ldots, a_n b_n, 0, 0, \ldots)$, a bounded operator whose range lies in $\operatorname{span}(e_1, \ldots, e_n)$, hence compact by 10.67. For $a \in \ell^2$,

$$\|(T - T_n)a\|^2 = \sum_{k>n} |a_k b_k|^2 \leq \Big(\sup_{k>n} |b_k|\Big)^2 \|a\|^2,$$

so $\|T - T_n\| \leq \sup_{k>n} |b_k| \to 0$. Thus T is a norm limit of compact operators, and $\mathcal{C}(\ell^2)$ is closed in $\mathcal{B}(\ell^2)$ by 10.69(a), so T is compact.
(ii) Suppose T is compact. Applying Exercise 1 of this section to the orthonormal sequence $e_1, e_2, \ldots$ gives

$$\lim_{n\to\infty} |b_n| = \lim_{n\to\infty} \|Te_n\| = 0.$$

Problem 10C.6. Suppose T is a bounded operator on a Hilbert space V. Prove that if there exists an orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V such that

$$\sum_{k\in\Gamma} \|Te_k\|^2 < \infty,$$

then T is compact.

Solution. Truncating the orthonormal basis produces finite-rank operators converging to T in norm, and $\mathcal{C}(V)$ is closed.
Write $S = \sum_{k\in\Gamma} \|Te_k\|^2 < \infty$; the terms are nonnegative, so S is the supremum of the sums over finite subsets of Γ (first bullet point following 8.53). Given $\varepsilon > 0$, choose a finite $\Omega \subseteq \Gamma$ with $\sum_{k\in\Omega} \|Te_k\|^2 > S - \varepsilon^2$; then for every finite $F \subseteq \Gamma \setminus \Omega$, disjointness of Ω and F gives

$$\sum_{k\in F} \|Te_k\|^2 = \sum_{k\in\Omega\cup F} \|Te_k\|^2 - \sum_{k\in\Omega} \|Te_k\|^2 \leq \varepsilon^2.$$

Define $P_\Omega\colon V \to V$ by

$$P_\Omega f = \sum_{k\in\Omega} \langle f, e_k\rangle e_k$$

(a finite sum), bounded because $\|P_\Omega f\|^2 = \sum_{k\in\Omega} |\langle f, e_k\rangle|^2 \le \|f\|^2$ by Bessel's inequality (8.57), with range in $\operatorname{span}\{e_k\}_{k\in\Omega}$. So TP_Ω has finite-dimensional range and is compact by 10.67.

Fix $f \in V$. Parseval's identity 8.63(a) gives $f = \sum_{k\in\Gamma}\langle f, e_k\rangle e_k$, so splitting off the finite piece indexed by Ω leaves the convergent unordered sum $f - P_\Omega f = \sum_{k\in\Gamma\setminus\Omega}\langle f, e_k\rangle e_k$ (Check!), and applying the bounded, hence continuous, operator T term by term yields

$$T(f - P_\Omega f) = \sum_{k\in\Gamma\setminus\Omega} \langle f, e_k\rangle\, Te_k.$$

For finite $F \subseteq \Gamma\setminus\Omega$, the triangle inequality, Cauchy–Schwarz, Bessel's inequality, and the tail bound above give

$$\begin{aligned}\Big\|\sum_{k\in F}\langle f, e_k\rangle Te_k\Big\| &\le \sum_{k\in F} |\langle f, e_k\rangle|\, \|Te_k\| \\ &\le \Big(\sum_{k\in F} |\langle f, e_k\rangle|^2\Big)^{1/2}\Big(\sum_{k\in F}\|Te_k\|^2\Big)^{1/2} \\ &\le \|f\|\cdot\varepsilon.\end{aligned}$$

Because $T(f - P_\Omega f)$ is the limit of such partial sums, this gives $\|(T - TP_\Omega)f\| \le \varepsilon\|f\|$ for every f, so $\|T - TP_\Omega\| \le \varepsilon$.

Taking $\varepsilon = 1/m$ yields finite sets Ω_m with TP_{Ω_m} compact and $\|T - TP_{\Omega_m}\| \le 1/m$. Since $\mathcal{C}(V)$ is closed in $\mathcal{B}(V)$ by 10.69(a), T is compact.

Problem 10C.7. Suppose T is a bounded operator on a Hilbert space V. Prove that if $\{e_k\}_{k\in\Gamma}$ and $\{f_j\}_{j\in\Omega}$ are orthonormal bases of V, then

$$\sum_{k\in\Gamma}\|Te_k\|^2 = \sum_{j\in\Omega}\|Tf_j\|^2.$$

Solution. Both sides equal $\sum_{j\in\Omega}\|T^*f_j\|^2$. All sums have nonnegative terms, so each is the supremum of its sums over finite index sets (first bullet point following 8.53) and lives in $[0,\infty]$, where the asserted equality is understood.

The interchange $\sum_{k\in\Gamma}\sum_{j\in\Omega} a_{k,j} = \sum_{j\in\Omega}\sum_{k\in\Gamma} a_{k,j}$, valid in $[0,\infty]$ for $a_{k,j}\ge 0$, is the only tool needed beyond Parseval. To see it, call the sides L and R, fix a finite $G\subseteq\Omega$, and for each $j\in G$ pick a real $M_j < \sum_{k\in\Gamma} a_{k,j}$ and a finite $F_j\subseteq\Gamma$ with $\sum_{k\in F_j} a_{k,j} > M_j$. With $F = \bigcup_{j\in G} F_j$, nonnegativity gives

$$\sum_{j\in G} M_j \le \sum_{j\in G}\sum_{k\in F} a_{k,j} = \sum_{k\in F}\sum_{j\in G} a_{k,j} \le L,$$

the middle step being a finite rearrangement. Supremum over the M_j, then over G, gives $R \le L$; symmetry gives $L \le R$.

Now let $\{e_k\}_{k\in\Gamma}$ and $\{f_j\}_{j\in\Omega}$ be any orthonormal bases. Parseval's identity 8.63(c) applied to Te_k, then the interchange, then $\langle Te_k, f_j\rangle = \langle e_k, T^*f_j\rangle$ with 8.63(c) applied to T^*f_j, give

$$\begin{aligned}\sum_{k\in\Gamma}\|Te_k\|^2 &= \sum_{k\in\Gamma}\sum_{j\in\Omega}|\langle Te_k, f_j\rangle|^2 \\ &= \sum_{j\in\Omega}\sum_{k\in\Gamma}|\langle T^*f_j, e_k\rangle|^2 = \sum_{j\in\Omega}\|T^*f_j\|^2.\end{aligned}$$

Call this $(\star)$. Using $(\star)$ with both bases equal to $\{f_j\}_{j\in\Omega}$ and then with the given pair,

$$\sum_{j\in\Omega}\|Tf_j\|^2 = \sum_{j\in\Omega}\|T^*f_j\|^2 = \sum_{k\in\Gamma}\|Te_k\|^2.$$

Problem 10C.8. Suppose T is a bounded operator on a Hilbert space. Prove that T is compact if and only if T^*T is compact.

Solution. (i) If T is compact, then T^*T is compact by 10.69(b), since T^* is bounded.
(ii) Suppose T^*T is compact, and let $f_1, f_2, \dots$ be a bounded sequence in V, say $M = \sup_n \|f_n\| < \infty$. Compactness gives a subsequence along which $T^*Tf_{n_1}, T^*Tf_{n_2}, \dots$ converges, hence is Cauchy. Writing $h = f_{n_j} - f_{n_k}$, so that $\|h\| \le 2M$, Cauchy–Schwarz gives

$$\|Tf_{n_j} - Tf_{n_k}\|^2 = \langle T^*Th, h\rangle \le \|T^*Th\|\,\|h\| \le 2M\,\|T^*Tf_{n_j} - T^*Tf_{n_k}\|.$$

The right side tends to 0 as $j, k \to \infty$, so $Tf_{n_1}, Tf_{n_2}, \dots$ is Cauchy and converges by completeness of V. Hence T is compact.

Problem 10C.9. Prove that if T is a compact operator on an infinite-dimensional Hilbert space, then $\|I - T\| \ge 1$.

Solution. Test $I - T$ on an orthonormal sequence, along which $\|Te_n\| \to 0$.
Since V is infinite-dimensional, no finite list spans it, so a linearly independent sequence $f_1, f_2, \dots$ may be chosen inductively; Gram–Schmidt (see the proof of 8.67) turns it into an orthonormal sequence $e_1, e_2, \dots$. By Exercise 1 of this section, $\lim_{n\to\infty} \|Te_n\| = 0$. Since $\|e_n\| = 1$,

$$\|I - T\| \ge \|(I - T)e_n\| \ge \|e_n\| - \|Te_n\| = 1 - \|Te_n\|,$$

and letting $n \to \infty$ gives $\|I - T\| \ge 1$.

Problem 10C.10. Show that if T is a surjective but not injective operator on a vector space V, then

$$\operatorname{null} T \subsetneq \operatorname{null} T^2 \subsetneq \operatorname{null} T^3 \subsetneq \cdots.$$

Solution. Pull a fixed nonzero null vector back through T^n. Fix $f \ne 0$ with $Tf = 0$, available because T is not injective.
The inclusions $\operatorname{null} T^n \subseteq \operatorname{null} T^{n+1}$ are immediate from $T^{n+1}f = T(T^n f)$. For strictness, fix $n \in \mathbb{Z}^+$; since T is surjective so is T^n, so there is $g \in V$ with $T^n g = f$. Then

$$T^{n+1}g = Tf = 0 \quad \text{while} \quad T^n g = f \ne 0,$$

so $g \in \operatorname{null} T^{n+1} \setminus \operatorname{null} T^n$. As n was arbitrary, every inclusion in the chain is strict.

Problem 10C.11. Suppose T is a compact operator on a Hilbert space and $\alpha \in \mathbf{F} \setminus \{0\}$.
(a) Prove that $\operatorname{range}(T - \alpha I)^{m-1} = \operatorname{range}(T - \alpha I)^m$ for some $m \in \mathbb{Z}^+$.
(b) Prove that $\operatorname{null}(T - \alpha I)^{n-1} = \operatorname{null}(T - \alpha I)^n$ for some $n \in \mathbb{Z}^+$.
(c) Show that the smallest positive integer m that works in (a) equals the smallest positive integer n that works in (b).

Solution. Write $S = T - \alpha I$ on the Hilbert space V, with $S^0 = I$. The engine is the Binomial Theorem: for $k \in \mathbb{Z}^+$,

$$S^k = C_k - \beta_k I, \qquad C_k = \sum_{j=1}^{k} \binom{k}{j} (-\alpha)^{k-j} T^j, \qquad \beta_k = -(-\alpha)^k \ne 0,$$

where each term of C_k is a bounded operator composed with the compact T, hence compact by 10.69(b), so C_k is compact by 10.69(a). Therefore $\operatorname{range} S^k$ is closed by 10.77 and $\operatorname{null} S^k$ is finite-dimensional by 10.82 (trivially so for $k = 0$).
(a) The ranges decrease, since $S^k f = S^{k-1}(Sf)$. Suppose every inclusion were strict. Fix n; then

$U = \operatorname{range} S^n$ is closed in the complete space V, hence a Hilbert space by 6.16(b), and $\operatorname{range} S^{n+1}$ is a proper closed subspace of U. Splitting some $f \in U \setminus \operatorname{range} S^{n+1}$ by 8.43 inside U and normalizing the (nonzero) component orthogonal to $\operatorname{range} S^{n+1}$ gives

$$f_n \in \operatorname{range} S^n \cap \left(\operatorname{range} S^{n+1}\right)^{\perp}, \qquad \|f_n\| = 1.$$

For $j < k$, writing $T = S + \alpha I$ gives $Tf_j - Tf_k = u + \alpha f_j$ with $u = Sf_j - Sf_k - \alpha f_k$. Here $Sf_j \in \operatorname{range} S^{j+1}$, while f_k and Sf_k lie in $\operatorname{range} S^k \subseteq \operatorname{range} S^{j+1}$ because $k \geq j+1$; so $u \perp f_j$ and the Pythagorean identity gives

$$\|Tf_j - Tf_k\|^2 = \|u\|^2 + |\alpha|^2 \geq |\alpha|^2 > 0.$$

Then $Tf_1, Tf_2, \ldots$ has no Cauchy subsequence though $f_1, f_2, \ldots$ is bounded, contradicting compactness of T.
(b) The nulls increase, and each $\operatorname{null} S^k$ is closed since S^k is bounded. Suppose every inclusion were strict. The same splitting, applied by 8.43 inside the Hilbert space $\operatorname{null} S^n$ to some $f \in \operatorname{null} S^n \setminus \operatorname{null} S^{n-1}$, gives

$$e_n \in \operatorname{null} S^n \cap \left(\operatorname{null} S^{n-1}\right)^{\perp}, \qquad \|e_n\| = 1.$$

For $j < k$, $Te_k - Te_j = v + \alpha e_k$ with $v = Se_k - Se_j - \alpha e_j$. Here $Se_k \in \operatorname{null} S^{k-1}$ because $S^{k-1}(Se_k) = 0$, while $e_j \in \operatorname{null} S^j \subseteq \operatorname{null} S^{k-1}$ and $Se_j \in \operatorname{null} S^{j-1} \subseteq \operatorname{null} S^{k-1}$; so $v \perp e_k$ and Pythagoras gives $\|Te_k - Te_j\| \geq |\alpha| > 0$, the same contradiction.
(c) Put $d_k = \dim \operatorname{null} S^k$ and $c_k = \dim(\operatorname{range} S^k)^{\perp}$. These agree and are finite: both vanish at $k = 0$, and for $k \geq 1$, 10.13(a) gives

$$\left(\operatorname{range} S^k\right)^{\perp} = \operatorname{null}(S^k)^* = \operatorname{null}\left(C_k^* - \overline{\beta_k} I\right),$$

with C_k^* compact by 10.73, so both spaces are finite-dimensional by 10.82, and 10.91 applied to the compact C_k and the nonzero β_k gives $d_k = c_k$.
Moreover $\operatorname{range} S^{k-1} = \operatorname{range} S^k$ exactly when $c_{k-1} = c_k$: the inclusion $(\operatorname{range} S^{k-1})^{\perp} \subseteq (\operatorname{range} S^k)^{\perp}$ between finite-dimensional spaces of equal dimension is an equality, and taking orthogonal complements of the closed ranges recovers them by 8.41. Likewise $\operatorname{null} S^{k-1} = \operatorname{null} S^k$ exactly when $d_{k-1} = d_k$, the nulls being nested and finite-dimensional. So the smallest m in (a) is the least k with $c_{k-1} = c_k$, the smallest n in (b) is the least k with $d_{k-1} = d_k$, and $d_k = c_k$ for all k forces $m = n$.

Problem 10C.12. Prove that if $f : [0,1] \to \mathbf{F}$ is a continuous function, then there exists a continuous function $g : [0,1] \to \mathbf{F}$ such that

$$f(x) = g(x) + \int_0^x g$$

for all $x \in [0,1]$.

Solution. Take

$$g(x) = f(x) - h(x), \qquad h(x) = e^{-x} \int_0^x e^t f(t)\, dt,$$

which is continuous because f is. Since f is continuous, the Fundamental Theorem of Calculus gives $h(0) = 0$ and

$$h'(x) = -e^{-x} \int_0^x e^t f(t)\, dt + e^{-x} e^x f(x) = f(x) - h(x) = g(x),$$

so $\int_0^x g = h(x) - h(0) = h(x)$ and therefore

$$g(x) + \int_0^x g = f(x) - h(x) + h(x) = f(x)$$

for all $x \in [0,1]$.

Method (2): let $\mathcal{V}$ be the Volterra operator $(\mathcal{V}h)(x) = \int_0^x h$ on $L^2([0,1])$ (Example 10.15), which is compact (comment after the proof of 10.70) with $\operatorname{sp}(\mathcal{V}) = \{0\}$ (Example 10.89). Hence $-1 \notin \operatorname{sp}(\mathcal{V})$, so $I + \mathcal{V}$ is invertible in $\mathcal{B}(L^2([0,1]))$, and since the continuous f is bounded on $[0,1]$, hence in $L^2([0,1])$, there is $g_0 \in L^2([0,1])$ with $g_0 + \mathcal{V}g_0 = f$. Cauchy–Schwarz on the measure-one interval $[0,1]$ gives $\|g_0\|_1 \le \|g_0\|_2 < \infty$, so g_0 is integrable and $x \mapsto \int_0^x g_0$ is continuous by the Dominated Convergence Theorem (3.31), the integrands $\chi_{[0,x_n]}g_0$ being dominated by $|g_0|$. Thus $g(x) = f(x) - \int_0^x g_0$ is continuous, equals g_0 almost everywhere by the equation above, and so satisfies $g(x) + \int_0^x g = g(x) + \int_0^x g_0 = f(x)$ at every $x \in [0,1]$.

Problem 10C.13. Suppose S is a bounded invertible operator on a Hilbert space V and T is a compact operator on V.
(a) Prove that $S + T$ has closed range.
(b) Prove that $S + T$ is injective if and only if $S + T$ is surjective.
(c) Prove that $\operatorname{null}(S + T)$ and $\operatorname{null}(S^* + T^*)$ are finite-dimensional.
(d) Prove that $\dim \operatorname{null}(S + T) = \dim \operatorname{null}(S^* + T^*)$.
(e) Prove that there exists $R \in \mathcal{B}(V)$ such that $\operatorname{range} R$ is finite-dimensional and $S + T + R$ is invertible.

Solution. Everything follows from the two factorizations

$$S + T = S(I + K) = (I + L)S, \qquad K = S^{-1}T, \quad L = TS^{-1},$$

in which K and L are compact by 10.69(b), since S^{-1} is bounded [10.18 with the Bounded Inverse Theorem 6.83], so S is a homeomorphism of V onto V. Also S^* is invertible with $(S^*)^{-1} = (S^{-1})^*$, and T^* is compact by 10.73.
(a) By 10.77 applied to the compact K and $\alpha = -1 \neq 0$, the subspace $\operatorname{range}(I + K)$ is closed, and

$$\operatorname{range}(S + T) = S\big(\operatorname{range}(I + K)\big)$$

is the inverse image of that closed set under the continuous S^{-1}, hence closed.
(b) Since S is bijective, $S + T = S(I + K)$ is injective (surjective) exactly when $I + K$ is. The Fredholm Alternative 10.85, applied to the compact K with $\alpha = -1$, makes "-1 is an eigenvalue of K" equivalent to "$K + I$ is not surjective", that is, $I + K$ fails injectivity exactly when it fails surjectivity.
(c) Injectivity of S gives $\operatorname{null}(S + T) = \operatorname{null}(K - (-1)I)$, finite-dimensional by 10.82 (compact K, $\alpha = -1 \neq 0$). The same argument with S^*, T^* in place of S, T gives $\operatorname{null}(S^* + T^*) = \operatorname{null}(I + (S^*)^{-1}T^*)$, finite-dimensional because $(S^*)^{-1}T^*$ is compact by 10.69(b).
(d) Chain the three identifications

$$\begin{aligned} \operatorname{null}(I + L) &= S\big(\operatorname{null}(S + T)\big), \\ \dim \operatorname{null}(I + L) &= \dim \operatorname{null}(I + L^*), \\ \operatorname{null}(S^* + T^*) &= \operatorname{null}(I + L^*). \end{aligned}$$

The first holds because $(I + L)Sh = (S + T)h$ and S is a linear isomorphism of V, so dimensions match; the second is 10.91 for the compact L with $\alpha = -1 = \overline{-1}$; the third holds because $(S + T)^* = S^*(I + L^*)$ with S^* injective. Hence $\dim \operatorname{null}(S + T) = \dim \operatorname{null}(S^* + T^*)$.
(e) Take $R = 0$ if $N = \operatorname{null}(S + T)$ is $\{0\}$, since then (b) makes $S + T$ bijective, hence invertible by 10.18 with bounded inverse by 6.83. Otherwise let $d = \dim N \ge 1$, finite by (c), and note that by (a) and 10.13(a),

$$\big(\operatorname{range}(S + T)\big)^{\perp} = \operatorname{null}(S^* + T^*),$$

which has dimension d by (d). Pick orthonormal bases $e_1, \dots, e_d$ of N (Gram–Schmidt) and $u_1, \dots, u_d$ of $(\operatorname{range}(S + T))^{\perp}$, and set

$$Rf = \sum_{j=1}^{d} \langle f, e_j \rangle u_j,$$

so $\|Rf\|^2 = \sum_{j=1}^{d} |\langle f, e_j\rangle|^2 \le \|f\|^2$ by the Pythagorean identity and Bessel's inequality (8.57), and range $R \subseteq \operatorname{span}\{u_1, \dots, u_d\}$ is finite-dimensional.
Then $S+T+R$ is injective: if $(S+T)f + Rf = 0$, the two summands are orthogonal, so both vanish; $Rf = 0$ forces $\langle f, e_j\rangle = 0$ for all j, while $(S+T)f = 0$ puts $f \in N$, whence $f = \sum_j \langle f, e_j\rangle e_j = 0$. Since R is compact by 10.67 and hence $T+R$ is compact by 10.69(a), part (b) applied to $T+R$ makes $S+T+R$ surjective as well, hence invertible by 10.18 and 6.83.

Problem 10C.14. Suppose T is a compact operator on a Hilbert space V. Prove that range T is a separable subspace of V.

Solution. Write range $T = \bigcup_{n=1}^{\infty} T(B_n)$, where $B_n = \{f \in V : \|f\| \le n\}$; it is enough that each $T(B_n)$ be separable, since a countable union of countable dense sets is one.
Each $T(B_n)$ is totally bounded. Otherwise some $\varepsilon > 0$ admits no finite cover of $T(B_n)$ by ε-balls centred in it, so one can choose $f_1, f_2, \dots \in B_n$ inductively (starting at $f_1 = 0$, with f_{k+1} outside the ε-balls about $Tf_1, \dots, Tf_k$) with

$$\|Tf_j - Tf_k\| \ge \varepsilon \qquad \text{whenever } j \ne k,$$

so $Tf_1, Tf_2, \dots$ has no convergent subsequence although $f_1, f_2, \dots$ is bounded, contradicting compactness (10.66).
A totally bounded set A is separable: choosing for each k a finite $F_k \subseteq A$ whose $1/k$-balls cover A makes $\bigcup_k F_k$ a countable dense subset (Check!). So each $T(B_n)$ has a countable dense subset D_n, and $D = \bigcup_n D_n$ is a countable dense subset of range T: any $g = Tf$ lies in $T(B_n)$ once $n \ge \|f\|$. Hence range T is separable.

Problem 10C.15. Suppose T is a compact operator on a Hilbert space V and $e_1, e_2, \dots$ is an orthonormal basis of range T. Let P_n denote the orthogonal projection of V onto $\operatorname{span}\{e_1, \dots, e_n\}$.
(a) Prove that $\lim_{n\to\infty} \|T - P_nT\| = 0$.
(b) Prove that a bounded operator on a Hilbert space V is compact if and only if it is the limit in $\mathcal{B}(V)$ of a sequence of bounded operators with finite-dimensional range.

Solution. Write $U = \overline{\operatorname{range} T}$, a closed subspace of V and hence a Hilbert space of which $e_1, e_2, \dots$ is an orthonormal basis.
(a) Two facts drive the argument. First, $\|I - P_n\| \le 1$, because $f - P_nf \perp P_nf$ makes $\|f\|^2 = \|P_nf\|^2 + \|f - P_nf\|^2$. Second, $\|h - P_nh\| \to 0$ for every $h \in U$: by 8.71 and Parseval,

$$\|h - P_nh\|^2 = \sum_{k>n} |\langle h, e_k\rangle|^2 \longrightarrow 0,$$

the series $\sum_k |\langle h, e_k\rangle|^2 = \|h\|^2$ being convergent.
Now if $\|T - P_nT\| \not\to 0$, there are $\varepsilon > 0$, indices $n_1 < n_2 < \cdots$, and vectors f_j with $\|f_j\| \le 1$ and $\|(I - P_{n_j})Tf_j\| > \varepsilon$. Compactness of T gives a subsequence with $Tf_{j_i} \to g$, and $g \in U$ because U is closed. Then

$$\varepsilon < \|(I - P_{n_{j_i}})Tf_{j_i}\| \le \|Tf_{j_i} - g\| + \|g - P_{n_{j_i}}g\|,$$

using $\|I - P_n\| \le 1$ on the first summand. Both terms on the right tend to 0 (the second because $g \in U$ and $n_{j_i} \to \infty$), a contradiction. Hence $\lim_{n\to\infty} \|T - P_nT\| = 0$.
(b) If $\|S - S_n\| \to 0$ with each S_n bounded of finite-dimensional range, then each S_n is compact by 10.67 and $\mathcal{C}(V)$ is closed in $\mathcal{B}(V)$ by 10.69(a), so S is compact.
Conversely let T be compact. By Exercise 14 of this section range T is separable, hence so is its closure U. If U is finite-dimensional, the constant sequence $T, T, \dots$ does the job. Otherwise 8.67 gives an orthonormal basis of U, necessarily countable (distinct basis vectors are $\sqrt{2}$ apart, so the balls of radius $\frac{1}{2}$ about them are disjoint and each captures a distinct point of a fixed countable dense set) and infinite; list it as $e_1, e_2, \dots$ and let P_n project onto $\operatorname{span}\{e_1, \dots, e_n\}$. Part (a) used only that $e_1, e_2, \dots$ is an orthonormal basis of U, so $\|T - P_nT\| \to 0$, and each P_nT is bounded with finite-dimensional range.

Problem 10C.16. Prove that if T is a compact operator on a Hilbert space V, then there exists a sequence $S_1, S_2, \ldots$ of invertible operators on V such that $\lim_{n\to\infty} \|T - S_n\| = 0$.

Solution. Take $S_n = T - \alpha_n I$, where $\alpha_n \notin \operatorname{sp}(T)$ satisfies $0 < |\alpha_n| < 1/n$. Such α_n exists because $\{\alpha \in \mathbf{F} : 0 < |\alpha| < 1/n\}$ contains the uncountable real interval $(0, 1/n)$, while $\operatorname{sp}(T)$ is countable: T is compact, so $\{\alpha \in \operatorname{sp}(T) : |\alpha| \geq 1/m\}$ is finite by 10.93, and

$$\operatorname{sp}(T) \subseteq \{0\} \cup \bigcup_{m=1}^{\infty} \Big\{\alpha \in \operatorname{sp}(T) : |\alpha| \geq \tfrac{1}{m}\Big\}$$

is a countable union of finite sets together with one point.
Each S_n is invertible, since $\alpha_n \notin \operatorname{sp}(T)$, and

$$\|T - S_n\| = |\alpha_n|\, \|I\| \leq |\alpha_n| < \tfrac{1}{n} \longrightarrow 0.$$

Problem 10C.17. Suppose T is a bounded operator on a Hilbert space such that $p(T)$ is compact for some nonzero polynomial p with coefficients in $\mathbf{F}$. Prove that $\operatorname{sp}(T)$ is a countable set.

Solution. $\operatorname{sp}(T)$ sits inside the p-preimage of the countable set $\operatorname{sp}(p(T))$, and every fibre of p is finite. If p is a nonzero constant c, then $I = c^{-1}p(T)$ is compact [$\mathcal{C}(V)$ is a subspace, 10.69(a)], so V is finite-dimensional (the identity on an infinite-dimensional Hilbert space is not compact; remark after 10.67), where $\operatorname{sp}(T)$ is the set of eigenvalues of T, at most $\dim V$ in number. So assume $d = \deg p \geq 1$.
(i) If $\alpha \in \operatorname{sp}(T)$ then $p(\alpha) \in \operatorname{sp}(p(T))$. Division by the root α gives a polynomial r over $\mathbf{F}$ with $p(z) - p(\alpha) = (z - \alpha)r(z)$, and polynomials in T commute, so

$$p(T) - p(\alpha)I = (T - \alpha I)\, r(T) = r(T)\, (T - \alpha I).$$

If the left side had inverse R, then $T - \alpha I$ would have right inverse $r(T)R$ and left inverse $R\,r(T)$; these coincide, since $B = B(T - \alpha I)A = A$ for a left inverse B and right inverse A, so $T - \alpha I$ would be invertible, contradicting $\alpha \in \operatorname{sp}(T)$.
(ii) $\operatorname{sp}(p(T))$ is countable, since $p(T)$ is compact makes each $\{\beta \in \operatorname{sp}(p(T)) : |\beta| \geq 1/m\}$ finite by 10.93, and $\operatorname{sp}(p(T))$ is contained in $\{0\}$ together with the union of those sets over $m \in \mathbf{Z}^+$.
(iii) For each β, the set $\{z \in \mathbf{F} : p(z) = \beta\}$ is the root set of the nonzero degree-d polynomial $p - \beta$, hence has at most d elements. By (i),

$$\operatorname{sp}(T) \subseteq \bigcup_{\beta \in \operatorname{sp}(p(T))} \{z \in \mathbf{F} : p(z) = \beta\},$$

a countable union of finite sets, hence countable.

Problem 10C.18. Suppose $T \in \mathcal{B}(\ell^2)$ is defined by $T(a_1, a_2, a_3, \ldots) = (a_2, a_3, a_4, \ldots)$. Suppose also that $\alpha \in \mathbf{F}$ and $|\alpha| < 1$.
(a) Show that the geometric multiplicity of α as an eigenvalue of T equals 1.
(b) Show that the algebraic multiplicity of α as an eigenvalue of T equals ∞.

Solution. Write $S = T - \alpha I$; the geometric multiplicity is $\dim \operatorname{null} S$ (10.81) and the algebraic multiplicity is $\dim \bigcup_n \operatorname{null}(S^n)$.
(a) The eigenvector is $w = (1, \alpha, \alpha^2, \ldots)$, which lies in ℓ^2 because $\sum_{n\geq 0} |\alpha|^{2n} = (1 - |\alpha|^2)^{-1} < \infty$, and satisfies $Tw = \alpha w$. Conversely, $Tf = \alpha f$ for $f = (a_1, a_2, \ldots)$ reads $a_{n+1} = \alpha a_n$, so $a_n = \alpha^{n-1}a_1$ and $f = a_1 w$. Hence $\operatorname{null} S = \operatorname{span}\{w\}$ has dimension 1.
(b) The operative fact is that S is surjective, with an explicit right inverse. Let R be the right shift $R(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$, an isometry, so $\|R^j\| \leq 1$ and $TR = I$, whence $TR^{j+1} = R^j$. Since $\sum_j |\alpha|^j < \infty$ and $\mathcal{B}(\ell^2)$ is a Banach space, $C = \sum_{j=0}^{\infty} \alpha^j R^{j+1}$ converges in $\mathcal{B}(\ell^2)$, and applying the

continuous T and αI term by term gives

$$\begin{aligned} SC &= \sum_{j=0}^{\infty} \alpha^j T R^{j+1} - \sum_{j=0}^{\infty} \alpha^{j+1} R^{j+1} \\ &= \sum_{j=0}^{\infty} \alpha^j R^j - \sum_{i=1}^{\infty} \alpha^i R^i = I. \end{aligned}$$

Now $\dim \operatorname{null}(S^n) = n$ by induction, the case $n = 1$ being (a). Given the case n, the map $\Phi f = Sf$ sends $\operatorname{null}(S^{n+1})$ into $\operatorname{null}(S^n)$ with null space $\operatorname{null} S$, of dimension 1, and is surjective because $g \in \operatorname{null}(S^n)$ has the preimage $f = Cg$, which satisfies $Sf = g$ and $S^{n+1} f = S^n g = 0$. So

$$\dim \operatorname{null}(S^{n+1}) = 1 + \dim \operatorname{null}(S^n) = n + 1.$$

Hence $\bigcup_n \operatorname{null}(S^n)$ contains subspaces of every finite dimension, so the algebraic multiplicity is ∞.

Problem 10C.19. Prove that the geometric multiplicity of an eigenvalue of a normal operator on a Hilbert space equals the algebraic multiplicity of that eigenvalue.

Solution. With $S = T - \alpha I$, the two multiplicities agree because $\operatorname{null}(S^n) = \operatorname{null}(S)$ for every n, so the increasing union $\bigcup_n \operatorname{null}(S^n)$ collapses to $\operatorname{null}(S)$.
S is normal: by 10.12(a)–(c), $S^* = T^* - \overline{\alpha} I$, and

$$\begin{aligned} SS^* &= TT^* - \overline{\alpha} T - \alpha T^* + |\alpha|^2 I, \\ S^* S &= T^* T - \overline{\alpha} T - \alpha T^* + |\alpha|^2 I, \end{aligned}$$

which agree because $TT^* = T^* T$.
Hence $\operatorname{null}(S^2) = \operatorname{null}(S)$. Indeed, if $S^2 f = 0$, then $g = Sf$ satisfies $Sg = 0$, so $S^* g = 0$ as well (trivially if $g = 0$, and by 10.56 applied to the normal S with eigenvalue 0 otherwise), whence

$$\|Sf\|^2 = \langle f, S^* S f \rangle = \langle f, S^* g \rangle = 0.$$

Induction finishes: if $\operatorname{null}(S^n) = \operatorname{null}(S)$ and $S^{n+1} f = 0$, then $Sf \in \operatorname{null}(S^n) = \operatorname{null}(S)$, so $S^2 f = 0$ and hence $Sf = 0$ by the previous paragraph.

Problem 10C.20. Prove that every nonzero eigenvalue of a compact operator on a Hilbert space has finite algebraic multiplicity.

Solution. With $S = T - \alpha I$ and $\operatorname{null}(S^0) = \{0\}$, the increasing chain $\operatorname{null}(S^k)$ stabilizes at a finite stage, and every term of it is finite-dimensional.
Finite-dimensionality: the Binomial Theorem gives $S^k = K_k - \beta_k I$, where $\beta_k = -(-\alpha)^k \neq 0$ and

$$K_k = \sum_{j=1}^{k} \binom{k}{j} (-\alpha)^{k-j} T^j$$

is compact [each T^j with $j \geq 1$ is compact by 10.69(b), and $\mathcal{C}(V)$ is a subspace of $\mathcal{B}(V)$ by 10.69(a)]. So 10.82, applied to K_k and $\beta_k \neq 0$, makes $\operatorname{null}(S^k)$ finite-dimensional.
Stabilization: suppose instead that every inclusion $\operatorname{null}(S^k) \subseteq \operatorname{null}(S^{k+1})$ were strict. Each null space is closed, S^k being continuous, so for every n we may take $h \in \operatorname{null}(S^n) \setminus \operatorname{null}(S^{n-1})$ and normalize $h - Ph$, where P projects onto $\operatorname{null}(S^{n-1})$, obtaining by 8.43

$$e_n \in \operatorname{null}(S^n) \cap \left(\operatorname{null}(S^{n-1})\right)^{\perp}, \qquad \|e_n\| = 1.$$

For $j < k$, writing $T = S + \alpha I$ gives $Te_j - Te_k = w - \alpha e_k$ with $w = Se_j - Se_k + \alpha e_j$. All three summands of w lie in $\operatorname{null}(S^{k-1})$: $Se_j \in \operatorname{null}(S^{j-1})$, $Se_k \in \operatorname{null}(S^{k-1})$, and $e_j \in \operatorname{null}(S^j)$, with $j \leq k - 1$. Since

$e_k \perp \text{null}(S^{k-1})$, the Pythagorean identity gives

$$\|Te_j - Te_k\|^2 = \|w\|^2 + |\alpha|^2 \geq |\alpha|^2 > 0,$$

so $Te_1, Te_2, \ldots$ has no convergent subsequence although $e_1, e_2, \ldots$ is bounded, contradicting compactness of T. Hence $\text{null}(S^n) = \text{null}(S^{n+1})$ for some $n \in \mathbf{Z}^+$.
The chain is then constant from n on: if $\text{null}(S^m) = \text{null}(S^{m+1})$ and $S^{m+2}f = 0$, then $Sf \in \text{null}(S^{m+1}) = \text{null}(S^m)$, so $S^{m+1}f = 0$. Therefore

$$\bigcup_{k=1}^{\infty} \text{null}(T - \alpha I)^k = \text{null}(S^n),$$

which is finite-dimensional.

Problem 10C.21. Prove that if T is a compact operator on a Hilbert space and α is a nonzero eigenvalue of T, then the algebraic multiplicity of α as an eigenvalue of T equals the algebraic multiplicity of $\overline{\alpha}$ as an eigenvalue of T^*.

Solution. Both multiplicities equal $\dim \text{null}(S^N)$ for a common stabilizing index N, where $S = T - \alpha I$, so that $S^* = T^* - \overline{\alpha} I$ by 10.12(a)–(c) and $(S^k)^* = (S^*)^k$ by 10.12(d).
Here T^* is compact by 10.73, and $\overline{\alpha}$ is a nonzero eigenvalue of T^*: $\alpha \in \text{sp}(T)$ gives $\overline{\alpha} \in \text{sp}(T^*)$ by 10.19, and the Fredholm Alternative 10.85 applied to the compact T^* with $\overline{\alpha} \neq 0$ upgrades that to an eigenvalue (remark following 10.89). So Exercise 20 of this section applies to both (T, α) and $(T^*, \overline{\alpha})$: the chains $\text{null}(S^k)$ and $\text{null}((S^*)^k)$ are each eventually constant, say from index N on, and

$$\bigcup_{k=1}^{\infty} \text{null}(S^k) = \text{null}(S^N), \qquad \bigcup_{k=1}^{\infty} \text{null}\big((S^*)^k\big) = \text{null}\big((S^*)^N\big).$$

The Binomial Theorem writes $S^N = K - \beta I$ with $\beta = -(-\alpha)^N \neq 0$ and $K = \sum_{j=1}^{N} \binom{N}{j} (-\alpha)^{N-j} T^j$ compact [10.69(a),(b)], so $(S^*)^N = (S^N)^* = K^* - \overline{\beta} I$ with K^* compact by 10.73. Now 10.91, applied to the compact K and the nonzero β, gives

$$\dim \text{null}(S^N) = \dim \text{null}(K - \beta I) = \dim \text{null}(K^* - \overline{\beta} I) = \dim \text{null}\big((S^*)^N\big).$$

Problem 10C.22. Prove that if V is a separable Hilbert space, then $\mathcal{C}(V)$, the Banach space of compact operators on V, is separable.

Solution. The countable dense set consists of the finite-rank operators

$$A_c f = \sum_{j,k \in \Gamma_n} c_{jk} \langle f, e_k \rangle e_j$$

whose matrix entries c_{jk} are rational. Assume $V \neq \{0\}$, the other case being trivial.
By 8.67 and its proof, V has an orthonormal basis $\{e_k\}_{k \in \Gamma}$ with $\Gamma \subseteq \mathbb{Z}^+$. Put $\Gamma_n = \{k \in \Gamma : k \leq n\}$, a finite set, and let P_n be the orthogonal projection onto the finite-dimensional, hence closed, subspace $U_n = \text{span}\{e_k\}_{k \in \Gamma_n}$, so that $P_n f = \sum_{k \in \Gamma_n} \langle f, e_k \rangle e_k$ by 8.71. Three properties are used below:
(i) $P_n f \to f$ for every $f \in V$: Parseval's identity 8.63(a) supplies a finite $G \subseteq \Gamma$ beyond which the partial sums are within ε of f, and $G \subseteq \Gamma_n$ for all large n.
(ii) $\|P_n\| \leq 1$ by 8.37(d), and $\|I - P_n\| \leq 1$ because $f - P_n f \perp P_n f$ [8.37(a)] and the Pythagorean Theorem 8.9.
(iii) $P_n^* = P_n$ (Check! from the displayed formula), so $I - P_n$ is self-adjoint too by 10.12(a)–(c).
Step (1): $\|T - TP_n\| \to 0$ for every $T \in \mathcal{C}(V)$. Otherwise there are $\varepsilon > 0$, indices $n_1 < n_2 < \cdots$, and vectors f_j with $\|f_j\| \leq 1$ and $\|Tg_j\| > \varepsilon$, where $g_j = (I - P_{n_j}) f_j$ has $\|g_j\| \leq 1$ by (ii). Compactness of T gives a subsequence with $Tg_{j_i} \to h$. For any $g \in V$, the self-adjointness (iii) of $I - P_n$ and

Cauchy–Schwarz give

$$|\langle Tg_{j_i}, g\rangle| = \left|\langle f_{j_i}, (I - P_{n_{j_i}})T^*g\rangle\right| \le \|T^*g - P_{n_{j_i}}T^*g\| \longrightarrow 0$$

by (i), so $\langle h, g\rangle = 0$ for every g, whence $h = 0$. But $\|h\| = \lim_i \|Tg_{j_i}\| \ge \varepsilon$, a contradiction.
Step (2): $\|T - P_nTP_n\| \to 0$. Split $T - P_nTP_n = (T - TP_n) + (I - P_n)TP_n$. For the second term, $\|P_n\| \le 1$ and 10.12(d) with (iii) and 10.11 give

$$\|(I - P_n)TP_n\| \le \|(I - P_n)T\| = \|T^* - T^*P_n\|.$$

Since T^* is compact by 10.73, Step (1) applies to T and to T^*, so

$$\|T - P_nTP_n\| \le \|T - TP_n\| + \|T^* - T^*P_n\| \longrightarrow 0.$$

Step (3): density. For any $\Gamma_n \times \Gamma_n$ matrix d over $\mathbf{F}$, orthonormality, Cauchy–Schwarz in $\mathbf{F}^{\Gamma_n}$, and Bessel's inequality 8.57 give

$$\begin{aligned}\|A_d f\|^2 &= \sum_{j\in\Gamma_n}\Big|\sum_{k\in\Gamma_n} d_{jk}\langle f, e_k\rangle\Big|^2\\ &\le \sum_{j\in\Gamma_n}\Big(\sum_{k\in\Gamma_n}|d_{jk}|^2\Big)\Big(\sum_{k\in\Gamma_n}|\langle f, e_k\rangle|^2\Big) \le \|f\|^2\sum_{j,k\in\Gamma_n}|d_{jk}|^2,\end{aligned}$$

so $\|A_d\| \le (\sum_{j,k}|d_{jk}|^2)^{1/2}$; in particular each A_c is bounded with range inside U_n, hence compact by 10.67. Let Q be $\mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$ according as $\mathbf{F}$ is $\mathbb{R}$ or $\mathbb{C}$, a countable dense subset of $\mathbf{F}$, and let D be the (countable) set of all A_c with c a $\Gamma_n \times \Gamma_n$ matrix over Q, $n \in \mathbb{Z}^+$.
Given $T \in \mathcal{C}(V)$ and $\varepsilon > 0$, Step (2) supplies n with $\|T - P_nTP_n\| < \varepsilon/2$, and expanding P_n on both sides shows $P_nTP_n = A_b$ with $b_{jk} = \langle Te_k, e_j\rangle$. Choosing c over Q with $\sum_{j,k\in\Gamma_n}|b_{jk} - c_{jk}|^2 < \varepsilon^2/4$, the estimate above and the linearity of $c \mapsto A_c$ give $\|P_nTP_n - A_c\| < \varepsilon/2$ and hence $\|T - A_c\| < \varepsilon$. Since $D \subseteq \mathcal{C}(V)$, which is closed in $\mathcal{B}(V)$ by 10.69(a), D is a countable dense subset of $\mathcal{C}(V)$, which is therefore separable (8.64).

10.4 Exercises 10D

Problem 10D.1. Prove that if T is a compact operator on a nonzero Hilbert space, then $\|T\|^2$ is an eigenvalue of T^*T.

Solution. By 10.96, $\|T\|^2 \in \operatorname{sp}(T^*T)$, and T^*T is compact by 10.69(b) (T compact, T^* bounded by 10.11 since T is bounded by 10.68).
(i) If $\|T\| \ne 0$, the Fredholm Alternative 10.85, applied to the compact T^*T and the nonzero scalar $\|T\|^2$, turns membership in the spectrum into eigenvalue-hood.
(ii) If $\|T\| = 0$, then $T^*T = 0$ and any nonzero $f \in V$ (one exists because $V \ne \{0\}$) satisfies $(T^*T)f = 0 = \|T\|^2 f$.

Problem 10D.2. Prove that if T is a self-adjoint operator on a nonzero Hilbert space V, then

$$\|T\| = \sup\{|\langle Tf, f\rangle| : f \in V \text{ and } \|f\| = 1\}.$$

Solution. Write M for the supremum in question; it is taken over the nonempty unit sphere of V, and $\|T\| < \infty$ because self-adjointness presupposes boundedness (10.44). Since $V \ne \{0\}$, homogeneity also gives $\|T\| = \sup\{\|Tf\| : \|f\| = 1\}$ (Check!, from the definition 6.43).
$M \le \|T\|$: for $\|f\| = 1$, Cauchy–Schwarz 8.11 gives $|\langle Tf, f\rangle| \le \|Tf\| \le \|T\|$. Rescaling g to $g/\|g\|$ then upgrades the definition of M to

$$|\langle Tg, g\rangle| \le M\|g\|^2 \qquad \text{for every } g \in V.$$

$\|T\| \leq M$: expanding $\langle T(g \pm h), g \pm h\rangle$ by additivity in each slot and subtracting, with $\langle Th, g\rangle = \langle h, Tg\rangle = \overline{\langle Tg, h\rangle}$ by self-adjointness,

$$4\,\mathrm{Re}\langle Tg, h\rangle = \langle T(g+h), g+h\rangle - \langle T(g-h), g-h\rangle,$$

so the previous display and the parallelogram equality 8.20 give

$$4|\,\mathrm{Re}\langle Tg, h\rangle| \leq M\|g+h\|^2 + M\|g-h\|^2 = 2M\big(\|g\|^2 + \|h\|^2\big).$$

Now take g with $\|g\| = 1$ and $Tg \neq 0$ (the case $Tg = 0$ being trivial), and set $h = Tg/\|Tg\|$, a unit vector with $\langle Tg, h\rangle = \|Tg\|$ real. The display becomes $4\|Tg\| \leq 4M$, so $\|Tg\| \leq M$; the supremum over unit g gives $\|T\| \leq M$.

Problem 10D.3. Suppose T is a bounded operator on a Hilbert space V and U is a closed subspace of V. Prove that the following are equivalent.
(a) U is an invariant subspace for T.
(b) $U^\perp$ is an invariant subspace for T^*.
(c) $TP_U = P_U T P_U$.

Solution. Prove (a) $\Leftrightarrow$ (b) and (a) $\Leftrightarrow$ (c), using 8.45(a) ($\mathrm{range}\, P_U = U$, $\mathrm{null}\, P_U = U^\perp$) and $P_U g = g$ for $g \in U$.
(a) $\Rightarrow$ (b). For $h \in U^\perp$ and $g \in U$, invariance puts $Tg \in U$, so

$$\langle T^* h, g\rangle = \langle h, Tg\rangle = 0,$$

and $g \in U$ was arbitrary, so $T^* h \in U^\perp$.
(b) $\Rightarrow$ (a). Apply the implication just proved to the bounded operator T^* and the closed subspace $U^\perp$ (closed by 8.40): it makes $(U^\perp)^\perp$ invariant for $(T^*)^*$. But $(T^*)^* = T$ by 10.11 and $(U^\perp)^\perp = U$ by 8.41, U being closed.
(a) $\Rightarrow$ (c). For $f \in V$, invariance gives $TP_U f \in U$, hence $P_U T P_U f = TP_U f$.
(c) $\Rightarrow$ (a). For $f \in U$ we have $P_U f = f$, so

$$Tf = TP_U f = P_U T P_U f = P_U(Tf) \in \mathrm{range}\, P_U = U.$$

Problem 10D.4. Suppose T is a bounded operator on a Hilbert space V and U is a closed subspace of V. Prove that the following are equivalent.
(a) U and $U^\perp$ are invariant subspaces for T.
(b) U and $U^\perp$ are invariant subspaces for T^*.
(c) $TP_U = P_U T$.

Solution. Throughout, 8.45 gives $\mathrm{range}\, P_U = U$, $\mathrm{null}\, P_U = U^\perp$, $P_U g = g$ for $g \in U$, and $I - P_U = P_{U^\perp}$.
(a) $\Leftrightarrow$ (b): Exercise 3 applied to the closed subspace U says U is invariant for T iff $U^\perp$ is invariant for T^*; applied to the closed subspace $U^\perp$ (closed by 8.40, with $(U^\perp)^\perp = U$ by 8.41 since U is closed) it says $U^\perp$ is invariant for T iff U is invariant for T^*. Conjoining gives (a) $\Leftrightarrow$ (b).
(a) $\Rightarrow$ (c): For $f \in V$, decompose $f = P_U f + P_{U^\perp} f$ by 8.43; invariance puts $TP_U f \in U$ and $TP_{U^\perp} f \in U^\perp = \mathrm{null}\, P_U$, so

$$P_U T f = P_U(TP_U f) + P_U(TP_{U^\perp} f) = TP_U f + 0.$$

(c) $\Rightarrow$ (a): For $f \in U$ we get $Tf = TP_U f = P_U T f \in \mathrm{range}\, P_U = U$; for $f \in U^\perp = \mathrm{null}\, P_U$ we get $P_U T f = TP_U f = 0$, so $Tf \in U^\perp$.

Problem 10D.5. Suppose T is a bounded operator on a nonseparable normed vector space V. Prove that T has a closed invariant subspace other than $\{0\}$ and V.

Solution. Take $U = \overline{W}$, where $W = \operatorname{span}\{T^n f : n \geq 0\}$ (finite linear combinations, $T^0 = I$) for any fixed $f \neq 0$; such f exists because $\{0\}$ is separable and V is not.
U is a closed subspace, since the closure of a subspace of a normed space is a subspace (Check!), and $U \neq \{0\}$ because $f \in W$.
U is invariant: $T(W) \subseteq W$ because T is linear and sends each spanning vector $T^n f$ to $T^{n+1} f$; and if $g_n \in W$ with $g_n \to g \in U$, then $Tg_n \to Tg$ by boundedness of T, so $Tg \in \overline{W} = U$.
$U \neq V$, because U is separable while V is not. Indeed, let $\mathbf{Q_F}$ be $\mathbf{Q}$ or $\mathbf{Q} + i\mathbf{Q}$ according as $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$, a countable dense subset of $\mathbf{F}$, and let $D \subseteq W$ consist of the sums $\sum_{n=0}^{N} c_n T^n f$ with $N \geq 0$ and all $c_n \in \mathbf{Q_F}$; then D is a countable union (over N) of images of the countable sets $(\mathbf{Q_F})^{N+1}$, hence countable. Given $w = \sum_{n=0}^{N} a_n T^n f \in W$ and $\varepsilon > 0$, put $C = 1 + \sum_{n=0}^{N} \|T^n f\|$ and pick $c_n \in \mathbf{Q_F}$ with $|a_n - c_n| < \varepsilon / C$; then

$$\Big\| w - \sum_{n=0}^{N} c_n T^n f \Big\| \leq \sum_{n=0}^{N} |a_n - c_n| \, \|T^n f\| < \varepsilon.$$

Thus D is dense in W, which is dense in U, so D is dense in U.

Problem 10D.6. Suppose T is an operator on a Banach space V with dimension greater than 2. Prove that T has an invariant subspace other than $\{0\}$ and V.
[For this exercise, T is not assumed to be bounded and the invariant subspace is not required to be closed.]

Solution. Fix $f \neq 0$ (possible as $\dim V > 2$) and set $U = \operatorname{span}\{T^n f : n \geq 0\}$, with $T^0 = I$. Then U is a subspace, $U \neq \{0\}$ since $f \in U$, and U is invariant because T carries each spanning vector $T^n f$ to $T^{n+1} f \in U$. Only properness is at issue.
(i) V infinite-dimensional. Then $U \neq V$, so U is the desired subspace. Otherwise $V = U = \bigcup_{n=1}^{\infty} V_n$ with $V_n = \operatorname{span}\{f, Tf, \ldots, T^{n-1} f\}$, since every element of U is a finite linear combination of the $T^k f$. Each V_n has dimension at most n, hence is closed (Exercise 9 in Section 6D) and proper; a proper subspace W has empty interior, for if $B(g, r) \subseteq W$ then every $h \neq 0$ satisfies $g + \frac{r}{2\|h\|} h \in W$, whence $h \in W$ and $W = V$. So the Banach space V would be a countable union of closed sets with empty interior, contradicting Baire's theorem 6.76(a).
(ii) $\dim V = n$ with $n \geq 3$. Assume $U = V$ (else take U). Let m be least with $f, Tf, \ldots, T^m f$ dependent; $m \leq n$ exists since any $n+1$ vectors are dependent. Then $f, \ldots, T^{m-1} f$ is independent and $T^m f \in W := \operatorname{span}\{f, \ldots, T^{m-1} f\}$, so $T^k f \in W$ for all k by induction ($T^{k+1} f \in T(W) \subseteq \operatorname{span}\{Tf, \ldots, T^m f\} \subseteq W$); hence $V = U = W$ and $m = n$, making $f, Tf, \ldots, T^{n-1} f$ a basis. Writing $T^n f = -(a_0 f + \cdots + a_{n-1} T^{n-1} f)$ and $p(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_0$ gives $p(T) f = 0$, hence $p(T) T^k f = T^k p(T) f = 0$ for all k and so $p(T) = 0$ on $V = U$.
Factor $p = q_1 q_2 \cdots q_r$ into monic irreducibles over $\mathbf{F}$, so $d := \deg q_1 \leq 2$ (degree 1 if $\mathbf{F} = \mathbf{C}$; degree 1 or 2 if $\mathbf{F} = \mathbf{R}$); since no polynomial of degree $n \geq 3$ is irreducible over $\mathbf{R}$ or $\mathbf{C}$, we have $r \geq 2$, and $s := q_2 \cdots q_r$ is monic with $1 \leq \deg s = n - d \leq n - 1$. Observe that a monic q with $1 \leq \deg q = e \leq n - 1$ has $q(T) f \neq 0$, because $q(T) f$ is a combination of $f, \ldots, T^e f$ with coefficient 1 on $T^e f$ and $f, \ldots, T^{n-1} f$ is independent.
Now $N := \operatorname{null} q_1(T)$ is invariant (if $q_1(T) g = 0$ then $q_1(T) T g = T q_1(T) g = 0$); $N \neq V$ since $q_1(T) f \neq 0$; and $N \neq \{0\}$ since $g := s(T) f \neq 0$ while $q_1(T) g = p(T) f = 0$, both by the observation.

Problem 10D.7. Suppose T is a self-adjoint compact operator on a Hilbert space that has only finitely many distinct eigenvalues. Prove that T has finite-dimensional range.

Solution. By 10.106(b) there are a countable Ω, an orthonormal family $\{e_k\}_{k \in \Omega}$ in the Hilbert space V, and $\{\alpha_k\}_{k \in \Omega}$ in $\mathbf{R} \setminus \{0\}$ with

$$Tf = \sum_{k \in \Omega} \alpha_k \langle f, e_k \rangle e_k \qquad \text{for all } f \in V;$$

it suffices to show Ω is finite, since then $\operatorname{range} T \subseteq \operatorname{span}\{e_k : k \in \Omega\}$, of dimension at most $\#\Omega$.

Taking $f = e_j$ and using orthonormality kills all terms but $k = j$, giving $Te_j = \alpha_j e_j$ with $e_j \neq 0$; so each α_k is a nonzero eigenvalue of T. By hypothesis these take only finitely many distinct values $\lambda_1, \ldots, \lambda_p$. For each j, the family $\{e_k : \alpha_k = \lambda_j\}$ is orthonormal, hence linearly independent, and lies in $\operatorname{null}(T - \lambda_j I)$, which is finite-dimensional by 10.82 since T is compact and $\lambda_j \neq 0$. Thus each $\{k : \alpha_k = \lambda_j\}$ is finite, and Ω is their union over $j \leq p$.

Problem 10D.8. (a) Prove that if T is a self-adjoint compact operator on a Hilbert space, then there exists a self-adjoint compact operator S such that $S^3 = T$.
(b) Prove that if T is a normal compact operator on a complex Hilbert space, then there exists a normal compact operator S such that $S^2 = T$.

Solution. In both parts S is the diagonal operator obtained by taking the corresponding root of each eigenvalue of T. The analytic content is the following lemma.
Lemma. Suppose $\{e_k\}_{k\in\Omega}$ is an orthonormal family in a Hilbert space V, where Ω is a countable set, and suppose $\{\beta_k\}_{k\in\Omega}$ is a family in $\mathbb{F}$ such that

$$\{k \in \Omega : |\beta_k| \geq \delta\} \text{ is finite for every } \delta > 0.$$

Then the formula

$$Sf = \sum_{k\in\Omega} \beta_k \langle f, e_k\rangle e_k$$

defines a compact operator S on V with $\|S\| \leq \sup_{k\in\Omega} |\beta_k|$, and

- (i) $S^* f = \sum_{k\in\Omega} \overline{\beta_k} \langle f, e_k\rangle e_k$ for all $f \in V$;
- (ii) $S^n f = \sum_{k\in\Omega} \beta_k^n \langle f, e_k\rangle e_k$ for all $f \in V$ and all $n \in \mathbb{Z}^+$;
- (iii) $Se_j = \beta_j e_j$ for each $j \in \Omega$, and $Sg = 0$ for every $g \in \left(\overline{\operatorname{span}}\{e_k\}_{k\in\Omega}\right)^\perp$.

Proof of Lemma. Taking $\delta = 1$ in the hypothesis bounds all but finitely many $|\beta_k|$ by 1; hence $M := \sup_{k\in\Omega} |\beta_k| < \infty$. For $f \in V$, Bessel's inequality 8.57 gives

$$\sum_{k\in\Omega} |\beta_k \langle f, e_k\rangle|^2 \leq M^2 \sum_{k\in\Omega} |\langle f, e_k\rangle|^2 \leq M^2 \|f\|^2 < \infty,$$

so the unordered sum defining Sf converges by 8.54(a), and 8.54(b) gives $\|Sf\|^2 = \sum_{k\in\Omega} |\beta_k|^2 |\langle f, e_k\rangle|^2 \leq M^2 \|f\|^2$; linearity is clear (Check!), so $\|S\| \leq M$.
Continuity of the inner product plus orthonormality give $\langle Sf, e_j\rangle = \beta_j \langle f, e_j\rangle$ for $j \in \Omega$. Taking $f = e_j$ there, and using $\langle g, e_k\rangle = 0$ for all k when $g \perp \{e_k\}$, gives (iii); feeding $S^n f$ into it gives (ii) by induction. For (i), let S' be the operator defined by $\{\overline{\beta_k}\}_{k\in\Omega}$, which satisfies the same hypothesis; expanding both inner products through the convergent unordered sums,

$$\begin{aligned}\langle Sf, g\rangle &= \sum_{k\in\Omega} \beta_k \langle f, e_k\rangle \langle e_k, g\rangle \\ &= \sum_{k\in\Omega} \langle f, e_k\rangle \overline{\overline{\beta_k} \langle g, e_k\rangle} = \langle f, S'g\rangle.\end{aligned}$$

valid for all $f, g \in V$, so $S^* = S'$.
For compactness: if Ω is finite then range S is finite-dimensional, so S is compact by 10.67; otherwise enumerate $\Omega = \{k_1, k_2, \ldots\}$, note $\beta_{k_j} \to 0$ by hypothesis, and set $S_n f = \sum_{j=1}^n \beta_{k_j} \langle f, e_{k_j}\rangle e_{k_j}$, compact by 10.67. Since $S - S_n$ is the diagonal operator for the family agreeing with $\{\beta_k\}$ on $\{k_j : j > n\}$ and 0 elsewhere, the norm estimate gives $\|S - S_n\| \leq \sup_{j>n} |\beta_{k_j}| \to 0$; and $\mathcal{C}(V)$ is closed in $\mathcal{B}(V)$ by 10.69(a).
Both parts also use: if $\{e_k\}_{k\in\Omega}$ is orthonormal with $Te_k = \alpha_k e_k$ and each $\alpha_k \neq 0$, then $\{k : |\alpha_k| \geq \delta\}$ is finite for every $\delta > 0$. Indeed eigenvalues lie in $\operatorname{sp}(T)$, so only finitely many distinct eigenvalues α satisfy $|\alpha| \geq \delta$ by 10.93 (T compact); and for each such α the orthonormal, hence independent, family $\{e_k : \alpha_k = \alpha\}$ lies in $\operatorname{null}(T - \alpha I)$, finite-dimensional by 10.82 since $\alpha \neq 0$.
(a) By 10.106(b) write $Tf = \sum_{k\in\Omega} \alpha_k \langle f, e_k\rangle e_k$ with Ω countable, $\{e_k\}_{k\in\Omega}$ orthonormal and $\alpha_k \in \mathbb{R}\backslash\{0\}$; taking $f = e_j$ gives $Te_j = \alpha_j e_j$, so the criterion above applies. Let β_k be the real cube root of α_k, so

$|\beta_k| \geq \delta \iff |\alpha_k| \geq \delta^3$ and the lemma's hypothesis holds. The resulting S is compact, is self-adjoint by (i) since each β_k is real, and satisfies $S^3 = T$ by (ii) since $\beta_k^3 = \alpha_k$.
(b) By 10.107 there is an orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V with $Te_k = \alpha_k e_k$, $\alpha_k \in \mathbb{C}$. Parseval 8.63(a) gives $f = \sum_{k\in\Gamma}\langle f, e_k\rangle e_k$, and applying the continuous operator T term by term yields $Tf = \sum_{k\in\Gamma}\alpha_k\langle f, e_k\rangle e_k = \sum_{k\in\Omega}\alpha_k\langle f, e_k\rangle e_k$, where $\Omega = \{k \in \Gamma : \alpha_k \neq 0\}$ is countable by the criterion above. Choose $\beta_k \in \mathbb{C}$ with $\beta_k^2 = \alpha_k$; then $|\beta_k| \geq \delta \iff |\alpha_k| \geq \delta^2$, so the lemma applies and its (ii) gives $S^2 = T$. By (iii), $Se_j = \beta_j e_j$ for $j \in \Omega$ and $Se_j = 0$ for $j \in \Gamma \setminus \Omega$ (such e_j being orthogonal to every e_k, $k \in \Omega$), so the orthonormal basis $\{e_k\}_{k\in\Gamma}$ consists of eigenvectors of S; hence S is normal by 10.103(b).

Problem 10D.9. Suppose T is a compact normal operator on a nonzero Hilbert space V. Prove that there is a subspace of V with dimension 1 or 2 that is an invariant subspace for T.
[If $\mathbb{F} = \mathbb{C}$, the desired result follows immediately from the Spectral Theorem for compact normal operators. Thus you can assume that $\mathbb{F} = \mathbb{R}$.]

Solution. If $\mathbb{F} = \mathbb{C}$, then 10.107 supplies an orthonormal basis of eigenvectors of T, nonempty since $V \neq \{0\}$; the span of any one of them is a 1-dimensional invariant subspace (10.101, second bullet). So assume $\mathbb{F} = \mathbb{R}$ and set $A = (T + T^*)/2$, $B = (T - T^*)/2$, so $T = A + B$, $A^* = A$, $B^* = -B$, and A, B are compact because T^* is compact (10.73) and $\mathcal{C}(V)$ is a subspace of $\mathcal{B}(V)$ (10.69(a)). Normality $TT^* = T^*T$ gives

$$TA = \tfrac{1}{2}(T^2 + TT^*) = \tfrac{1}{2}(T^2 + T^*T) = AT, \qquad T^*A = AT^*,$$

the second by the same computation, whence $BA = \frac{1}{2}(TA - T^*A) = \frac{1}{2}(AT - AT^*) = AB$.
By 10.99 applied to the self-adjoint compact A on $V \neq \{0\}$, either $\|A\|$ or $-\|A\|$ is an eigenvalue α of A; put $U = \operatorname{null}(A - \alpha I) \neq \{0\}$, a closed subspace since A is bounded, hence a nonzero Hilbert space. For $f \in U$, $(A - \alpha I)Bf = B(A - \alpha I)f = 0$, so U is invariant for B, and likewise for T since $AT = TA$. Let $C = B|_U$, compact on U because U is a closed invariant subspace of the compact B (discussion preceding 10.102). For $f, g \in U$, $\langle Cf, g\rangle = \langle f, B^*g\rangle = \langle f, -Cg\rangle$, so $C^* = -C$; hence $(C^2)^* = (-C)^2 = C^2$, and C^2 is compact by 10.69(b), with

$$\langle C^2 g, g\rangle = \langle Cg, C^*g\rangle = -\|Cg\|^2 \leq 0 \qquad (g \in U).$$

By 10.99 on $U \neq \{0\}$, pick an eigenvalue λ of C^2 and $g \neq 0$ in U with $C^2 g = \lambda g$; the display forces $\lambda\|g\|^2 \leq 0$, so $\lambda \leq 0$. Also $Tf = \alpha f + Cf$ for every $f \in U$.
(i) $\lambda = 0$. Then $\|Cg\|^2 = -\langle C^2 g, g\rangle = 0$, so $Tg = \alpha g$ and $\operatorname{span}\{g\}$ is invariant of dimension 1.
(ii) $\lambda < 0$. Take $W = \operatorname{span}\{g, Cg\}$. Then $\dim W = 2$, since $Cg = cg$ with $c \in \mathbb{R}$ would give $c^2 = \lambda < 0$; and W is invariant because

$$Tg = \alpha g + Cg \in W, \qquad T(Cg) = \alpha\, Cg + C^2 g = \lambda g + \alpha\, Cg \in W.$$

Problem 10D.10. Suppose T is a self-adjoint compact operator on a Hilbert space and $\|T\| \leq \frac{1}{4}$. Prove that there exists a self-adjoint compact operator S such that $S^2 + S = T$.

Solution. Take S to be the diagonal operator with eigenvalues $\varphi(\alpha_k)$, where α_k are the eigenvalues of T and

$$\varphi(t) = \frac{-1 + \sqrt{1 + 4t}}{2} = \frac{2t}{1 + \sqrt{1 + 4t}}$$

(the second form by rationalizing), defined for $t \in [-\frac{1}{4}, \frac{1}{4}]$ since there $1 + 4t \geq 0$. Writing $u = \sqrt{1 + 4t}$,

$$\varphi(t)^2 + \varphi(t) = \frac{u^2 - 2u + 1}{4} + \frac{2u - 2}{4} = \frac{u^2 - 1}{4} = t,$$

and $1 + u \geq 1$ on that interval gives $|\varphi(t)| \leq 2|t|$.
By 10.106(b) write $Tf = \sum_{k\in\Omega}\alpha_k\langle f, e_k\rangle e_k$ with Ω countable, $\{e_k\}$ orthonormal and $\alpha_k \in \mathbb{R} \setminus \{0\}$. Taking $f = e_j$ gives $Te_j = \alpha_j e_j$, so $|\alpha_k| \leq \|T\| \leq \frac{1}{4}$ and $\varphi(\alpha_k)$ is defined; moreover $\{k : |\alpha_k| \geq \delta\}$ is finite

for each $\delta > 0$, since 10.93 leaves only finitely many distinct such eigenvalues α and each $\operatorname{null}(T - \alpha I)$ is finite-dimensional by 10.82 ($\alpha \neq 0$) while $\{e_k : \alpha_k = \alpha\}$ is orthonormal, hence independent.
Set $\beta_k = \varphi(\alpha_k) \in \mathbb{R}$. Then $|\beta_k| \leq 2|\alpha_k|$, so $\{k : |\beta_k| \geq \delta\} \subseteq \{k : |\alpha_k| \geq \delta/2\}$ is finite for every $\delta > 0$, and the lemma of Exercise 8 produces a compact operator $Sf = \sum_{k\in\Omega} \beta_k \langle f, e_k\rangle e_k$, self-adjoint by part (i) of that lemma since the β_k are real. By part (ii) of the lemma and $\beta_k^2 + \beta_k = \alpha_k$,

$$(S^2 + S)f = \sum_{k\in\Omega} (\beta_k^2 + \beta_k)\langle f, e_k\rangle e_k = Tf.$$

Problem 10D.11. For $k \in \mathbb{Z}$, define $g_k \in L^2\big((-\pi, \pi]\big)$ and $h_k \in L^2\big((-\pi, \pi]\big)$ by

$$g_k(t) = \frac{1}{\sqrt{2\pi}}\, e^{it/2} e^{ikt} \qquad \text{and} \qquad h_k(t) = \frac{1}{\sqrt{2\pi}}\, e^{ikt};$$

here we are assuming that $\mathbb{F} = \mathbb{C}$.
(a) Use the conclusion of Example 10.108 to show that $\{g_k\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2\big((-\pi, \pi]\big)$.
(b) Use (a) to show that $\{h_k\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2\big((-\pi, \pi]\big)$.
(c) Use (b) to show that the orthonormal family in the third bullet point of Example 8.51 is an orthonormal basis of $L^2\big((-\pi, \pi]\big)$.

Solution. Each part transports the previous basis by a surjective linear isometry, which carries orthonormal bases to orthonormal bases: such a Ψ preserves inner products by polarization, and Ψ, Ψ^{-1} continuous gives $\Psi(\overline{\operatorname{span}} A) = \overline{\operatorname{span}}\,\Psi(A)$. We also use the change of variables $\int \varphi(ax + b)\, dx = \frac{1}{a} \int \varphi$ for $a > 0$ and $\varphi \geq 0$ measurable, which holds for $\varphi = \chi_E$ by translation invariance 2.7 and $|cE| = c|E|$, then extends by linearity and 3.11.
(a) Define $\Phi : L^2([0, 1]) \to L^2((-\pi, \pi])$ by $(\Phi f)(t) = \frac{1}{\sqrt{2\pi}} f(\frac{t}{2\pi} + \frac{1}{2})$; as t ranges over $(-\pi, \pi]$, the argument ranges over $(0, 1]$. The change of variables with $a = \frac{1}{2\pi}$, $b = \frac{1}{2}$, applied to $|f|^2$ extended by 0, gives

$$\|\Phi f\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Big| f\Big(\frac{t}{2\pi} + \frac{1}{2}\Big)\Big|^2 dt = \int_0^1 |f|^2 = \|f\|^2,$$

so Φ is a linear isometry, surjective since $(\Phi^{-1} v)(x) = \sqrt{2\pi}\, v(2\pi x - \pi)$ is a two-sided inverse. With $u_k(x) = e^{i(2k+1)\pi x}$, which form an orthonormal basis of $L^2([0, 1])$ by Example 10.108,

$$(\Phi u_k)(t) = \frac{1}{\sqrt{2\pi}} e^{i(2k+1)t/2} e^{i(2k+1)\pi/2} = c_k\, g_k(t), \qquad |c_k| = 1.$$

So $\{\Phi u_k\}$ is an orthonormal basis, and $g_k = \overline{c_k}\, \Phi u_k$ with $|c_k| = 1$ leaves orthonormality and the span unchanged.
(b) Multiplication $Mv = mv$ by $m(t) = e^{-it/2}$ is a linear isometry of $L^2((-\pi, \pi])$ since $|m| = 1$, bijective with inverse multiplication by $\overline{m}$; and $Mg_k = h_k$ (Check!), so (a) gives the conclusion.
(c) Those e_k are orthonormal by 8.51, so only the span is at issue. Here $e_0 = h_0$, and for $n \in \mathbb{Z}^+$, from $e^{\pm int} = \cos(nt) \pm i \sin(nt)$ and $\sqrt{2\pi}/(2\sqrt{\pi}) = 1/\sqrt{2}$,

$$e_n = \frac{1}{i\sqrt{2}}(h_n - h_{-n}), \qquad e_{-n} = \frac{1}{\sqrt{2}}(h_n + h_{-n}),$$

which solve to $h_{\pm n} = \frac{1}{\sqrt{2}}(e_{-n} \pm i\, e_n)$. Hence the two families have the same span, so the same closed span, which is $L^2((-\pi, \pi])$ by (b).

Problem 10D.12. Suppose T is a compact operator on a Hilbert space. Prove that $s_1(T) = \|T\|$.

Solution. The operator T^*T on V is compact (10.73 and 10.69(b)) and self-adjoint, and its eigenvalues obey two bounds.

Upper: if $T^*Tf = \alpha f$ with $f \neq 0$, then

$$\alpha\|f\|^2 = \langle T^*Tf, f\rangle = \|Tf\|^2 \leq \|T\|^2\|f\|^2,$$

so $\alpha \leq \|T\|^2$ (and $\alpha \geq 0$).
Lower: if $\|T\| \neq 0$, then $\|T\|^2 \in \operatorname{sp}(T^*T)$ by 10.96, and since $\|T\|^2 \neq 0$ and T^*T is compact, the Fredholm Alternative 10.85 makes $\|T\|^2$ a positive eigenvalue of T^*T.
(i) T^*T has no positive eigenvalue. Then $s_1(T) = 0$ by 10.116 (second bullet), and $\|T\| \neq 0$ would contradict the lower bound, so $\|T\| = 0 = s_1(T)$.
(ii) T^*T has a positive eigenvalue. The set P of positive eigenvalues has a maximum: fixing $\alpha_0 \in P$, the set $\{\alpha \in P : \alpha \geq \alpha_0\}$ is finite by 10.93 and nonempty, and its largest element dominates all of P. By 10.116 (first bullet) the singular values are the positive square roots of the elements of P in decreasing order, so $s_1(T) = \sqrt{\max P}$. Here $T \neq 0$, so the lower bound gives $\|T\|^2 \in P$ and hence $\|T\|^2 \leq \max P$, while the upper bound applied to $\max P$ gives $\max P \leq \|T\|^2$. Thus $s_1(T)^2 = \|T\|^2$ with both quantities nonnegative.

Problem 10D.13. Suppose T is a compact operator on a Hilbert space and $n \in \mathbb{Z}^+$. Prove that $\dim \operatorname{range} T < n$ if and only if $s_n(T) = 0$.

Solution. Both sides say $|\Omega| < n$, where Ω indexes a singular value decomposition of T. Since T^*T is compact (10.73, 10.69(b)), self-adjoint, and has nonnegative eigenvalues (start of the proof of 10.113), 10.106(b) supplies a countable Ω, an orthonormal $\{e_k\}_{k\in\Omega}$ in V and positive reals $\{\alpha_k\}$ with $(T^*T)f = \sum_{k\in\Omega} \alpha_k\langle f, e_k\rangle e_k$; setting $s_k = \sqrt{\alpha_k}$ and $h_k = Te_k/s_k$, the proof of 10.113 makes $\{h_k\}_{k\in\Omega}$ orthonormal with

$$Tf = \sum_{k\in\Omega} s_k\langle f, e_k\rangle h_k \qquad (f \in V).$$

First, $\dim \operatorname{range} T = |\Omega|$: each $h_k = Te_k/s_k \in \operatorname{range} T$ and $\{h_k\}$ is orthonormal, hence independent, giving $\geq$; and if Ω is finite the display puts $\operatorname{range} T \subseteq \operatorname{span}\{h_k\}$, giving $\leq$.
Second, for every $\alpha > 0$,

$$\operatorname{null}(T^*T - \alpha I) = \operatorname{span}\{e_k : k \in \Omega,\ \alpha_k = \alpha\}.$$

Indeed $\supseteq$ holds since $(T^*T)e_j = \alpha_j e_j$; and if $(T^*T)f = \alpha f$, pairing the formula for $(T^*T)f$ with e_j gives $(\alpha_j - \alpha)\langle f, e_j\rangle = 0$, so $\alpha f = \sum_{\alpha_k=\alpha} \alpha\langle f, e_k\rangle e_k$ and $f = \sum_{\alpha_k=\alpha}\langle f, e_k\rangle e_k$, a finite sum because $\{e_k : \alpha_k = \alpha\}$ is orthonormal inside $\operatorname{null}(T^*T - \alpha I)$, finite-dimensional by 10.82.
Consequently the positive eigenvalues of T^*T are exactly the values taken by $\{\alpha_k\}$ (otherwise the display gives a zero null space), and each has geometric multiplicity equal to $\#\{k : \alpha_k = \alpha\}$. So the list of positive singular values in 10.116 is a decreasing rearrangement of $\{s_k\}_{k\in\Omega}$, and the number N of positive singular values equals $|\Omega|$, with $s_m(T) > 0 \iff m \leq N$. Hence

$$s_n(T) = 0 \iff N < n \iff |\Omega| < n \iff \dim \operatorname{range} T < n.$$

Problem 10D.14. Suppose T is a compact operator on a Hilbert space V with singular value decomposition

$$Tf = \sum_{k=1}^{\infty} s_k(T)\langle f, e_k\rangle h_k$$

for all $f \in V$. For $n \in \mathbb{Z}^+$, define $T_n : V \to V$ by

$$T_n f = \sum_{k=1}^{n} s_k(T)\langle f, e_k\rangle h_k.$$

Prove that $\lim_{n\to\infty} \|T - T_n\| = 0$.

[This exercise gives another proof, in addition to the proof suggested by Exercise 15 in Section 10C, that an operator on a Hilbert space is compact if and only if it is the limit of bounded operators with finite-dimensional range.]

Solution. It suffices to prove $\|T-T_n\| \le s_{n+1}(T)$ and $s_n(T) \to 0$; here $\{e_k\}$ and $\{h_k\}$ are the orthonormal families of 10.113, and only the decreasing ordering $s_1(T) \ge s_2(T) \ge \cdots$ of 10.116 is used.
For the estimate, fix n and $f \in V$. Then $(T-T_n)f = \sum_{k>n} s_k(T)\langle f, e_k\rangle h_k$, so orthonormality of $\{h_k\}_{k>n}$ and 8.54(b), then $s_k(T) \le s_{n+1}(T)$ for $k > n$ and Bessel's inequality 8.57 for $\{e_k\}$, give

$$\|(T-T_n)f\|^2 = \sum_{k>n} s_k(T)^2 |\langle f, e_k\rangle|^2 \le s_{n+1}(T)^2 \|f\|^2.$$

For the limit: T^*T is compact (10.73, 10.69(b)) and self-adjoint. If it has only finitely many positive eigenvalues, then $s_k(T) = 0$ for large k by 10.116 (second bullet). Otherwise every $s_k(T)$ is, by 10.116 (first bullet), the positive square root of a positive eigenvalue α of T^*T, listed $\dim \operatorname{null}(T^*T - \alpha I)$ times. Fix $\delta > 0$: only finitely many eigenvalues satisfy $\alpha \ge \delta^2$ by 10.93, and each such $\operatorname{null}(T^*T - \alpha I)$ is finite-dimensional by 10.82 since $\alpha \ne 0$. So $\{k : s_k(T) \ge \delta\}$ is finite, and $s_n(T) \to 0$.

Problem 10D.15. Suppose T is a compact operator on a Hilbert space V and $n \in \mathbb{Z}^+$. Prove that

$$\inf\{\|T - S\| : S \in \mathcal{B}(V) \text{ and } \dim \operatorname{range} S < n\} = s_n(T).$$

Solution. The infimum I is attained by the truncation of T after $n-1$ singular values. Fix a singular value decomposition $Tf = \sum_{k\in\Omega} s_k \langle f, e_k\rangle h_k$ (10.113), with $\{e_k\}, \{h_k\}$ orthonormal and $s_k > 0$. By the uniqueness discussed before 10.116 the s_k are the positive square roots of the positive eigenvalues of T^*T repeated by their (finite, by 10.82) multiplicities, so we may reindex by a bijection onto an initial segment of $\mathbb{Z}^+$ putting them in decreasing order, which alters neither orthonormality nor the unordered sums; thus Ω is an initial segment, $s_k = s_k(T)$ for $k \in \Omega$, and $s_j(T) = 0$ for $j \notin \Omega$.
$I \le s_n(T)$: take $Sf = \sum_{k\in\Omega, k\le n-1} s_k\langle f, e_k\rangle h_k$, a sum of at most $n-1$ terms, so $\dim \operatorname{range} S < n$. Then $(T-S)f = \sum_{k\in\Omega, k\ge n} s_k \langle f, e_k\rangle h_k$, so 8.54(b) and Bessel 8.57 give

$$\|(T-S)f\|^2 = \sum_{k\in\Omega,\, k\ge n} s_k^2 |\langle f, e_k\rangle|^2 \le s_n(T)^2\|f\|^2,$$

the middle step because $s_k \le s_n = s_n(T)$ for $k \ge n$ in Ω when $n \in \Omega$, while if $n \notin \Omega$ no such k exists (Ω is an initial segment) and both sides vanish.
$I \ge s_n(T)$: assume $s_n(T) > 0$, so $n \in \Omega$ and $e_1, \dots, e_n$ are orthonormal. Given S with $\dim \operatorname{range} S < n$, the restriction of S to the n-dimensional $U = \operatorname{span}\{e_1, \dots, e_n\}$ has $\dim \operatorname{null}(S|_U) \ge n - (n-1) = 1$ by the Fundamental Theorem of Linear Maps, so some $f = \sum_{k=1}^n a_k e_k \in U$ has $\|f\| = 1$ and $Sf = 0$. Here $\langle f, e_k\rangle = a_k$ for $k \le n$ and 0 for $k \in \Omega$ with $k > n$, so $Tf = \sum_{k=1}^n s_k a_k h_k$ and

$$\|(T-S)f\|^2 = \sum_{k=1}^n s_k^2 |a_k|^2 \ge s_n(T)^2 \sum_{k=1}^n |a_k|^2 = s_n(T)^2.$$

Problem 10D.16. Suppose T is a compact operator on a Hilbert space V and $n \in \mathbb{Z}^+$. Prove that

$$s_n(T) = \inf\{\|T|_{U^\perp}\| : U \text{ is a subspace of } V \text{ with } \dim U < n\}.$$

Solution. The infimum I is attained by $U = \operatorname{span}\{e_k : k \in \Omega,\ k \le n-1\}$, for the singular value decomposition $Tf = \sum_{k\in\Omega} s_k\langle f, e_k\rangle h_k$ normalized exactly as in Exercise 15: Ω an initial segment of $\mathbb{Z}^+$, $s_k = s_k(T)$ for $k \in \Omega$, and $s_j(T) = 0$ otherwise.
$I \le s_n(T)$: with that U, which has $\dim U \le n-1$, any $f \in U^\perp$ has $\langle f, e_k\rangle = 0$ for $k \le n-1$ in Ω, so

$Tf = \sum_{k\in\Omega, k\geq n} s_k\langle f, e_k\rangle h_k$ and, by 8.54(b) then Bessel 8.57,

$$\|Tf\|^2 = \sum_{k\in\Omega,\, k\geq n} s_k^2|\langle f, e_k\rangle|^2 \leq s_n(T)^2\|f\|^2,$$

the middle step because $s_k \leq s_n = s_n(T)$ for such k when $n \in \Omega$, and the sum is empty otherwise.
$I \geq s_n(T)$: assume $s_n(T) > 0$, so $n \in \Omega$. Let U be any subspace with $d := \dim U < n$, with basis $u_1, \ldots, u_d$, and let $W = \operatorname{span}\{e_1, \ldots, e_n\}$, of dimension n. The linear map $\varphi : W \to \mathbb{F}^d$, $\varphi(f) = (\langle f, u_1\rangle, \ldots, \langle f, u_d\rangle)$, has null $\varphi \neq \{0\}$ by the Fundamental Theorem of Linear Maps since $n > d$; pick $f = \sum_{k=1}^n a_k e_k$ in it with $\|f\| = 1$. Then $f \perp u_j$ for each j, hence $f \in U^\perp$, and since $\langle f, e_k\rangle = a_k$ for $k \leq n$ and 0 for $k \in \Omega$ with $k > n$,

$$\|Tf\|^2 = \sum_{k=1}^n s_k^2|a_k|^2 \geq s_n(T)^2\sum_{k=1}^n |a_k|^2 = s_n(T)^2.$$

Problem 10D.17. Suppose T is a compact operator on a Hilbert space V with singular value decomposition

$$Tf = \sum_{k\in\Omega} s_k\langle f, e_k\rangle h_k$$

for all $f \in V$. Prove that

$$T^*f = \sum_{k\in\Omega} s_k\langle f, h_k\rangle e_k$$

for all $f \in V$.

Solution. Write $Sf = \sum_{k\in\Omega} s_k\langle f, h_k\rangle e_k$; it suffices to show that S is a bounded operator with $\langle Tf, g\rangle = \langle f, Sg\rangle$ for all $f, g \in V$. Here $\{e_k\}_{k\in\Omega}$, $\{h_k\}_{k\in\Omega}$ are orthonormal and the s_k are positive, as in 10.113. Taking $f = e_j$ in the decomposition gives $Te_j = s_jh_j$, so $s_j = \|Te_j\| \leq \|T\|$; hence Bessel's inequality 8.57 for $\{h_k\}_{k\in\Omega}$ gives

$$\sum_{k\in\Omega}|s_k\langle f, h_k\rangle|^2 \leq \|T\|^2\sum_{k\in\Omega}|\langle f, h_k\rangle|^2 \leq \|T\|^2\|f\|^2 < \infty,$$

so the unordered sum defining Sf converges by 8.54(a) and $\|Sf\|^2 \leq \|T\|^2\|f\|^2$ by 8.54(b); linearity is clear (Check!), so $\|S\| \leq \|T\|$.
A bounded linear functional φ passes through an unordered sum $\sum_{k\in\Omega} x_k = x$, because $\big|\varphi(x) - \sum_{k\in F}\varphi(x_k)\big| \leq \|\varphi\|\,\|x - \sum_{k\in F} x_k\|$ for every finite $F \subseteq \Omega$. Applying this to Tf with $\varphi = \langle\,\cdot\,, g\rangle$, and to Sg with $\varphi = \langle\,\cdot\,, f\rangle$ followed by conjugation,

$$\langle Tf, g\rangle = \sum_{k\in\Omega} s_k\langle f, e_k\rangle\overline{\langle g, h_k\rangle} = \overline{\langle Sg, f\rangle} = \langle f, Sg\rangle.$$

By 10.1 we also have $\langle Tf, g\rangle = \langle f, T^*g\rangle$, so $\langle f, T^*g - Sg\rangle = 0$ for all $f, g \in V$; taking $f = T^*g - Sg$ gives $T^* = S$.

Problem 10D.18. Suppose that T is an operator on a finite-dimensional Hilbert space V with $\dim V = n$.
(a) Prove that T is invertible if and only if $s_n(T) \neq 0$.
(b) Suppose T is invertible and T has a singular value decomposition

$$Tf = s_1(T)\langle f, e_1\rangle h_1 + \cdots + s_n(T)\langle f, e_n\rangle h_n$$

for all $f \in V$. Show that

$$T^{-1}f = \frac{\langle f, h_1\rangle}{s_1(T)}e_1 + \cdots + \frac{\langle f, h_n\rangle}{s_n(T)}e_n$$

for all $f \in V$.

Solution. Here T is compact, being a bounded operator with finite-dimensional range (10.67), so singular values apply.
(a) T is invertible iff $\dim \operatorname{range} T = n$, since an operator on a finite-dimensional space is invertible iff it is surjective, and by Exercise 13 that fails exactly when $s_n(T) = 0$.
(b) Define $Rf = \sum_{j=1}^{n} \frac{\langle f,h_j\rangle}{s_j(T)} e_j$, which makes sense because $s_1(T) \geq \cdots \geq s_n(T) > 0$ by (a). The orthonormal lists $e_1, \ldots, e_n$ and $h_1, \ldots, h_n$ are independent and of length $n = \dim V$, hence bases of V, and evaluating both formulas on them gives $Te_j = s_j(T)h_j$ and $Rh_j = e_j/s_j(T)$. Therefore

$$(RT)e_j = e_j, \qquad (TR)h_j = h_j \qquad (j = 1, \ldots, n),$$

so the linear maps RT and TR agree with I on a basis, giving $R = T^{-1}$.

Problem 10D.19. Suppose T is a compact operator on a Hilbert space V. Prove that

$$\sum_{k\in\Gamma} \|Te_k\|^2 = \sum_{n=1}^{\infty} \big(s_n(T)\big)^2$$

for every orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V.

Solution. The left side is independent of the orthonormal basis, and one basis built from a singular value decomposition evaluates it. All sums are of nonnegative terms, taken in $[0, \infty]$ as $\sum_{i\in I} a_i = \sup\{\sum_{i\in F} a_i : F \subseteq I \text{ finite}\}$, which agrees with the unordered sum when the latter converges; the asserted equality is an equality in $[0, \infty]$.
Such double sums may be interchanged: with $A = \sum_{k\in\Gamma}\sum_{j\in\Lambda} a_{k,j}$ and $B = \sum_{j\in\Lambda}\sum_{k\in\Gamma} a_{k,j}$, any finite $F \subseteq \Gamma$, $G \subseteq \Lambda$ satisfy $\sum_{j\in G}\sum_{k\in F} a_{k,j} = \sum_{k\in F}\sum_{j\in G} a_{k,j} \leq A$; and given a finite G and $c < \sum_{j\in G}\sum_{k\in\Gamma} a_{k,j}$, pick $t_j < \sum_{k\in\Gamma} a_{k,j}$ with $\sum_{j\in G} t_j > c$, then finite F_j with $\sum_{k\in F_j} a_{k,j} > t_j$, and put $F = \bigcup_{j\in G} F_j$; nonnegativity gives $\sum_{j\in G}\sum_{k\in F} a_{k,j} > c$, so $c < A$. Hence $B \leq A$, and $A \leq B$ by symmetry.
Independence of the basis: for orthonormal bases $\{e_k\}_{k\in\Gamma}$ and $\{g_j\}_{j\in\Lambda}$ of V, Parseval 8.63(c) applied twice, the interchange, and 10.1 give

$$\begin{aligned}\sum_{k\in\Gamma} \|Te_k\|^2 &= \sum_{k\in\Gamma}\sum_{j\in\Lambda} |\langle Te_k, g_j\rangle|^2 = \sum_{j\in\Lambda}\sum_{k\in\Gamma} |\langle e_k, T^*g_j\rangle|^2 \\ &= \sum_{j\in\Lambda} \|T^*g_j\|^2,\end{aligned}$$

whose right side does not mention $\{e_k\}$.
Evaluation: write $Tf = \sum_{k\in\Omega} s_k(T)\langle f, u_k\rangle v_k$ as in Exercise 15, with Ω an initial segment of $\mathbb{Z}^+$, $s_k(T) > 0$ on Ω and $s_j(T) = 0$ off it. Put $U = \overline{\operatorname{span}}\{u_k\}_{k\in\Omega}$, so $\{u_k\}_{k\in\Omega}$ is an orthonormal basis of U by 8.61, and let $\{w_i\}_{i\in\Lambda}$ be an orthonormal basis of the Hilbert space $U^\perp$ (8.75). Their union is orthonormal, and its closed span contains $U + U^\perp = V$ by 8.43, so it is an orthonormal basis of V. Here $Tu_j = s_j(T)v_j$ gives $\|Tu_j\|^2 = s_j(T)^2$, while $w_i \perp u_k$ for all k gives $Tw_i = 0$. Splitting the sum over the disjoint union,

$$\sum_{k\in\Omega} \|Tu_k\|^2 + \sum_{i\in\Lambda} \|Tw_i\|^2 = \sum_{k\in\Omega} s_k(T)^2 = \sum_{n=1}^{\infty} \big(s_n(T)\big)^2,$$

the last equality because $s_n(T) = 0$ off Ω and nonnegative unordered sums over subsets of $\mathbb{Z}^+$ agree with the series.

Problem 10D.20. Use the result of Example 10.124 to evaluate $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Solution. $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$.
Write $S = \sum_{n=1}^{\infty} 1/n^2$, which is finite because $\sum_{n=1}^{N} n^{-2} \le 1 + \sum_{n=2}^{N}\bigl(\frac{1}{n-1} - \frac{1}{n}\bigr) < 2$ for every N. Splitting the first $2N$ terms by parity,

$$\sum_{n=1}^{2N} \frac{1}{n^2} = \sum_{k=1}^{N} \frac{1}{(2k-1)^2} + \frac{1}{4}\sum_{k=1}^{N} \frac{1}{k^2},$$

and letting $N \to \infty$, with $\sum_{k=1}^{\infty}(2k-1)^{-2} = \pi^2/8$ by Example 10.124, gives $S = \pi^2/8 + S/4$, so $S = \frac{4}{3}\cdot\frac{\pi^2}{8} = \pi^2/6$.

Problem 10D.21. Suppose T is a normal compact operator on a complex Hilbert space. Prove that the following are equivalent.
(a) $\operatorname{range} T$ is finite-dimensional.
(b) $\operatorname{sp}(T)$ is a finite set.
(c) $s_n(T) = 0$ for some $n \in \mathbb{Z}$.

Solution. Since 10.116 defines $s_n(T)$ only for $n \in \mathbb{Z}^+$, the printed $n \in \mathbb{Z}$ in (c) is read as $n \in \mathbb{Z}^+$. We prove (a) $\Leftrightarrow$ (c), then (a) $\Rightarrow$ (b) $\Rightarrow$ (a).
(a) $\Leftrightarrow$ (c): Exercise 13 gives $s_n(T) = 0 \iff \dim\operatorname{range} T < n$, and $\operatorname{range} T$ is finite-dimensional precisely when $\dim\operatorname{range} T < n$ for some $n \in \mathbb{Z}^+$.
(a) $\Rightarrow$ (b): let $m = \dim\operatorname{range} T < \infty$. Eigenvectors $f_1, \dots, f_j$ for distinct nonzero eigenvalues $\alpha_1, \dots, \alpha_j$ satisfy $f_i = Tf_i/\alpha_i \in \operatorname{range} T$, and are pairwise orthogonal by 10.57 (T normal), hence independent; so $j \le m$ and T has at most m nonzero eigenvalues. Every nonzero $\alpha \in \operatorname{sp}(T)$ is an eigenvalue by the Fredholm Alternative 10.85, so $\operatorname{sp}(T)$ has at most $m+1$ elements.
(b) $\Rightarrow$ (a): by 10.107 there is an orthonormal basis $\{f_k\}_{k\in\Gamma}$ of V with $Tf_k = \lambda_k f_k$. Each λ_k lies in the finite set $\operatorname{sp}(T)$, since $T - \lambda_k I$ is not injective; and for each nonzero value λ the orthonormal, hence independent, family $\{f_k : \lambda_k = \lambda\}$ lies in $\operatorname{null}(T - \lambda I)$, finite-dimensional by 10.82. So $\Lambda = \{k : \lambda_k \ne 0\}$ is finite. By Parseval 8.63(a) and continuity of T,

$$Tf = \sum_{k\in\Gamma} \lambda_k \langle f, f_k\rangle f_k = \sum_{k\in\Lambda} \lambda_k \langle f, f_k\rangle f_k,$$

so $\operatorname{range} T \subseteq \operatorname{span}\{f_k\}_{k\in\Lambda}$, which is finite-dimensional.

Problem 10D.22. Find the singular values of the Volterra operator.
[Your answer, when combined with Exercise 12, should show that the norm of the Volterra operator is $\frac{2}{\pi}$. This appearance of π can be surprising because the definition of the Volterra operator does not involve π.]

Solution. The singular values of the Volterra operator V, $(Vf)(x) = \int_0^x f$ (Example 10.15), are

$$s_n(V) = \frac{2}{(2n-1)\pi} \qquad (n \in \mathbb{Z}^+),$$

each occurring exactly once; with Exercise 12 this gives $\|V\| = s_1(V) = 2/\pi$.
V is the integral operator with kernel $K = \chi_{\{y<x\}} \in L^2(\mu \times \mu)$, hence compact (paragraph after the proof of 10.70), and 10.16 gives $(V^*f)(x) = \int_x^1 f$; so V^*V is compact and self-adjoint, and by 10.116 the singular values of V are the positive square roots of the positive eigenvalues of V^*V, with multiplicity. Suppose $V^*Vf = \alpha f$ with $\alpha > 0$ and $f \ne 0$. Since $\|f\|_1 \le \|f\|_2$ by Hölder 7.9 on the finite measure

space $[0,1]$, Vf is defined, and it is continuous because $|(Vf)(x)-(Vf)(x')| \le \int_{[x',x]} |f| \to 0$ by 3.31. So the Fundamental Theorem of Calculus makes V^*Vf continuously differentiable with $(V^*Vf)' = -Vf$; then $f = \frac{1}{\alpha}V^*Vf$ agrees a.e. with a C^1 function, and after replacing f by that representative, $(Vf)' = f$, so f is C^2 with

$$\alpha f''(x) = -(Vf)'(x) = -f(x) \qquad (0 \le x \le 1).$$

Both sides of $V^*Vf = \alpha f$ and of $(V^*Vf)' = -Vf$ are now continuous, so they hold pointwise; evaluating the first at $x=1$ and the second at $x=0$ gives $f(1)=0$ and $f'(0)=0$.
Write $\alpha = 1/\lambda^2$ with $\lambda > 0$, so $f''+\lambda^2 f = 0$ and hence $f(x) = A\cos(\lambda x) + B\sin(\lambda x)$. (If $u''+\lambda^2 u = 0$ with $u(0)=u'(0)=0$ then $(u')^2+\lambda^2u^2$ has zero derivative and vanishes at 0, so $u \equiv 0$ for real u, and for complex u apply this to its real and imaginary parts; subtracting the combination with $A = f(0)$, $B = f'(0)/\lambda$ gives the claim.) Now $f'(0)=0$ forces $B=0$ and so $A \ne 0$, and $f(1)=0$ forces $\cos\lambda = 0$, i.e. $\lambda = \lambda_k := \frac{(2k+1)\pi}{2}$ with $k \ge 0$, since $\lambda > 0$. Thus $\alpha = 4/((2k+1)^2\pi^2)$ and its eigenspace lies in the span of $f_k(x) = \cos(\lambda_k x)$.
Each such α is indeed an eigenvalue: $(Vf_k)(x) = \sin(\lambda_k x)/\lambda_k$, so, using $\cos\lambda_k = 0$,

$$(V^*Vf_k)(x) = \int_x^1 \frac{\sin(\lambda_k t)}{\lambda_k}\,dt = \frac{\cos(\lambda_k x) - \cos\lambda_k}{\lambda_k^2} = \frac{1}{\lambda_k^2} f_k(x).$$

By the previous paragraph each has geometric multiplicity 1 and there are no others, so listing $1/\lambda_k$ in decreasing order gives $s_n(V) = 1/\lambda_{n-1} = \frac{2}{(2n-1)\pi}$.

11 Fourier Analysis

Contents

11.1 Exercises 11A

Problem 11A.1. Prove that $\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}$ for all $f \in L^1(\partial\mathbf{D})$ and all $n \in \mathbf{Z}$.

Solution. On $\partial\mathbf{D}$ we have $\overline{z} = 1/z$, so $\overline{z^n} = z^{-n}$ and $\overline{z^{-n}} = z^n$; hence $\overline{f(z)}\,\overline{z^n} = \overline{f(z)}\,z^{-n} = \overline{f(z)\overline{z^{-n}}}$. Since conjugation commutes with integration, 11.7 gives

$$\begin{aligned}\widehat{\overline{f}}(n) &= \int_{\partial\mathbf{D}} \overline{f(z)\,\overline{z^{-n}}}\,d\sigma(z)\\ &= \overline{\int_{\partial\mathbf{D}} f(z)\,\overline{z^{-n}}\,d\sigma(z)} = \overline{\widehat{f}(-n)}.\end{aligned}$$

Problem 11A.2. Suppose $1 \le p \le \infty$ and $n \in \mathbf{Z}$.
(a) Show that the function $f \mapsto \widehat{f}(n)$ is a bounded linear functional on $L^p(\partial\mathbf{D})$ with norm 1.
(b) Find all $f \in L^p(\partial\mathbf{D})$ such that $\|f\|_p = 1$ and $|\widehat{f}(n)| = 1$.

Solution. Write $\varphi(f) = \widehat{f}(n)$. Since $\sigma(\partial\mathbf{D}) = 1$, we have $L^p(\partial\mathbf{D}) \subseteq L^1(\partial\mathbf{D})$ with

$$\|f\|_1 \le \|f\|_p \qquad (*)$$

for all $p \in [1,\infty]$: this is 7.10 for $1 < p < \infty$ (the constant being $\sigma(\partial\mathbf{D})^{(p-1)/p} = 1$), and $\int |f|\,d\sigma \le \|f\|_\infty$ for $p = \infty$. So φ is defined on $L^p(\partial\mathbf{D})$.
(a) φ is linear by 11.9(a),(b), and $|\varphi(f)| \le \|f\|_1 \le \|f\|_p$ by 11.9(c) and $(*)$, so $\|\varphi\| \le 1$. Equality holds because $g(z) = z^n$ has $|g| = 1$, hence $\|g\|_p = 1$, while $\varphi(g) = \int_{\partial\mathbf{D}} z^n\overline{z^n}\,d\sigma = 1$ by 11.6.
(b) For $1 < p \le \infty$ these are exactly $f(z) = \lambda z^n$ with $|\lambda| = 1$; for $p = 1$ they are exactly $f(z) = \lambda h(z) z^n$ with $|\lambda| = 1$ and $h \ge 0$ in $L^1(\partial\mathbf{D})$ with $\int h\,d\sigma = 1$.
These work: such an f has $\|f\|_1 = \int h\,d\sigma = 1$ and $\widehat{f}(n) = \lambda\int h\,d\sigma = \lambda$; and $h = 1$ gives $f(z) = \lambda z^n$ with $\|f\|_p = 1$ for all p.
No others: if $\|f\|_p = 1$ and $|\widehat{f}(n)| = 1$, then 11.9(c) and $(*)$ give

$$1 = |\widehat{f}(n)| \le \|f\|_1 \le \|f\|_p = 1,$$

so both are equalities. Put $g(z) = f(z)\overline{z^n}$, so $|g| = |f|$ and $\widehat{f}(n) = \int g\,d\sigma$, and pick $|\lambda| = 1$ with $\overline{\lambda}\int g\,d\sigma = |\int g\,d\sigma|$. Then $\int(|g| - \operatorname{Re}(\overline{\lambda}g))\,d\sigma = \|f\|_1 - |\widehat{f}(n)| = 0$ with nonnegative integrand, so by 3.43 $\operatorname{Re}(\overline{\lambda}g) = |g| = |\overline{\lambda}g|$ a.e., forcing $\overline{\lambda}g = |g|$ a.e. and hence

$$f(z) = \lambda|f(z)|z^n \quad \text{a.e.} \qquad (**)$$

With $h = |f|$ this is the claimed form for $p = 1$.
For $1 < p < \infty$ put $g = |f|$, so $\int g\,d\sigma = \int g^p\,d\sigma = 1$, and let $\psi(t) = t^p - pt + p - 1$; then $\psi'(t) = p(t^{p-1} - 1)$ is negative on $(0,1)$ and positive on $(1,\infty)$, so $\psi \ge 0$ with equality only at $t = 1$. Since

$$\int_{\partial\mathbf{D}} \psi(g)\,d\sigma = 1 - p + (p-1) = 0,$$

we get $g = 1$ a.e. For $p = \infty$, $|f| \le 1$ a.e. and $\int(1 - |f|)\,d\sigma = 0$ give the same. Substituting $|f| = 1$ into $(**)$ yields $f(z) = \lambda z^n$ a.e.

Problem 11A.3. Show that if $0 \le r < 1$ and $t \in \mathbf{R}$, then

$$P_r(e^{it}) = \frac{1 - r^2}{1 - 2r\cos t + r^2}.$$

Solution. By 11.14, $P_r(e^{it}) = (1-r^2)/|1-re^{it}|^2$, so it suffices to evaluate the denominator. Since r is real, $\overline{1-re^{it}} = 1-re^{-it}$, so

$$|1-re^{it}|^2 = (1-re^{it})(1-re^{-it}) = 1-r(e^{it}+e^{-it})+r^2 = 1-2r\cos t+r^2.$$

Problem 11A.4. Suppose $f \in L^1(\partial\mathbf{D})$, $z \in \partial\mathbf{D}$, and f is continuous at z. Prove that

$$\lim_{r\uparrow 1}(\mathcal{P}_r f)(z) = f(z).$$

[The result in this exercise differs from 11.18 because here we are assuming continuity only at a single point and we are not even assuming that f is bounded, as compared to 11.18, which assumed continuity at all points of $\partial\mathbf{D}$.]

Solution. Since $\int_{\partial\mathbf{D}} P_r(z\overline{w})\,d\sigma(w) = \int_{\partial\mathbf{D}} P_r\,d\sigma = 1$ by 11.17 and 11.16(b), the integral formula 11.15 gives

$$(\mathcal{P}_r f)(z) - f(z) = \int_{\partial\mathbf{D}} \big(f(w)-f(z)\big)P_r(z\overline{w})\,d\sigma(w).$$

Also $1-z\overline{w} = \overline{w}(w-z)$ for $w \in \partial\mathbf{D}$, so $|1-z\overline{w}| = |w-z|$.
Given $\varepsilon > 0$, continuity at z supplies $\delta > 0$ with $|f(w)-f(z)| < \varepsilon$ whenever $|w-z| < \delta$; let A_δ be that set of w, which is measurable, being relatively open. Since $P_r > 0$ by 11.16(a), the part of the integral over A_δ is at most $\varepsilon \int_{\partial\mathbf{D}} P_r(z\overline{w})\,d\sigma(w) = \varepsilon$.
Off A_δ we bound the kernel instead of f, which need not be bounded: there $|1-z\overline{w}| \ge \delta$, so if $1-r < \delta/2$ then $|1-rz\overline{w}| \ge |1-z\overline{w}| - (1-r) > \delta/2$, whence by 11.14

$$P_r(z\overline{w}) = \frac{1-r^2}{|1-rz\overline{w}|^2} < \frac{4(1-r^2)}{\delta^2},$$

and therefore, using $\sigma(\partial\mathbf{D}) = 1$ and $f \in L^1(\partial\mathbf{D})$,

$$\int_{\partial\mathbf{D}\setminus A_\delta} |f(w)-f(z)|\,P_r(z\overline{w})\,d\sigma(w) \le \frac{4(1-r^2)}{\delta^2}\big(\|f\|_1 + |f(z)|\big).$$

Since δ depends only on ε and $1-r^2 \to 0$, we get $\limsup_{r\uparrow 1} |(\mathcal{P}_r f)(z) - f(z)| \le \varepsilon$ for every $\varepsilon > 0$.

Problem 11A.5. Suppose $a, b \in \mathbf{C}$, $f \in L^1(\partial\mathbf{D})$, $z \in \partial\mathbf{D}$, $\lim_{t\downarrow 0} f(e^{it}z) = a$, and $\lim_{t\uparrow 0} f(e^{it}z) = b$. Prove that

$$\lim_{r\uparrow 1}(\mathcal{P}_r f)(z) = \frac{a+b}{2}.$$

[If $a \ne b$, then f is said to have a jump discontinuity at z.]

Solution. The Poisson kernel splits its unit mass evenly between the two sides of z, so the one-sided limits are averaged with equal weights. Write $p_r(t) = P_r(e^{it}) = \frac{1-r^2}{1-2r\cos t+r^2}$ by Exercise 3, an even function of t; in particular $P_r(e^{-it}) = p_r(t)$.
Rotation invariance $\int_{\partial\mathbf{D}} G(zw)\,d\sigma(w) = \int_{\partial\mathbf{D}} G\,d\sigma$ holds for $G \in L^1(\sigma)$: apply 11.17 to $h(\zeta) = G(\overline{\zeta})$ with $\overline{z}$ in place of z, giving $\int G(zw)\,d\sigma(w) = \int G(\overline{\zeta})\,d\sigma(\zeta)$, then 11.17 with $h = G$ and $z = 1$. Applying this to $G(w) = f(w)P_r(z\overline{w})$ in 11.15, with $z\overline{zw} = \overline{w}$, and transferring to $(-\pi, \pi]$ by 11.4,

$$(\mathcal{P}_r f)(z) = \int_{\partial\mathbf{D}} f(zw)P_r(\overline{w})\,d\sigma(w) = \int_{-\pi}^{\pi} f(e^{it}z)\,p_r(t)\,\frac{dt}{2\pi},$$

all integrals being defined since the same invariance gives $\int_{-\pi}^{\pi} |f(e^{it}z)|\frac{dt}{2\pi} = \|f\|_1 < \infty$.

By 11.16(b) and 11.4, $\int_{-\pi}^{\pi} p_r(t)\frac{dt}{2\pi} = 1$, so evenness of p_r gives $\int_0^{\pi} p_r \frac{dt}{2\pi} = \int_{-\pi}^{0} p_r \frac{dt}{2\pi} = \frac{1}{2}$ and hence

$$\begin{aligned}(\mathcal{P}_r f)(z) - \frac{a+b}{2} &= \int_0^{\pi} \big(f(e^{it}z) - a\big)p_r(t)\frac{dt}{2\pi} \\ &\quad + \int_{-\pi}^{0} \big(f(e^{it}z) - b\big)p_r(t)\frac{dt}{2\pi}.\end{aligned}$$

Given $\varepsilon > 0$, the one-sided limits supply $\delta \in (0,\pi)$ with $|f(e^{it}z)-a| < \varepsilon$ for $0 < t < \delta$ and $|f(e^{it}z)-b| < \varepsilon$ for $-\delta < t < 0$; split both integrals at $\pm\delta$. Since $p_r > 0$, the two inner pieces contribute at most $2 \cdot \varepsilon \cdot \frac{1}{2} = \varepsilon$. For $\delta \le |t| \le \pi$ we have $\cos t \le \cos\delta$, so

$$1 - 2r\cos t + r^2 = (1-r)^2 + 2r(1-\cos t) \ge 2r(1-\cos\delta),$$

giving $p_r(t) \le \frac{1-r^2}{1-\cos\delta}$ for $\frac{1}{2} \le r < 1$ (note $1-\cos\delta > 0$); hence the two outer pieces total at most $\frac{1-r^2}{1-\cos\delta}(\|f\|_1 + |a| + |b|)$. As δ depends only on ε and $1-r^2 \to 0$, the limsup as $r \uparrow 1$ is at most ε for every $\varepsilon > 0$.

Problem 11A.6. Prove that for each $p \in [1,\infty)$, there exists $f \in L^1(\partial\mathbf{D})$ such that

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^p = \infty.$$

Solution. Suppose not, so that $\sum_n |\widehat{f}(n)|^p < \infty$ for every $f \in L^1(\partial\mathbf{D})$; the Dirichlet kernel will contradict this. Then $Tf = (\widehat{f}(n))_{n\in\mathbf{Z}}$ maps the Banach space $L^1(\partial\mathbf{D})$ into the Banach space $\ell^p(\mathbf{Z})$ (both by 7.24), and is linear by 11.9(a),(b).
T has closed graph: if $f_k \to f$ in L^1 and $Tf_k \to c$ in ℓ^p, then $|\widehat{f_k}(n) - \widehat{f}(n)| \le \|f_k - f\|_1 \to 0$ by 11.9(a)–(c), while $|\widehat{f_k}(n) - c_n| \le \|Tf_k - c\|_p \to 0$, so $c = Tf$. By the Closed Graph Theorem 6.85 there is $C < \infty$ with

$$\Big(\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^p\Big)^{1/p} \le C\|f\|_1 \quad \text{for all } f \in L^1(\partial\mathbf{D}).$$

Apply this to $D_N(w) = \sum_{n=-N}^{N} w^n$. By 11.6 we get $\widehat{D_N}(m) = 1$ for $|m| \le N$ and 0 otherwise, so the left side is $(2N+1)^{1/p}$. For the right side, summing the geometric series and multiplying numerator and denominator by $e^{-it/2}$,

$$D_N(e^{it}) = e^{-iNt}\,\frac{e^{i(2N+1)t} - 1}{e^{it} - 1} = \frac{\sin\big((N+\frac{1}{2})t\big)}{\sin(t/2)} \qquad (t \neq 0),$$

so $|D_N(e^{it})| \le 1/|\sin(t/2)| \le \pi/|t|$ for $0 < |t| \le \pi$, using $\sin u \ge 2u/\pi$ on $[0,\pi/2]$ by concavity; also $|D_N| \le 2N+1$ from the definition. Using the second bound on $|t| \le \frac{1}{N}$ and the first on $\frac{1}{N} \le |t| \le \pi$,

$$\|D_N\|_1 \le \frac{1}{2\pi}\Big(\frac{2(2N+1)}{N} + 2\pi\ln(\pi N)\Big) \le 3 + \ln N,$$

since $\frac{1}{2\pi}(4 + \frac{2}{N}) < 1$ and $\ln\pi < 1.2$. Thus $(2N+1)^{1/p} \le C(3 + \ln N)$ for every N, which fails for large N because $p < \infty$.

Problem 11A.7. Suppose $\zeta \in \partial\mathbf{D}$. Show that the function

$$w \mapsto \frac{1-|w|^2}{|1-\overline{\zeta}w|^2}$$

is harmonic on $\mathbf{C} \setminus \{\zeta\}$ by finding an analytic function on $\mathbf{C} \setminus \{\zeta\}$ whose real part is the function above.

Solution. Take
$$g(w) = \frac{1+\overline{\zeta}w}{1-\overline{\zeta}w}.$$
This is a quotient of polynomials in w, so it is analytic off the zero set of the denominator; and $1-\overline{\zeta}w = 0$ exactly when $w = 1/\overline{\zeta} = \zeta$, the last equality because $\zeta\overline{\zeta} = |\zeta|^2 = 1$. Thus g is analytic on $\mathbf{C} \setminus \{\zeta\}$.
Write $u = \overline{\zeta}w$, so $u \neq 1$ and $|u| = |w|$. Multiplying numerator and denominator by $\overline{1-u}$,
$$\frac{1+u}{1-u} = \frac{(1+u)(1-\overline{u})}{|1-u|^2} = \frac{(1-|u|^2)+(u-\overline{u})}{|1-u|^2}.$$
Here $u-\overline{u}$ is purely imaginary while $1-|u|^2$ and $|1-u|^2$ are real, so taking real parts and substituting $u = \overline{\zeta}w$ gives
$$(\operatorname{Re} g)(w) = \frac{1-|w|^2}{|1-\overline{\zeta}w|^2} \qquad \text{for } w \in \mathbf{C} \setminus \{\zeta\}.$$
The real part of a function analytic on an open subset of $\mathbf{C}$ is harmonic there, by the first bullet point of 11.20, so the given function is harmonic on $\mathbf{C} \setminus \{\zeta\}$.

Problem 11A.8. Suppose $f : \partial\mathbb{D} \to \mathbb{R}$ is the function defined by
$$f(x,y) = x^4 y$$
for $(x,y) \in \mathbb{R}^2$ with $x^2+y^2 = 1$. Find a polynomial u of two variables x, y such that u is harmonic on $\mathbb{R}^2$ and $u|_{\partial\mathbb{D}} = f$.
[Of course, $u|_{\mathbb{D}}$ is the Poisson integral of f. However, here you are asked to find an explicit formula for u in closed form, without involving or computing an integral. It may help to think of f as defined by $f(z) = (\operatorname{Re} z)^4(\operatorname{Im} z)$ for $z \in \partial\mathbb{D}$.]

Solution. The answer is
$$u(x,y) = \tfrac{1}{16}\big(5x^4y - 10x^2y^3 + y^5 + 9x^2y - 3y^3 + 2y\big),$$
which is exactly $u(w) = \frac{1}{16}\operatorname{Im}\big(w^5 + 3w^3 + 2w\big)$ under the usual identification $w = x+yi$.
Since $g(w) = \frac{1}{16}(w^5+3w^3+2w)$ is a polynomial, hence analytic on $\mathbb{C}$, the function $u = \operatorname{Im} g$ is harmonic on $\mathbb{R}^2$ by the first bullet point of 11.20. It remains to check $u|_{\partial\mathbb{D}} = f$.
For $z \in \partial\mathbb{D}$ we have $\operatorname{Re} z = \frac{z+\overline{z}}{2}$, $\operatorname{Im} z = \frac{z-\overline{z}}{2i}$ and $z\overline{z} = 1$, so
$$f(z) = \frac{(z+\overline{z})^4(z-\overline{z})}{32i}, \qquad (z+\overline{z})^4 = z^4 + 4z^2 + 6 + 4\overline{z}^2 + \overline{z}^4,$$
the second identity by the binomial theorem with $z\overline{z} = 1$. Multiplying out, again using $z\overline{z} = 1$ (so $z^4\overline{z} = z^3$, $z\overline{z}^2 = \overline{z}$, and so on),
$$\begin{aligned}(z+\overline{z})^4(z-\overline{z}) &= \big(z^5 + 4z^3 + 6z + 4\overline{z} + \overline{z}^3\big)\\ &\qquad - \big(z^3 + 4z + 6\overline{z} + 4\overline{z}^3 + \overline{z}^5\big)\\ &= \big(z^5 - \overline{z}^5\big) + 3\big(z^3 - \overline{z}^3\big) + 2\big(z - \overline{z}\big)\\ &= 2i\operatorname{Im}\big(z^5 + 3z^3 + 2z\big).\end{aligned}$$
Hence $f(z) = \frac{1}{16}\operatorname{Im}(z^5 + 3z^3 + 2z) = u(z)$ on $\partial\mathbb{D}$, as required.
Writing $w = x+yi$ and picking out the odd-index binomial terms gives $\operatorname{Im}(w^3) = 3x^2y - y^3$ and $\operatorname{Im}(w^5) = 5x^4y - 10x^2y^3 + y^5$, which assembles into the displayed polynomial.

Problem 11A.9. Find a formula (in closed form, not as an infinite sum) for $\mathcal{P}_r f$, where f is the function in the second bullet point of Example 11.8.

Solution. For $f(z) = 1/|3-z|^2$, every $r \in [0,1)$ and every $z \in \partial\mathbb{D}$,

$$(\mathcal{P}_r f)(z) = \frac{9-r^2}{8\,|3-rz|^2} = \frac{9-r^2}{8\,(9-6r\operatorname{Re} z + r^2)}.$$

Indeed, the second bullet point of Example 11.8 gives $\widehat{f}(n) = \frac{1}{8}3^{-|n|}$, so by the definition 11.11, with $\rho := r/3 \in [0, \frac{1}{3})$,

$$(\mathcal{P}_r f)(z) = \sum_{n=-\infty}^{\infty} r^{|n|}\widehat{f}(n)z^n = \frac{1}{8}\sum_{n=-\infty}^{\infty} \rho^{|n|}z^n = \frac{1}{8}\cdot\frac{1-\rho^2}{|1-\rho z|^2},$$

the last equality by 11.13 (legitimate since $\rho \in [0,1)$ and $|z| = 1$, which also makes the series absolutely convergent). Multiplying numerator and denominator by 9, using $|1 - \frac{r}{3}z|^2 = \frac{1}{9}|3-rz|^2$, yields the first form; and

$$|3-rz|^2 = 9 - 3r(z+\overline{z}) + r^2|z|^2 = 9 - 6r\operatorname{Re} z + r^2$$

yields the second.

Problem 11A.10. Suppose $f : \partial\mathbb{D} \to \mathbb{C}$ is three times continuously differentiable. Prove that

$$f^{[1]}(z) = i\sum_{n=-\infty}^{\infty} n\,\widehat{f}(n)\,z^n$$

for all $z \in \partial\mathbb{D}$.

Solution. Apply 11.27 to $f^{[1]}$ and then 11.26 to f.
The hypothesis of 11.27 holds for $f^{[1]}$: with $\widetilde{g}(t) = g(e^{it})$ as in 11.24, the second bullet point of 11.24 gives $\widetilde{f^{[1]}} = \widetilde{f}^{(1)}$, so $\big(\widetilde{f^{[1]}}\big)^{(2)} = \widetilde{f}^{(3)}$, which exists on $\mathbb{R}$ and is continuous by hypothesis; hence $f^{[1]}$ is twice continuously differentiable. Thus 11.27 gives, for $z \in \partial\mathbb{D}$,

$$f^{[1]}(z) = \sum_{n=-\infty}^{\infty} \widehat{f^{[1]}}(n)\,z^n,$$

the series converging absolutely since $\sum_n |\widehat{f^{[1]}}(n)| < \infty$ by 11.29 applied to $f^{[1]}$. Now f is continuously differentiable, so 11.26 with $k=1$ gives $\widehat{f^{[1]}}(n) = in\widehat{f}(n)$, and absolute convergence lets i come out of the sum:

$$f^{[1]}(z) = \sum_{n=-\infty}^{\infty} i\,n\,\widehat{f}(n)\,z^n = i\sum_{n=-\infty}^{\infty} n\,\widehat{f}(n)\,z^n.$$

Problem 11A.11. Let $C(\partial\mathbb{D})$ denote the Banach space of continuous functions from $\partial\mathbb{D}$ to $\mathbb{C}$, with the supremum norm. For $M \in \mathbb{Z}^+$, define a linear functional $\varphi_M : C(\partial\mathbb{D}) \to \mathbb{C}$ by

$$\varphi_M(f) = \sum_{n=-M}^{M} \widehat{f}(n).$$

Thus $\varphi_M(f)$ is a partial sum of the Fourier series $\sum_{n=-\infty}^{\infty} \widehat{f}(n)z^n$, evaluated at $z = 1$.

(a) Show that

$$\varphi_M(f) = \int_{-\pi}^{\pi} f(e^{it})\,\frac{\sin\left(M+\frac{1}{2}\right)t}{\sin\frac{t}{2}}\,\frac{dt}{2\pi}$$

for every $f \in C(\partial\mathbb{D})$ and every $M \in \mathbb{Z}^+$.
(b) Show that

$$\lim_{M\to\infty} \int_{-\pi}^{\pi} \left|\frac{\sin\left(M+\frac{1}{2}\right)t}{\sin\frac{t}{2}}\right| \frac{dt}{2\pi} = \infty.$$

(c) Show that $\lim_{M\to\infty} \|\varphi_M\| = \infty$.
(d) Show that there exists $f \in C(\partial\mathbb{D})$ such that $\displaystyle\lim_{M\to\infty} \sum_{n=-M}^{M} \widehat{f}(n)$ does not exist (as an element of $\mathbb{C}$).
[Because the sum in (d) is a partial sum of the Fourier series evaluated at $z = 1$, part (d) shows that the Fourier series of a continuous function on $\partial\mathbb{D}$ need not converge pointwise on $\partial\mathbb{D}$.
The family of functions (one for each $M \in \mathbb{Z}^+$) on $\partial\mathbb{D}$ defined by

$$e^{it} \mapsto \frac{\sin\left(M+\frac{1}{2}\right)t}{\sin\frac{t}{2}}$$

is called the Dirichlet kernel.]

Solution. Set $D_M(t) = \sum_{n=-M}^{M} e^{int}$, a continuous real-valued 2π-periodic function (pair n with $-n$), and $L_M = \int_{-\pi}^{\pi} |D_M(t)|\,\frac{dt}{2\pi}$. For $t \in [-\pi,\pi]\setminus\{0\}$, summing the geometric series of ratio $e^{it} \neq 1$ and then multiplying numerator and denominator by $e^{-it/2}$,

$$D_M(t) = \frac{e^{i(M+1)t} - e^{-iMt}}{e^{it}-1} = \frac{2i\sin\left(M+\frac{1}{2}\right)t}{2i\sin\frac{t}{2}} = \frac{\sin\left(M+\frac{1}{2}\right)t}{\sin\frac{t}{2}}.$$

So D_M agrees with the Dirichlet kernel off the null set $\{0\}$, and may replace it inside any integral over $[-\pi,\pi]$.
(a) By the definition 11.7, linearity of the integral across the finite sum, and the reindexing $n \mapsto -n$ of $\{-M,\dots,M\}$,

$$\varphi_M(f) = \sum_{n=-M}^{M} \int_{-\pi}^{\pi} f(e^{it})e^{-int}\,\frac{dt}{2\pi} = \int_{-\pi}^{\pi} f(e^{it})\,D_M(t)\,\frac{dt}{2\pi};$$

now insert the closed form.
(b) The integral in question is L_M. Using $\left|\sin\frac{t}{2}\right| \leq \frac{|t|}{2}$, evenness of the integrand, and the substitution $s = \left(M+\frac{1}{2}\right)t$,

$$\begin{aligned} L_M &\geq \frac{2}{\pi}\int_0^{(M+\frac{1}{2})\pi} \frac{|\sin s|}{s}\,ds \geq \frac{2}{\pi}\sum_{k=0}^{M-1} \frac{1}{(k+1)\pi}\int_{k\pi}^{(k+1)\pi} |\sin s|\,ds \\ &= \frac{4}{\pi^2}\sum_{k=1}^{M}\frac{1}{k}, \end{aligned}$$

bounding $\frac{1}{s} \geq \frac{1}{(k+1)\pi}$ on $[k\pi,(k+1)\pi]$ and using $\int_{k\pi}^{(k+1)\pi} |\sin s|\,ds = 2$. The harmonic series diverges, so $L_M \to \infty$.
(c) In fact $\|\varphi_M\| = L_M$. The bound $|\varphi_M(f)| \leq \|f\|_\infty L_M$ is immediate from (a). Conversely, fix $\varepsilon > 0$ and put $f_\varepsilon(e^{it}) = D_M(t)/(|D_M(t)|+\varepsilon)$; this is continuous on $\partial\mathbb{D}$ with $\|f_\varepsilon\|_\infty \leq 1$, since D_M is continuous and 2π-periodic and the denominator is at least ε. As $a^2/(|a|+\varepsilon) \geq |a|-\varepsilon$ for real a, part (a) gives

$$\|\varphi_M\| \geq \varphi_M(f_\varepsilon) = \int_{-\pi}^{\pi} \frac{D_M(t)^2}{|D_M(t)|+\varepsilon}\,\frac{dt}{2\pi} \geq L_M - \varepsilon,$$

and $\varepsilon > 0$ was arbitrary. With (b), $\|\varphi_M\| \to \infty$.
(d) Otherwise $\lim_M \varphi_M(f)$ exists, hence $\sup_M |\varphi_M(f)| < \infty$, for every f in the Banach space $C(\partial\mathbb{D})$; since each φ_M is a bounded linear functional by (c), the Principle of Uniform Boundedness 6.86 would force $\sup_M \|\varphi_M\| < \infty$, contradicting (c).

Problem 11A.12. Define $f : \partial\mathbb{D} \to \mathbb{R}$ by

$$f(z) = \begin{cases} 1 & \text{if } \operatorname{Im} z > 0, \\ -1 & \text{if } \operatorname{Im} z < 0, \\ 0 & \text{if } \operatorname{Im} z = 0. \end{cases}$$

(a) Show that if $n \in \mathbb{Z}$, then

$$\widehat{f}(n) = \begin{cases} -\dfrac{2i}{n\pi} & \text{if } n \text{ is odd}, \\ 0 & \text{if } n \text{ is even}. \end{cases}$$

(b) Show that

$$(\mathcal{P}_r f)(z) = \frac{2}{\pi} \arctan \frac{2r \operatorname{Im} z}{1 - r^2}$$

for every $r \in [0, 1)$ and every $z \in \partial\mathbb{D}$.
(c) Verify that $\lim_{r\uparrow 1} (\mathcal{P}_r f)(z) = f(z)$ for every $z \in \partial\mathbb{D}$.
(d) Prove that $\mathcal{P}_r f$ does not converge uniformly to f on $\partial\mathbb{D}$ as $r \uparrow 1$.

Solution. (a) With $f(e^{it}) = 1$ on $(0, \pi)$, -1 on $(-\pi, 0)$, and 0 on the null set $\{0, \pi\}$, the definition 11.7 gives, for $n \neq 0$,

$$\begin{aligned} \widehat{f}(n) &= \frac{1}{2\pi} \left(\int_0^\pi e^{-int}\, dt - \int_{-\pi}^0 e^{-int}\, dt \right) \\ &= \frac{1}{2\pi} \cdot \frac{2\big(1 - (-1)^n\big)}{in} = \frac{1 - (-1)^n}{in\pi}, \end{aligned}$$

using $e^{\mp in\pi} = (-1)^n$. This vanishes for n even and equals $-\frac{2i}{n\pi}$ for n odd (as $\frac{1}{i} = -i$); and $\widehat{f}(0) = 0$, both integrals being π.
(b) Put $w = rz$. By 11.11 and (a) the series $\sum_n r^{|n|} \widehat{f}(n) z^n$ converges absolutely (terms dominated by $\frac{2}{\pi} r^{|n|}$), so we may pair n with $-n$; since $z^{-n} = \overline{z}^n$ and $z^n - \overline{z}^n = 2i \operatorname{Im}(z^n)$, each pair with n odd and positive contributes $\frac{4}{n\pi} \operatorname{Im}(w^n)$. Hence

$$(\mathcal{P}_r f)(z) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\operatorname{Im}\big(w^{2k+1}\big)}{2k+1} = \frac{4}{\pi} \operatorname{Im} g(w), \qquad g(w) := \sum_{k=0}^{\infty} \frac{w^{2k+1}}{2k+1}.$$

That power series has radius of convergence 1, so g is analytic on $\mathbb{D}$ with $g(0) = 0$ and $g'(w) = \sum_k w^{2k} = \frac{1}{1-w^2}$. Setting $\psi(w) = \frac{1+w}{1-w}$,

$$\psi(w) = \frac{(1+w)(1-\overline{w})}{|1-w|^2} = \frac{\big(1 - |w|^2\big) + 2i \operatorname{Im} w}{|1-w|^2},$$

so $\operatorname{Re} \psi > 0$ on $\mathbb{D}$; hence ψ lands in the right half-plane, where Log is analytic, and $h := \frac{1}{2} \operatorname{Log} \circ \psi$ is analytic on $\mathbb{D}$ with $h(0) = 0$ and $h'(w) = \frac{1}{2} \psi'(w)/\psi(w) = \frac{1}{1-w^2}$. Two analytic functions on the connected set $\mathbb{D}$ with the same derivative that agree at 0 coincide, so $g = h$. Taking imaginary parts, with $\operatorname{Im} \operatorname{Log}(a + bi) = \arctan \frac{b}{a}$ for $a > 0$,

$$\operatorname{Im} g(w) = \tfrac{1}{2} \arctan \frac{2 \operatorname{Im} w}{1 - |w|^2}, \qquad \text{so} \qquad (\mathcal{P}_r f)(z) = \frac{2}{\pi} \arctan \frac{2r \operatorname{Im} z}{1 - r^2}.$$

(c) As $r \uparrow 1$ we have $1 - r^2 \downarrow 0$ through positive values. (i) If $\operatorname{Im} z > 0$ the argument of arctan tends to $+\infty$, so $(\mathcal{P}_r f)(z) \to \frac{2}{\pi} \cdot \frac{\pi}{2} = 1 = f(z)$. (ii) If $\operatorname{Im} z < 0$ it tends to $-\infty$, giving $-1 = f(z)$. (iii) If $\operatorname{Im} z = 0$ the argument is 0 for every r, so $(\mathcal{P}_r f)(z) = 0 = f(z)$.

(d) For $r \in [\sqrt{2}-1,\, 1)$ set $s_r = \frac{1-r^2}{2r}$, which lies in $(0, 1]$ since $s_r \le 1$ is equivalent to $r^2 + 2r - 1 \ge 0$, and put $z_r = \sqrt{1 - s_r^2} + i s_r \in \partial\mathbb{D}$. Then $f(z_r) = 1$ while (b) gives $(\mathcal{P}_r f)(z_r) = \frac{2}{\pi}\arctan 1 = \frac{1}{2}$, so

$$\|f - \mathcal{P}_r f\|_\infty \ge \big|f(z_r) - (\mathcal{P}_r f)(z_r)\big| = \tfrac{1}{2}$$

for all such r, ruling out uniform convergence.

11.2 Exercises 11B

Problem 11B.1. Show that the family $\{e_k\}_{k\in\mathbb{Z}}$ of trigonometric functions defined by 11.1 is an orthonormal basis of $L^2\big((-\pi, \pi]\big)$.

Solution. Transport the basis $\{z^n\}_{n\in\mathbb{Z}}$ of 11.30 across the unitary map

$$U\colon L^2(\partial\mathbb{D}) \to L^2\big((-\pi,\pi]\big), \qquad (Uf)(t) = \frac{1}{\sqrt{2\pi}}\, f(e^{it}).$$

That U is unitary is the transfer of integration recorded just after 11.4: it gives $\langle Uf, Uh\rangle = \int_{\partial\mathbb{D}} f\overline{h}\, d\sigma = \langle f, h\rangle$, and surjectivity because $g \mapsto \sqrt{2\pi}\, g(\,\cdot\,) \circ \varphi^{-1}$ inverts it, where $\varphi(t) = e^{it}$ is the bijection of 11.2 (measurability matches by the definition 11.4). Hence $f_n := U(z^n)$, that is $f_n(t) = e^{int}/\sqrt{2\pi}$, is an orthonormal basis of $L^2\big((-\pi,\pi]\big)$.
By 11.1, $e_0 = \frac{1}{\sqrt{2\pi}} = f_0$ and, for $n \in \mathbb{Z}^+$, $e_n(t) = \frac{\sin(nt)}{\sqrt{\pi}}$ and $e_{-n}(t) = \frac{\cos(nt)}{\sqrt{\pi}}$. So

$$f_{\pm n} = \frac{e_{-n} \pm i e_n}{\sqrt{2}}, \qquad \text{equivalently} \qquad e_{-n} = \frac{f_n + f_{-n}}{\sqrt{2}}, \quad e_n = \frac{f_n - f_{-n}}{i\sqrt{2}}.$$

Thus $\operatorname{span}\{e_k\} = \operatorname{span}\{f_n\}$, which is dense. Orthonormality of $\{e_k\}$ follows from that of $\{f_n\}$: $\|e_{\pm n}\|^2 = \frac{1}{2}(\|f_n\|^2 + \|f_{-n}\|^2) = 1$ and $\langle e_n, e_{-n}\rangle = \frac{1}{2i}(\|f_n\|^2 - \|f_{-n}\|^2) = 0$, while for $|j| \ne |k|$ the vectors e_j and e_k are combinations of the disjoint pairs $\{f_{|j|}, f_{-|j|}\}$ and $\{f_{|k|}, f_{-|k|}\}$, so $\langle e_j, e_k\rangle = 0$. By the definition 8.61, $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal basis.

Problem 11B.2. Use the result of Exercise 12(a) in Section 11A to show that

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots = \frac{\pi^2}{8}.$$

Solution. Apply Parseval's identity to the function f of Exercise 12 in Section 11A, for which $\widehat{f}(n) = -\frac{2i}{n\pi}$ for n odd and $\widehat{f}(n) = 0$ for n even by part (a) of that exercise.
Since $\{z \in \partial\mathbb{D} : \operatorname{Im} z = 0\} = \{1, -1\}$ has σ-measure 0, we have $|f| = 1$ almost everywhere, so $f \in L^2(\partial\mathbb{D})$ and $\|f\|_2^2 = \sigma(\partial\mathbb{D}) = 1$. By 11.30 the family $\{z^n\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\partial\mathbb{D})$, and $\langle f, z^n\rangle = \widehat{f}(n)$ by the definition 11.7, so 8.63(c) gives

$$1 = \|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 = 2\sum_{k=0}^{\infty} \frac{4}{(2k+1)^2\pi^2} = \frac{8}{\pi^2}\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2},$$

the regrouping into pairs $\pm n$ being legitimate for a family of nonnegative numbers. Hence the sum equals $\pi^2/8$.

Problem 11B.3. Use techniques similar to Example 11.32 to evaluate $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4}$.

[If you feel industrious, you may also want to evaluate $\sum_{n=1}^{\infty} 1/n^6$. Similar techniques work to evaluate $\sum_{n=1}^{\infty} 1/n^k$ for each positive even integer k. You can become famous if you figure out how to evaluate $\sum_{n=1}^{\infty} 1/n^3$, which currently is an open question.]

Solution. $\sum_{n=1}^{\infty} 1/n^4 = \pi^4/90$. Apply Parseval to $f(e^{it}) = t^2$ on $(-\pi, \pi]$, which is bounded and measurable, hence in $L^2(\partial\mathbb{D})$.
Directly, $\widehat{f}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} t^2\,dt = \frac{\pi^2}{3}$. For $n \neq 0$, oddness of $t^2\sin(nt)$ and the antiderivative

$$\int t^2\cos(nt)\,dt = \frac{t^2\sin(nt)}{n} + \frac{2t\cos(nt)}{n^2} - \frac{2\sin(nt)}{n^3}$$

(which vanishes at 0 and equals $2\pi(-1)^n/n^2$ at π) give

$$\widehat{f}(n) = \frac{1}{\pi}\int_0^{\pi} t^2\cos(nt)\,dt = \frac{2(-1)^n}{n^2}.$$

Since $\{z^n\}_{n\in\mathbb{Z}}$ is an orthonormal basis by 11.30 and $\langle f, z^n\rangle = \widehat{f}(n)$, Parseval 8.63(c) gives

$$\frac{\pi^4}{9} + 8\sum_{n=1}^{\infty}\frac{1}{n^4} = \sum_{n=-\infty}^{\infty}|\widehat{f}(n)|^2 = \|f\|_2^2 = \int_{-\pi}^{\pi} t^4\,\frac{dt}{2\pi} = \frac{\pi^4}{5},$$

whence $\sum_{n\geq 1} n^{-4} = \frac{1}{8}\pi^4(\frac{1}{5} - \frac{1}{9}) = \pi^4/90$.
For $\sum_{n=1}^{\infty} 1/n^6$ take $g(e^{it}) = t^3$. Then $\widehat{g}(0) = 0$ by oddness, and for $n \neq 0$, using $\int t^3\sin(nt)\,dt = -\frac{t^3\cos(nt)}{n} + \frac{3t^2\sin(nt)}{n^2} + \frac{6t\cos(nt)}{n^3} - \frac{6\sin(nt)}{n^4}$ (Check!),

$$\widehat{g}(n) = \frac{-i}{\pi}\int_0^{\pi} t^3\sin(nt)\,dt = i(-1)^n\left(\frac{\pi^2}{n} - \frac{6}{n^3}\right).$$

Summing $|\widehat{g}(n)|^2 = \frac{\pi^4}{n^2} - \frac{12\pi^2}{n^4} + \frac{36}{n^6}$ over $n \neq 0$ (three absolutely convergent series) with $\zeta(2) = \pi^2/6$ from Example 11.32 and $\zeta(4) = \pi^4/90$, Parseval gives

$$\frac{\pi^6}{15} + 72\sum_{n=1}^{\infty}\frac{1}{n^6} = \|g\|_2^2 = \int_{-\pi}^{\pi} t^6\,\frac{dt}{2\pi} = \frac{\pi^6}{7},$$

so $\sum_{n\geq 1} n^{-6} = \frac{1}{72}\pi^6(\frac{1}{7} - \frac{1}{15}) = \pi^6/945$.

Problem 11B.4. Suppose $f, g\colon \partial\mathbb{D} \to \mathbb{C}$ are measurable functions. Prove that the function $(w, z) \mapsto f(w)\,g(z\overline{w})$ is a measurable function from $\partial\mathbb{D} \times \partial\mathbb{D}$ to $\mathbb{C}$.
[Here the σ-algebra on $\partial\mathbb{D} \times \partial\mathbb{D}$ is the usual product σ-algebra as defined in 5.2.]

Solution. Everything reduces to one claim: the map $m(w, z) = z\overline{w}$ from $\partial\mathbb{D} \times \partial\mathbb{D}$ to $\partial\mathbb{D}$ is $\mathcal{S} \otimes \mathcal{S}$-measurable, where $\mathcal{S}$ is the σ-algebra of 11.4.
Write $\varphi(t) = e^{it}$ for the bijection $(-\pi, \pi] \to \partial\mathbb{D}$ and $\mathcal{B}$ for the Borel subsets of $(-\pi, \pi]$; by 11.4, $\mathcal{S} = \{E : \varphi^{-1}(E) \in \mathcal{B}\} = \{\varphi(B) : B \in \mathcal{B}\}$. Let $\psi(s, t)$ be the representative of $t - s$ in $(-\pi, \pi]$, so that $\varphi \circ \psi = m \circ (\varphi \times \varphi)$ by 2π-periodicity of $e^{i\theta}$.
(i) $\psi^{-1}(B) \in \mathcal{B} \otimes \mathcal{B}$ for $B \in \mathcal{B}$. Each $h_c(s, t) = t - s + c$ is $\mathcal{B} \otimes \mathcal{B}$-measurable by 2.46(a) (the coordinate maps are, their inverse images being measurable rectangles), and since $t - s \in (-2\pi, 2\pi)$,

$$\psi^{-1}(B) = \bigcup_{j=1}^{3}\bigl(D_j \cap h_{c_j}^{-1}(B)\bigr), \qquad (c_1, c_2, c_3) = (2\pi, 0, -2\pi),$$

where D_1, D_2, D_3 are the h_0-inverse images of $(-\infty, -\pi]$, $(-\pi, \pi]$, (π, ∞).
(ii) $C \in \mathcal{B} \otimes \mathcal{B}$ implies $(\varphi \times \varphi)(C) \in \mathcal{S} \otimes \mathcal{S}$. The collection of such C is a σ-algebra, because $\varphi \times \varphi$ is a bijection, so taking images commutes with complements and unions; and it contains every measurable rectangle, since $(\varphi \times \varphi)(B_1 \times B_2) = \varphi(B_1) \times \varphi(B_2)$.

For $E \in \mathcal{S}$ put $B = \varphi^{-1}(E) \in \mathcal{B}$. Bijectivity of $\varphi \times \varphi$ and the identity $\varphi \circ \psi = m \circ (\varphi \times \varphi)$ give $m^{-1}(E) = (\varphi \times \varphi)\big(\psi^{-1}(B)\big)$, which lies in $\mathcal{S} \otimes \mathcal{S}$ by (i) and (ii). This proves the claim.
Hence $(w, z) \mapsto g(z\overline{w}) = (g \circ m)(w, z)$ is measurable, and so is $(w, z) \mapsto f(w)$ (inverse images of its real and imaginary parts are measurable rectangles). Their product is measurable by 2.46(a), applied to the real and imaginary parts of $f(w)g(z\overline{w})$.

Problem 11B.5. Where does the proof of 11.42 fail when $p = \infty$?

Solution. At the very first step: the choice of a continuous g on $\partial\mathbb{D}$ with $\|f - g\|_p < \varepsilon$. Continuous functions are dense in $L^p(\partial\mathbb{D})$ for $p \in [1, \infty)$ but not in $L^\infty(\partial\mathbb{D})$.
Nothing else in the proof uses $p < \infty$:

- the triangle inequality $\|f - P_r f\|_\infty \leq \|f - g\|_\infty + \|g - P_r g\|_\infty + \|P_r g - P_r f\|_\infty$ holds in any normed space;
- 11.18 supplies $R \in [0, 1)$ with $\|g - P_r g\|_\infty < \varepsilon$ for $r \in (R, 1)$, for continuous g, with no reference to p;
- $P_r(g - f) = P_r * (g - f) = (g - f) * P_r$ by 11.35 and commutativity of convolution 11.41, which needs only that both functions lie in $L^1(\partial\mathbb{D})$;
- 11.38 is proved for all $p \in [1, \infty]$, giving $\|P_r * (g - f)\|_\infty \leq \|P_r\|_1 \|g - f\|_\infty$;
- $\|P_r\|_1 = 1$ by 11.16(a) and 11.16(b).

The gap is not an artifact: the conclusion of 11.42 itself is false for $p = \infty$, by Exercise 12(d) in Section 11A.

Problem 11B.6. Suppose $f \in L^1(\partial\mathbb{D})$. Prove that f is real valued (almost everywhere) if and only if $\widehat{f}(-n) = \overline{\widehat{f}(n)}$ for every $n \in \mathbb{Z}$.

Solution. Both directions follow from Exercise 1 in Section 11A, which states that $\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}$ for every $n \in \mathbb{Z}$ (note $\overline{f} \in L^1(\partial\mathbb{D})$ since $|\overline{f}| = |f|$).
If f is real valued almost everywhere then $\overline{f} = f$ in $L^1(\partial\mathbb{D})$, so

$$\widehat{f}(n) = \widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)},$$

and conjugating gives $\widehat{f}(-n) = \overline{\widehat{f}(n)}$.
Conversely, if $\widehat{f}(-n) = \overline{\widehat{f}(n)}$ for all n, then conjugating gives $\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)} = \widehat{f}(n)$, so by linearity of $h \mapsto \widehat{h}(n)$ every Fourier coefficient of $\overline{f} - f \in L^1(\partial\mathbb{D})$ vanishes. By 11.43, $\overline{f} = f$ almost everywhere.

Problem 11B.7. Suppose $f \in L^1(\partial\mathbb{D})$. Show that $f \in L^2(\partial\mathbb{D})$ if and only if $\displaystyle\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 < \infty$.

Solution. If $f \in L^2(\partial\mathbb{D})$ then $\sum_n |\widehat{f}(n)|^2 = \|f\|_2^2 < \infty$ by Parseval 8.63(c), since $\{z^n\}_{n\in\mathbb{Z}}$ is an orthonormal basis by 11.30 and $\widehat{f}(n) = \langle f, z^n \rangle$.
Conversely, suppose $f \in L^1(\partial\mathbb{D})$ with $\sum_n |\widehat{f}(n)|^2 < \infty$. As $L^2(\partial\mathbb{D})$ is a Hilbert space (complete by 7.24) in which $\{z^n\}$ is orthonormal (11.6), 8.58(a) makes the unordered sum

$$g := \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, z^n$$

converge in $L^2(\partial\mathbb{D})$. For $m \in \mathbb{Z}$ and finite $\Omega \ni m$, orthonormality gives $\langle \sum_{n\in\Omega} \widehat{f}(n) z^n, z^m \rangle = \widehat{f}(m)$, so by Cauchy–Schwarz

$$\big|\langle g, z^m \rangle - \widehat{f}(m)\big| \leq \Big\|g - \sum_{n\in\Omega} \widehat{f}(n) z^n\Big\|_2,$$

which can be made arbitrarily small; hence $\widehat{g}(m) = \widehat{f}(m)$ for all m. Now $g \in L^2(\partial\mathbb{D}) \subseteq L^1(\partial\mathbb{D})$, the inclusion by 7.9 with $\sigma(\partial\mathbb{D}) = 1$, so $f - g \in L^1(\partial\mathbb{D})$ has all Fourier coefficients 0; by 11.43, $f = g$ almost everywhere, and so $f \in L^2(\partial\mathbb{D})$.

Problem 11B.8. Suppose $f \in L^2(\partial D)$. Prove that $|f(z)| = 1$ for almost every $z \in \partial D$ if and only if

$$\sum_{k=-\infty}^{\infty} \hat{f}(k)\,\overline{\hat{f}(k-n)} = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \neq 0 \end{cases}$$

for all $n \in \mathbb{Z}$.

Solution. Both directions come from the identity

$$\widehat{|f|^2}(n) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\overline{\hat{f}(k-n)} \qquad (n \in \mathbb{Z}),$$

together with $\hat{1}(n) = \langle z^0, z^n \rangle$, which is 1 for $n = 0$ and 0 otherwise by orthonormality 11.6.
To prove the identity, fix n and set $g = z^n f \in L^2(\partial D)$ (as $|z^n| = 1$). Since $z^n \overline{z}^k = \overline{z}^{k-n}$ on ∂D, we get $\hat{g}(k) = \hat{f}(k-n)$. The family $\{z^k\}_{k \in \mathbb{Z}}$ is an orthonormal basis by 11.30, so Parseval 8.63(b) applied to the pair f, g gives

$$\sum_{k=-\infty}^{\infty} \hat{f}(k)\,\overline{\hat{f}(k-n)} = \langle f, g \rangle = \int_{\partial D} |f|^2\, \overline{z}^n \, d\sigma = \widehat{|f|^2}(n),$$

where $|f|^2 \in L^1(\partial D)$ by Cauchy–Schwarz, and the series converges absolutely by Cauchy–Schwarz in $\ell^2(\mathbb{Z})$ using Bessel 8.57.
(i) If $|f| = 1$ almost everywhere then $|f|^2 = 1$, so the left side equals $\hat{1}(n)$ for every n, which is the displayed condition. (ii) If the condition holds then $|f|^2 - 1 \in L^1(\partial D)$ has all Fourier coefficients 0 by 11.9(a) and 11.9(b), so $|f|^2 = 1$ almost everywhere by 11.43.

Problem 11B.9. For this exercise, for each $r \in [0, 1)$ think of $\mathcal{P}_r$ as an operator on $L^2(\partial D)$.

- (a) Show that $\mathcal{P}_r$ is a self-adjoint compact operator for each $r \in [0, 1)$.
- (b) For each $r \in [0, 1)$, find all eigenvalues and eigenvectors of $\mathcal{P}_r$.
- (c) Prove or disprove: $\lim_{r \uparrow 1} \|I - \mathcal{P}_r\| = 0$.

Solution. In the orthonormal basis $\{z^n\}_{n \in \mathbb{Z}}$ of $L^2(\partial D)$ (11.30), $\mathcal{P}_r$ is the diagonal operator

$$\widehat{\mathcal{P}_r f}(n) = r^{|n|} \hat{f}(n) \qquad (n \in \mathbb{Z}),$$

since the series of 11.11 converges uniformly on ∂D, hence in L^2 as σ is a probability measure. Parseval 8.63(c) then gives $\|\mathcal{P}_r f\|_2 \le \|f\|_2$.
(a) Self-adjoint: by Parseval 8.63(b),

$$\langle \mathcal{P}_r f, g \rangle = \sum_{n=-\infty}^{\infty} r^{|n|} \hat{f}(n) \overline{\hat{g}(n)} = \langle f, \mathcal{P}_r g \rangle,$$

each $r^{|n|}$ being real. Compact: let $T_M f = \sum_{|n| \le M} r^{|n|} \hat{f}(n) z^n$, which is bounded with finite-dimensional range, hence compact by 10.67. Parseval gives $\|(\mathcal{P}_r - T_M) f\|_2^2 = \sum_{|n| > M} r^{2|n|} |\hat{f}(n)|^2 \le r^{2(M+1)} \|f\|_2^2$, so $\|\mathcal{P}_r - T_M\| \le r^{M+1} \to 0$ as $r < 1$; now 10.69(a).
(b) Comparing Fourier coefficients, $\mathcal{P}_r f = \lambda f$ forces $\hat{f}(n) = 0$ whenever $r^{|n|} \neq \lambda$, so any $\lambda \notin \{r^{|n|}\}$ gives $f = 0$ by 11.43; and each $r^{|n|}$ is an eigenvalue with eigenvector z^n.

- (i) $r \in (0, 1)$: the eigenvalues are $1, r, r^2, \ldots$, with eigenvectors the nonzero constants for $\lambda = 1$ and the nonzero elements $a z^m + b z^{-m}$ of $\operatorname{span}\{z^m, z^{-m}\}$ for $\lambda = r^m$, $m \in \mathbb{Z}^+$. Here 0 is not an eigenvalue.

- (ii) $r=0$: $\mathcal{P}_0 f = \langle f, 1\rangle 1$ is the orthogonal projection onto the constants, with eigenvalue 1 (eigenvectors the nonzero constants) and eigenvalue 0 (eigenvectors the nonzero f with $\hat{f}(0)=0$).

(c) False: $\|I-\mathcal{P}_r\| = 1$ for every $r\in[0,1)$. Indeed $I-\mathcal{P}_r$ is diagonal with entries $1-r^{|n|}\in[0,1]$, so $\|I-\mathcal{P}_r\|\le 1$ by Parseval, while $\|(I-\mathcal{P}_r)z^n\|_2 = 1-r^{|n|}\to 1$ as $n\to\infty$ and $\|z^n\|_2=1$.

Problem 11B.10. Suppose $f\in L^1(\partial D)$. Define $T\colon L^2(\partial D)\to L^2(\partial D)$ by $Tg = f*g$.

- (a) Show that T is a compact operator on $L^2(\partial D)$.
- (b) Prove that T is injective if and only if $\hat{f}(n)\ne 0$ for every $n\in\mathbb{Z}$.
- (c) Find a formula for T^*.
- (d) Prove: T is self-adjoint if and only if all Fourier coefficients of f are real.
- (e) Show that T is a normal operator.

Solution. By 11.44, T is the diagonal operator $\widehat{Tg}(n)=\hat{f}(n)\hat{g}(n)$ in the orthonormal basis $\{z^n\}_{n\in\mathbb{Z}}$ of 11.30; it maps $L^2(\partial D)$ into itself because $L^2(\partial D)\subseteq L^1(\partial D)$ (σ is finite) and 11.38 with $p=2$ gives $\|Tg\|_2\le\|f\|_1\|g\|_2$. Throughout we use Parseval 8.63(c) and the fact that Fourier coefficients determine an element of $L^1(\partial D)$ (11.43).

(a) Set $\varepsilon_M=\sup_{|n|>M}|\hat{f}(n)|$, which tends to 0 by the Riemann–Lebesgue Lemma 11.10. With Q_M the orthogonal projection onto $\operatorname{span}\{z^n : |n|\le M\}$, the operator $Q_M T$ has finite-dimensional range, hence is compact by 10.67, and Parseval gives

$$\|(T-Q_MT)g\|_2^2=\sum_{|n|>M}|\hat{f}(n)|^2|\hat{g}(n)|^2\le\varepsilon_M^2\|g\|_2^2.$$

So $\|T-Q_MT\|\le\varepsilon_M\to 0$, and T is compact by 10.69(a).

(b) If $\hat{f}(m)=0$ then all Fourier coefficients of Tz^m vanish (those with $n\ne m$ because $\widehat{z^m}(n)=0$ by 11.6), so $Tz^m=0$ by 11.43 with $z^m\ne 0$. Conversely, if every $\hat{f}(n)\ne 0$ and $Tg=0$, then $\hat{g}(n)=0$ for all n, so $g=0$ by 11.43.

(c) $T^*h=f^{\#}*h$, where $f^{\#}(z)=\overline{f(\bar{z})}$; equivalently $\widehat{T^*h}(n)=\overline{\hat{f}(n)}\hat{h}(n)$. Taking $z=1$ in 11.17 gives $\int_{\partial D}h(\bar{w})\,d\sigma(w)=\int_{\partial D}h\,d\sigma$; applied to $|f|$ this puts $f^{\#}$ in $L^1(\partial D)$, and applied to $\overline{f(\zeta)}\zeta^n$ it gives $\widehat{f^{\#}}(n)=\overline{\hat{f}(n)}$. Hence, by Parseval 8.63(b) and 11.44 twice,

$$\langle Tg,h\rangle=\sum_{n=-\infty}^{\infty}\hat{g}(n)\overline{\overline{\hat{f}(n)}\hat{h}(n)}=\langle g, f^{\#}*h\rangle,$$

the sums converging absolutely since $\{\hat{f}(n)\}$ is bounded by 11.9(c) while $\{\hat{g}(n)\},\{\hat{h}(n)\}\in\ell^2(\mathbb{Z})$. As $h\mapsto f^{\#}*h$ is bounded on $L^2(\partial D)$, uniqueness of the adjoint gives the formula; unwinding with 11.17 once more,

$$(T^*h)(z)=\int_{\partial D}\overline{f(w)}\,h(zw)\,d\sigma(w).$$

(d) If $T=T^*$, comparing the m-th Fourier coefficients of Tz^m and T^*z^m (and $\widehat{z^m}(m)=1$ by 11.6) gives $\hat{f}(m)=\overline{\hat{f}(m)}$, so every $\hat{f}(m)$ is real. Conversely, if every $\hat{f}(n)$ is real then $\widehat{f^{\#}}(n)=\hat{f}(n)$, so $f^{\#}=f$ by 11.43 and $T^*=T$.

(e) By (c) and 11.44, $\widehat{TT^*g}(n)=|\hat{f}(n)|^2\hat{g}(n)=\widehat{T^*Tg}(n)$ for every n, so $TT^*g=T^*Tg$ by 11.43.

Problem 11B.11. Show that if $f,g\in L^1(\partial D)$ then

$$(f*g)^{\sim}(t)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\tilde{f}(x)\,\tilde{g}(t-x)\,dx,$$

for those $t\in\mathbb{R}$ such that $(f*g)(e^{it})$ makes sense; here $(f*g)^{\sim}$, $\tilde{f}$, and $\tilde{g}$ denote the transfers to the real line as defined in 11.24.

Solution. Put $z = e^{it}$ and apply the transfer of integration recorded after 11.4 to $w \mapsto f(w)g(z\bar{w})$, the integrand in the definition 11.36 of convolution. Writing $w = e^{ix}$ gives $z\bar{w} = e^{i(t-x)}$, so the integrand transfers to $\tilde{f}(x)\,\tilde{g}(t-x)$ and

$$(f * g)^{\sim}(t) = \int_{\partial D} f(w)\,g(z\bar{w})\,d\sigma(w) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{f}(x)\,\tilde{g}(t-x)\,dx,$$

for exactly those t at which $(f * g)(e^{it})$ makes sense, which by 11.37 is almost every t. (Here $\tilde{g}(t-x)$ is unambiguous because $\tilde{g}$ is 2π-periodic, and $w \mapsto g(z\bar{w})$ is measurable since $w \mapsto z\bar{w}$ corresponds under 11.2 to $x \mapsto t - x$ modulo 2π.)

Problem 11B.12. Suppose $1 \le p \le \infty$. Prove that if $f \in L^p(\partial D)$ and $g \in L^{p'}(\partial D)$, then $f * g$ is a continuous function on ∂D.

Solution. Everything follows from continuity of translation: writing $h_z(w) = h(z\bar{w})$, the map $z \mapsto h_z$ is continuous from ∂D into $L^q(\partial D)$ whenever $q < \infty$ and $h \in L^q(\partial D)$.
Note first that $\|h_z\|_q = \|h\|_q$ for every $q \in [1,\infty]$, by 11.17 applied to $|h|^q$ when $q < \infty$ and because $w \mapsto z\bar{w}$ preserves null sets when $q = \infty$. Hence 7.9 gives, for each z,

$$\int_{\partial D} |f(w)g(z\bar{w})|\,d\sigma(w) \le \|f\|_p\,\|g_z\|_{p'} = \|f\|_p\,\|g\|_{p'} < \infty,$$

so $(f * g)(z)$ is defined at every point of ∂D.
For the claim, fix z_0 and $\varepsilon > 0$, and (continuous functions being dense in $L^q(\partial D)$ for $q < \infty$, as used in the proof of 11.42) pick φ continuous with $\|h - \varphi\|_q < \varepsilon$. Since $\|h_z - \varphi_z\|_q = \|h - \varphi\|_q$,

$$\|h_z - h_{z_0}\|_q \le 2\varepsilon + \|\varphi_z - \varphi_{z_0}\|_q \le 2\varepsilon + \|\varphi_z - \varphi_{z_0}\|_\infty,$$

using $\sigma(\partial D) = 1$. As $|z\bar{w} - z_0\bar{w}| = |z - z_0|$ for all w, and φ is uniformly continuous on the compact set ∂D, the last term tends to 0 as $z \to z_0$. Since ε was arbitrary, $\|h_z - h_{z_0}\|_q \to 0$.
At least one of p, p' is finite. (i) If $p' < \infty$, then 7.9 gives

$$|(f * g)(z) - (f * g)(z_0)| \le \|f\|_p\,\|g_z - g_{z_0}\|_{p'} \longrightarrow 0$$

as $z \to z_0$, by the claim with $h = g$, $q = p'$. (ii) If $p' = \infty$ then $p = 1$, and since $f * g = g * f$ by 11.41 (applicable as $L^q(\partial D) \subseteq L^1(\partial D)$ for finite σ), the same estimate with the roles reversed gives $|(g * f)(z) - (g * f)(z_0)| \le \|g\|_\infty \|f_z - f_{z_0}\|_1 \to 0$. In both cases $f * g$ is continuous.

Problem 11B.13. Suppose $g \in L^1(\partial D)$ is such that $\hat{g}(n) \ne 0$ for infinitely many $n \in \mathbb{Z}$. Prove that if $f \in L^1(\partial D)$, then $f * g \ne g$.

Solution. Suppose $f * g = g$ for some $f \in L^1(\partial D)$. Taking Fourier coefficients, 11.44 gives

$$\big(\hat{f}(n) - 1\big)\hat{g}(n) = 0 \qquad (n \in \mathbb{Z}),$$

so $\hat{f}(n) = 1$ on the set $S = \{n : \hat{g}(n) \ne 0\}$, which is infinite and hence unbounded. Choosing $n_k \in S$ with $|n_k| \to \infty$ contradicts the Riemann–Lebesgue Lemma 11.10, which forces $\hat{f}(n_k) \to 0$.

Problem 11B.14. Show that there exists a two-sided sequence $\ldots, b_{-2}, b_{-1}, b_0, b_1, b_2, \ldots$ such that $\lim_{n\to\pm\infty} b_n = 0$ but there does not exist $f \in L^1(\partial D)$ with $\hat{f}(n) = b_n$ for all $n \in \mathbb{Z}$.

Solution. Such a sequence exists because the map $\Phi(f) = (\hat{f}(n))_{n\in\mathbb{Z}}$ from $L^1(\partial D)$ to $c_0(\mathbb{Z})$ is not surjective, where $c_0(\mathbb{Z})$ is the space of two-sided sequences tending to 0 at $\pm\infty$ under the supremum norm.

Both spaces are Banach: $L^1(\partial D)$ by 7.24, and $c_0(\mathbb{Z})$ as a closed subspace of $\ell^\infty(\mathbb{Z})$ (a uniform limit of sequences vanishing at infinity vanishes at infinity). The map Φ lands in $c_0(\mathbb{Z})$ by the Riemann–Lebesgue Lemma 11.10, is linear by 11.9(a) and 11.9(b), is bounded by 11.9(c), and is injective by 11.43. So if Φ were surjective, the Bounded Inverse Theorem 6.83 would supply $c \in (0, \infty)$ with

$$\|f\|_1 \le c \sup_{n\in\mathbb{Z}} |\hat{f}(n)| \qquad \text{for all } f \in L^1(\partial D).$$

The Dirichlet kernels $D_M(z) = \sum_{n=-M}^{M} z^n$ refute this. They are continuous, and by 11.6 their Fourier coefficients are 1 for $|n| \le M$ and 0 otherwise, so the right side above would be c for every M. But summing the geometric series gives $D_M(e^{it}) = \sin\big((M+\frac{1}{2})t\big)/\sin\frac{t}{2}$, so by Exercise 11(b) in Section 11A,

$$\|D_M\|_1 = \int_{-\pi}^{\pi} \left| \frac{\sin\big((M+\frac{1}{2})t\big)}{\sin\frac{t}{2}} \right| \frac{dt}{2\pi} \longrightarrow \infty$$

as $M \to \infty$. Hence Φ is not surjective, and any $b \in c_0(\mathbb{Z})$ outside its range is the required sequence.

Problem 11B.15. Prove that if $f, g \in L^2(\partial D)$, then

$$\widehat{fg}(n) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\hat{g}(n-k)$$

for every $n \in \mathbb{Z}$.

Solution. Fix n and set $h(z) = \overline{g(z)}\, z^n$, so that $h \in L^2(\partial D)$ with $\overline{h(z)} = g(z)\bar{z}^n$ and hence

$$\widehat{fg}(n) = \int_{\partial D} f(z)g(z)\bar{z}^n \, d\sigma(z) = \langle f, h\rangle.$$

(Here $fg \in L^1(\partial D)$ by 7.9.) Since $z\bar{z} = 1$ on ∂D,

$$\hat{h}(k) = \int_{\partial D} \overline{g(z)\,\bar{z}^{n-k}}\, d\sigma(z) = \overline{\hat{g}(n-k)}.$$

The family $\{z^k\}_{k\in\mathbb{Z}}$ is an orthonormal basis by 11.30, so Parseval 8.63(b) gives

$$\widehat{fg}(n) = \langle f, h\rangle = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\overline{\hat{h}(k)} = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\hat{g}(n-k),$$

the series converging absolutely by Cauchy–Schwarz in $\ell^2(\mathbb{Z})$ together with Bessel 8.57, since $k \mapsto n-k$ is a bijection of $\mathbb{Z}$.

Problem 11B.16. Suppose $f \in L^1(\partial D)$. Prove that $\mathcal{P}_r(\mathcal{P}_s f) = \mathcal{P}_{rs} f$ for all $r, s \in [0, 1)$.

Solution. Both sides have the same Fourier coefficients, namely $(rs)^{|n|}\hat{f}(n)$.
Indeed, the series $\sum_m s^{|m|}\hat{f}(m)z^m$ of 11.11 converges uniformly on ∂D (its terms are dominated by $\|f\|_1 s^{|m|}$), so it may be integrated termwise against $\bar{z}^n$; orthonormality of $\{z^m\}_{m\in\mathbb{Z}}$ (11.6) then leaves one term:

$$\widehat{\mathcal{P}_s f}(n) = s^{|n|}\hat{f}(n) \qquad (n \in \mathbb{Z}).$$

Uniform convergence also makes $\mathcal{P}_s f$ continuous, hence in $L^1(\partial D)$, so 11.11 applies to it and gives, for $z \in \partial D$,

$$\big(\mathcal{P}_r(\mathcal{P}_s f)\big)(z) = \sum_{n=-\infty}^{\infty} r^{|n|}s^{|n|}\hat{f}(n)z^n = (\mathcal{P}_{rs} f)(z),$$

by 11.11 again with $rs \in [0,1)$.

Problem 11B.17. Suppose $p \in [1,\infty]$ and $f \in L^p(\partial D)$. Prove that if $0 \le r < s < 1$, then

$$\|\mathcal{P}_r f\|_p \le \|\mathcal{P}_s f\|_p.$$

Solution. Factor $\mathcal{P}_r$ through $\mathcal{P}_s$: since $0 \le r < s < 1$ we have $t := r/s \in [0,1)$, and Exercise 16 in this section gives

$$\mathcal{P}_t(\mathcal{P}_s f) = \mathcal{P}_{ts} f = \mathcal{P}_r f.$$

(Here $f \in L^1(\partial D)$ because $\|f\|_1 \le \|f\|_p$ by 7.9, σ being a probability measure.)
Now $\mathcal{P}_t$ does not increase the L^p-norm. Indeed 11.35 says $\mathcal{P}_t g = g * P_t$, which equals $P_t * g$ by 11.41; the Poisson kernel is continuous and positive with $\|P_t\|_1 = 1$ by 11.16(a) and 11.16(b); so 11.38 gives

$$\|\mathcal{P}_t g\|_p = \|P_t * g\|_p \le \|P_t\|_1 \, \|g\|_p = \|g\|_p.$$

Apply this with $g = \mathcal{P}_s f$, which lies in $L^p(\partial D)$ because the series of 11.11 converges uniformly on ∂D, making $\mathcal{P}_s f$ continuous hence bounded. Combining the two displays gives $\|\mathcal{P}_r f\|_p \le \|\mathcal{P}_s f\|_p$.

Problem 11B.18. Prove Wirtinger's inequality: If $f\colon \mathbb{R} \to \mathbb{R}$ is a continuously differentiable 2π-periodic function and $\int_{-\pi}^{\pi} f(t)\,dt = 0$, then

$$\int_{-\pi}^{\pi} \big(f(t)\big)^2\,dt \le \int_{-\pi}^{\pi} \big(f'(t)\big)^2\,dt,$$

with equality if and only if $f(t) = a\sin(t) + b\cos(t)$ for some constants a, b.

Solution. Transfer to the circle: let $F(e^{it}) = f(t)$, well defined by 2π-periodicity, so $\widetilde{F} = f$ and F is once continuously differentiable in the sense of 11.24, with $F^{[1]}(e^{it}) = f'(t)$. Both are continuous, hence in $L^2(\partial D)$, and by the transfer of integration after 11.4,

$$\|F\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)^2\,dt, \qquad \big\|F^{[1]}\big\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f'(t)^2\,dt.$$

The hypotheses read $\hat{F}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f = 0$ and, by 11.26 with $k=1$, $\widehat{F^{[1]}}(n) = in\hat{F}(n)$.
Since $\{z^n\}_{n\in\mathbb{Z}}$ is an orthonormal basis by 11.30, Parseval 8.63(c) applied to F and to $F^{[1]}$ gives, both sums being finite,

$$\big\|F^{[1]}\big\|_2^2 - \|F\|_2^2 = \sum_{n\ne 0}(n^2-1)\big|\hat{F}(n)\big|^2 \ge 0,$$

using $\hat{F}(0) = 0$ to drop the $n = 0$ term from $\|F\|_2^2$. Multiplying by 2π gives Wirtinger's inequality.
Equality forces every term to vanish, so $\hat{F}(n) = 0$ for $|n| \ge 2$ as well as for $n = 0$. By 11.31 the Fourier series converges to F in $L^2(\partial D)$, so $F(z) = \hat{F}(1)z + \hat{F}(-1)\bar{z}$ almost everywhere, hence everywhere: both sides are continuous, so the set where they differ is relatively open, and a nonempty relatively open subset of ∂D has positive σ-measure by 11.4. Since f is real valued, $\hat{F}(-1) = \int_{\partial D} F(z)z\,d\sigma(z) = \overline{\hat{F}(1)}$, so writing $\hat{F}(1) = \alpha - i\beta$,

$$f(t) = 2\,\mathrm{Re}\big(\hat{F}(1)e^{it}\big) = 2\beta\sin t + 2\alpha\cos t.$$

Conversely, for $f(t) = a\sin t + b\cos t$ we have $f'(t) = a\cos t - b\sin t$, and $\int_{-\pi}^{\pi}\sin^2 = \int_{-\pi}^{\pi}\cos^2 = \pi$ with $\int_{-\pi}^{\pi}\sin t\cos t\,dt = 0$ make both sides equal $\pi(a^2+b^2)$.

11.3 Exercises 11C

Problem 11C.1. Suppose $f \in L^1(\mathbf{R})$. Prove that $\|\hat{f}\|_\infty = \|f\|_1$ if and only if there exists $\zeta \in \partial\mathbf{D}$ and $t \in \mathbf{R}$ such that $\zeta f(x)e^{-itx} \geq 0$ for almost every $x \in \mathbf{R}$.

Solution. Suppose first that $\zeta f(x)e^{-itx} \geq 0$ almost everywhere for some $\zeta \in \partial\mathbf{D}$, $t \in \mathbf{R}$. Since $|\zeta f(x)e^{-itx}| = |f(x)|$, that hypothesis says $\zeta f(x)e^{-itx} = |f(x)|$ almost everywhere, so with $s_0 = t/(2\pi)$,

$$|\hat{f}(s_0)| = \left|\bar{\zeta}\int_{-\infty}^{\infty} \zeta f(x)e^{-itx}\,dx\right| = \|f\|_1.$$

Hence $\|\hat{f}\|_\infty \geq \|f\|_1$, and the reverse inequality is 11.49.
Conversely, suppose $\|\hat{f}\|_\infty = \|f\|_1$; we may assume $\|f\|_1 > 0$, since otherwise $\zeta = 1$, $t = 0$ works. As $\hat{f}$ is continuous, $\|\hat{f}\|_\infty = \sup_s |\hat{f}(s)|$ (a continuous function exceeding c somewhere exceeds c on a set of positive measure). By 11.49 there is M with $|\hat{f}(s)| < \frac{1}{2}\|f\|_1$ for $|s| > M$, so the supremum is attained on the compact interval $[-M, M]$, say at s_0, giving $|\hat{f}(s_0)| = \|f\|_1$.
Write $\hat{f}(s_0) = \lambda\|f\|_1$ with $|\lambda| = 1$, and set $\zeta = \bar{\lambda}$, $t = 2\pi s_0$, $g(x) = \zeta f(x)e^{-itx}$. Then $|g| = |f|$ and

$$\int_{-\infty}^{\infty} g(x)\,dx = \bar{\lambda}\,\hat{f}(s_0) = \|f\|_1 = \int_{-\infty}^{\infty} |g(x)|\,dx.$$

Taking real parts, the nonnegative function $h = |g| - \operatorname{Re} g \in L^1(\mathbf{R})$ has $\int h = 0$, so $h = 0$ almost everywhere by Markov's inequality 4.1 applied with $c = 1/n$ and countable subadditivity. Where $\operatorname{Re} g(x) = |g(x)|$ we get $(\operatorname{Im} g(x))^2 = 0$, hence $g(x) = |g(x)| \geq 0$; that is, $\zeta f(x)e^{-itx} \geq 0$ almost everywhere.

Problem 11C.2. Suppose $f(x) = xe^{-\pi x^2}$ for all $x \in \mathbf{R}$. Show that $\hat{f} = -if$.

Solution. Apply 11.50 to $\varphi(x) = e^{-\pi x^2}$, taking $g(x) = x\varphi(x) = f(x)$; both φ and f lie in $L^1(\mathbf{R})$, since $\int_{-\infty}^{\infty} |x|e^{-\pi x^2}\,dx = 1/\pi < \infty$. As $\hat{\varphi} = \varphi$ by Example 11.51, for all $t \in \mathbf{R}$

$$\begin{aligned} -2\pi i\,\hat{f}(t) &= (\hat{\varphi})'(t) = \varphi'(t) \\ &= -2\pi t\,e^{-\pi t^2}. \end{aligned}$$

Dividing by $-2\pi i$ and using $1/i = -i$ gives $\hat{f}(t) = -i\,te^{-\pi t^2} = -if(t)$.

Problem 11C.3. Suppose $f(x) = 4\pi x^2e^{-\pi x^2} - e^{-\pi x^2}$ for all $x \in \mathbf{R}$. Show that $\hat{f} = -f$.

Solution. Write $f = 4\pi f_2 - \varphi$, where $\varphi(x) = e^{-\pi x^2}$ and $f_k(x) = x^k e^{-\pi x^2}$; each lies in $L^1(\mathbf{R})$, being continuous with decay faster than any power of $1/|x|$. By 11C.2, $\hat{f}_1(t) = -i\,te^{-\pi t^2}$, so applying 11.50 to f_1 with $g = f_2 \in L^1(\mathbf{R})$,

$$\begin{aligned} -2\pi i\,\hat{f}_2(t) &= (\hat{f}_1)'(t) \\ &= \left(-i + 2\pi i t^2\right)e^{-\pi t^2}, \end{aligned}$$

whence $\hat{f}_2(t) = \left(\frac{1}{2\pi} - t^2\right)e^{-\pi t^2}$. Since $\hat{\varphi} = \varphi$ by Example 11.51, linearity of the Fourier transform gives, for all $t \in \mathbf{R}$,

$$\begin{aligned} \hat{f}(t) &= 4\pi\,\hat{f}_2(t) - \hat{\varphi}(t) \\ &= \left(2 - 4\pi t^2\right)e^{-\pi t^2} - e^{-\pi t^2} \\ &= -\left(4\pi t^2 e^{-\pi t^2} - e^{-\pi t^2}\right) = -f(t). \end{aligned}$$

Problem 11C.4. Find $f \in L^1(\mathbb{R})$ such that $f \neq 0$ and $\hat{f} = if$.

Solution. Take $f(x) = (4\pi x^3 - 3x)e^{-\pi x^2}$, which is in $L^1(\mathbb{R})$ (continuous, decaying faster than any power of $1/|x|$) and nonzero, as $f(1) = (4\pi - 3)e^{-\pi} \neq 0$.
With $f_k(x) = x^k e^{-\pi x^2}$, all in $L^1(\mathbb{R})$, we have $f = 4\pi f_3 - 3f_1$. By 11C.2 and the computation in 11C.3, $\hat{f}_1(t) = -i\, te^{-\pi t^2}$ and $\hat{f}_2(t) = \big(\frac{1}{2\pi} - t^2\big)e^{-\pi t^2}$. Applying 11.50 to f_2 with $g = f_3 \in L^1(\mathbb{R})$,

$$\begin{aligned} -2\pi i\, \hat{f}_3(t) &= (\hat{f}_2)'(t) \\ &= \big(2\pi t^3 - 3t\big)e^{-\pi t^2}, \end{aligned}$$

so $\hat{f}_3(t) = \frac{i}{2\pi}(2\pi t^3 - 3t)e^{-\pi t^2}$. Hence for all $t \in \mathbb{R}$,

$$\begin{aligned} \hat{f}(t) &= 4\pi\, \hat{f}_3(t) - 3\, \hat{f}_1(t) \\ &= i\big(4\pi t^3 - 6t\big)e^{-\pi t^2} + 3ite^{-\pi t^2} \\ &= i\big(4\pi t^3 - 3t\big)e^{-\pi t^2} = if(t). \end{aligned}$$

Problem 11C.5. Prove that if p is a polynomial on $\mathbb{R}$ with complex coefficients and $f : \mathbb{R} \to \mathbb{C}$ is defined by $f(x) = p(x)e^{-\pi x^2}$, then there exists a polynomial q on $\mathbb{R}$ with complex coefficients such that $\deg q = \deg p$ and $\hat{f}(t) = q(t)e^{-\pi t^2}$ for all $t \in \mathbb{R}$.

Solution. Set $g_n(x) = x^n e^{-\pi x^2}$, each in $L^1(\mathbb{R})$ by Gaussian decay; we prove by induction that $\hat{g}_n(t) = q_n(t)e^{-\pi t^2}$ for a polynomial q_n of degree n with leading coefficient $(-i)^n$.
For $n = 0$, Example 11.51 gives $q_0 = 1$. Assuming the claim for n, apply 11.50 to g_n with $g = g_{n+1} \in L^1(\mathbb{R})$:

$$\begin{aligned} (\hat{g}_n)'(t) &= \big(q_n'(t) - 2\pi t\, q_n(t)\big)e^{-\pi t^2} \\ &= -2\pi i\, \hat{g}_{n+1}(t), \end{aligned}$$

so $\hat{g}_{n+1}(t) = q_{n+1}(t)e^{-\pi t^2}$ with $q_{n+1} = \frac{i}{2\pi}\big(q_n' - 2\pi t\, q_n\big)$, a polynomial. Since q_n' contributes no term of degree $n+1$ while $t\, q_n(t)$ has degree $n+1$ with leading coefficient $(-i)^n$, the polynomial q_{n+1} has degree $n+1$ with leading coefficient $\frac{i}{2\pi}(-2\pi)(-i)^n = (-i)^{n+1} \neq 0$.
Now let p be a polynomial. If $p = 0$ take $q = 0$. Otherwise write $p(x) = \sum_{k=0}^n a_k x^k$ with $a_n \neq 0$; then $f = \sum_{k=0}^n a_k g_k \in L^1(\mathbb{R})$, so linearity of the Fourier transform gives $\hat{f}(t) = q(t)e^{-\pi t^2}$ with $q = \sum_{k=0}^n a_k q_k$. As $\deg q_k = k \leq n$, no power above t^n occurs in q, and the coefficient of t^n is $a_n(-i)^n \neq 0$; hence $\deg q = n = \deg p$.

Problem 11C.6. Suppose

$$f(x) = \begin{cases} xe^{-2\pi x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

Show that $\hat{f}(t) = \dfrac{1}{4\pi^2(1+it)^2}$ for all $t \in \mathbb{R}$.

Solution. Everything comes from the formula $\int_0^\infty xe^{-ax}\,dx = 1/a^2$, valid for every $a \in \mathbb{C}$ with $c = \operatorname{Re} a > 0$, applied with $a = 2\pi(1+it)$.
To prove the formula, put $F(x) = -\frac{x}{a}e^{-ax} - \frac{1}{a^2}e^{-ax}$; then $F'(x) = xe^{-ax}$ (Check!), and since $|e^{-ax}| = e^{-cx}$,

$$|F(x)| \leq \Big(\frac{x}{|a|} + \frac{1}{|a|^2}\Big)e^{-cx} \longrightarrow 0 \quad \text{as } x \to \infty.$$

The same antiderivative with a replaced by c gives $\int_0^M xe^{-cx}\,dx \nearrow 1/c^2$, so by 3.11 applied to $\chi_{(0,k)}(x)\, xe^{-cx}$ the dominating function $|xe^{-ax}| = xe^{-cx}$ is integrable on $(0,\infty)$. Hence the Fundamental Theorem of Calculus on $[0,M]$ (applied to the real and imaginary parts of F) together with 3.31 yields

$$\int_0^\infty xe^{-ax}\,dx = \lim_{M\to\infty}\bigl(F(M)-F(0)\bigr) = \frac{1}{a^2}.$$

Taking $a = 2\pi$ shows $\int_{-\infty}^\infty |f| = 1/(4\pi^2) < \infty$, so $f \in L^1(\mathbb{R})$. Now fix $t \in \mathbb{R}$ and take $a = 2\pi(1+it)$, which has $\operatorname{Re} a = 2\pi > 0$; as f vanishes on $(-\infty, 0]$,

$$\hat{f}(t) = \int_0^\infty xe^{-2\pi(1+it)x}\,dx = \frac{1}{4\pi^2(1+it)^2}.$$

Problem 11C.7. Prove the formulas in 11.55 for the Fourier transforms of translations, rotations, and dilations.

Solution. Each formula is the change of variable that matches it; fix $f \in L^1(\mathbb{R})$, $b, t \in \mathbb{R}$.
(i) $g(x) = f(x-b)$. Translation invariance of Lebesgue measure (2.7) gives $\int_{-\infty}^\infty h(u+b)\,du = \int_{-\infty}^\infty h(x)\,dx$ for every $h \in L^1(\mathbb{R})$: true for characteristic functions of measurable sets by 2.7 (translation carries Borel sets to Borel sets and null sets to null sets, hence measurable sets to measurable sets by 2.70), then for simple functions by linearity, for nonnegative measurable functions by 3.11, and for $L^1(\mathbb{R})$ by splitting into real and imaginary, positive and negative, parts. In particular $\int |g| = \int |f| < \infty$, so $g \in L^1(\mathbb{R})$. Applying this to $h(x) = f(x-b)e^{-2\pi itx}$, for which $h(u+b) = e^{-2\pi ibt}f(u)e^{-2\pi itu}$,

$$\hat{g}(t) = \int_{-\infty}^\infty h(x)\,dx = e^{-2\pi ibt}\hat{f}(t).$$

(ii) $g(x) = e^{2\pi ibx}f(x)$. Here $|g| = |f|$, so $g \in L^1(\mathbb{R})$, and since $e^{2\pi ibx}e^{-2\pi itx} = e^{-2\pi i(t-b)x}$,

$$\hat{g}(t) = \int_{-\infty}^\infty f(x)\,e^{-2\pi i(t-b)x}\,dx = \hat{f}(t-b).$$

(iii) $b \neq 0$, $g(x) = f(bx)$. For $c \neq 0$ the map $I \mapsto cI$ is a bijection of the open intervals onto themselves with $\ell(cI) = |c|\,\ell(I)$, carrying covers of E onto covers of cE, so the definition of outer measure (2.2) gives $|cE| = |c|\,|E|$ for every $E \subseteq \mathbb{R}$; and cE is measurable whenever E is, since by 2.70 there is a Borel $B \subseteq E$ with $|E \setminus B| = 0$ and then $cB \subseteq cE$ is Borel with $|cE \setminus cB| = |c|\cdot 0 = 0$. Hence $x \mapsto \psi(bx)$ is measurable when ψ is, because $\{x : \psi(bx) \in B\} = \frac{1}{b}\{u : \psi(u) \in B\}$, and for $\psi = \chi_E$ we get $\psi(b\,\cdot) = \chi_{(1/b)E}$, so

$$\int_{-\infty}^\infty \psi(bx)\,dx = \left|\tfrac{1}{b}E\right| = \frac{1}{|b|}\int_{-\infty}^\infty \psi(u)\,du,$$

which extends to nonnegative simple functions by linearity, to nonnegative measurable ψ by 2.89 and 3.11, and to $L^1(\mathbb{R})$ by splitting into parts. In particular $\int |g| = \frac{1}{|b|}\int |f| < \infty$, so $g \in L^1(\mathbb{R})$; applying the rule to $\psi(u) = f(u)e^{-2\pi i(t/b)u}$, for which $\psi(bx) = f(bx)e^{-2\pi itx}$,

$$\hat{g}(t) = \frac{1}{|b|}\int_{-\infty}^\infty f(u)\,e^{-2\pi i(t/b)u}\,du = \frac{1}{|b|}\hat{f}\Bigl(\frac{t}{b}\Bigr).$$

Problem 11C.8. Suppose $f \in L^1(\mathbb{R})$ and $n \in \mathbb{Z}^+$. Define $g\colon \mathbb{R} \to \mathbb{C}$ by $g(x) = x^n f(x)$. Prove that if $g \in L^1(\mathbb{R})$, then $\hat{f}$ is n times continuously differentiable on $\mathbb{R}$ and

$$(\hat{f})^{(n)}(t) = (-2\pi i)^n \hat{g}(t)$$

for all $t \in \mathbb{R}$.

Solution. Iterate 11.50 along $g_k(x) = x^k f(x)$ for $k \in \{0, 1, \ldots, n\}$, so that $g_0 = f$ and $g_n = g$. Each g_k lies in $L^1(\mathbb{R})$, because $|x|^k \le 1 + |x|^n$ for all x gives $|g_k| \le |f| + |g|$, an integrable bound.

We show by induction that $\hat{f}$ is k times continuously differentiable with $(\hat{f})^{(k)} = (-2\pi i)^k \widehat{g_k}$. For $k = 0$ this reads $\hat{f} = \widehat{g_0}$, and $\hat{f}$ is continuous by 11.49. Suppose it holds for some $k < n$. Both g_k and $x \mapsto x g_k(x) = g_{k+1}(x)$ lie in $L^1(\mathbb{R})$, so 11.50 makes $\widehat{g_k}$ continuously differentiable with $(\widehat{g_k})' = -2\pi i\, \widehat{g_{k+1}}$, and hence

$$(\hat{f})^{(k+1)}(t) = (-2\pi i)^k (\widehat{g_k})'(t) = (-2\pi i)^{k+1} \widehat{g_{k+1}}(t),$$

which is continuous since $g_{k+1} \in L^1(\mathbb{R})$ (11.49). Taking $k = n$ gives $(\hat{f})^{(n)}(t) = (-2\pi i)^n \hat{g}(t)$ for all $t \in \mathbb{R}$.

Problem 11C.9. Suppose $n \in \mathbb{Z}^+$ and $f \in L^1(\mathbb{R})$ is n times continuously differentiable and $f^{(k)} \in L^1(\mathbb{R})$ for $k = 1, \ldots, n$. Prove that if $t \in \mathbb{R}$, then

$$\widehat{f^{(n)}}(t) = (2\pi i t)^n \hat{f}(t).$$

Solution. Iterate 11.54: we show by induction that $\widehat{f^{(k)}}(t) = (2\pi i t)^k \hat{f}(t)$ for $k \in \{0, 1, \ldots, n\}$, the case $k = 0$ reading $\hat{f} = \hat{f}$.
Suppose the formula holds for some $k < n$. Then $f^{(k)} \in L^1(\mathbb{R})$ (by hypothesis, or by $f^{(0)} = f$ when $k = 0$); since $k + 1 \le n$, the function $f^{(k)}$ is differentiable with derivative $f^{(k+1)}$, which is continuous because f is n times continuously differentiable; and $f^{(k+1)} \in L^1(\mathbb{R})$ by hypothesis. These are exactly the hypotheses of 11.54, which therefore gives

$$\widehat{f^{(k+1)}}(t) = 2\pi i t\, \widehat{f^{(k)}}(t) = (2\pi i t)^{k+1} \hat{f}(t)$$

for all $t \in \mathbb{R}$. The case $k = n$ is the assertion.

Problem 11C.10. Suppose $1 \le p \le \infty$, $f \in L^p(\mathbb{R})$, and $g \in L^{p'}(\mathbb{R})$. Prove that $f * g$ is a uniformly continuous function on $\mathbb{R}$.

Solution. Everything follows from the estimate $|(f * g)(x) - (f * g)(x')| \le \|g\|_{p'} \|f - f_{-h}\|_p$, where $h = x' - x$ and $f_h(u) = f(u - h)$ as in Exercise 23 in Section 7A.
Lebesgue measure is invariant under $t \mapsto c - t$ for each $c \in \mathbb{R}$ (translation invariance is 2.7, and $\ell(-I) = \ell(I)$ for open intervals gives invariance under $t \mapsto -t$; both maps preserve measurability), so $\|u(c - \cdot)\|_q = \|u\|_q$ for every $q \in [1, \infty]$. Hence for each x, since $\|g(x - \cdot)\|_{p'} = \|g\|_{p'}$, Hölder's inequality (7.9) puts $t \mapsto f(t) g(x - t)$ in $L^1(\mathbb{R})$ with integral of modulus at most $\|f\|_p \|g\|_{p'}$; thus $(f * g)(x)$ is defined for every x.
Because $\frac{1}{p} + \frac{1}{p'} = 1$, at most one of p, p' equals ∞, and $f * g = g * f$ by 11.65, so interchanging (f, p) with (g, p') if necessary we may assume $p < \infty$. For $x, x' \in \mathbb{R}$ and $h = x' - x$ we have $f(x' - t) = f_{-h}(x - t)$, so 11.65 followed by 7.9 and the invariance above gives

$$\begin{aligned} \big|(f * g)(x) - (f * g)(x')\big| &= \Big| \int_{-\infty}^{\infty} g(t)\, (f - f_{-h})(x - t)\, dt \Big| \\ &\le \|g\|_{p'}\, \|f - f_{-h}\|_p. \end{aligned}$$

Now let $\varepsilon > 0$, and assume $\|g\|_{p'} > 0$ (otherwise $f * g = 0$). Since $1 \le p < \infty$, Exercise 23(a) in Section 7A makes $h \mapsto \|f - f_h\|_p$ uniformly continuous, and it vanishes at $h = 0$; so there is $\delta > 0$ with $\|f - f_{-h}\|_p < \varepsilon / \|g\|_{p'}$ whenever $|h| < \delta$. The display then gives $|(f * g)(x) - (f * g)(x')| < \varepsilon$ whenever $|x - x'| < \delta$, with δ independent of x.

Problem 11C.11. Suppose $f \in L^\infty(\mathbb{R})$, $x \in \mathbb{R}$, and f is continuous at x. Prove that

$$\lim_{y \downarrow 0} (\mathcal{P}_y f)(x) = f(x).$$

Solution. Since $f \in L^\infty(\mathbb{R})$ and $P_y \in L^1(\mathbb{R})$, the convolution $\mathcal{P}_y f = f * P_y$ is defined at every point (11.70 with $p = \infty$) and 11.65 writes it as $(\mathcal{P}_y f)(x) = \int_{-\infty}^{\infty} P_y(t) f(x-t)\,dt$.
Continuity at x forces $|f(x)| \le \|f\|_\infty$: given $\varepsilon > 0$, continuity supplies $\delta > 0$ with $|f(z) - f(x)| < \varepsilon$ for $|z - x| < \delta$, and since $(x-\delta, x+\delta)$ has positive measure while $|f| \le \|f\|_\infty$ almost everywhere, some such z has $|f(z)| \le \|f\|_\infty$, whence $|f(x)| \le \|f\|_\infty + \varepsilon$. Hence $|f(x-t) - f(x)| \le 2\|f\|_\infty$ for almost every t.
Fix $\varepsilon > 0$ and take $\delta > 0$ with $|f(x-t) - f(x)| < \varepsilon$ whenever $|t| < \delta$. Using $P_y > 0$ and $\int_{-\infty}^{\infty} P_y = 1$ (11.69(a),(b)) and splitting at $|t| = \delta$,

$$\begin{aligned}\big|(\mathcal{P}_y f)(x) - f(x)\big| &\le \int_{-\infty}^{\infty} |f(x-t) - f(x)|\, P_y(t)\,dt \\ &\le \varepsilon + 2\|f\|_\infty \int_{\{t\,:\,|t|\ge\delta\}} P_y(t)\,dt\end{aligned}$$

for every $y > 0$. The last integral tends to 0 as $y \downarrow 0$ by 11.69(c), so the limit superior of the left side is at most ε; as ε was arbitrary, $\lim_{y\downarrow 0}(\mathcal{P}_y f)(x) = f(x)$.

Problem 11C.12. Suppose $p \in [1,\infty]$ and $f \in L^p(\mathbb{R})$. Prove that $\mathcal{P}_y(\mathcal{P}_{y'} f) = \mathcal{P}_{y+y'} f$ for all $y, y' > 0$.

Solution. The identity is the semigroup law $P_y * P_{y'} = P_{y+y'}$ transported through Fubini; p' is the conjugate exponent.
First, $P_y \in L^q(\mathbb{R})$ for every $q \in [1,\infty]$: by 11.68, $0 < P_y(x) \le \frac{1}{\pi y}$, while $\|P_y\|_1 = 1$ by 11.69(a),(b), so $\|P_y\|_q^q \le \|P_y\|_\infty^{q-1}\|P_y\|_1 < \infty$ when $q < \infty$. Hence for $g \in L^p(\mathbb{R})$, 11.64 and 11.65 give

$$\|\mathcal{P}_y g\|_p = \|P_y * g\|_p \le \|P_y\|_1\,\|g\|_p = \|g\|_p,$$

so $\mathcal{P}_{y'} f \in L^p(\mathbb{R})$ and $\mathcal{P}_y(\mathcal{P}_{y'} f)$ is defined everywhere ($P_y \in L^{p'}(\mathbb{R})$ with 7.9).
For the semigroup law, $P_y * P_{y'}$ and $P_{y+y'}$ both lie in $L^1(\mathbb{R})$ (11.64 with $p = 1$), and 11.66 with 11.79 gives

$$\widehat{P_y * P_{y'}}(t) = e^{-2\pi y|t|} e^{-2\pi y'|t|} = \widehat{P_{y+y'}}(t),$$

so their difference has Fourier transform 0 and 11.80 makes them equal almost everywhere. Both are continuous ($P_{y+y'}$ visibly; $P_y * P_{y'}$ by 11C.10 with exponents 1 and ∞), hence equal everywhere.
Fix $x \in \mathbb{R}$. The nonnegative $F(s,t) = |f(t)| P_{y'}(s-t) P_y(x-s)$ is product measurable (a function of t alone times continuous functions of (s,t)), so Tonelli (5.28) and then Hölder (7.9) give

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(s,t)\,dt\,ds &= \int_{-\infty}^{\infty} P_y(x-s)\,\big(|f| * P_{y'}\big)(s)\,ds \\ &\le \|P_y\|_{p'}\,\|f\|_p < \infty,\end{aligned}$$

using $\big\||f| * P_{y'}\big\|_p \le \|f\|_p$ and $\|P_y(x-\cdot)\|_{p'} = \|P_y\|_{p'}$ (P_y is even and 2.7 applies). Hence Fubini (5.32) permits the interchange below, and the substitution $r = s - t$ (again 2.7) identifies the inner integral:

$$\begin{aligned}\big(\mathcal{P}_y(\mathcal{P}_{y'} f)\big)(x) &= \int_{-\infty}^{\infty} f(t)\Big(\int_{-\infty}^{\infty} P_{y'}(s-t)\,P_y(x-s)\,ds\Big)dt \\ &= \int_{-\infty}^{\infty} f(t)\,(P_{y'} * P_y)(x-t)\,dt \\ &= \int_{-\infty}^{\infty} f(t)\,P_{y+y'}(x-t)\,dt = \big(\mathcal{P}_{y+y'} f\big)(x),\end{aligned}$$

the last line by the semigroup law together with 11.65.

Problem 11C.13. Suppose $p \in [1, \infty]$ and $f \in L^p(\mathbb{R})$. Prove that if $0 < y < y'$, then

$$\|\mathcal{P}_y f\|_p \geq \|\mathcal{P}_{y'} f\|_p.$$

Solution. Set $s = y' - y > 0$, so that $\mathcal{P}_{y'} f = \mathcal{P}_s(\mathcal{P}_y f)$ by 11C.12.
Each $\mathcal{P}_s$ is a contraction on $L^p(\mathbb{R})$: since $\|P_s\|_1 = 1$ by 11.69(a),(b), results 11.64 and 11.65 give $\mathcal{P}_s g = P_s * g \in L^p(\mathbb{R})$ with $\|\mathcal{P}_s g\|_p \leq \|P_s\|_1 \|g\|_p = \|g\|_p$ for every $g \in L^p(\mathbb{R})$. Applying this with $g = \mathcal{P}_y f \in L^p(\mathbb{R})$,

$$\begin{aligned}\|\mathcal{P}_{y'} f\|_p &= \|\mathcal{P}_s(\mathcal{P}_y f)\|_p \\ &\leq \|\mathcal{P}_y f\|_p.\end{aligned}$$

Problem 11C.14. Suppose $f \in L^1(\mathbb{R})$.
(a) Prove that $\widehat{\overline{f}}(t) = \overline{\hat{f}(-t)}$ for all $t \in \mathbb{R}$.
(b) Prove that $f(x) \in \mathbb{R}$ for almost every $x \in \mathbb{R}$ if and only if $\hat{f}(t) = \overline{\hat{f}(-t)}$ for all $t \in \mathbb{R}$.

Solution. (a) Conjugation commutes with the integral, and $\overline{e^{-2\pi i(-t)x}} = e^{-2\pi i t x}$, so for $t \in \mathbb{R}$

$$\begin{aligned}\widehat{\overline{f}}(t) &= \int_{-\infty}^{\infty} \overline{f(x)\, e^{-2\pi i(-t)x}}\, dx \\ &= \overline{\int_{-\infty}^{\infty} f(x)\, e^{-2\pi i(-t)x}\, dx} = \overline{\hat{f}(-t)}.\end{aligned}$$

(b) If f is real almost everywhere, then $\overline{f} = f$ as elements of $L^1(\mathbb{R})$, so $\widehat{\overline{f}} = \hat{f}$ and (a) gives $\hat{f}(t) = \overline{\hat{f}(-t)}$ for all t.
Conversely, if $\hat{f}(t) = \overline{\hat{f}(-t)}$ for all t, then (a) rewrites this as $\hat{f} = \widehat{\overline{f}}$; by linearity the function $f - \overline{f} \in L^1(\mathbb{R})$ has Fourier transform identically 0, so 11.80 gives $f(x) = \overline{f(x)}$ for almost every x, that is, $f(x) \in \mathbb{R}$ for almost every x.

Problem 11C.15. Define $f \in L^1(\mathbb{R})$ by $f(x) = e^{-x^4}\chi_{[0,\infty)}(x)$. Show that $\widehat{f} \notin L^1(\mathbb{R})$.

Solution. If $\widehat{f}$ were in $L^1(\mathbb{R})$, then f would agree almost everywhere with a continuous function, which is impossible because f jumps by 1 at the origin.
In detail, suppose $\widehat{f} \in L^1(\mathbb{R})$ and set $h(x) = \big(\widehat{f}\big)\widehat{\ }(-x) = \int_{-\infty}^{\infty} \widehat{f}(t) e^{2\pi i x t}\, dt$, continuous by 11.49 applied to $\widehat{f} \in L^1(\mathbb{R})$. Since $f, \widehat{f} \in L^1(\mathbb{R})$, the Fourier Inversion Formula (11.76) gives $f = h$ on a set E with $|\mathbb{R} \setminus E| = 0$. For each $n \in \mathbb{Z}^+$ the intervals $(0, \frac{1}{n})$ and $(-\frac{1}{n}, 0)$ have positive measure, hence meet E; choose a_n and b_n in them. Then $a_n, b_n \to 0$ while

$$h(a_n) = f(a_n) = e^{-a_n^4} \longrightarrow 1, \qquad h(b_n) = f(b_n) = 0,$$

so continuity of h at 0 forces $1 = h(0) = 0$, a contradiction.

Problem 11C.16. Suppose $f \in L^1(\mathbb{R})$ and $\widehat{f} \in L^1(\mathbb{R})$. Prove that $f \in L^2(\mathbb{R})$ and $\widehat{f} \in L^2(\mathbb{R})$.

Solution. The hypotheses force f to be bounded, and a bounded function in $L^1(\mathbb{R})$ lies in $L^2(\mathbb{R})$.
Since f and $\widehat{f}$ are both in $L^1(\mathbb{R})$, the Fourier Inversion Formula (11.76) gives $f(x) = \big(\widehat{f}\big)\widehat{\ }(-x)$ for almost every x, while the definition of the Fourier transform bounds $\big|\big(\widehat{f}\big)\widehat{\ }(s)\big| \leq \|\widehat{f}\|_1$ for every s. Hence $\|f\|_\infty \leq \|\widehat{f}\|_1 < \infty$, so

$$\|f\|_2^2 = \int_{-\infty}^{\infty} |f(x)|\, |f(x)|\, dx \leq \|f\|_\infty \|f\|_1 < \infty,$$

giving $f \in L^2(\mathbb{R})$. Thus $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, so Plancherel's Theorem (11.82) gives $\|\widehat{f}\|_2 = \|f\|_2 < \infty$, whence $\widehat{f} \in L^2(\mathbb{R})$.

Problem 11C.17. Prove there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $\lim\limits_{t\to\pm\infty} g(t) = 0$ and $g \notin \{\widehat{f} : f \in L^1(\mathbb{R})\}$.

Solution. Take the odd function

$$g(t) = \begin{cases} \dfrac{1}{\log t} & \text{if } t \geq 2, \\ \dfrac{t}{2\log 2} & \text{if } 0 \leq t < 2, \\ -g(-t) & \text{if } t < 0. \end{cases}$$

The two formulas agree at $t = 2$ and $g(0) = 0$, so g is continuous and real valued on $\mathbb{R}$, with $g(t) \to 0$ as $t \to \pm\infty$. Its operative feature is the divergence property

$$\lim_{T\to\infty} \int_1^T \frac{g(t)}{t}\, dt = \infty,$$

since $g(t)/t$ is bounded on $[1, 2]$ while $\int_2^\infty \frac{dt}{t\log t} = \lim_{T\to\infty}(\log\log T - \log\log 2) = \infty$.
We need one lemma: $S(A) = \int_0^A \frac{\sin u}{u}\, du$ (integrand 1 at $u = 0$) has $M := \sup_{A\geq 0} |S(A)| < \infty$ and a finite limit L at ∞. Indeed $|S(A)| \leq 1$ for $A \leq 1$, while for $A > 1$ integration by parts gives

$$\int_1^A \frac{\sin u}{u}\, du = \cos 1 - \frac{\cos A}{A} - \int_1^A \frac{\cos u}{u^2}\, du,$$

of absolute value at most $1 + 1 + \int_1^\infty u^{-2}\, du = 3$, so $|S| \leq 4$; and the right side converges as $A \to \infty$ because $\frac{\cos A}{A} \to 0$ and $\int_1^\infty \frac{\cos u}{u^2}\, du$ converges absolutely.
Suppose now $f \in L^1(\mathbb{R})$ satisfies $\widehat{f} = g$. Reflection invariance of Lebesgue measure gives $\int_{-\infty}^\infty f(-x)e^{-2\pi itx}\, dx = \widehat{f}(-t)$, so $f_1(x) = \frac{f(x)-f(-x)}{2}$ is an odd element of $L^1(\mathbb{R})$ with $\widehat{f_1}(t) = \frac{g(t)-g(-t)}{2} = g(t)$. Write $f_1 = u + iv$ with $u, v \in L^1(\mathbb{R})$ real valued and odd. For any real odd $w \in L^1(\mathbb{R})$,

$$\widehat{w}(t) = -2i \int_0^\infty w(x) \sin(2\pi tx)\, dx,$$

because $x \mapsto w(x)\cos(2\pi tx)$ is odd and integrable (so integrates to 0) while $x \mapsto w(x)\sin(2\pi tx)$ is even. Hence $g = \widehat{u} + i\widehat{v} = -2iA(t) + 2B(t)$ with A, B the real numbers $\int_0^\infty u \sin(2\pi tx)\, dx$ and $\int_0^\infty v \sin(2\pi tx)\, dx$; as g is real valued, taking real parts gives

$$g(t) = 2\int_0^\infty v(x)\sin(2\pi tx)\, dx \qquad \text{for all } t \in \mathbb{R}.$$

Fix $T > 1$. The function $(t, x) \mapsto \frac{v(x)\sin(2\pi tx)}{t}$ is product measurable on $[1, T] \times (0, \infty)$ (continuous in (t, x) times measurable in x), Lebesgue measure is σ-finite, and Tonelli (5.28) bounds its absolute integral by $\|v\|_1 \log T < \infty$; so Fubini (5.32) permits the interchange

$$\int_1^T \frac{g(t)}{t}\, dt = 2\int_0^\infty v(x)\, \Phi_T(x)\, dx, \qquad \Phi_T(x) = \int_1^T \frac{\sin(2\pi tx)}{t}\, dt.$$

For $x > 0$ the substitution $u = 2\pi tx$ turns $\Phi_T(x)$ into $S(2\pi Tx) - S(2\pi x)$, so $|\Phi_T| \leq 2M$ and $\Phi_T(x) \to L - S(2\pi x)$ as $T \to \infty$. Hence for any sequence $T_n \to \infty$, the Dominated Convergence Theorem (3.31) with dominating function $2M|v| \in L^1(\mathbb{R})$ gives

$$\lim_{n\to\infty} \int_1^{T_n} \frac{g(t)}{t}\, dt = 2\int_0^\infty v(x)\big(L - S(2\pi x)\big)\, dx,$$

a real number of modulus at most $4M\|v\|_1$. This contradicts the divergence property, so no $f \in L^1(\mathbb{R})$ has $\widehat{f} = g$.

Problem 11C.18. Prove that if $f \in L^1(\mathbf{R})$, then $\|\widehat{f}\|_2 = \|f\|_2$.
[This exercise slightly improves Plancherel's Theorem (11.82) because here we have the weaker hypothesis that $f \in L^1(\mathbf{R})$ instead of $f \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$. Because of Plancherel's Theorem, here you need only prove that if $f \in L^1(\mathbf{R})$ and $\|f\|_2 = \infty$, then $\|\widehat{f}\|_2 = \infty$.]

Solution. Plancherel's Theorem (11.82) settles the case $\|f\|_2 < \infty$, so it suffices to prove $\|f\|_2 \le \|\widehat{f}\|_2$, with $\|\cdot\|_2 \in [0, \infty]$; the case $\|f\|_2 = \infty$ then forces $\|\widehat{f}\|_2 = \infty$.
Fix $y > 0$ and let P_y be the Poisson kernel (11.68). Then $\|P_y\|_1 = 1$ by 11.69(a),(b), and $P_y \in L^2(\mathbf{R})$ because $0 < P_y \le \frac{1}{\pi y}$ and $P_y(x) \le \frac{y}{\pi x^2}$ for $x \neq 0$; so 11.64 applied twice with f in the L^1 slot gives $\|f * P_y\|_1 \le \|f\|_1 \|P_y\|_1$ and $\|f * P_y\|_2 \le \|f\|_1 \|P_y\|_2$, both finite. Hence 11.82 applies to $f * P_y \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$, and since $(f * P_y)\widehat{\ }(t) = \widehat{f}(t) e^{-2\pi y|t|}$ by 11.85 (that is, 11.66 with 11.79),

$$\begin{aligned}\|f * P_y\|_2 &= \Big(\int_{-\infty}^{\infty} |\widehat{f}(t)|^2 e^{-4\pi y|t|}\, dt\Big)^{1/2} \\ &\le \|\widehat{f}\|_2.\end{aligned}$$

By 11.74 with $p = 1$ we have $\|f - f * P_y\|_1 \to 0$ as $y \downarrow 0$, so choosing $y_n \to 0$ and passing to a subsequence by 7.23 gives $(f * P_{y_n})(x) \to f(x)$ for almost every x. Fatou's Lemma (Exercise 17 in Section 3A), applied to the nonnegative functions $|f * P_{y_n}|^2$, therefore yields

$$\|f\|_2^2 \le \liminf_{n\to\infty} \|f * P_{y_n}\|_2^2 \le \|\widehat{f}\|_2^2.$$

Problem 11C.19. Suppose $y > 0$. Define an operator T on $L^2(\mathbf{R})$ by $Tf = f * P_y$.
(a) Show that T is a self-adjoint operator on $L^2(\mathbf{R})$.
(b) Show that $\operatorname{sp}(T) = [0, 1]$.
[Because the spectrum of each compact operator is a countable set (by 10.93), part (b) above implies that T is not a compact operator. This conclusion differs from the situation on the unit circle—see Exercise 9 in Section 11B.]

Solution. (a) The Poisson kernel $P_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2}$ (11.68) is real valued and even, with $\|P_y\|_1 = 1$ by 11.69(a),(b); so 11.65 and 11.64 give $\|Tf\|_2 = \|P_y * f\|_2 \le \|P_y\|_1 \|f\|_2 = \|f\|_2$, making T a bounded linear operator on $L^2(\mathbf{R})$ with $\|T\| \le 1$.
Let $f, g \in L^2(\mathbf{R})$. The function $(t, x) \mapsto f(t) P_y(x - t)\overline{g(x)}$ is product measurable (measurable in each variable alone times a continuous function of (t, x)), and Tonelli (5.28) with 11.64, 11.65, and Hölder (7.9) bounds its absolute integral by $\big\||f| * P_y\big\|_2 \|g\|_2 \le \|f\|_2 \|g\|_2 < \infty$. Hence Fubini (5.32) permits the interchange below, and $P_y(x - t) = P_y(t - x)$ is real, so

$$\begin{aligned}\langle Tf, g\rangle &= \int_{-\infty}^{\infty} f(t)\Big(\int_{-\infty}^{\infty} P_y(t - x)\, \overline{g(x)}\, dx\Big) dt \\ &= \int_{-\infty}^{\infty} f(t)\, \overline{(g * P_y)(t)}\, dt = \langle f, Tg\rangle.\end{aligned}$$

Thus $T^* = T$.
(b) $\operatorname{sp}(T) = [0, 1]$. Put $m(t) = e^{-2\pi y|t|} \in (0, 1]$ and let $M_m h = mh$, bounded since $|m| \le 1$. For $f \in L^2(\mathbf{R})$, Exercise 11C.20 with 11.65 and 11.79 gives

$$\mathcal{F}(Tf) = \mathcal{F}(P_y * f) = \widehat{P_y} \cdot \mathcal{F}f = M_m(\mathcal{F}f),$$

so $\mathcal{F}T = M_m\mathcal{F}$; as $\mathcal{F}$ is unitary by 11.87(a), $T - \lambda I = \mathcal{F}^{-1}(M_m - \lambda I)\mathcal{F}$ for every λ, and hence $\operatorname{sp}(T) = \operatorname{sp}(M_m)$.
(i) $\lambda \notin [0, 1]$. Then $d = \inf_t |m(t) - \lambda| \ge \operatorname{dist}(\lambda, [0, 1]) > 0$, so $\frac{1}{m - \lambda}$ is bounded by $\frac{1}{d}$ and $M_{1/(m-\lambda)}$ is a bounded two-sided inverse of $M_m - \lambda I$; thus $\lambda \notin \operatorname{sp}(M_m)$.

(ii) $\lambda \in [0,1]$. Choose t_n with $m(t_n) \to \lambda$: for $\lambda \in (0,1]$ the Intermediate Value Theorem gives a single t with $m(t) = \lambda$, since m is continuous with $m(0) = 1$ and $m(t) \to 0$; for $\lambda = 0$ take $t_n = n$. Continuity of m at t_n supplies $\delta_n > 0$ with $|m(s) - m(t_n)| < \frac{1}{n}$ on $I_n = (t_n - \delta_n, t_n + \delta_n)$, and the unit vectors $h_n = \chi_{I_n}/\sqrt{2\delta_n}$ satisfy

$$\|(M_m - \lambda I)h_n\|_2^2 \le \sup_{s \in I_n} |m(s) - \lambda|^2 \le \Big(\tfrac{1}{n} + |m(t_n) - \lambda|\Big)^2 \longrightarrow 0.$$

A bounded inverse S would then give $1 = \|h_n\|_2 \le \|S\|\,\|(M_m - \lambda I)h_n\|_2 \to 0$, so $\lambda \in \operatorname{sp}(M_m)$.

Problem 11C.20. Prove that if $f \in L^1(\mathbb{R})$ and $g \in L^2(\mathbb{R})$, then $\mathcal{F}(f * g) = \widehat{f}\,\mathcal{F}g$.

Solution. Truncating g reduces the identity to 11.66, and the truncations converge.
Put $g_n = g\chi_{[-n,n]}$. Hölder's inequality (7.9) gives $\|g_n\|_1 \le \sqrt{2n}\,\|g\|_2 < \infty$, so $g_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, while $|g - g_n|^2 \to 0$ pointwise under the dominating function $|g|^2 \in L^1(\mathbb{R})$, so $\|g - g_n\|_2 \to 0$ by 3.31. Since $f, g_n \in L^1(\mathbb{R})$, result 11.66 gives $(f * g_n)\widehat{\ } = \widehat{f}\,\widehat{g_n}$; as $f * g_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ by 11.64 with $p = 1$ and with $p = 2$, the definition 11.86 turns this into

$$\mathcal{F}(f * g_n) = \widehat{f}\,\mathcal{F}g_n.$$

Now let $n \to \infty$. By 11.64 with $p = 2$ we have $f * g \in L^2(\mathbb{R})$ and $\|f * g - f * g_n\|_2 \le \|f\|_1\|g - g_n\|_2 \to 0$, so, $\mathcal{F}$ being an isometry on $L^2(\mathbb{R})$ (11.82 with 11.86), $\mathcal{F}(f * g_n) \to \mathcal{F}(f * g)$. On the other side, $\|\widehat{f}\,\|_\infty \le \|f\|_1$ by 11.49, so

$$\|\widehat{f}\,\mathcal{F}g - \widehat{f}\,\mathcal{F}g_n\|_2 \le \|\widehat{f}\,\|_\infty\|g - g_n\|_2 \to 0.$$

Uniqueness of limits in $L^2(\mathbb{R})$ gives $\mathcal{F}(f * g) = \widehat{f}\,\mathcal{F}g$.

Problem 11C.21. Prove that if $f, g \in L^2(\mathbb{R})$, then $\widehat{fg} = (\mathcal{F}f) * (\mathcal{F}g)$.

Solution. The identity is unitarity of $\mathcal{F}$ (11.87(a)) tested against $h_t(x) = \overline{g(x)}\,e^{2\pi i t x}$.
Two identities for $\mathcal{F}$ on $L^2(\mathbb{R})$ are needed; each holds on the dense subspace $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and extends by continuity, both sides being norm-preserving (conjugate) linear maps of $L^2(\mathbb{R})$ into itself (11.82, 11.86).
(A) $\mathcal{F}(\overline{h})(s) = \overline{(\mathcal{F}h)(-s)}$ for almost every s; on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ this is 11C.14(a).
(B) $\mathcal{F}(e^{2\pi i b\,\cdot}\,h)(s) = (\mathcal{F}h)(s - b)$ for almost every s; on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ this is 11.55(b), the modulated function having the same modulus as h.
Fix $t \in \mathbb{R}$. Then $|h_t| = |g|$ puts h_t in $L^2(\mathbb{R})$, and since $fg \in L^1(\mathbb{R})$ by Hölder (7.9),

$$\begin{aligned} \langle f, h_t \rangle &= \int_{-\infty}^{\infty} f(x)g(x)e^{-2\pi i t x}\,dx \\ &= \widehat{fg}(t). \end{aligned}$$

Applying (B) with $b = t$ and $h = \overline{g}$, then (A) with $h = g$, gives $(\mathcal{F}h_t)(s) = \overline{(\mathcal{F}g)(t - s)}$ for almost every s. As $\mathcal{F}$ is unitary it preserves inner products, so

$$\begin{aligned} \widehat{fg}(t) = \langle f, h_t \rangle = \langle \mathcal{F}f, \mathcal{F}h_t \rangle &= \int_{-\infty}^{\infty} (\mathcal{F}f)(s)\,(\mathcal{F}g)(t - s)\,ds \\ &= \big((\mathcal{F}f) * (\mathcal{F}g)\big)(t), \end{aligned}$$

the integral being unaffected by altering $\mathcal{F}h_t$ on a null set.

12 Probability Measures

Contents

12.1 Exercises

Problem 12.1. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $A \in \mathcal{F}$. Prove that A and $\Omega \setminus A$ are independent if and only if $P(A) = 0$ or $P(A) = 1$.

Solution. Because $A \cap (\Omega \setminus A) = \emptyset$ has probability 0, and $P(\Omega \setminus A) = 1 - P(A)$ by finite additivity, the definition of independence (12.7) says that A and $\Omega \setminus A$ are independent if and only if

$$0 = P(A)\big(1 - P(A)\big).$$

A product of real numbers vanishes exactly when a factor does, so this holds if and only if $P(A) = 0$ or $P(A) = 1$.

Problem 12.2. Suppose P is Lebesgue measure on $[0, 1]$. Give an example of two disjoint Borel subsets A and B of $[0, 1]$ such that $P(A) = P(B) = \frac{1}{2}$, $[0, \frac{1}{2}]$ and A are independent, and $[0, \frac{1}{2}]$ and B are independent.

Solution. Take

$$A = \left[0, \tfrac{1}{4}\right] \cup \left(\tfrac{1}{2}, \tfrac{3}{4}\right], \qquad B = \left(\tfrac{1}{4}, \tfrac{1}{2}\right] \cup \left(\tfrac{3}{4}, 1\right].$$

These are Borel, and they are disjoint because the four intervals $\left[0, \frac{1}{4}\right]$, $\left(\frac{1}{4}, \frac{1}{2}\right]$, $\left(\frac{1}{2}, \frac{3}{4}\right]$, $\left(\frac{3}{4}, 1\right]$ partition $[0, 1]$, with A the union of the first and third and B the union of the second and fourth. Each of the four has measure $\frac{1}{4}$ (2.14 with 2.68; single points are null, so included endpoints do not matter), so $P(A) = P(B) = \frac{1}{2} = P\big(\left[0, \frac{1}{2}\right]\big)$. Finally

$$\begin{aligned} P\big(\left[0, \tfrac{1}{2}\right] \cap A\big) &= P\big(\left[0, \tfrac{1}{4}\right]\big) = \tfrac{1}{4} = \tfrac{1}{2} \cdot \tfrac{1}{2}, \\ P\big(\left[0, \tfrac{1}{2}\right] \cap B\big) &= P\big(\left(\tfrac{1}{4}, \tfrac{1}{2}\right]\big) = \tfrac{1}{4} = \tfrac{1}{2} \cdot \tfrac{1}{2}, \end{aligned}$$

so $\left[0, \frac{1}{2}\right]$ is independent of each of A and B.

Problem 12.3. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $A, B \in \mathcal{F}$. Prove that the following are equivalent.

- A and B are independent events.
- A and $\Omega \setminus B$ are independent events.
- $\Omega \setminus A$ and B are independent events.
- $\Omega \setminus A$ and $\Omega \setminus B$ are independent events.

Solution. All four are equivalent by one lemma: for $C, D \in \mathcal{F}$, the events C and D are independent if and only if $\Omega \setminus C$ and D are. Indeed D is the disjoint union of $C \cap D$ and $(\Omega \setminus C) \cap D$, so $P((\Omega \setminus C) \cap D) = P(D) - P(C \cap D)$; and $P(\Omega \setminus C)P(D) = P(D) - P(C)P(D)$. Subtracting,

$$P\big((\Omega \setminus C) \cap D\big) - P(\Omega \setminus C)P(D) = P(C)P(D) - P(C \cap D),$$

so one side vanishes exactly when the other does. Independence is symmetric in its two events ($C \cap D = D \cap C$), so the lemma may equally complement the second event.
Now the lemma with $C = A$, $D = B$ gives (first) $\iff$ (third); with $C = B$, $D = A$, plus symmetry, it gives (first) $\iff$ (second); and with $C = A$, $D = \Omega \setminus B$ it gives (second) $\iff$ (fourth).

Problem 12.4. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $\{A_k\}_{k \in \Gamma}$ is a family of events. Prove the family $\{A_k\}_{k \in \Gamma}$ is independent if and only if the family $\{\Omega \setminus A_k\}_{k \in \Gamma}$ is independent.

Solution. Assume $\{A_k\}_{k \in \Gamma}$ is independent and prove the stronger mixed statement: for all distinct

$k_1, \dots, k_n \in \Gamma$ and all $m \in \{0, \dots, n\}$,

$$P\Big(\bigcap_{j=1}^{m}(\Omega \setminus A_{k_j}) \cap \bigcap_{j=m+1}^{n} A_{k_j}\Big) = \prod_{j=1}^{m}\big(1 - P(A_{k_j})\big) \prod_{j=m+1}^{n} P(A_{k_j}),$$

with empty intersections read as Ω and empty products as 1. We induct on m; the case $m = 0$ is the independence hypothesis (12.7) when $n \geq 1$, and reads $P(\Omega) = 1$ when $n = 0$.

Let $m \geq 1$ and set $C = \bigcap_{j=1}^{m-1}(\Omega \setminus A_{k_j}) \cap \bigcap_{j=m+1}^{n} A_{k_j}$. The induction hypothesis applies to C (its $n-1$ distinct indices, the first $m-1$ complemented) and also to $C \cap A_{k_m}$ (its n distinct indices with the same $m-1$ complemented, after a reordering that changes neither intersection nor product). Since C is the disjoint union of $C \cap A_{k_m}$ and $C \cap (\Omega \setminus A_{k_m})$, additivity gives

$$\begin{aligned} P\big(C \cap (\Omega \setminus A_{k_m})\big) &= P(C) - P(C \cap A_{k_m}) \\ &= \prod_{j=1}^{m-1}\big(1 - P(A_{k_j})\big)\big(1 - P(A_{k_m})\big) \prod_{j=m+1}^{n} P(A_{k_j}), \end{aligned}$$

which is the claim for m, because $C \cap (\Omega \setminus A_{k_m})$ is the required intersection.

Taking $m = n$, and using $P(\Omega \setminus A_k) = 1 - P(A_k)$, shows that $\{\Omega \setminus A_k\}_{k \in \Gamma}$ is independent. Applying that implication to the family $\{\Omega \setminus A_k\}_{k \in \Gamma}$, whose complements are the A_k, gives the converse.

Problem 12.5. Give an example of a probability space $(\Omega, \mathcal{F}, P)$ and events A, B_1, B_2 such that A and B_1 are independent, A and B_2 are independent, but A and $B_1 \cup B_2$ are not independent.

Solution. Take $\Omega = \{1, 2, 3, 4\}$ with $\mathcal{F}$ all subsets of Ω and $P(E) = |E|/4$, and

$$A = \{1, 2\}, \qquad B_1 = \{1, 3\}, \qquad B_2 = \{1, 4\}.$$

Each of these has probability $\frac{1}{2}$, and $A \cap B_1 = A \cap B_2 = \{1\}$ has probability $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$, so A is independent of B_1 and of B_2. But $B_1 \cup B_2 = \{1, 3, 4\}$ while $A \cap (B_1 \cup B_2) = \{1\}$, so

$$P\big(A \cap (B_1 \cup B_2)\big) = \tfrac{1}{4} \neq \tfrac{3}{8} = P(A)\, P(B_1 \cup B_2).$$

Problem 12.6. Give an example of a probability space $(\Omega, \mathcal{F}, P)$ and events A_1, A_2, A_3 such that A_1 and A_2 are independent, A_1 and A_3 are independent, and A_2 and A_3 are independent, but the family A_1, A_2, A_3 is not independent.

Solution. Take $\Omega = \{H, T\}^2$ (two tosses of a fair coin), $\mathcal{F} = 2^{\Omega}$, and P counting measure on Ω divided by 4, with

$$\begin{aligned} A_1 &= \{(H,H), (H,T)\} \quad \text{(first toss heads)}, \\ A_2 &= \{(H,H), (T,H)\} \quad \text{(second toss heads)}, \\ A_3 &= \{(H,H), (T,T)\} \quad \text{(the tosses agree)}. \end{aligned}$$

Each A_j has two elements, so $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$, and each of the three pairwise intersections equals $\{(H,H)\}$ (Check!). Hence for $j \neq k$,

$$P(A_j \cap A_k) = \tfrac{1}{4} = \tfrac{1}{2} \cdot \tfrac{1}{2} = P(A_j) \cdot P(A_k),$$

so the three pairs are independent. But $A_1 \cap A_2 \cap A_3 = \{(H,H)\}$, so

$$P(A_1 \cap A_2 \cap A_3) = \tfrac{1}{4} \neq \tfrac{1}{8} = P(A_1) \cdot P(A_2) \cdot P(A_3),$$

and thus A_1, A_2, A_3 fails the condition in 12.7 for the index set $\{1, 2, 3\}$.

Problem 12.7. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space, $A \in \mathcal{F}$, and $B_1 \subseteq B_2 \subseteq \cdots$ is an increasing sequence of events such that A and B_n are independent events for each $n \in \mathbf{Z}^+$. Show that A and $\bigcup_{n=1}^{\infty} B_n$ are independent.

Solution. Write $B = \bigcup_{n=1}^{\infty} B_n$; both $\{B_n\}$ and $\{A \cap B_n\}$ are increasing sequences of events, with $\bigcup_{n=1}^{\infty}(A \cap B_n) = A \cap B$. Hence two applications of 2.59 and the hypothesis $P(A \cap B_n) = P(A) \cdot P(B_n)$ give

$$\begin{aligned} P(A \cap B) &= \lim_{n\to\infty} P(A \cap B_n) = P(A) \cdot \lim_{n\to\infty} P(B_n) \\ &= P(A) \cdot P(B), \end{aligned}$$

the constant $P(A)$ coming out of the limit because $\lim_{n\to\infty} P(B_n)$ exists (it is $P(B)$, again by 2.59). By 12.7, A and B are independent.

Problem 12.8. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $\{A_t\}_{t\in\mathbb{R}}$ is an independent family of events such that $P(A_t) < 1$ for each $t \in \mathbb{R}$. Prove that there exists a sequence $t_1, t_2, \ldots$ in $\mathbb{R}$ such that

$$P\Big(\bigcap_{n=1}^{\infty} A_{t_n}\Big) = 0.$$

Solution. Take $t_1, t_2, \ldots$ distinct in $S_m = \{t \in \mathbb{R} : P(\Omega \setminus A_t) > \frac{1}{m}\}$, for an $m \in \mathbb{Z}^+$ with S_m infinite. Such an m exists: every t has $P(\Omega \setminus A_t) = 1 - P(A_t) > 0$ and so lies in S_m for large m, whence $\mathbb{R} = \bigcup_{m=1}^{\infty} S_m$; were every S_m finite, $\mathbb{R}$ would be countable.
Fix $N \in \mathbb{Z}^+$. The $t_1, \ldots, t_N$ are distinct indices of the independent family, so 12.7 and the bound $P(A_{t_n}) = 1 - P(\Omega \setminus A_{t_n}) \le 1 - \frac{1}{m}$ give

$$P\Big(\bigcap_{n=1}^{\infty} A_{t_n}\Big) \le P\Big(\bigcap_{n=1}^{N} A_{t_n}\Big) = \prod_{n=1}^{N} P(A_{t_n}) \le \Big(1 - \tfrac{1}{m}\Big)^N,$$

the first inequality by monotonicity of P (2.57). Since $0 \le 1 - \frac{1}{m} < 1$, letting $N \to \infty$ forces $P\big(\bigcap_{n=1}^{\infty} A_{t_n}\big) = 0$.

Problem 12.9. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $B_1, \ldots, B_n \in \mathcal{F}$ are such that $P(B_1 \cap \cdots \cap B_n) > 0$. Prove that

$$P(A \cap B_1 \cap \cdots \cap B_n) = P(B_1) \cdot P_{B_1}(B_2) \cdots P_{B_1\cap\cdots\cap B_{n-1}}(B_n) \cdot P_{B_1\cap\cdots\cap B_n}(A)$$

for every event $A \in \mathcal{F}$.

Solution. The right side telescopes. Put $C_k = B_1 \cap \cdots \cap B_k$, so that $C_1 \supseteq \cdots \supseteq C_n$ and hence $P(C_k) \ge P(C_n) > 0$ by 2.57(a); thus each P_{C_k} is defined (12.23). For $k \in \{2, \ldots, n\}$, since $B_k \cap C_{k-1} = C_k$, definition 12.23 gives

$$P_{C_{k-1}}(B_k) = \frac{P(B_k \cap C_{k-1})}{P(C_{k-1})} = \frac{P(C_k)}{P(C_{k-1})}, \qquad P_{C_n}(A) = \frac{P(A \cap C_n)}{P(C_n)}.$$

All denominators are nonzero, so multiplying these together with $P(B_1) = P(C_1)$ gives

$$\begin{aligned} &P(B_1) \cdot P_{C_1}(B_2) \cdots P_{C_{n-1}}(B_n) \cdot P_{C_n}(A) \\ &\quad = P(C_1) \cdot \frac{P(C_2)}{P(C_1)} \cdots \frac{P(C_n)}{P(C_{n-1})} \cdot \frac{P(A \cap C_n)}{P(C_n)} = P(A \cap C_n), \end{aligned}$$

each $P(C_k)$ occurring once in a numerator and once in a denominator. As $C_n = B_1 \cap \cdots \cap B_n$, this is the asserted identity.

Problem 12.10. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $A \in \mathcal{F}$ is an event such that $0 < P(A) < 1$. Prove that

$$P(B) = P_A(B) \cdot P(A) + P_{\Omega \setminus A}(B) \cdot P(\Omega \setminus A)$$

for every event $B \in \mathcal{F}$.

Solution. Both conditional measures are defined: $P(A) > 0$ by hypothesis and $P(\Omega \setminus A) = 1 - P(A) > 0$ by additivity and $P(A) < 1$, so 12.23 applies to each. Fix $B \in \mathcal{F}$. Then $B \cap A$ and $B \cap (\Omega \setminus A)$ are disjoint with union B, so additivity of P and 12.23 give

$$\begin{aligned} P(B) &= P(B \cap A) + P\big(B \cap (\Omega \setminus A)\big) \\ &= \frac{P(B \cap A)}{P(A)} \cdot P(A) + \frac{P\big(B \cap (\Omega \setminus A)\big)}{P(\Omega \setminus A)} \cdot P(\Omega \setminus A) \\ &= P_A(B) \cdot P(A) + P_{\Omega \setminus A}(B) \cdot P(\Omega \setminus A). \end{aligned}$$

Problem 12.11. Give an example of a probability space $(\Omega, \mathcal{F}, P)$ and $X, Y \in \mathcal{L}^2(P)$ such that $\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y)$ but X and Y are not independent random variables.

Solution. Take $\Omega = \{-1, 0, 1\}$ with $\mathcal{F} = 2^\Omega$ and P counting measure divided by 3, and set $X(\omega) = \omega$, $Y(\omega) = \omega^2$; every function on this finite set is bounded and measurable, so $X, Y \in \mathcal{L}^2(P)$. Each singleton has probability $\frac{1}{3}$, so (using $\omega^4 = \omega^2$ on Ω)

$$EX = 0, \quad E(X^2) = \tfrac{2}{3}, \quad EY = \tfrac{2}{3}, \quad E(Y^2) = \tfrac{2}{3},$$

and $X + Y$ takes the values $0, 0, 2$ at $-1, 0, 1$, so $E(X+Y) = \frac{2}{3}$ and $E((X+Y)^2) = \frac{4}{3}$. Hence 12.20 gives

$$\begin{aligned} \sigma^2(X) + \sigma^2(Y) &= \left(\tfrac{2}{3} - 0\right) + \left(\tfrac{2}{3} - \tfrac{4}{9}\right) = \tfrac{8}{9} \\ &= \tfrac{4}{3} - \tfrac{4}{9} = \sigma^2(X+Y). \end{aligned}$$

But with the Borel sets $U = V = \{0\}$ we have $\{X \in U\} = \{Y \in V\} = \{0\}$, so

$$P\big(\{X \in U\} \cap \{Y \in V\}\big) = \tfrac{1}{3} \neq \tfrac{1}{9} = P(X \in U) \cdot P(Y \in V),$$

and therefore X and Y are not independent (12.14).

Problem 12.12. Suppose $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$ are probability spaces, X is a random variable on Ω, Y is a random variable on Ω', and $\widetilde{X} = \widetilde{Y}$. Prove that $P_X = P'_Y$.

Solution. The Monotone Class Theorem upgrades the hypothesis from half-lines to $\mathcal{B}$. Write $H = \widetilde{X} = \widetilde{Y}$; by 12.27 the hypothesis says $P_X((-\infty, s]) = H(s) = P'_Y((-\infty, s])$ for all $s \in \mathbb{R}$. Let $\mathcal{A}$ be the collection of finite unions of half-open intervals, meaning sets

$$I(\alpha, \beta) = \{x \in \mathbb{R} : \alpha < x \leq \beta\}, \quad \alpha \in \mathbb{R} \cup \{-\infty\},\ \beta \in \mathbb{R} \cup \{\infty\}.$$

(i) $\mathcal{A}$ is an algebra (5.10). Intersections stay inside the class, since $I(\alpha_1, \beta_1) \cap I(\alpha_2, \beta_2) = I(\max \alpha_i, \min \beta_i)$, and each complement $\mathbb{R} \setminus I(\alpha, \beta)$ is the union of at most the two half-open intervals $I(-\infty, \alpha)$ and $I(\beta, \infty)$; so De Morgan plus distributivity turns the complement of a finite union into a finite union again (Check!).
(ii) The smallest σ-algebra $\mathcal{S}$ containing $\mathcal{A}$ is $\mathcal{B}$. Half-open intervals are Borel, so $\mathcal{S} \subseteq \mathcal{B}$; conversely $(a, b) = \bigcup_{k=1}^\infty (a, b - \frac{1}{k}]$ and $(-\infty, b) = \bigcup_{k=1}^\infty (-k, b - \frac{1}{k}]$ lie in $\mathcal{S}$, so $\mathcal{S}$ contains every open interval, hence every open set (a countable union of such), hence $\mathcal{B}$.
(iii) P_X and P'_Y agree on $\mathcal{A}$. Additivity gives

$$P_X\big((a, b]\big) = H(b) - H(a), \qquad P_X\big((a, \infty)\big) = 1 - H(a),$$

along with $P_X(\mathbb{R}) = 1$ and $P_X(\emptyset) = 0$, and identically for P'_Y; so the two agree on every half-open interval. Given $E = I(\alpha_1, \beta_1) \cup \cdots \cup I(\alpha_m, \beta_m)$, cutting $\mathbb{R}$ at the finitely many real numbers among the α_j, β_j exhibits E as a finite union of pairwise disjoint half-open intervals, so finite additivity gives $P_X(E) = P'_Y(E)$.
(iv) $\mathcal{M} = \{B \in \mathcal{B} : P_X(B) = P'_Y(B)\}$ is a monotone class (5.15): for increasing sequences apply 2.59 to each measure, and for decreasing sequences 2.60, whose finiteness hypothesis holds because $P_X(E_1) \le 1$ and $P'_Y(E_1) \le 1$.
By (iii), $\mathcal{M}$ contains $\mathcal{A}$, hence the smallest monotone class containing $\mathcal{A}$, which by 5.17 is $\mathcal{S} = \mathcal{B}$. Thus $\mathcal{M} = \mathcal{B}$, that is, $P_X = P'_Y$.

Problem 12.13. Suppose $H : \mathbb{R} \to (0, 1)$ is a continuous one-to-one function satisfying conditions (a) through (d) of 12.29. Show that the function $X : (0, 1) \to \mathbb{R}$ produced in the proof of 12.29 is the inverse function of H.

Solution. The function is $X(\omega) = \sup\{t \in \mathbb{R} : H(t) < \omega\}$ for $\omega \in (0, 1)$ (12.30), and the set inside the supremum is the half-line $(-\infty, H^{-1}(\omega))$.
First, H is strictly increasing: for $s < t$, condition (a) gives $H(s) \le H(t)$ and one-to-oneness rules out equality. Next, H maps $\mathbb{R}$ onto $(0, 1)$: given $u \in (0, 1)$, conditions (b) and (c) supply a, b with $H(a) < u < H(b)$, hence $a < b$, and the Intermediate Value Theorem applied to the continuous H on $[a, b]$ gives r with $H(r) = u$. So $H^{-1}\colon (0, 1) \to \mathbb{R}$ exists.
Now fix $\omega \in (0, 1)$ and put $r = H^{-1}(\omega)$. If $t < r$ then $H(t) < H(r) = \omega$, and if $t \ge r$ then $H(t) \ge \omega$; hence

$$\{t \in \mathbb{R} : H(t) < \omega\} = (-\infty, r),$$

whose supremum is r. Therefore $X(\omega) = r = H^{-1}(\omega)$ for every $\omega \in (0, 1)$.

Problem 12.14. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and X is a random variable. Prove that the following are equivalent.

- $\widetilde{X}$ is a continuous function on $\mathbb{R}$.
- $\widetilde{X}$ is a uniformly continuous function on $\mathbb{R}$.
- $P(X = t) = 0$ for every $t \in \mathbb{R}$.
- $(\widetilde{X} \circ X)^{\sim}(s) = s$ for all $s \in [0, 1]$.

Solution. Label the statements (A), (B), (C), (D); we prove (A) $\Leftrightarrow$ (C), (A) $\Leftrightarrow$ (B), (A) $\Rightarrow$ (D), (D) $\Rightarrow$ (C). The proof of 12.29 gives that $\widetilde{X}$ is increasing and right continuous with limits 0 at $-\infty$ and 1 at ∞; write $Z = \widetilde{X} \circ X$, a random variable with values in $[0, 1]$ because $\widetilde{X}$ is increasing, hence Borel measurable by 2.39, and 2.44 applies.
Jump identity: the events $\{X \le s - \frac{1}{k}\}$ increase to $\{X < s\}$, so 2.59 and monotonicity of $\widetilde{X}$ give $\lim_{t \uparrow s} \widetilde{X}(t) = P(X < s)$, whence by additivity

$$\widetilde{X}(s) - \lim_{t \uparrow s} \widetilde{X}(t) = P(X \le s) - P(X < s) = P(X = s).$$

(A) $\Leftrightarrow$ (C). By right continuity, $\widetilde{X}$ is continuous at s exactly when its left limit at s equals $\widetilde{X}(s)$, which by the jump identity is exactly $P(X = s) = 0$.
(B) $\Rightarrow$ (A) is immediate. For (A) $\Rightarrow$ (B), let $\varepsilon > 0$ and choose $a < b$ with $\widetilde{X}(a) < \frac{\varepsilon}{2}$ and $\widetilde{X}(b) > 1 - \frac{\varepsilon}{2}$; continuity on the compact interval $[a - 1, b + 1]$ gives $\delta \in (0, 1)$ with $|\widetilde{X}(u) - \widetilde{X}(v)| < \varepsilon$ whenever u, v lie there with $|u - v| < \delta$. Now take $u \le v$ with $v - u < \delta$.

- (i) $v \le a$: then $0 \le \widetilde{X}(v) - \widetilde{X}(u) \le \widetilde{X}(a) < \varepsilon$, as $\widetilde{X} \ge 0$ is increasing.
- (ii) $u \ge b$: then $0 \le \widetilde{X}(v) - \widetilde{X}(u) \le 1 - \widetilde{X}(b) < \varepsilon$, as $\widetilde{X} \le 1$.
- (iii) otherwise $v > a$ and $u < b$, so $u > a - 1$ and $v < b + 1$ and the choice of δ applies.

(A) $\Rightarrow$ (D). Let $s \in [0, 1]$ and $T_s = \{t \in \mathbb{R} : \widetilde{X}(t) = s\}$. For $s = 1$, $Z \le 1$ gives $\widetilde{Z}(1) = 1$. For $s \in [0, 1)$ with $T_s = \emptyset$: the Intermediate Value Theorem and the limits $0, 1$ force $s = 0$, and $Z \ge 0$ gives $\{Z \le 0\} = \{X \in T_0\} = \emptyset$, so $\widetilde{Z}(0) = 0$. Otherwise T_s is nonempty, closed (continuity) and bounded above (since $\widetilde{X} \to 1 > s$ when $s < 1$), so $b_s = \max T_s$ satisfies $\widetilde{X}(b_s) = s$, and

$$\{Z \le s\} = \{X \le b_s\},$$

since $X(\omega) \le b_s$ forces $Z(\omega) \le \widetilde{X}(b_s) = s$, while $X(\omega) > b_s$ forces $\widetilde{X}(X(\omega)) \ge s$ with equality excluded by $b_s = \max T_s$. Hence $\widetilde{Z}(s) = \widetilde{X}(b_s) = s$.
(D) $\Rightarrow$ (C), by contraposition. Suppose $P(X = t_0) = c > 0$ and put $s_0 = \widetilde{X}(t_0)$. The jump identity gives $\lim_{t\uparrow t_0} \widetilde{X}(t) = s_0 - c \ge 0$, so $\widetilde{X}(t) \le s_0 - c$ for $t < t_0$ and $\widetilde{X}(t) \ge s_0$ for $t \ge t_0$; thus $\widetilde{X}$, and hence Z, omits the interval $(s_0 - c, s_0)$. So $\{Z \le s\} = \{Z \le s_0 - c\}$ for all $s \in [s_0 - c, s_0) \subseteq [0, 1]$, and (D) would force

$$s_0 - \tfrac{c}{2} = \widetilde{Z}\Big(s_0 - \tfrac{c}{2}\Big) = \widetilde{Z}(s_0 - c) = s_0 - c,$$

contradicting $c > 0$.

Problem 12.15. Suppose $\alpha > 0$ and

$$h(x) = \begin{cases} 0 & \text{if } x < 0, \\ \alpha^2 x e^{-\alpha x} & \text{if } x \ge 0. \end{cases}$$

Let $P = h\,d\lambda$ and let X be the random variable defined by $X(x) = x$ for $x \in \mathbb{R}$.
(a) Verify that $\int_{-\infty}^{\infty} h\,d\lambda = 1$.
(b) Find a formula for the distribution function $\widetilde{X}$.
(c) Find a formula (in terms of α) for EX.
(d) Find a formula (in terms of α) for $\sigma(X)$.

Solution. The answers are 1 in (a), $EX = \frac{2}{\alpha}$ in (c), $\sigma(X) = \frac{\sqrt{2}}{\alpha}$ in (d), and in (b)

$$\widetilde{X}(s) = \begin{cases} 0 & \text{if } s < 0, \\ 1 - (1 + \alpha s)e^{-\alpha s} & \text{if } s \ge 0. \end{cases}$$

All four come from the moment integrals

$$J_n := \int_{[0,\infty)} x^n e^{-\alpha x}\,d\lambda(x) = \frac{n!}{\alpha^{n+1}} \qquad (n \ge 0).$$

To see this, note that $x^n e^{-\alpha x}$ is continuous and nonnegative, so on each $[0, b]$ its Lebesgue and Riemann integrals agree (3.34); integration by parts gives

$$I_0(b) = \frac{1 - e^{-\alpha b}}{\alpha}, \qquad I_n(b) = -\frac{b^n e^{-\alpha b}}{\alpha} + \frac{n}{\alpha} I_{n-1}(b)$$

for $I_n(b) = \int_0^b x^n e^{-\alpha x}\,dx$. The functions $x^n e^{-\alpha x}\chi_{[0,k]}$ increase pointwise to $x^n e^{-\alpha x}\chi_{[0,\infty)}$, so 3.11 gives $J_n = \lim_{k\to\infty} I_n(k)$; since $b^n e^{-\alpha b} \to 0$, the recursions pass to $J_0 = \frac{1}{\alpha}$ and $J_n = \frac{n}{\alpha} J_{n-1}$, and induction finishes.
Part (a). Since $h \ge 0$ is Borel and vanishes on $(-\infty, 0)$,

$$\int_{-\infty}^{\infty} h\,d\lambda = \alpha^2 J_1 = \alpha^2 \cdot \frac{1}{\alpha^2} = 1.$$

In particular $h \in L^1(\mathbb{R})$, so 12.33 applies to $P = h\,d\lambda$ and $X(x) = x$.
Part (b). By 12.33, h is the density of X, so $\widetilde{X}(s) = \int_{-\infty}^{s} h\,d\lambda$, which is 0 for $s < 0$. For $s \ge 0$ the integrand is continuous on $[0, s]$, so 3.34 and the antiderivative $-(1 + \alpha x)e^{-\alpha x}$ of $\alpha^2 x e^{-\alpha x}$ give

$$\widetilde{X}(s) = \int_0^s \alpha^2 x e^{-\alpha x}\,dx = 1 - (1 + \alpha s)e^{-\alpha s}.$$

Part (c). For every Borel $f \geq 0$ we have $\int f\,dP = \int fh\,d\lambda$: this is the definition of P for $f = \chi_B$, extends to nonnegative simple f by 3.20 and 3.16, and then to general $f \geq 0$ by 2.89 and two applications of 3.11 (one for P, one for λ). Taking $f = |X|$,

$$\int_{\mathbb{R}} |X|\,dP = \alpha^2 J_2 = \frac{2}{\alpha} < \infty,$$

so $X \in L^1(P)$ and the mean formula of 12.33 gives $EX = \alpha^2 J_2 = \frac{2}{\alpha}$.
Part (d). Likewise $\int_{\mathbb{R}} |X|^2\,dP = \alpha^2 J_3 = \frac{6}{\alpha^2} < \infty$, so $X \in L^2(P)$ and the variance formula of 12.33 gives

$$\sigma^2(X) = \frac{6}{\alpha^2} - \frac{4}{\alpha^2} = \frac{2}{\alpha^2}, \qquad \sigma(X) = \frac{\sqrt{2}}{\alpha}.$$

Problem 12.16. Suppose $\mathcal{B}$ is the σ-algebra of Borel subsets of $[0,1)$ and P is Lebesgue measure on $\big([0,1),\mathcal{B}\big)$. Let $\{e_k\}_{k\in\mathbb{Z}^+}$ be the family of functions defined by the fourth bullet point of Example 8.51 (notice that $k=0$ is excluded). Show that the family $\{e_k\}_{k\in\mathbb{Z}^+}$ is an i.i.d.

Solution. By 8.51, e_k equals 1 on the dyadic interval $I_{k,n} = [\frac{n-1}{2^k}, \frac{n}{2^k})$ for n odd and -1 for n even; for fixed $k \in \mathbb{Z}^+$ the 2^k intervals $I_{k,1}, \dots, I_{k,2^k}$ are disjoint with union $[0,1)$ and $P(I_{k,n}) = 2^{-k}$, so e_k is a linear combination of their characteristic functions and hence a random variable.
(i) Marginals. Exactly 2^{k-1} of the $n \in \{1,\dots,2^k\}$ are odd (here $k \geq 1$ is used), so additivity gives $P(e_k = 1) = P(e_k = -1) = 2^{k-1}\cdot 2^{-k} = \frac{1}{2}$. Hence

$$\widetilde{e_k}(s) = \begin{cases} 0 & \text{if } s < -1, \\ \frac{1}{2} & \text{if } -1 \leq s < 1, \\ 1 & \text{if } s \geq 1, \end{cases}$$

a formula not involving k; so the family is identically distributed (12.35). (For $k=0$ we would have $e_0 \equiv 1$, which is why $k=0$ is excluded.)
(ii) e_k reads the k-th binary digit. For $x \in [0,1)$ and $j \in \mathbb{Z}^+$ put $d_j(x) = \lfloor 2^j x\rfloor - 2\lfloor 2^{j-1}x\rfloor$; with $a = 2^{j-1}x$ this is $\lfloor 2(a - \lfloor a\rfloor)\rfloor \in \{0,1\}$. Since $\lfloor x\rfloor = 0$, the sum $\sum_{j=1}^k d_j(x)2^{k-j}$ telescopes to $\lfloor 2^k x\rfloor$, and $x \in I_{k,n}$ exactly when $n-1 = \lfloor 2^k x\rfloor$; hence the index n with $x \in I_{k,n}$ is

$$n = 1 + \sum_{j=1}^{k} d_j(x)2^{k-j}.$$

Every term with $j<k$ is even, so n is odd if and only if $d_k(x) = 0$, that is, $e_k(x) = 1 \iff d_k(x) = 0$. The same formula shows $x \mapsto (d_1(x),\dots,d_N(x))$ is constant on each $I_{N,n}$ and, since $(a_1,\dots,a_N) \mapsto 1 + \sum_j a_j 2^{N-j}$ is a bijection $\{0,1\}^N \to \{1,\dots,2^N\}$, attains each element of $\{0,1\}^N$ on exactly one of the 2^N intervals.
(iii) Prescribed values. Let $k_1 < \cdots < k_m$, let $\varepsilon_i \in \{-1,1\}$, put $\delta_i = \frac{1-\varepsilon_i}{2}$ and $N = k_m$. By (ii), $\bigcap_{i=1}^m \{e_{k_i} = \varepsilon_i\}$ is the union of those $I_{N,n}$ whose digit string has $a_{k_i} = \delta_i$, and exactly 2^{N-m} of the 2^N strings do; the intervals being disjoint of measure 2^{-N},

$$P\Big(\bigcap_{i=1}^m \{e_{k_i} = \varepsilon_i\}\Big) = 2^{N-m}\cdot 2^{-N} = 2^{-m} = \prod_{i=1}^m P(e_{k_i} = \varepsilon_i).$$

(iv) Independence. Let $\{U_k\}$ be Borel subsets of $\mathbb{R}$ and $k_1 < \cdots < k_m$ distinct indices (permuting indices changes neither side). Put $V_i = U_{k_i} \cap \{-1,1\}$; since e_{k_i} takes only the values ± 1, $\{e_{k_i} \in U_{k_i}\}$ is the disjoint union of $\{e_{k_i} = \varepsilon\}$ over $\varepsilon \in V_i$, so $P(e_{k_i} \in U_{k_i}) = \frac{|V_i|}{2}$ by (i). Distributing the intersection over these unions gives a disjoint union over the $|V_1|\cdots|V_m|$ tuples in $V_1 \times \cdots \times V_m$, each of probability 2^{-m} by (iii), so

$$P\Big(\bigcap_{i=1}^m \{e_{k_i} \in U_{k_i}\}\Big) = |V_1|\cdots|V_m|\cdot 2^{-m} = \prod_{i=1}^m P(e_{k_i} \in U_{k_i}).$$

So these events are independent (12.7) for every choice of $\{U_k\}$, which by 12.14 makes $\{e_k\}_{k\in\mathbb{Z}^+}$ an independent family; with (i) it is an i.i.d. family (12.35).

Problem 12.17. Suppose $\mathcal{B}$ is the σ-algebra of Borel subsets of $(-\pi,\pi]$ and P is Lebesgue measure on $\big((-\pi,\pi],\mathcal{B}\big)$ divided by 2π. Let $\{e_k\}_{k\in\mathbb{Z}\setminus\{0\}}$ be the family of trigonometric functions defined by the third bullet point of Example 8.51 (notice that $k=0$ is excluded).
(a) Show that $\{e_k\}_{k\in\mathbb{Z}\setminus\{0\}}$ is not an independent family of random variables.
(b) Show that $\{e_k\}_{k\in\mathbb{Z}\setminus\{0\}}$ is an identically distributed family.

Solution. By 8.51, $e_k(t)=\frac{1}{\sqrt{\pi}}\sin(kt)$ for $k>0$ and $\frac{1}{\sqrt{\pi}}\cos(kt)$ for $k<0$; each is continuous, hence a random variable on $((-\pi,\pi],\mathcal{B})$, and $P(A)=\frac{\lambda(A)}{2\pi}$ where λ is Lebesgue measure.
Part (a). Take $k=1$ and $k=-1$ with $U_1=U_{-1}=U=\{x\in\mathbb{R}:x^2>\frac{1}{2\pi}\}$ and $U_k=\mathbb{R}$ otherwise. The Pythagorean identity $e_1(t)^2+e_{-1}(t)^2=\frac{1}{\pi}$ makes $\{e_1\in U\}$ and $\{e_{-1}\in U\}$ disjoint, since a common point would give $\frac{1}{\pi}>\frac{1}{2\pi}+\frac{1}{2\pi}$. Each has probability $\frac12$: the condition $e_1(t)^2>\frac{1}{2\pi}$ says $|\sin t|>\frac{1}{\sqrt{2}}$, and $e_{-1}(t)^2>\frac{1}{2\pi}$ says $|\cos t|>\frac{1}{\sqrt2}$, so

$$\{e_1\in U\}=\Big(-\tfrac{3\pi}{4},-\tfrac{\pi}{4}\Big)\cup\Big(\tfrac{\pi}{4},\tfrac{3\pi}{4}\Big),$$
$$\{e_{-1}\in U\}=\Big(-\pi,-\tfrac{3\pi}{4}\Big)\cup\Big(-\tfrac{\pi}{4},\tfrac{\pi}{4}\Big)\cup\Big(\tfrac{3\pi}{4},\pi\Big],$$

both of measure π. Hence

$$P\big(\{e_1\in U_1\}\cap\{e_{-1}\in U_{-1}\}\big)=0\neq\tfrac14=P(e_1\in U_1)\cdot P(e_{-1}\in U_{-1}),$$

so these events fail 12.7 and by 12.14 the family is not independent.
Part (b). Fix $s\in\mathbb{R}$ and put

$$S=\{u\in\mathbb{R}:\sin u\le s\sqrt{\pi}\},\qquad T=\{u\in\mathbb{R}:\cos u\le s\sqrt{\pi}\};$$

both are Borel (preimages of the closed set $(-\infty,s\sqrt{\pi}\,]$ under continuous maps) and periodic, meaning $S+2\pi=S$. Two facts about a periodic Borel set S:

- (A) $\lambda(S\cap(a,a+2\pi])=\lambda(S\cap(-\pi,\pi])$ for every $a\in\mathbb{R}$. Shifting a by a multiple of 2π changes neither side (periodicity and translation invariance 2.7), so take $-\pi<a\le\pi$; then $S\cap(a,a+2\pi]$ is the disjoint union of $S\cap(a,\pi]$ and $(S\cap(-\pi,a])+2\pi$, whose measures add to $\lambda(S\cap(-\pi,\pi])$ by 2.7.
- (B) $\lambda(\{t\in(-\pi,\pi]:mt\in S\})=\lambda(S\cap(-\pi,\pi])$ for $m\in\mathbb{Z}^+$. That set is $\frac1m(S\cap(-m\pi,m\pi])$, Borel with measure $\frac1m\lambda(S\cap(-m\pi,m\pi])$ by 5.41; and $(-m\pi,m\pi]$ splits into m windows of length 2π, each contributing $\lambda(S\cap(-\pi,\pi])$ by (A).

For $k>0$, $e_k(t)\le s$ means $kt\in S$, so (B) with $m=k$ gives $P(e_k\le s)=\frac{1}{2\pi}\lambda(S\cap(-\pi,\pi])$, a quantity free of k. For $k<0$, evenness of cosine gives $e_k(t)=\frac{1}{\sqrt{\pi}}\cos(mt)$ with $m=-k$, so (B) gives $P(e_k\le s)=\frac{1}{2\pi}\lambda(T\cap(-\pi,\pi])$. Finally $\cos u=\sin(u+\frac{\pi}{2})$ gives $T=S-\frac{\pi}{2}$, so 2.7 and then (A) on the window $(-\frac{\pi}{2},\frac{3\pi}{2}]$ give

$$\lambda\big(T\cap(-\pi,\pi]\big)=\lambda\Big(S\cap\Big(-\tfrac{\pi}{2},\tfrac{3\pi}{2}\Big]\Big)=\lambda\big(S\cap(-\pi,\pi]\big).$$

Thus $\widetilde{e_k}(s)$ does not depend on k, so the family is identically distributed (12.35).