Solutions to Tao's An Introduction to Measure Theory

Aayush Bajaj

2026-08-31T01:00:00+10:00

Contents

1 Notation **2**
1.1 §0.0 — Notation . . . 3

2 Measure Theory **6**
2.1 §1.1 — Prologue: The Problem of Measure . . . 7
2.2 §1.2 — Lebesgue Measure . . . 41
2.3 §1.3 — The Lebesgue Integral . . . 71
2.4 §1.4 — Abstract Measure Spaces . . . 106
2.5 §1.5 — Modes of Convergence . . . 166
2.6 §1.6 — Differentiation Theorems . . . 185
2.7 §1.7 — Outer Measures, Pre-measures, and Product Measures . . . 259

3 Related Articles **298**
3.1 §2.2 — The Rademacher Differentiation Theorem . . . 299
3.2 §2.3 — Probability Spaces . . . 305
3.3 §2.4 — Infinite Product Spaces and the Kolmogorov Extension Theorem . . . 307

Solutions to every exercise in Terence Tao's *An Introduction to Measure Theory* (AMS GSM 126, 2011) — 237 exercises, woven inline through the text as Tao builds the theory through them. The book is the MATH5825 companion text and is filed at An Introduction to Measure Theory (Tao).

1 Notation

Contents

1.1 §0.0 — Notation 3

1.1 §0.0 — Notation

Problem 0.0.1. If $(x_\alpha)_{\alpha\in A}$ is a collection of numbers $x_\alpha \in [0,+\infty]$ such that $\sum_{\alpha\in A} x_\alpha < \infty$, show that $x_\alpha = 0$ for all but at most countably many $\alpha \in A$, even if A itself is uncountable.

Solution. Write $S := \sum_{\alpha\in A} x_\alpha$, which by hypothesis is a finite non-negative real number. Recall the definition (0.1) of the sum over an arbitrary index set:

$$\sum_{\alpha\in A} x_\alpha = \sup_{F\subset A,\ F \text{ finite}} \sum_{\alpha\in F} x_\alpha.$$

In particular, we have the bound

$$\sum_{\alpha\in F} x_\alpha \le S \qquad \text{for every finite } F \subset A,$$

since S is by definition an upper bound for all such finite subsums. Call this bound $(\star)$.
For each natural number $n \ge 1$ define

$$A_n := \{\alpha \in A : x_\alpha > 1/n\}.$$

We claim that each A_n is finite, with $|A_n| \le nS$. Suppose not, so that A_n contains at least k distinct elements for some integer $k > nS$. Choose a finite subset $F \subset A_n$ with $|F| = k$. Since $x_\alpha > 1/n$ for every $\alpha \in F$, and a finite sum of elements of $[0,+\infty]$ is monotone in each summand, we get

$$\sum_{\alpha\in F} x_\alpha \ge \frac{k}{n} > S,$$

contradicting $(\star)$. Hence $|A_n| \le nS < \infty$ for every n, so each A_n is a finite set.
Now observe that

$$\{\alpha \in A : x_\alpha \neq 0\} = \bigcup_{n=1}^{\infty} A_n.$$

Indeed, the inclusion $\supset$ is immediate since every element of some A_n has $x_\alpha > 1/n > 0$. Conversely, if $x_\alpha \neq 0$ then $x_\alpha \in (0,+\infty]$, so there is a natural number n with $x_\alpha > 1/n$ (take any $n > 1/x_\alpha$ if x_α is finite, and $n = 1$ if $x_\alpha = +\infty$), whence $\alpha \in A_n$.
Thus $\{\alpha \in A : x_\alpha \neq 0\}$ is a countable union of finite sets, and therefore is at most countable. (Enumerating such a union requires a countable sequence of choices, which is legitimate here since the axiom of choice, Axiom 0.0.4, is assumed throughout the text; concretely, since the A_n are increasing, one enumerates A_1, then $A_2 \setminus A_1$, then $A_3 \setminus A_2$, and so on, choosing an ordering of each of these finite sets.) Equivalently, $x_\alpha = 0$ for all but at most countably many $\alpha \in A$. Note that no assumption on the cardinality of A was used, so this holds even when A is uncountable.
(As a side remark, the same argument shows that if $\sum_{\alpha\in A} x_\alpha$ is finite then no x_α can equal $+\infty$, since a singleton $F = \{\alpha\}$ in $(\star)$ would force $+\infty \le S$.)

Problem 0.0.2. (Tonelli's theorem for series over arbitrary sets). Let A, B be sets (possibly infinite or uncountable), and $(x_{n,m})_{n\in A, m\in B}$ be a doubly infinite sequence of extended non-negative reals $x_{n,m} \in [0,+\infty]$ indexed by A and B. Show that

$$\sum_{(n,m)\in A\times B} x_{n,m} = \sum_{n\in A}\sum_{m\in B} x_{n,m} = \sum_{m\in B}\sum_{n\in A} x_{n,m}.$$

(Hint: although not strictly necessary, you may find it convenient to first establish the fact that if $\sum_{n\in A} x_n$ is finite, then x_n is non-zero for at most countably many n.)

Solution. Throughout, all sums over arbitrary index sets are understood in the sense of the definition (0.1),

$$\sum_{\alpha\in A} x_\alpha = \sup_{F\subset A,\ F \text{ finite}} \sum_{\alpha\in F} x_\alpha,$$

where the supremum ranges over all finite subsets F of A (including $F=\emptyset$, whose sum is 0, so the supremum is over a non-empty collection and is always well defined in $[0,+\infty]$). We take the route that does not use the hint; the countability fact of Exercise 0.0.1 is not needed.

We first record three elementary facts.

Fact 1 (consistency with finite sums). If A is itself finite, then $\sum_{\alpha\in A} x_\alpha$ as defined by (0.1) coincides with the ordinary finite sum. Indeed $F=A$ is one of the admissible finite subsets, and for any $F\subset A$ we have $\sum_{\alpha\in F} x_\alpha \le \sum_{\alpha\in A} x_\alpha$ because the omitted terms are non-negative; so the supremum is attained at $F=A$. In particular all the usual commutativity and associativity rules for finite sums of elements of $[0,+\infty]$ apply, and finite sums are monotone in each summand.

Fact 2 (monotonicity in the index set). If $A'\subset A$ then $\sum_{\alpha\in A'} x_\alpha \le \sum_{\alpha\in A} x_\alpha$, since every finite subset of A' is also a finite subset of A, so the supremum defining the left-hand side is over a subcollection of the quantities defining the right-hand side.

Fact 3 (upward continuity of addition). Let $k\ge 1$ and let $P_1,\dots,P_k$ be non-empty subsets of $[0,+\infty]$. Then

$$\sup\{p_1+\cdots+p_k : p_1\in P_1,\dots,p_k\in P_k\} = \sup P_1+\cdots+\sup P_k.$$

By induction it suffices to treat $k=2$. The inequality $\le$ is clear from monotonicity of addition. For $\ge$: if some $\sup P_j=+\infty$, say $\sup P_1=+\infty$, then fixing any $p_2\in P_2$ and letting $p_1\in P_1$ be arbitrarily large makes p_1+p_2 arbitrarily large, so the left-hand side is $+\infty$ and the inequality holds. Otherwise $\sup P_1, \sup P_2$ are both finite; given $\varepsilon>0$ choose $p_j\in P_j$ with $p_j>\sup P_j-\varepsilon/2$, so that $p_1+p_2>\sup P_1+\sup P_2-\varepsilon$. Letting $\varepsilon\to 0$ gives the claim.

Now set

$$S := \sum_{(n,m)\in A\times B} x_{n,m}, \qquad y_n := \sum_{m\in B} x_{n,m}\ \ (n\in A), \qquad T := \sum_{n\in A} y_n = \sum_{n\in A}\sum_{m\in B} x_{n,m}.$$

Step 1: $S\le T$. Let $F\subset A\times B$ be an arbitrary finite set. Let

$$G := \{n\in A : (n,m)\in F \text{ for some } m\in B\}, \qquad F_n := \{m\in B : (n,m)\in F\}\ \ (n\in G),$$

so that G is a finite subset of A, each F_n is a finite subset of B, and F is the disjoint union of the sets $\{n\}\times F_n$ over $n\in G$. Regrouping the finite sum over F according to this partition (legitimate by Fact 1) gives

$$\sum_{(n,m)\in F} x_{n,m} = \sum_{n\in G}\sum_{m\in F_n} x_{n,m}.$$

By Fact 2, $\sum_{m\in F_n} x_{n,m} \le \sum_{m\in B} x_{n,m} = y_n$ for each $n\in G$; since finite sums are monotone in each summand, the right-hand side is at most $\sum_{n\in G} y_n$, which in turn is at most T by Fact 2 again (applied to $G\subset A$ and the family $(y_n)_{n\in A}$). Hence $\sum_{(n,m)\in F} x_{n,m}\le T$ for every finite $F\subset A\times B$, and taking the supremum over such F in (0.1) yields $S\le T$.

Step 2: $T\le S$. It suffices, by the definition (0.1) of T as a supremum, to show that

$$\sum_{n\in G} y_n \le S \qquad \text{for every finite } G\subset A.$$

Fix such a G, say $G=\{n_1,\dots,n_k\}$ with the n_j distinct (the case $G=\emptyset$ being trivial). Let $H_1,\dots,H_k$ be arbitrary finite subsets of B, and put $H := H_1\cup\cdots\cup H_k$, again a finite subset of B. Then $G\times H$ is a finite subset of $A\times B$, so by Fact 1 (regrouping the finite sum over the rectangle $G\times H$ into rows, which is legitimate since the n_j are distinct) together with the definition of S,

$$\sum_{j=1}^{k}\sum_{m\in H_j} x_{n_j,m} \le \sum_{j=1}^{k}\sum_{m\in H} x_{n_j,m} = \sum_{(n,m)\in G\times H} x_{n,m} \le S,$$

the first inequality holding because $H_j \subset H$ for each j and the discarded terms are non-negative. Thus the quantity $\sum_{j=1}^{k} \sum_{m \in H_j} x_{n_j,m}$ is bounded above by S as the finite sets $H_1, \dots, H_k \subset B$ vary independently. Taking the supremum over these independent choices and applying Fact 3 with

$$P_j := \Big\{ \sum_{m \in H_j} x_{n_j,m} : H_j \subset B \text{ finite} \Big\},$$

which is non-empty (it contains the empty sum 0, taking $H_j = \emptyset$) and whose supremum is precisely y_{n_j} by (0.1), we conclude

$$\sum_{j=1}^{k} y_{n_j} = \sup P_1 + \cdots + \sup P_k \leq S,$$

that is, $\sum_{n \in G} y_n \leq S$. As G was an arbitrary finite subset of A, we get $T \leq S$.
Combining Steps 1 and 2,

$$\sum_{(n,m) \in A \times B} x_{n,m} = \sum_{n \in A} \sum_{m \in B} x_{n,m}.$$

Step 3: the third expression. Define $\tilde{x}_{m,n} := x_{n,m}$ for $(m,n) \in B \times A$. Applying what we have just proved, with the roles of A and B interchanged and the family $(\tilde{x}_{m,n})_{m \in B, n \in A}$ in place of $(x_{n,m})$, gives

$$\sum_{(m,n) \in B \times A} \tilde{x}_{m,n} = \sum_{m \in B} \sum_{n \in A} \tilde{x}_{m,n} = \sum_{m \in B} \sum_{n \in A} x_{n,m}.$$

On the other hand the map $\phi : B \times A \to A \times B$ given by $\phi(m,n) := (n,m)$ is a bijection, and $x_{\phi(m,n)} = x_{n,m} = \tilde{x}_{m,n}$, so the change of variables formula (0.2) gives

$$\sum_{(n,m) \in A \times B} x_{n,m} = \sum_{(m,n) \in B \times A} x_{\phi(m,n)} = \sum_{(m,n) \in B \times A} \tilde{x}_{m,n}.$$

Chaining the last two displays with the identity $S = T$ obtained in Steps 1 and 2 yields

$$\sum_{(n,m) \in A \times B} x_{n,m} = \sum_{n \in A} \sum_{m \in B} x_{n,m} = \sum_{m \in B} \sum_{n \in A} x_{n,m},$$

as required. Note that no countability assumption on A or B was used anywhere; the argument only used the definition (0.1), the relabelling formula (0.2), and the arithmetic of $[0, +\infty]$. Taking $A = B = \mathbb{N}$ recovers Theorem 0.0.2, since for $A = \mathbb{N}$ the sum $\sum_{n \in \mathbb{N}} x_n$ defined by (0.1) agrees with $\lim_{N \to \infty} \sum_{n=1}^{N} x_n$, as noted in the text.

2 Measure Theory

Contents

2.1 §1.1 — Prologue: The Problem of Measure 7
2.2 §1.2 — Lebesgue Measure 41
2.3 §1.3 — The Lebesgue Integral 71
2.4 §1.4 — Abstract Measure Spaces 106
2.5 §1.5 — Modes of Convergence 166
2.6 §1.6 — Differentiation Theorems 185
2.7 §1.7 — Outer Measures, Pre-measures, and Product Measures 259

2.1 §1.1 — Prologue: The Problem of Measure

Problem 1.1.1. (Boolean closure). Show that if $E, F \subset \mathbb{R}^d$ are elementary sets, then the union $E \cup F$, the intersection $E \cap F$, and the set theoretic difference $E\backslash F := \{x \in E : x \notin F\}$, and the symmetric difference $E\Delta F := (E\backslash F)\cup(F\backslash E)$ are also elementary. If $x \in \mathbb{R}^d$, show that the translate $E + x := \{y + x : y \in E\}$ is also an elementary set.

Solution. Throughout, recall Definition 1.1.1: an interval is a subset of $\mathbb{R}$ of one of the forms $[a, b], [a, b), (a, b], (a, b)$ with $a \leq b$ real (so degenerate intervals, and the empty set, are allowed); a box in $\mathbb{R}^d$ is a Cartesian product $B = I_1 \times \cdots \times I_d$ of d intervals; and an elementary set is a finite union of boxes. We first record three facts about intervals and boxes.

Step 1: intervals are closed under intersection, and the difference of two intervals is a union of at most two intervals.

Let $I, J \subset \mathbb{R}$ be intervals. Each is a convex subset of $\mathbb{R}$ which is bounded (possibly empty), hence so is $I \cap J$: indeed if $x < y < z$ with $x, z \in I \cap J$ then $y \in I$ and $y \in J$ by convexity of each, so $y \in I \cap J$. A bounded convex subset of $\mathbb{R}$ is either empty or is exactly of one of the four listed forms, with $a = \inf$ and $b = \sup$ of the set; and the empty set is the degenerate interval (a, a). Hence $I \cap J$ is an interval. For the difference, write J as having endpoints $c \leq d$. Then

$$I\backslash J = I \cap (\mathbb{R}\backslash J) = (I \cap L) \cup (I \cap R),$$

where $L := \{x \in \mathbb{R} : x < c\}$ if $c \in J$ and $L := \{x \in \mathbb{R} : x \leq c\}$ if $c \notin J$, and likewise $R := \{x \in \mathbb{R} : x > d\}$ if $d \in J$ and $R := \{x \in \mathbb{R} : x \geq d\}$ if $d \notin J$. These satisfy $\mathbb{R}\backslash J = L \cup R$: a point $x < c$ or $x > d$ lies outside J (as $J \subset [c, d]$) and lies in L or R respectively; a point with $c < x < d$ lies in J, since J is an interval with endpoints $c \leq d$ and so contains (c, d); and the two endpoints $x = c$, $x = d$ are placed into L, resp. R, exactly when they are not in J. Note L and R are convex. Each of $I \cap L$ and $I \cap R$ is a bounded convex subset of $\mathbb{R}$, hence (by the previous paragraph) an interval. Thus $I\backslash J$ is a union of at most two intervals. (If $J = \emptyset$ we may simply take $I\backslash J = I$.)

Step 2: the intersection of two boxes is a box, and the difference of two boxes is elementary.

Let $B = I_1 \times \cdots \times I_d$ and $B' = J_1 \times \cdots \times J_d$ be boxes. Then

$$B \cap B' = (I_1 \cap J_1) \times \cdots \times (I_d \cap J_d),$$

since $x \in B \cap B'$ if and only if $x_k \in I_k$ and $x_k \in J_k$ for every k. By Step 1 each factor is an interval, so $B \cap B'$ is a box.

For the difference, I claim

$$B\backslash B' = \bigcup_{k=1}^{d} (I_1 \cap J_1) \times \cdots \times (I_{k-1} \cap J_{k-1}) \times (I_k\backslash J_k) \times I_{k+1} \times \cdots \times I_d.$$

Indeed, a point x lies in $B\backslash B'$ precisely when $x_j \in I_j$ for all j and $x_k \notin J_k$ for at least one k; letting k be the least such index gives $x_j \in I_j \cap J_j$ for $j < k$, $x_k \in I_k\backslash J_k$, and $x_j \in I_j$ for $j > k$, which is membership in the k-th set on the right. Conversely each set on the right is visibly contained in B and misses B' (its k-th coordinate avoids J_k). By Step 1, $I_k\backslash J_k$ is a union of at most two intervals, so expanding that union distributes each summand into at most two boxes. Hence $B\backslash B'$ is a finite union of boxes, i.e. elementary. (In fact the d sets displayed are pairwise disjoint, since they are distinguished by the least index k at which the J-condition fails, so this even expresses $B\backslash B'$ as a disjoint union of at most $2d$ boxes.)

Step 3: the Boolean operations.

Write $E = B_1 \cup \cdots \cup B_n$ and $F = B'_1 \cup \cdots \cup B'_m$ as finite unions of boxes.

Union: $E \cup F = B_1 \cup \cdots \cup B_n \cup B'_1 \cup \cdots \cup B'_m$ is a finite union of boxes, hence elementary.

Intersection: distributing,

$$E \cap F = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} (B_i \cap B'_j),$$

and each $B_i \cap B'_j$ is a box by Step 2. So $E \cap F$ is a finite union of boxes, hence elementary. By induction, any finite intersection of elementary sets is elementary.
Difference: distributing again,

$$E\backslash F = \bigcup_{i=1}^{n}(B_i\backslash F) = \bigcup_{i=1}^{n}\bigcap_{j=1}^{m}(B_i\backslash B'_j).$$

(The first equality holds because $x \in E\backslash F$ iff x lies in some B_i and in no B'_j; the second because failing to lie in F means failing to lie in each B'_j separately.) Each $B_i\backslash B'_j$ is elementary by Step 2; each inner finite intersection is therefore elementary by the previous paragraph; and the outer finite union of elementary sets is elementary since a finite union of finite unions of boxes is a finite union of boxes. Hence $E\backslash F$ is elementary. (If $m = 0$, i.e. $F = \emptyset$, then $E\backslash F = E$ is elementary trivially.)
Symmetric difference: $E\Delta F = (E\backslash F) \cup (F\backslash E)$ is the union of two elementary sets by the previous paragraph, hence elementary by the union case.
Step 4: translations.
Let $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. If $I = [a, b]$ is an interval and $t \in \mathbb{R}$, then $I + t = [a + t, b + t]$ is again an interval of the same type (and of the same length), and similarly for the half-open and open cases. Consequently, for a box $B = I_1 \times \cdots \times I_d$ we have

$$B + x = (I_1 + x_1) \times \cdots \times (I_d + x_d),$$

because $y + x \in B + x$ iff $y_k \in I_k$ for all k iff $(y + x)_k \in I_k + x_k$ for all k. So $B + x$ is a box. Finally, translation commutes with unions: if $E = B_1 \cup \cdots \cup B_n$ then

$$E + x = (B_1 + x) \cup \cdots \cup (B_n + x),$$

a finite union of boxes. Hence $E + x$ is elementary. □

Problem 1.1.2. Give an alternate proof of Lemma 1.1.2(ii) by showing that any two partitions of E into boxes admit a mutual refinement into boxes that arise from taking Cartesian products of elements from finite collections of disjoint intervals.

Solution. We must show: if an elementary set $E \subset \mathbb{R}^d$ is partitioned into disjoint boxes in two ways, $E = B_1 \cup \cdots \cup B_k$ and $E = B'_1 \cup \cdots \cup B'_{k'}$, then $|B_1| + \cdots + |B_k| = |B'_1| + \cdots + |B'_{k'}|$. We do this without the discretisation argument of the text, by building a common grid refinement.
Step 1: the one-dimensional refinement.
Let $I_1, \ldots, I_n$ be finitely many intervals in $\mathbb{R}$. Discard the empty ones (they will be handled by the empty subcollection). List the distinct endpoints of the remaining intervals in increasing order as $a_0 < a_1 < \cdots < a_p$. Define the **atoms** of this family to be the intervals

$$\{a_0\}, \{a_1\}, \ldots, \{a_p\}, \quad (a_0, a_1), (a_1, a_2), \ldots, (a_{p-1}, a_p).$$

These are $2p + 1$ pairwise disjoint intervals (singletons are degenerate intervals $[a_s, a_s]$, allowed by Definition 1.1.1). I claim each I_r is the union of the subcollection of atoms it contains. Let I_r have endpoints $a_u \le a_v$. Every atom contained in I_r is a subset of I_r, so one inclusion is trivial. Conversely let $x \in I_r$, so $a_u \le x \le a_v$. If $x = a_s$ for some s, then $\{a_s\}$ is an atom contained in I_r (it contains x, and $\{x\} \subset I_r$), and x lies in it. Otherwise x is not an endpoint, so $a_s < x < a_{s+1}$ for a unique s with $u \le s \le v - 1$; since I_r is an interval containing $a_u \le a_s < x$ and $x < a_{s+1} \le a_v$, and $I_r \supset (a_u, a_v)$, the whole atom (a_s, a_{s+1}) lies in $(a_u, a_v) \subset I_r$. Either way x lies in an atom contained in I_r, proving the claim.
Moreover the lengths add up correctly:

$$\sum_{J \text{ atom}, J \subset I_r} |J| = \sum_{s=u}^{v-1}(a_{s+1} - a_s) = a_v - a_u = |I_r|,$$

since the singleton atoms have length 0 and the open atoms contained in I_r are precisely (a_s, a_{s+1}) for $u \le s \le v-1$ (such an atom lies in I_r as shown above; conversely an open atom (a_s, a_{s+1}) meeting I_r forces $u \le s \le v-1$). The middle sum telescopes. For $I_r = \emptyset$ both sides are 0.
Step 2: the mutual refinement in $\mathbb{R}^d$.
Consider all $k + k'$ boxes $B_1, \dots, B_k, B'_1, \dots, B'_{k'}$ at once, and write each as a product of intervals, $B_i = I_{i,1} \times \cdots \times I_{i,d}$ and $B'_j = I'_{j,1} \times \cdots \times I'_{j,d}$. Fix a coordinate $\ell \in \{1, \dots, d\}$ and apply Step 1 to the finite family of ℓ-th factor intervals $I_{1,\ell}, \dots, I_{k,\ell}, I'_{1,\ell}, \dots, I'_{k',\ell}$; this produces a finite collection $\mathcal{J}_\ell$ of pairwise disjoint intervals such that every one of those factor intervals is the union of the members of $\mathcal{J}_\ell$ it contains, with lengths adding as in Step 1.
Now define the grid boxes to be the Cartesian products

$$Q = J_1 \times \cdots \times J_d, \qquad J_\ell \in \mathcal{J}_\ell.$$

There are finitely many of them and they are pairwise disjoint: two distinct grid boxes differ in some coordinate ℓ, where their factors are distinct hence disjoint members of $\mathcal{J}_\ell$, so the boxes are disjoint.
Each of our $k + k'$ boxes is a disjoint union of grid boxes. Indeed, for $B = I_1 \times \cdots \times I_d$ one of them, writing $I_\ell = \bigcup \{J \in \mathcal{J}_\ell : J \subset I_\ell\}$ (Step 1) and distributing the product over these unions gives

$$B = \bigcup \big\{ J_1 \times \cdots \times J_d \ : \ J_\ell \in \mathcal{J}_\ell,\ J_\ell \subset I_\ell \ (1 \le \ell \le d) \big\},$$

a disjoint union of grid boxes, namely exactly the grid boxes contained in B. Furthermore, by Step 1 and the distributive law for finite sums,

$$\sum_{Q \text{ grid box}, Q \subset B} |Q| \ = \ \sum_{J_\ell \in \mathcal{J}_\ell,\ J_\ell \subset I_\ell} \prod_{\ell=1}^{d} |J_\ell| \ = \ \prod_{\ell=1}^{d} \Big(\sum_{J \in \mathcal{J}_\ell,\, J \subset I_\ell} |J| \Big) \ = \ \prod_{\ell=1}^{d} |I_\ell| \ = \ |B|. \qquad (*)$$

This is the promised mutual refinement: the two given partitions of E are both refined by the single family of grid boxes, which are products of elements of the finite disjoint interval collections $\mathcal{J}_1, \dots, \mathcal{J}_d$.
Step 3: conclusion.
Let $\mathcal{Q}_E$ be the set of grid boxes contained in E. First note that a grid box Q which meets E lies in $\mathcal{Q}_E$ and in exactly one B_i: since Q meets $E = B_1 \cup \cdots \cup B_k$, it meets some B_i, and B_i is a union of grid boxes by Step 2, so Q meets some grid box $Q' \subset B_i$; distinct grid boxes are disjoint, so $Q = Q' \subset B_i$; and the B_i are disjoint, so this i is unique. Conversely every grid box contained in some B_i meets E (unless it is empty, in which case it contributes 0 to every sum and lies in every B_i vacuously; to avoid this nuisance simply discard empty grid boxes at the outset, which changes no sum). Hence the grid boxes contained in the various B_i are precisely the members of $\mathcal{Q}_E$, each occurring for exactly one i. Therefore, using $(*)$ for each B_i,

$$\sum_{i=1}^{k} |B_i| = \sum_{i=1}^{k} \sum_{Q \subset B_i} |Q| = \sum_{Q \in \mathcal{Q}_E} |Q|.$$

The identical argument applied to the second partition gives $\sum_{j=1}^{k'} |B'_j| = \sum_{Q \in \mathcal{Q}_E} |Q|$, since $\mathcal{Q}_E$ depends only on E and on the grid, not on which partition we started from. Comparing the two displays,

$$|B_1| + \cdots + |B_k| = |B'_1| + \cdots + |B'_{k'}|,$$

which is Lemma 1.1.2(ii). □

Problem 1.1.3. (Uniqueness of elementary measure). Let $d \ge 1$. Let $m' : \mathcal{E}(\mathbb{R}^d) \to \mathbb{R}^+$ be a map from the collection $\mathcal{E}(\mathbb{R}^d)$ of elementary subsets of $\mathbb{R}^d$ to the nonnegative reals that obeys the non-negativity, finite additivity, and translation invariance properties. Show that there exists a constant $c \in \mathbb{R}^+$ such that $m'(E) = c\,m(E)$ for all elementary sets E. In particular, if we impose

the additional normalisation $m'([0,1)^d) = 1$, then $m' \equiv m$. (Hint: Set $c := m'([0,1)^d)$, and then compute $m'([0,\frac{1}{n})^d)$ for any positive integer n.)

Solution. Set $c := m'([0,1)^d) \in \mathbb{R}^+$, so $c \geq 0$. We show $m' = c\,m$ on $\mathcal{E}(\mathbb{R}^d)$.
Step 0: elementary consequences of the three axioms.
Taking $E = F = \emptyset$ in finite additivity gives $m'(\emptyset) = 2m'(\emptyset)$, so $m'(\emptyset) = 0$. If $E \subset F$ are elementary, then $F\backslash E$ is elementary by Exercise 1.1.1 and $F = E \cup (F\backslash E)$ is a disjoint union, so by finite additivity and non-negativity

$$m'(F) = m'(E) + m'(F\backslash E) \;\geq\; m'(E),$$

i.e. m' is monotone. By induction, finite additivity extends to any finite disjoint union: $m'(E_1 \cup \cdots \cup E_k) = m'(E_1) + \cdots + m'(E_k)$ for pairwise disjoint elementary $E_1, \ldots, E_k$.
Step 1: the small half-open cubes.
Fix a positive integer n and write $Q_n := [0,\frac{1}{n})^d$. The unit cube decomposes as

$$[0,1)^d = \bigcup_{k \in \{0,1,\ldots,n-1\}^d} \Big(Q_n + \tfrac{k}{n}\Big),$$

and this union is disjoint: a point $x \in [0,1)^d$ lies in the piece indexed by the unique k with $k_j = \lfloor nx_j \rfloor \in \{0,\ldots,n-1\}$. There are n^d pieces, each a translate of Q_n, so by finite additivity and translation invariance

$$c = m'([0,1)^d) = n^d\, m'(Q_n), \qquad \text{hence} \qquad m'(Q_n) = \frac{c}{n^d} = c\,m(Q_n),$$

using $m(Q_n) = |Q_n| = n^{-d}$. Consequently, if S is a disjoint union of N translates of Q_n, then by finite additivity and translation invariance $m'(S) = Nc/n^d$, while by Lemma 1.1.2 $m(S) = N/n^d$; hence

$$m'(S) = c\,m(S) \qquad \text{for every such } S. \tag{1}$$

Step 2: $m'(B) = c|B|$ for every box B.
Let $B = I_1 \times \cdots \times I_d$ be a box, with I_j an interval of endpoints $a_j \leq b_j$, and let $L_j := b_j - a_j = |I_j|$. For a positive integer n, the half-open cubes $Q_n + \frac{k}{n}$ with $k \in \mathbb{Z}^d$ are pairwise disjoint and cover $\mathbb{R}^d$ (again by $k_j = \lfloor nx_j \rfloor$). Define

$$A_n := \bigcup\{Q_n + \tfrac{k}{n} \;:\; k \in \mathbb{Z}^d,\; Q_n + \tfrac{k}{n} \subset B\}, \qquad C_n := \bigcup\{Q_n + \tfrac{k}{n} \;:\; k \in \mathbb{Z}^d,\; (Q_n + \tfrac{k}{n}) \cap B \neq \emptyset\}.$$

Since B is bounded, only finitely many grid cubes meet it, so both unions are finite: A_n and C_n are elementary, are disjoint unions of translates of Q_n, and satisfy $A_n \subset B \subset C_n$.
Count the cubes in each coordinate. A grid cube lies in B exactly when $[\frac{k_j}{n}, \frac{k_j+1}{n}) \subset I_j$ for every j, and it meets B exactly when $[\frac{k_j}{n}, \frac{k_j+1}{n}) \cap I_j \neq \emptyset$ for every j. For a fixed j:

- If $[\frac{k}{n}, \frac{k+1}{n}) \subset I_j$ then $\frac{k}{n} \geq a_j$ and $\frac{k+1}{n} \leq b_j$, so $na_j \leq k \leq nb_j - 1$: at most nL_j such k. Conversely if $na_j < k$ and $k+1 < nb_j$ then $[\frac{k}{n}, \frac{k+1}{n}) \subset (a_j, b_j) \subset I_j$; the number of integers k with $na_j < k < nb_j - 1$ is at least $nL_j - 2$. So the count $\alpha_{j,n}$ of admissible k satisfies $\max(0, nL_j - 2) \leq \alpha_{j,n} \leq nL_j$.
- If $[\frac{k}{n}, \frac{k+1}{n}) \cap I_j \neq \emptyset$, pick x in the intersection: then $\frac{k}{n} \leq x \leq b_j$ and $\frac{k+1}{n} > x \geq a_j$, so $na_j - 1 < k \leq nb_j$; the count $\gamma_{j,n}$ of such k satisfies $\gamma_{j,n} \leq nL_j + 1$.

Since the grid cubes making up A_n (resp. C_n) are indexed by tuples k whose coordinates range independently over the admissible sets, A_n consists of $\prod_j \alpha_{j,n}$ cubes and C_n of $\prod_j \gamma_{j,n}$ cubes, whence

$$m(A_n) = \frac{1}{n^d} \prod_{j=1}^{d} \alpha_{j,n} \;\geq\; \prod_{j=1}^{d} \max\Big(0, L_j - \tfrac{2}{n}\Big), \qquad m(C_n) = \frac{1}{n^d} \prod_{j=1}^{d} \gamma_{j,n} \;\leq\; \prod_{j=1}^{d} \Big(L_j + \tfrac{1}{n}\Big).$$

Both bounds converge to $\prod_j L_j = |B|$ as $n \to \infty$, and $m(A_n) \leq |B| \leq m(C_n)$ by monotonicity of m; hence $m(C_n) - m(A_n) \to 0$.

Now apply monotonicity of m' (Step 0) together with (1):

$$c\,m(A_n) = m'(A_n) \;\le\; m'(B) \;\le\; m'(C_n) = c\,m(C_n).$$

Since $c\,m(A_n)$ and $c\,m(C_n)$ both converge to $c|B|$, we conclude $m'(B) = c|B| = c\,m(B)$.
Step 3: general elementary sets.
Let E be elementary. By Lemma 1.1.2(i), $E = B_1 \cup \cdots \cup B_k$ for disjoint boxes B_i, and by Lemma 1.1.2(ii) $m(E) = |B_1| + \cdots + |B_k|$. By finite additivity of m' and Step 2,

$$m'(E) = \sum_{i=1}^{k} m'(B_i) = \sum_{i=1}^{k} c|B_i| = c\,m(E).$$

This proves the claim with $c = m'([0,1)^d) \in \mathbb{R}^+$.
Finally, if we normalise by $m'([0,1)^d) = 1$, then $c = 1$ and the identity reads $m'(E) = m(E)$ for every elementary E, i.e. $m' \equiv m$. $\square$
(Remark: translation invariance and finite additivity cannot be dropped. Without translation invariance, $m'(E) := m(E \cap [0,1]^d)$ is non-negative and finitely additive but is not a multiple of m; without finite additivity, $m'(E) := m(E)^2$ is non-negative and translation invariant but is not a multiple of m. Non-negativity enters the argument only once, through the monotonicity of m' established in Step 0, which is what licenses the squeeze $m'(A_n) \le m'(B) \le m'(C_n)$ in Step 2; without it the argument above gives no conclusion, since Step 1 pins down m' only on finite disjoint unions of dyadic-type cubes.)

Problem 1.1.4. Let $d_1, d_2 \ge 1$, and let $E_1 \subset \mathbb{R}^{d_1}$, $E_2 \subset \mathbb{R}^{d_2}$ be elementary sets. Show that $E_1 \times E_2 \subset \mathbb{R}^{d_1+d_2}$ is elementary, and $m_{d_1+d_2}(E_1 \times E_2) = m_{d_1}(E_1) \times m_{d_2}(E_2)$.

Solution. Step 1: products of boxes are boxes, and volumes multiply.
Let $B = I_1 \times \cdots \times I_{d_1} \subset \mathbb{R}^{d_1}$ and $B' = I'_1 \times \cdots \times I'_{d_2} \subset \mathbb{R}^{d_2}$ be boxes. Identifying $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with $\mathbb{R}^{d_1+d_2}$ in the obvious way, we have

$$B \times B' = I_1 \times \cdots \times I_{d_1} \times I'_1 \times \cdots \times I'_{d_2},$$

a Cartesian product of $d_1 + d_2$ intervals, hence a box in $\mathbb{R}^{d_1+d_2}$. Its volume is, straight from Definition 1.1.1,

$$|B \times B'| = |I_1| \cdots |I_{d_1}| \cdot |I'_1| \cdots |I'_{d_2}| = |B| \cdot |B'|.$$

Step 2: the product set is elementary.
By Lemma 1.1.2(i) we may write $E_1 = B_1 \cup \cdots \cup B_n$ and $E_2 = B'_1 \cup \cdots \cup B'_m$ where the B_i are disjoint boxes in $\mathbb{R}^{d_1}$ and the B'_j are disjoint boxes in $\mathbb{R}^{d_2}$. Then

$$E_1 \times E_2 = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} (B_i \times B'_j),$$

since a point (x, y) lies in $E_1 \times E_2$ exactly when x lies in some B_i and y lies in some B'_j. This is a finite union of boxes by Step 1, so $E_1 \times E_2$ is elementary.
Moreover the boxes $B_i \times B'_j$ are pairwise disjoint: for $(i, j) \neq (i', j')$,

$$(B_i \times B'_j) \cap (B_{i'} \times B'_{j'}) = (B_i \cap B_{i'}) \times (B'_j \cap B'_{j'}),$$

and either $i \neq i'$, making the first factor empty, or $j \neq j'$, making the second factor empty; either way the intersection is empty.
Step 3: the measure identity.
Since $\{B_i \times B'_j\}_{i,j}$ is a partition of $E_1 \times E_2$ into disjoint boxes, Lemma 1.1.2(ii) lets us compute $m_{d_1+d_2}(E_1 \times E_2)$ from this partition, and using Step 1 and then the distributive law for finite sums,

$$m_{d_1+d_2}(E_1 \times E_2) = \sum_{i=1}^{n} \sum_{j=1}^{m} |B_i \times B'_j| = \sum_{i=1}^{n} \sum_{j=1}^{m} |B_i|\,|B'_j| = \Big(\sum_{i=1}^{n} |B_i| \Big) \Big(\sum_{j=1}^{m} |B'_j| \Big).$$

By Lemma 1.1.2(ii) applied to the two given partitions, the two bracketed sums are $m_{d_1}(E_1)$ and $m_{d_2}(E_2)$ respectively. Hence

$$m_{d_1+d_2}(E_1 \times E_2) = m_{d_1}(E_1)\, m_{d_2}(E_2),$$

as required. □

Problem 1.1.5. (Characterisation of Jordan measurability). Let $E \subset \mathbb{R}^d$ be bounded. Show that the following are equivalent:

1. E is Jordan measurable.
2. For every $\varepsilon > 0$, there exist elementary sets $A \subset E \subset B$ such that $m(B\backslash A) \leq \varepsilon$.
3. For every $\varepsilon > 0$, there exists an elementary set A such that $m^{*,(J)}(A\Delta E) \leq \varepsilon$.

Solution. Throughout, E is bounded, so it is contained in some box B_0; hence the family of elementary sets containing E is non-empty and $m^{*,(J)}(E) \leq |B_0| < \infty$, while the family of elementary sets contained in E contains $\emptyset$, so $m_{*,(J)}(E) \geq 0$. We record two standing facts.
(a) $m_{*,(J)}(E) \leq m^{*,(J)}(E)$. If $A \subset E \subset B$ with A, B elementary, then $A \subset B$, so $m(A) \leq m(B)$ by monotonicity of elementary measure. Taking the supremum over A and the infimum over B gives the claim.
(b) $m^{*,(J)}$ is monotone, and $m^{*,(J)}(C) \leq m(C)$ for elementary C. If $S \subset T$ are bounded, every elementary set containing T contains S, so the infimum defining $m^{*,(J)}(S)$ is over a larger family: $m^{*,(J)}(S) \leq m^{*,(J)}(T)$. And for elementary C we may take $B = C$ in the infimum, giving $m^{*,(J)}(C) \leq m(C)$.
(1) $\Rightarrow$ (2). Suppose $m_{*,(J)}(E) = m^{*,(J)}(E)$, and let $\varepsilon > 0$. By definition of the supremum there is an elementary $A \subset E$ with $m(A) \geq m_{*,(J)}(E) - \varepsilon/2$, and by definition of the infimum there is an elementary $B \supset E$ with $m(B) \leq m^{*,(J)}(E) + \varepsilon/2$. Then $A \subset B$, and $B\backslash A$ is elementary by Exercise 1.1.1, with $B = A \cup (B\backslash A)$ a disjoint union; finite additivity of elementary measure gives

$$m(B\backslash A) = m(B) - m(A) \leq \left(m^{*,(J)}(E) + \tfrac{\varepsilon}{2}\right) - \left(m_{*,(J)}(E) - \tfrac{\varepsilon}{2}\right) = \varepsilon,$$

using $m^{*,(J)}(E) = m_{*,(J)}(E)$.
(2) $\Rightarrow$ (1). Let $\varepsilon > 0$ and take elementary $A \subset E \subset B$ with $m(B\backslash A) \leq \varepsilon$. By finite additivity, $m(B) = m(A) + m(B\backslash A) \leq m(A) + \varepsilon$. Since A is a competitor for the inner supremum and B for the outer infimum,

$$m^{*,(J)}(E) \leq m(B) \leq m(A) + \varepsilon \leq m_{*,(J)}(E) + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $m^{*,(J)}(E) \leq m_{*,(J)}(E)$; combined with (a) this gives equality, i.e. E is Jordan measurable.
(2) $\Rightarrow$ (3). Let $\varepsilon > 0$ and take elementary $A \subset E \subset B$ with $m(B\backslash A) \leq \varepsilon$. Since $A \subset E$ we have $A\backslash E = \emptyset$, so

$$A\Delta E = E\backslash A \subset B\backslash A.$$

By the monotonicity in (b), followed by the second half of (b) applied to the elementary set $B\backslash A$,

$$m^{*,(J)}(A\Delta E) \leq m^{*,(J)}(B\backslash A) \leq m(B\backslash A) \leq \varepsilon.$$

(3) $\Rightarrow$ (2). Let $\varepsilon > 0$. Apply (3) with $\varepsilon/3$ to obtain an elementary set A with $m^{*,(J)}(A\Delta E) \leq \varepsilon/3$. (Note $A\Delta E$ is bounded, being contained in the union of the bounded sets A and E, so its Jordan outer measure is defined.) By definition of the infimum there is an elementary set $C \supset A\Delta E$ with

$$m(C) \leq m^{*,(J)}(A\Delta E) + \tfrac{\varepsilon}{3} \leq \tfrac{2\varepsilon}{3} \leq \varepsilon.$$

Define

$$A' := A\backslash C, \qquad B' := A \cup C,$$

which are elementary by Exercise 1.1.1. We check $A' \subset E \subset B'$.

For $A' \subset E$: let $x \in A$ with $x \notin C$. Since $A\backslash E \subset A\Delta E \subset C$ and $x \notin C$, we get $x \notin A\backslash E$; as $x \in A$, this forces $x \in E$.
For $E \subset B'$: let $x \in E$. If $x \in A$ then $x \in B'$. If $x \notin A$ then $x \in E\backslash A \subset A\Delta E \subset C \subset B'$.
Finally $B'\backslash A' \subset C$: if $x \in A \cup C$ and $x \notin A\backslash C$, then either $x \in A$, in which case $x \notin A\backslash C$ forces $x \in C$; or $x \notin A$, in which case $x \in C$. Hence by monotonicity of elementary measure

$$m(B'\backslash A') \leq m(C) \leq \varepsilon,$$

which is exactly (2) for the pair $A' \subset E \subset B'$.
This closes the cycle: (1) $\Leftrightarrow$ (2) and (2) $\Leftrightarrow$ (3), so all three are equivalent. $\square$
(As the text notes, criterion (2) immediately shows every elementary set E is Jordan measurable (take $A = B = E$, so $m(B\backslash A) = m(\emptyset) = 0 \leq \varepsilon$), and then $m_{*,(J)}(E) = m^{*,(J)}(E) = m(E)$, since E is a competitor in both the supremum and the infimum. So Jordan measure extends elementary measure, and $m(\emptyset) = 0$ persists.)

Problem 1.1.6. Let $E, F \subset \mathbb{R}^d$ be Jordan measurable sets.
1. (Boolean closure) Show that $E \cup F$, $E \cap F$, $E\backslash F$, and $E\Delta F$ are Jordan measurable.
2. (Non-negativity) $m(E) \geq 0$.
3. (Finite additivity) If E, F are disjoint, then $m(E \cup F) = m(E) + m(F)$.
4. (Monotonicity) If $E \subset F$, then $m(E) \leq m(F)$.
5. (Finite subadditivity) $m(E \cup F) \leq m(E) + m(F)$.
6. (Translation invariance) For any $x \in \mathbb{R}^d$, $E + x$ is Jordan measurable, and $m(E + x) = m(E)$.

Solution. Recall that Jordan measurable sets are by convention bounded, so $E \cup F$, $E \cap F$, $E\backslash F$, $E\Delta F$ and $E + x$ are all bounded and their Jordan inner and outer measures are defined. We use throughout the criterion of Exercise 1.1.5, in the form: a bounded set S is Jordan measurable if and only if for every $\varepsilon > 0$ there are elementary $A \subset S \subset B$ with $m(B\backslash A) \leq \varepsilon$.
(1) Boolean closure. Fix $\varepsilon > 0$. By Exercise 1.1.5 applied to E and to F with the parameter $\varepsilon/2$, choose elementary sets

$$A \subset E \subset B, \qquad A' \subset F \subset B', \qquad m(B\backslash A) \leq \tfrac{\varepsilon}{2}, \quad m(B'\backslash A') \leq \tfrac{\varepsilon}{2}.$$

All the sets built below from A, A', B, B' by Boolean operations are elementary, by Exercise 1.1.1. In each case we exhibit an elementary inner and outer approximation whose difference is contained in $(B\backslash A) \cup (B'\backslash A')$; since elementary measure is monotone and finitely subadditive, that difference then has elementary measure at most $\frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, and Exercise 1.1.5 gives Jordan measurability.
Union. Clearly $A \cup A' \subset E \cup F \subset B \cup B'$. If $x \in (B \cup B')\backslash(A \cup A')$ then $x \notin A$ and $x \notin A'$, and x lies in B or in B'; in the first case $x \in B\backslash A$, in the second $x \in B'\backslash A'$. So $(B \cup B')\backslash(A \cup A') \subset (B\backslash A) \cup (B'\backslash A')$.
Intersection. Clearly $A \cap A' \subset E \cap F \subset B \cap B'$. If $x \in (B \cap B')\backslash(A \cap A')$ then $x \in B$, $x \in B'$, and $x \notin A$ or $x \notin A'$; correspondingly $x \in B\backslash A$ or $x \in B'\backslash A'$. So $(B \cap B')\backslash(A \cap A') \subset (B\backslash A) \cup (B'\backslash A')$.
Difference. We claim $A\backslash B' \subset E\backslash F \subset B\backslash A'$. Indeed if $x \in A$ and $x \notin B'$ then $x \in E$, and $x \notin F$ because $F \subset B'$; and if $x \in E\backslash F$ then $x \in B$, while $x \notin F \supset A'$ gives $x \notin A'$. Next, suppose $x \in (B\backslash A')\backslash(A\backslash B')$, so $x \in B$, $x \notin A'$, and either $x \notin A$ or $x \in B'$. In the first case $x \in B\backslash A$; in the second case $x \in B'$ and $x \notin A'$, i.e. $x \in B'\backslash A'$. So again the difference set is contained in $(B\backslash A) \cup (B'\backslash A')$.
Symmetric difference. By the previous paragraph both $E\backslash F$ and $F\backslash E$ are Jordan measurable, and hence so is their union $E\Delta F = (E\backslash F) \cup (F\backslash E)$ by the union case.
(2) Non-negativity. The empty set is elementary and $\emptyset \subset E$, so it is a competitor in the supremum defining the inner measure; hence

$$m(E) = m_{*,(J)}(E) \geq m(\emptyset) = 0.$$

(3) Finite additivity. Suppose $E \cap F = \emptyset$; by (1) the set $E \cup F$ is Jordan measurable, so $m(E \cup F)$ is defined.

Lower bound. Let $A \subset E$ and $A' \subset F$ be elementary. Since E and F are disjoint, so are A and A'; and $A \cup A'$ is elementary with $A \cup A' \subset E \cup F$. By finite additivity of elementary measure,

$$m_{*,(J)}(E \cup F) \geq m(A \cup A') = m(A) + m(A').$$

Taking the supremum over A and then over A' (the two families are independent of one another) gives $m(E \cup F) \geq m_{*,(J)}(E) + m_{*,(J)}(F) = m(E) + m(F)$.
Upper bound. Let $B \supset E$ and $B' \supset F$ be elementary. Then $B \cup B'$ is elementary and contains $E \cup F$, so by finite subadditivity of elementary measure,

$$m^{*,(J)}(E \cup F) \leq m(B \cup B') \leq m(B) + m(B').$$

Taking the infimum over B and over B' gives $m(E \cup F) \leq m^{*,(J)}(E) + m^{*,(J)}(F) = m(E) + m(F)$.
Combining the two bounds, $m(E \cup F) = m(E) + m(F)$.
(4) Monotonicity. If $E \subset F$, then every elementary $A \subset E$ also satisfies $A \subset F$, so the supremum defining $m_{*,(J)}(F)$ is over a larger family than that defining $m_{*,(J)}(E)$. Hence $m(E) = m_{*,(J)}(E) \leq m_{*,(J)}(F) = m(F)$.
(5) Finite subadditivity. By (1), $F \backslash E$ and $E \cup F$ are Jordan measurable. Since $E \cup F = E \cup (F \backslash E)$ is a disjoint union, part (3) gives $m(E \cup F) = m(E) + m(F \backslash E)$, and since $F \backslash E \subset F$, part (4) gives $m(F \backslash E) \leq m(F)$. Therefore

$$m(E \cup F) \leq m(E) + m(F).$$

(6) Translation invariance. Fix $x \in \mathbb{R}^d$. For any set S and elementary A we have, by Exercise 1.1.1, that $A + x$ is elementary, and $m(A + x) = m(A)$ by translation invariance of elementary measure. Also $A \subset S$ if and only if $A + x \subset S + x$, and $S \subset B$ if and only if $S + x \subset B + x$, since $y \mapsto y + x$ is a bijection of $\mathbb{R}^d$.
Given $\varepsilon > 0$, take elementary $A \subset E \subset B$ with $m(B \backslash A) \leq \varepsilon$ (Exercise 1.1.5). Then $A + x \subset E + x \subset B + x$ are elementary and

$$(B + x) \backslash (A + x) = (B \backslash A) + x,$$

(because $y + x \in B + x$ and $y + x \notin A + x$ iff $y \in B$ and $y \notin A$), so $m\big((B + x) \backslash (A + x)\big) = m(B \backslash A) \leq \varepsilon$. By Exercise 1.1.5, $E + x$ is Jordan measurable.
Finally, $A \mapsto A + x$ is a bijection from the elementary subsets of E onto the elementary subsets of $E + x$ preserving elementary measure, so the two suprema agree:

$$m(E + x) = m_{*,(J)}(E + x) = \sup_{A \subset E,\ A \text{ elementary}} m(A + x) = \sup_{A \subset E,\ A \text{ elementary}} m(A) = m(E). \qquad \square$$

Problem 1.1.7. (Regions under graphs are Jordan measurable). Let B be a closed box in $\mathbb{R}^d$, and let $f : B \to \mathbb{R}$ be a continuous function.

1. Show that the graph $\{(x, f(x)) : x \in B\} \subset \mathbb{R}^{d+1}$ is Jordan measurable in $\mathbb{R}^{d+1}$ with Jordan measure zero. (Hint: on a compact metric space, continuous functions are uniformly continuous.)
2. Show that the set $\{(x, t) : x \in B;\ 0 \leq t \leq f(x)\} \subset \mathbb{R}^{d+1}$ is Jordan measurable.

Solution. Write $B = [a_1, b_1] \times \cdots \times [a_d, b_d]$. Being closed and bounded, B is a compact subset of $\mathbb{R}^d$, so f is bounded on B (say $|f| \leq M$) and, as the hint recalls, f is uniformly continuous on B.
A common construction: fine partitions of B.
Fix $\delta > 0$. For a positive integer n and each coordinate j, partition $[a_j, b_j]$ into n intervals of length $(b_j - a_j)/n$, taking them half-open on the right except for the last, which is closed:

$$[t_{j,0}, t_{j,1}),\ [t_{j,1}, t_{j,2}),\ \ldots,\ [t_{j,n-2}, t_{j,n-1}),\ [t_{j,n-1}, t_{j,n}], \qquad t_{j,r} := a_j + \tfrac{r}{n}(b_j - a_j)$$

(if $a_j = b_j$, simply use the single degenerate interval $[a_j, a_j]$). These are disjoint intervals with union $[a_j, b_j]$. Taking Cartesian products over j yields finitely many pairwise disjoint boxes $B_1, \ldots, B_N$

(discard the empty ones) whose union is B; each has side lengths $(b_j - a_j)/n$, hence diameter

$$\Big(\sum_{j=1}^{d} \big(\tfrac{b_j - a_j}{n}\big)^2\Big)^{1/2} = \frac{\operatorname{diam}(B)}{n}.$$

Choosing n large enough that $\operatorname{diam}(B)/n \le \delta$, we obtain a partition of B into finitely many disjoint nonempty boxes $B_1, \dots, B_N$ each of diameter at most δ, and by Lemma 1.1.2(ii),

$$\sum_{i=1}^{N} |B_i| = m(B) = |B|. \tag{$\dagger$}$$

Pick a point $x_i \in B_i$ for each i and set $c_i := f(x_i)$.
(1) The graph has Jordan measure zero.
Let $G := \{(x, f(x)) : x \in B\}$. It is bounded, since $G \subset B \times [-M, M]$.
Let $\varepsilon > 0$. By uniform continuity choose $\delta > 0$ such that $|f(x) - f(y)| \le \varepsilon$ whenever $x, y \in B$ satisfy $|x - y| \le \delta$. Take a partition $B_1, \dots, B_N$ as above with all diameters at most δ, and points $x_i \in B_i$, $c_i = f(x_i)$. For $x \in B_i$ we have $|x - x_i| \le \operatorname{diam}(B_i) \le \delta$, hence

$$c_i - \varepsilon \le f(x) \le c_i + \varepsilon.$$

Therefore the portion of the graph lying over B_i is contained in the box $B_i \times [c_i - \varepsilon, c_i + \varepsilon] \subset \mathbb{R}^{d+1}$, of volume $|B_i| \cdot 2\varepsilon$ (Exercise 1.1.4, or directly from Definition 1.1.1). Consequently

$$G \subset S := \bigcup_{i=1}^{N} B_i \times [c_i - \varepsilon, c_i + \varepsilon],$$

an elementary subset of $\mathbb{R}^{d+1}$, and by finite subadditivity of elementary measure together with ($\dagger$),

$$m(S) \le \sum_{i=1}^{N} 2\varepsilon |B_i| = 2\varepsilon |B|.$$

Since S is a competitor in the infimum defining the Jordan outer measure, $m^{*,(J)}(G) \le 2\varepsilon|B|$. As $\varepsilon > 0$ was arbitrary and $|B|$ is a fixed finite constant, $m^{*,(J)}(G) = 0$. On the other hand $\emptyset \subset G$ is elementary, so $m_{*,(J)}(G) \ge 0$, while $m_{*,(J)}(G) \le m^{*,(J)}(G)$ (fact (a) in the solution to Exercise 1.1.5). Hence

$$m_{*,(J)}(G) = m^{*,(J)}(G) = 0,$$

so G is Jordan measurable with $m(G) = 0$.
(2) The region under the graph is Jordan measurable.
Let

$$E := \{(x, t) : x \in B,\ 0 \le t \le f(x)\} \subset \mathbb{R}^{d+1}.$$

(Note that for those x with $f(x) < 0$ the corresponding slice is empty, so E is the region between the graph and the hyperplane $t = 0$ over the set where $f \ge 0$.) It is bounded: $E \subset B \times [0, M]$.
Let $\varepsilon > 0$, choose δ by uniform continuity as in part (1), and take the partition $B_1, \dots, B_N$ of B into disjoint boxes of diameter at most δ, with $x_i \in B_i$ and $c_i = f(x_i)$, so that

$$c_i - \varepsilon \le f(x) \le c_i + \varepsilon \qquad \text{for all } x \in B_i. \tag{$*$}$$

Define the index sets $I := \{i : c_i \ge \varepsilon\}$ and $I' := \{i : c_i \ge -\varepsilon\}$ (so $I \subset I'$), and put

$$L := \bigcup_{i \in I} B_i \times [0, c_i - \varepsilon], \qquad U := \bigcup_{i \in I'} B_i \times [0, c_i + \varepsilon].$$

Both are finite unions of boxes in $\mathbb{R}^{d+1}$ (the intervals $[0, c_i - \varepsilon]$ for $i \in I$ and $[0, c_i + \varepsilon]$ for $i \in I'$ are genuine intervals since their right endpoints are ≥ 0), hence elementary; and $L \subset U$ since $I \subset I'$ and $[0, c_i - \varepsilon] \subset [0, c_i + \varepsilon]$.

$L \subset E$. Let $(x,t) \in B_i \times [0, c_i - \varepsilon]$ with $i \in I$. Then $x \in B$, $t \geq 0$, and by $(*)$, $t \leq c_i - \varepsilon \leq f(x)$. So $(x,t) \in E$.
$E \subset U$. Let $(x,t) \in E$. Since the B_i partition B, there is a (unique) i with $x \in B_i$, and by $(*)$, $0 \leq t \leq f(x) \leq c_i + \varepsilon$. In particular $c_i + \varepsilon \geq 0$, i.e. $i \in I'$, and $(x,t) \in B_i \times [0, c_i + \varepsilon] \subset U$.
Estimating $m(U\backslash L)$. Because the B_i are pairwise disjoint, so are the boxes $B_i \times [0, c_i \pm \varepsilon]$ appearing in L and in U; hence by Lemma 1.1.2(ii),

$$m(L) = \sum_{i \in I} |B_i|(c_i - \varepsilon), \qquad m(U) = \sum_{i \in I'} |B_i|(c_i + \varepsilon),$$

using $|B_i \times [0,s]| = |B_i| \cdot s$. Since $L \subset U$ are elementary, finite additivity gives $m(U\backslash L) = m(U) - m(L)$, so

$$m(U\backslash L) = \sum_{i \in I} |B_i| \big[(c_i + \varepsilon) - (c_i - \varepsilon)\big] \; + \sum_{i \in I' \backslash I} |B_i|(c_i + \varepsilon).$$

The first sum equals $\sum_{i \in I} 2\varepsilon |B_i|$. For $i \in I'\backslash I$ we have $-\varepsilon \leq c_i < \varepsilon$, hence $0 \leq c_i + \varepsilon < 2\varepsilon$, so the second sum is at most $\sum_{i \in I'\backslash I} 2\varepsilon |B_i|$. Adding and using $(\dagger)$,

$$m(U\backslash L) \leq 2\varepsilon \sum_{i=1}^{N} |B_i| = 2\varepsilon |B|.$$

Thus for every $\varepsilon > 0$ we have produced elementary sets $L \subset E \subset U$ with $m(U\backslash L) \leq 2|B|\varepsilon$. Given an arbitrary $\eta > 0$, applying the above with $\varepsilon := \eta/(2|B| + 1)$ yields elementary $L \subset E \subset U$ with $m(U\backslash L) \leq \eta$. By Exercise 1.1.5 (criterion (2)), E is Jordan measurable. $\square$

Problem 1.1.8. Let A, B, C be three points in $\mathbb{R}^2$.
1. Show that the solid triangle with vertices A, B, C is Jordan measurable.
2. Show that the Jordan measure of the solid triangle is equal to $\frac{1}{2}|(B - A) \wedge (C - A)|$, where $|(a,b) \wedge (c,d)| := |ad - bc|$.

(Hint: It may help to first do the case when one of the edges, say AB, is horizontal.)

Solution. Throughout, the solid triangle with vertices A, B, C means the convex hull

$$T(A,B,C) := \{\alpha A + \beta B + \gamma C : \alpha, \beta, \gamma \geq 0, \; \alpha + \beta + \gamma = 1\},$$

and we abbreviate

$$\Delta(A,B,C) := (B - A) \wedge (C - A) = (b_1 - a_1)(c_2 - a_2) - (b_2 - a_2)(c_1 - a_1)$$

for $A = (a_1, a_2)$, $B = (b_1, b_2)$, $C = (c_1, c_2)$. Expanding,

$$\Delta(A,B,C) = (b_1 c_2 - b_2 c_1) - (a_1 c_2 - a_2 c_1) + (a_1 b_2 - a_2 b_1),$$

which is (up to sign) the 3×3 determinant with columns $(1, a_1, a_2)$, $(1, b_1, b_2)$, $(1, c_1, c_2)$. Hence Δ is alternating in its three arguments, so $|\Delta(A,B,C)|$ is unchanged by permuting A, B, C; and Δ visibly depends only on the differences $B - A$, $C - A$, so it is invariant under translating all three points by a common vector. Both T and $|\Delta|$ are also unchanged in the appropriate sense by the reflection $\rho(x,y) := (x, -y)$: $\rho(T(A,B,C)) = T(\rho A, \rho B, \rho C)$ and $|\Delta(\rho A, \rho B, \rho C)| = |\Delta(A,B,C)|$. Since ρ maps boxes to boxes of the same volume, it maps elementary sets to elementary sets of the same measure, and therefore preserves Jordan inner and outer measure; so it suffices to prove the claims after applying ρ. The same remark applies to translations, by Exercise 1.1.6(6).
Step 0: the degenerate case. Suppose A, B, C are collinear, equivalently $B - A$ and $C - A$ are linearly dependent, equivalently $\Delta(A,B,C) = 0$. Then T is the convex hull of three collinear points, which is a (possibly degenerate) closed line segment S. If S is vertical or a point, it is contained in a box $\{x_0\} \times [c,d]$ of volume 0; that box is elementary of measure 0, so $m^{*,(J)}(S) = 0$ and S is Jordan measurable with $m(S) = 0$. If S is not vertical, then $S = \{(x, f(x)) : x \in [\alpha, \beta]\}$ for an affine (hence

continuous) function f on the closed box $[\alpha,\beta]\subset\mathbb{R}^1$, so by Exercise 1.1.7(1) S is Jordan measurable with $m(S)=0$. In both cases $m(T)=0=\frac{1}{2}|\Delta(A,B,C)|$, as required. We may therefore assume from now on that A,B,C are not collinear.

Step 1: the case of a horizontal edge. Suppose two of the vertices have the same y-coordinate. Translating (which changes neither T's measurability and measure, by Exercise 1.1.6(6), nor $|\Delta|$), and using the permutation symmetry of $|\Delta|$, we may assume the vertices are

$$P=(p,0),\qquad Q=(q,0),\qquad R=(r,s)$$

with $p<q$ (the two base vertices are distinct, else the three points would be collinear) and $s\neq 0$; applying ρ if necessary, assume $s>0$.

First we describe $T:=T(P,Q,R)$ by horizontal slices. A point $z=\alpha P+\beta Q+\gamma R$ with $\alpha,\beta,\gamma\geq 0$ summing to 1 has y-coordinate γs; writing $t:=\gamma\in[0,1]$, its x-coordinate is $\alpha p+\beta q+tr$ with $\alpha+\beta=1-t$ and $\alpha,\beta\geq 0$, which ranges over exactly the interval $[(1-t)p+tr,\ (1-t)q+tr]$. Hence, setting

$$\ell(y):=(1-y/s)p+(y/s)r,\qquad u(y):=(1-y/s)q+(y/s)r,$$

we have

$$T=\{(x,y):0\leq y\leq s,\ \ell(y)\leq x\leq u(y)\},$$

and $u(y)-\ell(y)=(1-y/s)(q-p)\geq 0$ on $[0,s]$. Note ℓ,u are affine in y with slopes $(r-p)/s$ and $(r-q)/s$; put $K:=|r-p|+|r-q|$, so that over any y-interval of length s/N each of ℓ,u varies by at most K/N in total (i.e. the sum of the two variations is at most K/N).

Fix $N\geq 1$ and set $y_j:=js/N$ for $0\leq j\leq N$. For $1\leq j\leq N$ write

$$M_j:=\max_{[y_{j-1},y_j]}u,\quad m_j:=\min_{[y_{j-1},y_j]}u,\quad M_j':=\max_{[y_{j-1},y_j]}\ell,\quad m_j':=\min_{[y_{j-1},y_j]}\ell$$

(all attained, since ℓ,u are continuous on a compact interval). Define the boxes

$$B_j^-:=[M_j',m_j]\times[y_{j-1},y_j],\qquad B_j^+:=[m_j',M_j]\times[y_{j-1},y_j],$$

with the convention that B_j^- is to be read as the empty set (which is elementary) when $m_j<M_j'$, so that $[M_j',m_j]$ is only ever formed when it is a genuine interval, and the elementary sets $A_N:=\bigcup_{j=1}^N B_j^-$, $C_N:=\bigcup_{j=1}^N B_j^+$.

If $(x,y)\in B_j^-$ then $y\in[y_{j-1},y_j]\subset[0,s]$ and $\ell(y)\leq M_j'\leq x\leq m_j\leq u(y)$, so $(x,y)\in T$; thus $A_N\subset T$. Conversely if $(x,y)\in T$, pick j with $y\in[y_{j-1},y_j]$; then $m_j'\leq\ell(y)\leq x\leq u(y)\leq M_j$, so $(x,y)\in B_j^+$; thus $T\subset C_N$. The boxes $B_1^\pm,\dots,B_N^\pm$ are not quite disjoint, since consecutive slabs share the horizontal line $y=y_j$; but replacing each $[y_{j-1},y_j]$ by $[y_{j-1},y_j)$ for $j<N$ produces genuinely disjoint boxes of the same volumes, whose union differs from A_N (resp. C_N) by a subset of the $N-1$ segments $\{(x,y_j):x\in\mathbb{R},\ |x|\leq M\}$ for a suitable M, each of which is a box of volume 0. Hence by finite additivity together with finite subadditivity and monotonicity of elementary measure,

$$m(A_N)=\frac{s}{N}\sum_{j=1}^N(m_j-M_j')_+,\qquad m(C_N)=\frac{s}{N}\sum_{j=1}^N(M_j-m_j'),$$

where $(\cdot)_+$ denotes the positive part.

We estimate the gap slab by slab. If $m_j-M_j'\geq 0$ then

$$(M_j-m_j')-(m_j-M_j')=(M_j-m_j)+(M_j'-m_j')\leq K/N,$$

by the choice of K. If $m_j-M_j'<0$ then $(M_j-m_j')-0=(m_j-M_j')+(M_j-m_j)+(M_j'-m_j')<K/N$ as well. Hence in all cases the j-th term of $m(C_N)$ exceeds that of $m(A_N)$ by at most $\frac{s}{N}\cdot\frac{K}{N}$, and so

$$0\leq m(C_N)-m(A_N)\leq\frac{sK}{N}.$$

Since $A_N \subset T \subset C_N$ with A_N, C_N elementary, the definition of Jordan inner and outer measure gives

$$0 \le m^{*,(J)}(T) - m_{*,(J)}(T) \le m(C_N) - m(A_N) \le \frac{sK}{N}$$

for every N; letting $N \to \infty$ shows $m^{*,(J)}(T) = m_{*,(J)}(T)$, i.e. T is Jordan measurable, and moreover $m(A_N) \le m(T) \le m(C_N)$ for every N.
To compute $m(T)$, note that for each j we have $M_j - m'_j \ge u(y_j) - \ell(y_j)$ and, by the variation bound again, $M_j - m'_j \le u(y_j) - \ell(y_j) + K/N$. Therefore

$$\frac{s}{N}\sum_{j=1}^{N}\big(u(y_j)-\ell(y_j)\big) \le m(C_N) \le \frac{s}{N}\sum_{j=1}^{N}\big(u(y_j)-\ell(y_j)\big) + \frac{sK}{N}.$$

Now $u(y_j) - \ell(y_j) = (1 - j/N)(q-p)$, so

$$\frac{s}{N}\sum_{j=1}^{N}\big(u(y_j)-\ell(y_j)\big) = \frac{s(q-p)}{N}\sum_{j=1}^{N}\Big(1-\frac{j}{N}\Big) = \frac{s(q-p)}{N}\cdot\frac{N-1}{2} \xrightarrow[N\to\infty]{} \frac{s(q-p)}{2}.$$

Hence $m(C_N) \to \frac{1}{2}s(q-p)$, and since $m(C_N) - \frac{sK}{N} \le m(A_N) \le m(T) \le m(C_N)$, we conclude

$$m(T) = \tfrac{1}{2}s(q-p) = \tfrac{1}{2}\,(\text{base length})\cdot(\text{height}).$$

Finally $\Delta(P,Q,R) = (q-p,0)\wedge(r-p,s) = (q-p)s - 0\cdot(r-p) = (q-p)s$, so indeed $m(T) = \frac{1}{2}|\Delta(P,Q,R)|$. This settles both parts whenever some edge of the triangle is horizontal.
Step 2: reduction of the general case to Step 1. Let A, B, C be non-collinear and relabel them as P_1, P_2, P_3 with y-coordinates $y_1 \le y_2 \le y_3$; this is legitimate since T does not depend on the labelling and $|\Delta|$ is permutation-invariant. Non-collinearity forces $y_1 < y_3$. If $y_1 = y_2$ or $y_2 = y_3$, an edge is horizontal and Step 1 applies. So assume $y_1 < y_2 < y_3$.
Set $\lambda := \frac{y_2-y_1}{y_3-y_1} \in (0,1)$ and $Q := P_1 + \lambda(P_3 - P_1) = (1-\lambda)P_1 + \lambda P_3$, a point of the segment P_1P_3 with y-coordinate y_2. Since P_2 is not on the line through P_1, P_3, we have $P_2 \ne Q$, and P_2Q is a horizontal segment. Let $T_1 := T(P_1, P_2, Q)$ and $T_2 := T(P_2, P_3, Q)$. Each has a horizontal edge P_2Q and a third vertex off that horizontal line ($y_1 < y_2$, resp. $y_3 > y_2$), so Step 1 applies to each: T_1, T_2 are Jordan measurable with

$$m(T_1) = \tfrac{1}{2}\,h\,(y_2 - y_1), \qquad m(T_2) = \tfrac{1}{2}\,h\,(y_3 - y_2), \qquad h := |x(P_2) - x(Q)| > 0.$$

I claim $T_1 = T \cap \{y \le y_2\}$ and $T_2 = T \cap \{y \ge y_2\}$, where $T := T(P_1, P_2, P_3)$. Since $Q \in T$ and T is convex, $T_1 \subset T$; all three vertices of T_1 have $y \le y_2$ and the half-plane $\{y \le y_2\}$ is convex, so $T_1 \subset T \cap \{y \le y_2\}$. Conversely, let $z = \alpha P_1 + \beta P_2 + \gamma P_3 \in T$ ($\alpha, \beta, \gamma \ge 0$, summing to 1) with $y(z) \le y_2$. From $P_3 = \lambda^{-1}\big(Q - (1-\lambda)P_1\big)$ we get

$$z = \Big(\alpha - \tfrac{\gamma(1-\lambda)}{\lambda}\Big)P_1 + \beta P_2 + \tfrac{\gamma}{\lambda}Q,$$

and the three coefficients sum to 1. The last two are ≥ 0; for the first, the hypothesis $\alpha y_1 + \beta y_2 + \gamma y_3 \le y_2 = (\alpha+\beta+\gamma)y_2$ rearranges to $\gamma(y_3 - y_2) \le \alpha(y_2 - y_1)$, and since $\frac{1-\lambda}{\lambda} = \frac{y_3-y_2}{y_2-y_1}$ this says exactly $\frac{\gamma(1-\lambda)}{\lambda} \le \alpha$. Hence $z \in T_1$. The proof for T_2 is symmetric.
Consequently $T = T_1 \cup T_2$, and $T_1 \cap T_2 \subset T \cap \{y = y_2\}$, which is a horizontal line segment and hence a Jordan null set by Step 0. By Exercise 1.1.6(1) T is Jordan measurable, and by Exercise 1.1.6(3) applied to the disjoint decomposition $T = T_1 \cup (T_2 \setminus T_1)$ together with $m(T_2 \setminus T_1) = m(T_2) - m(T_1 \cap T_2) = m(T_2)$ (using Exercise 1.1.6(1),(3) and $m(T_1 \cap T_2) = 0$, which follows from monotonicity, Exercise 1.1.6(4), and Step 0),

$$m(T) = m(T_1) + m(T_2) = \tfrac{1}{2}h\big((y_2 - y_1) + (y_3 - y_2)\big) = \tfrac{1}{2}h(y_3 - y_1).$$

It remains to identify this with $\frac{1}{2}|\Delta(P_1, P_2, P_3)|$. Since $Q - P_1 = \lambda(P_3 - P_1)$ and the wedge is bilinear,

$$\Delta(P_1, P_2, Q) = (P_2 - P_1) \wedge (Q - P_1) = \lambda\,(P_2 - P_1) \wedge (P_3 - P_1) = \lambda\,\Delta(P_1, P_2, P_3).$$

By Step 1 applied to T_1 (whose horizontal edge is P_2Q of length h and whose height is $y_2 - y_1$), $|\Delta(P_1, P_2, Q)| = 2m(T_1) = h(y_2 - y_1)$. Therefore

$$|\Delta(P_1, P_2, P_3)| = \frac{1}{\lambda} h(y_2 - y_1) = \frac{y_3 - y_1}{y_2 - y_1} h(y_2 - y_1) = h(y_3 - y_1) = 2m(T).$$

Thus $m(T) = \frac{1}{2}|(B - A) \wedge (C - A)|$ in all cases, completing both parts.

Problem 1.1.9. Show that every compact convex polytope in $\mathbb{R}^d$ is Jordan measurable.
[A closed convex polytope is a subset of $\mathbb{R}^d$ formed by intersecting together finitely many closed half-spaces of the form $\{x \in \mathbb{R}^d : x \cdot v \leq c\}$, where $v \in \mathbb{R}^d$, $c \in \mathbb{R}$, and $\cdot$ denotes the usual dot product on $\mathbb{R}^d$. A compact convex polytope is a closed convex polytope which is also bounded.]

Solution. We prove the stronger statement that every compact convex subset C of $\mathbb{R}^d$ is Jordan measurable. A compact convex polytope is an intersection of finitely many closed half-spaces, hence closed and convex, and it is bounded by hypothesis, hence compact; so it is a set of this type and the claim follows.
We first record two elementary facts.
Scaling of elementary measure. For $\theta > 0$ and a box $B = I_1 \times \cdots \times I_d$, the dilate $\theta B := \{\theta x : x \in B\}$ is again a box, namely $(\theta I_1) \times \cdots \times (\theta I_d)$, and $|\theta B| = \theta^d |B|$ since each interval has its length scaled by θ. If E is elementary, write $E = B_1 \cup \cdots \cup B_k$ as a finite union of disjoint boxes (Lemma 1.1.2(i)); then $\theta E = \theta B_1 \cup \cdots \cup \theta B_k$ is again such a union of disjoint boxes, so θE is elementary and

$$m(\theta E) = \sum_{i=1}^{k} |\theta B_i| = \theta^d \sum_{i=1}^{k} |B_i| = \theta^d m(E).$$

Bounded subsets of hyperplanes are Jordan null. Let $H = \{x \in \mathbb{R}^d : x \cdot v = c\}$ with $v \neq 0$, and let $S \subset H$ be bounded. Pick j with $v_j \neq 0$. Permuting the coordinates (a permutation of coordinates carries boxes to boxes of the same volume, hence elementary sets to elementary sets of the same measure, and so preserves Jordan inner and outer measure), we may assume $j = d$. Then H is the graph $\{(x, f(x)) : x \in \mathbb{R}^{d-1}\}$ of the affine, hence continuous, function $f(x) := (c - \sum_{i<d} x_i v_i)/v_d$. Since S is bounded, $S \subset \{(x, f(x)) : x \in B\}$ for a sufficiently large closed box $B \subset \mathbb{R}^{d-1}$, and by Exercise 1.1.7(1) that graph is Jordan measurable with Jordan measure zero. By monotonicity of Jordan outer measure, $m^{*,(J)}(S) = 0$, so S is Jordan measurable with $m(S) = 0$. (For $d = 1$ a hyperplane is a point, which is a box of volume zero, so the conclusion is immediate.)
Now let $C \subset \mathbb{R}^d$ be compact and convex. If $C = \emptyset$ there is nothing to prove, so assume $C \neq \emptyset$.
Case 1: C has empty interior. I claim C is contained in an affine hyperplane. Indeed, if C contained $d+1$ affinely independent points $p_0, \ldots, p_d$, then by convexity C would contain their convex hull, the simplex $\{p_0 + \sum_{i=1}^{d} t_i(p_i - p_0) : t_i \geq 0, \sum t_i \leq 1\}$; as the vectors $p_i - p_0$ form a basis of $\mathbb{R}^d$, the linear map sending the standard basis to this basis is a homeomorphism of $\mathbb{R}^d$, and it carries the (nonempty) interior of the standard simplex onto a nonempty open subset of that simplex, contradicting $\mathrm{int}(C) = \emptyset$. So C contains at most d affinely independent points, hence is contained in an affine subspace of dimension at most $d-1$, and therefore in some affine hyperplane H. Since C is bounded, the previous paragraph gives that C is Jordan measurable (with measure zero).
Case 2: C has nonempty interior. By translation invariance of Jordan measurability (Exercise 1.1.6(6); translations carry boxes to boxes of equal volume) we may assume $0 \in \mathrm{int}(C)$, say $B(0, r) \subset C$ for some $r > 0$.
Sub-claim: for every $\theta \in (0, 1)$ we have $\theta C \subset \mathrm{int}(C)$. Indeed, let $x \in C$ and let w satisfy $|w - \theta x| < (1 - \theta)r$. Write $w = \theta x + (1 - \theta)z$ with $z := (w - \theta x)/(1 - \theta)$, so $|z| < r$ and hence $z \in B(0, r) \subset C$. By convexity $w \in C$. Thus the open ball of radius $(1 - \theta)r$ about θx lies in C, i.e. $\theta x \in \mathrm{int}(C)$.
Fix $\theta \in (0, 1)$. The set θC is compact (a continuous image of the compact set C) and is contained in the open set $\mathrm{int}(C)$. For each $y \in \theta C$ choose $\varepsilon_y > 0$ so small that the closed cube Q_y of side $2\varepsilon_y$ centred at y is contained in $\mathrm{int}(C)$. The open cubes $\mathrm{int}(Q_y)$ cover the compact set θC, so finitely many suffice, say those centred at $y_1, \ldots, y_k$. Put $A := Q_{y_1} \cup \cdots \cup Q_{y_k}$, a finite union of boxes and hence an

elementary set, with

$$\theta C \subset A \subset C.$$

From $A \subset C$ and the definition of Jordan inner measure, $m(A) \leq m_{*,(J)}(C)$. From $\theta C \subset A$ we get $C \subset \theta^{-1}A$, and $\theta^{-1}A$ is elementary with $m(\theta^{-1}A) = \theta^{-d}m(A)$ by the scaling fact above. Hence, by the definition of Jordan outer measure,

$$m^{*,(J)}(C) \leq m(\theta^{-1}A) = \theta^{-d}m(A) \leq \theta^{-d}m_{*,(J)}(C).$$

Since C is bounded it lies in some box, so $m_{*,(J)}(C) \leq m^{*,(J)}(C) < \infty$. Letting $\theta \to 1^-$ in the displayed inequality gives $m^{*,(J)}(C) \leq m_{*,(J)}(C)$. The reverse inequality $m_{*,(J)}(C) \leq m^{*,(J)}(C)$ holds for every bounded set (if $A \subset C \subset B$ with A, B elementary then $m(A) \leq m(B)$ by monotonicity of elementary measure). Therefore $m_{*,(J)}(C) = m^{*,(J)}(C)$ and C is Jordan measurable.
In both cases C is Jordan measurable, and in particular every compact convex polytope in $\mathbb{R}^d$ is Jordan measurable.

Problem 1.1.10.

1. Show that all open and closed Euclidean balls $B(x,r) := \{y \in \mathbb{R}^d : |y-x| < r\}$, $\overline{B(x,r)} := \{y \in \mathbb{R}^d : |y-x| \leq r\}$ in $\mathbb{R}^d$ are Jordan measurable, with Jordan measure $c_d r^d$ for some constant $c_d > 0$ depending only on d.
2. Establish the crude bounds
$$\left(\frac{2}{\sqrt{d}}\right)^d \leq c_d \leq 2^d.$$

(An exact formula for c_d is $c_d = \frac{1}{d}\omega_d$, where $\omega_d := \frac{2\pi^{d/2}}{\Gamma(d/2)}$ is the volume of the unit sphere $S^{d-1} \subset \mathbb{R}^d$ and Γ is the Gamma function, but we will not derive this formula here.)

Solution. Part 1. We may assume $r > 0$; for $r = 0$ the open ball is empty and the closed ball is the single point x, which is a box of volume zero, so both are Jordan measurable with measure $0 = c_d \cdot 0^d$.
Dilation. For $\theta > 0$ and $E \subset \mathbb{R}^d$ write $\theta E := \{\theta y : y \in E\}$. As noted in the solution to Exercise 1.1.9, if E is elementary then so is θE, with $m(\theta E) = \theta^d m(E)$ (a box $I_1 \times \cdots \times I_d$ dilates to the box $\theta I_1 \times \cdots \times \theta I_d$, whose volume is θ^d times the original, and a partition of E into disjoint boxes dilates to such a partition of θE). The map $E \mapsto \theta E$ is a bijection of the elementary sets onto themselves (with inverse $F \mapsto \theta^{-1}F$) and preserves inclusions in both directions. Hence for any bounded $E \subset \mathbb{R}^d$,

$$m_{*,(J)}(\theta E) = \theta^d\, m_{*,(J)}(E), \qquad m^{*,(J)}(\theta E) = \theta^d\, m^{*,(J)}(E),$$

since $A \subset E \subset B$ with A, B elementary holds if and only if $\theta A \subset \theta E \subset \theta B$ with $\theta A, \theta B$ elementary. In particular E is Jordan measurable if and only if θE is, in which case $m(\theta E) = \theta^d m(E)$.
Closed balls. The closed ball $\overline{B(0,1)}$ is bounded and closed, hence compact, and it is convex (if $|y|, |z| \leq 1$ and $t \in [0,1]$ then $|ty + (1-t)z| \leq t|y| + (1-t)|z| \leq 1$). The solution to Exercise 1.1.9 established the stronger statement that every compact convex subset of $\mathbb{R}^d$ is Jordan measurable, so $\overline{B(0,1)}$ is Jordan measurable. Define

$$c_d := m\big(\overline{B(0,1)}\big).$$

Since $\overline{B(0,r)} = r\,\overline{B(0,1)}$, the dilation identity gives that $\overline{B(0,r)}$ is Jordan measurable with $m(\overline{B(0,r)}) = c_d r^d$; and since $\overline{B(x,r)} = \overline{B(0,r)} + x$, translation invariance (Exercise 1.1.6(6)) gives that $\overline{B(x,r)}$ is Jordan measurable with

$$m\big(\overline{B(x,r)}\big) = c_d r^d.$$

Positivity of c_d. If $|y_i| \leq 1/\sqrt{d}$ for all i then $|y|^2 = \sum_i y_i^2 \leq d \cdot \frac{1}{d} = 1$, so the cube $[-1/\sqrt{d}, 1/\sqrt{d}]^d$ is contained in $\overline{B(0,1)}$. This cube is elementary of measure $(2/\sqrt{d})^d$, so by the definition of Jordan inner measure,

$$c_d = m_{*,(J)}\big(\overline{B(0,1)}\big) \geq \Big(\frac{2}{\sqrt{d}}\Big)^d > 0.$$

Open balls. Again by translation invariance it suffices to treat $B(0,r)$. For $0 < \theta < 1$ we have $\overline{B(0,\theta r)} \subset B(0,r) \subset \overline{B(0,r)}$. Jordan inner and outer measure are monotone (an elementary subset of the smaller set is an elementary subset of the larger, and an elementary superset of the larger is one of the smaller), so

$$c_d(\theta r)^d = m_{*,(J)}\big(\overline{B(0,\theta r)}\big) \le m_{*,(J)}\big(B(0,r)\big) \le m^{*,(J)}\big(B(0,r)\big) \le m^{*,(J)}\big(\overline{B(0,r)}\big) = c_d r^d.$$

Letting $\theta \to 1^-$, the extreme left-hand side tends to $c_d r^d$, so all four quantities equal $c_d r^d$. In particular $m_{*,(J)}(B(0,r)) = m^{*,(J)}(B(0,r))$, so $B(0,r)$ is Jordan measurable with $m(B(0,r)) = c_d r^d$, and hence so is $B(x,r)$ with the same measure.
Part 2. The lower bound $c_d \ge (2/\sqrt{d})^d$ was proved above from the inscribed cube $[-1/\sqrt{d}, 1/\sqrt{d}]^d \subset \overline{B(0,1)}$.
For the upper bound, if $|y| \le 1$ then $|y_i| \le |y| \le 1$ for each i, so $\overline{B(0,1)} \subset [-1,1]^d$. The latter is elementary of measure 2^d, so by the definition of Jordan outer measure

$$c_d = m^{*,(J)}\big(\overline{B(0,1)}\big) \le 2^d.$$

Combining,

$$\Big(\frac{2}{\sqrt{d}}\Big)^d \le c_d \le 2^d,$$

as required.

Problem 1.1.11. This exercise assumes familiarity with linear algebra. Let $L : \mathbb{R}^d \to \mathbb{R}^d$ be a linear transformation.

1. Show that there exists a non-negative real number D such that $m(L(E)) = D\, m(E)$ for every elementary set E (note from previous exercises that $L(E)$ is Jordan measurable). (Hint: apply Exercise 1.1.3 to the map $E \mapsto m(L(E))$.)
2. Show that if E is Jordan measurable, then $L(E)$ is also, and $m(L(E)) = D\, m(E)$.
3. Show that $D = |\det L|$. (Hint: Work first with the case when L is an elementary transformation, using Gaussian elimination. Alternatively, work with the cases when L is a diagonal transformation or an orthogonal transformation, using the unit ball in the latter case, and use the polar decomposition.)

Solution. We use repeatedly the fact, established in the solution to Exercise 1.1.9, that a bounded subset of an affine hyperplane of $\mathbb{R}^d$ is Jordan measurable with measure zero, and that every compact convex subset of $\mathbb{R}^d$ is Jordan measurable.
Part 1. First suppose L is not invertible. Then $L(\mathbb{R}^d)$ is a proper linear subspace, hence contained in a hyperplane through the origin. For elementary (so bounded) E, the set $L(E)$ is bounded and lies in that hyperplane, hence is Jordan measurable with $m(L(E)) = 0$. Thus $D := 0$ works.
Now suppose L is invertible. We first check that $L(E)$ is Jordan measurable for every elementary E. Let B be a box, $\overline{B}$ its closure and B° its interior. The set $\overline{B}$ is a compact convex polytope, so $L(\overline{B})$ is compact (continuous image of a compact set) and convex (linear images of convex sets are convex), and hence Jordan measurable by Exercise 1.1.9 (in the strengthened form proved there). Next, $\overline{B} \setminus B^\circ$ is contained in the union of the $2d$ faces of $\overline{B}$, each of which lies in a hyperplane; so $L(\overline{B} \setminus B^\circ)$ is a bounded set contained in a union of $2d$ hyperplanes (images of hyperplanes under the invertible map L are hyperplanes), and therefore has Jordan outer measure zero by finite subadditivity of Jordan outer measure. Since

$$L(\overline{B}) \setminus L(B) \subset L(\overline{B} \setminus B) \subset L(\overline{B} \setminus B^\circ)$$

(if $y = L(z)$ with $z \in \overline{B}$ and $y \notin L(B)$, then $z \notin B$), the set $N := L(\overline{B}) \setminus L(B)$ has Jordan outer measure zero, hence is Jordan measurable of measure zero. Consequently $L(B) = L(\overline{B}) \setminus N$ is Jordan measurable by Exercise 1.1.6(1), with $m(L(B)) = m(L(\overline{B}))$ by Exercise 1.1.6(3) and (4). Finally, an elementary set E is a finite union of boxes, and $L(E)$ is then the corresponding finite union of the Jordan measurable sets $L(B_i)$, hence Jordan measurable by Exercise 1.1.6(1).

We may therefore define $m' : \mathcal{E}(\mathbb{R}^d) \to \mathbb{R}^+$ by $m'(E) := m(L(E))$, and we verify the three hypotheses of Exercise 1.1.3.
Non-negativity is Exercise 1.1.6(2). For finite additivity, let E, F be disjoint elementary sets; since L is injective, $L(E)$ and $L(F)$ are disjoint and $L(E \cup F) = L(E) \cup L(F)$, so $m'(E \cup F) = m(L(E)) + m(L(F)) = m'(E) + m'(F)$ by Exercise 1.1.6(3). For translation invariance, $L(E + x) = L(E) + L(x)$ by linearity, so $m'(E + x) = m(L(E) + L(x)) = m(L(E)) = m'(E)$ by Exercise 1.1.6(6).
By Exercise 1.1.3 there is a constant $D \in \mathbb{R}^+$ with $m'(E) = D\, m(E)$, i.e. $m(L(E)) = D\, m(E)$, for every elementary set E. This proves Part 1. Note that taking $E := [0,1)^d$ gives $D = m(L([0,1)^d))$.
Part 2. If L is not invertible, then $D = 0$ and, for bounded E, $L(E)$ is a bounded subset of a hyperplane, hence Jordan measurable with $m(L(E)) = 0 = D\, m(E)$.
Suppose L is invertible and E is Jordan measurable. Let $\varepsilon > 0$. By Exercise 1.1.5 there are elementary sets $A \subset E \subset B$ with $m(B \setminus A) \le \varepsilon$. Applying L (a bijection) gives

$$L(A) \subset L(E) \subset L(B), \qquad L(B) \setminus L(A) = L(B \setminus A),$$

and all three of $L(A), L(B), L(B \setminus A)$ are Jordan measurable by Part 1, with

$$m(L(B)) - m(L(A)) = m(L(B) \setminus L(A)) = m(L(B \setminus A)) = D\, m(B \setminus A) \le D\varepsilon,$$

using Exercise 1.1.6(1),(3) for the first equality. By monotonicity of Jordan inner and outer measure,

$$m(L(A)) = m_{*,(J)}(L(A)) \le m_{*,(J)}(L(E)) \le m^{*,(J)}(L(E)) \le m^{*,(J)}(L(B)) = m(L(B)),$$

so $m^{*,(J)}(L(E)) - m_{*,(J)}(L(E)) \le D\varepsilon$. As $\varepsilon > 0$ was arbitrary, $L(E)$ is Jordan measurable. Moreover the same chain gives

$$D\, m(A) = m(L(A)) \le m(L(E)) \le m(L(B)) = D\, m(B),$$

while $m(A) \le m(E) \le m(B)$ and $m(B) - m(A) = m(B \setminus A) \le \varepsilon$. Hence $|m(L(E)) - D\, m(E)| \le D\varepsilon$ for every $\varepsilon > 0$, so $m(L(E)) = D\, m(E)$.
Part 3. Write $D(L)$ for the constant produced above, so $D(L) = m(L([0,1)^d))$.
Multiplicativity. For linear L_1, L_2 and any Jordan measurable E, Part 2 gives that $L_2(E)$ is Jordan measurable and then that $L_1(L_2(E))$ is; taking $E := [0,1)^d$,

$$D(L_1 L_2) = m\big(L_1(L_2([0,1)^d))\big) = D(L_1)\, m\big(L_2([0,1)^d)\big) = D(L_1)D(L_2).$$

Also $D(\mathrm{id}) = m([0,1)^d) = 1$.
If L is not invertible then $D(L) = 0 = |\det L|$, so assume L is invertible. By Gaussian elimination, every invertible $d \times d$ matrix is a finite product of elementary matrices, which are of the following three types. Since both $L \mapsto D(L)$ and $L \mapsto |\det L|$ are multiplicative, it suffices to check $D = |\det|$ on each type.
Type 1 (scaling): S multiplies the i-th coordinate by some $c \neq 0$ and fixes the others. Then $S([0,1)^d)$ is the box whose i-th side is $[0, c)$ if $c > 0$ and $(c, 0]$ if $c < 0$, and whose other sides are $[0, 1)$; its volume is $|c|$. Hence $D(S) = |c| = |\det S|$.
Type 2 (transposition): P swaps coordinates i and j. Then $P([0,1)^d) = [0,1)^d$, so $D(P) = 1 = |-1| = |\det P|$.
Type 3 (shear): $T = T_{ij}(c)$ sends x to $x + cx_j e_i$, where $i \neq j$ and $c \in \mathbb{R}$; here $\det T = 1$, so we must show $D(T) = 1$. Note that $T_{ij}(c)T_{ij}(c') = T_{ij}(c + c')$ (since $i \neq j$, the j-th coordinate is untouched), so $T^n = T_{ij}(nc)$ for all $n \ge 1$ and $T^{-1} = T_{ij}(-c)$. Now for any c', every point of $T_{ij}(c')([0,1)^d)$ has all coordinates in $[0,1)$ except the i-th, which lies in the interval between $\min(0, c')$ and $1 + \max(0, c')$, of length $1 + |c'|$. Hence $T_{ij}(c')([0,1)^d)$ is contained in a box of volume $1 + |c'|$, and by monotonicity

$$D(T_{ij}(c')) = m\big(T_{ij}(c')([0,1)^d)\big) \le 1 + |c'|.$$

Applying this with $c' = nc$ and using multiplicativity, $D(T)^n = D(T^n) \le 1 + n|c|$ for every $n \ge 1$. If $D(T) > 1$ the left side grows geometrically while the right side grows only linearly, a contradiction;

hence $D(T) \leq 1$. The same argument applied to $T^{-1} = T_{ij}(-c)$ gives $D(T^{-1}) \leq 1$. But $D(T)D(T^{-1}) = D(\mathrm{id}) = 1$, so both must equal 1. Thus $D(T) = 1 = |\det T|$.
Writing an invertible L as a product $M_1 \cdots M_k$ of elementary matrices, we conclude

$$D(L) = \prod_{r=1}^{k} D(M_r) = \prod_{r=1}^{k} |\det M_r| = |\det L|,$$

completing Part 3.

Problem 1.1.12. Define a Jordan null set to be a Jordan measurable set of Jordan measure zero. Show that any subset of a Jordan null set is a Jordan null set.

Solution. Let $N \subset \mathbb{R}^d$ be a Jordan null set and let $S \subset N$. Since N is Jordan measurable it is bounded (by convention, unbounded sets are not Jordan measurable), hence S is bounded and its Jordan inner and outer measures are defined.
Jordan outer measure is monotone: if $S \subset N$ and $B \supset N$ is elementary, then $B \supset S$ is elementary as well, so the infimum defining $m^{*,(J)}(S)$ is taken over a superset of the collection of elementary sets used for N; therefore

$$m^{*,(J)}(S) \leq m^{*,(J)}(N) = m(N) = 0.$$

On the other hand $m_{*,(J)}(S) \geq 0$, since the empty set is elementary, is contained in S, and has $m(\emptyset) = 0$, so the supremum defining $m_{*,(J)}(S)$ is over a nonempty set of non-negative numbers. Finally $m_{*,(J)}(S) \leq m^{*,(J)}(S)$: if $A \subset S \subset B$ with A, B elementary, then $A \subset B$ and hence $m(A) \leq m(B)$ by monotonicity of elementary measure, and taking the supremum over A and the infimum over B gives the inequality.
Combining, $0 \leq m_{*,(J)}(S) \leq m^{*,(J)}(S) \leq 0$, so $m_{*,(J)}(S) = m^{*,(J)}(S) = 0$. Thus S is Jordan measurable with $m(S) = 0$, i.e. S is a Jordan null set.

Problem 1.1.13. Show that (1.1) holds for all Jordan measurable $E \subset \mathbb{R}^d$; that is, show that

$$m(E) = \lim_{N\to\infty} \frac{1}{N^d} \#\Big(E \cap \frac{1}{N}\mathbb{Z}^d\Big).$$

Solution. Since E is Jordan measurable it is bounded, so $E \cap \frac{1}{N}\mathbb{Z}^d$ is a finite set for every $N \geq 1$ and the quantities appearing above make sense.
We first record that (1.1) holds for elementary sets; this is exactly the discretisation argument in the proof of Lemma 1.1.2(ii), whose one-dimensional step we recall for completeness. Let I be an interval with endpoints $a \leq b$. Then $\#(I \cap \frac{1}{N}\mathbb{Z})$ is the number of integers n with $n/N \in I$, i.e. the number of integers in an interval of length $(b-a)N$; such integers are consecutive, so their number is at most $(b-a)N + 1$, and it is at least $(b-a)N - 1$ (the open interval (aN, bN) has length $(b-a)N$ and so contains at least $(b-a)N - 1$ integers, all of which qualify). Dividing by N and letting $N \to \infty$ gives $|I| = \lim_{N\to\infty} \frac{1}{N}\#(I \cap \frac{1}{N}\mathbb{Z})$. Taking Cartesian products (the grid points of $\frac{1}{N}\mathbb{Z}^d$ lying in a box $Q = I_1 \times \cdots \times I_d$ are exactly the d-tuples whose j-th entry lies in $I_j \cap \frac{1}{N}\mathbb{Z}$) yields $|Q| = \lim_{N\to\infty} \frac{1}{N^d}\#(Q \cap \frac{1}{N}\mathbb{Z}^d)$ for every box Q, and hence, summing over a partition of an elementary set F into disjoint boxes $B_1, \ldots, B_k$ (Lemma 1.1.2(i)),

$$m(F) = |B_1| + \cdots + |B_k| = \lim_{N\to\infty} \frac{1}{N^d} \#\Big(F \cap \frac{1}{N}\mathbb{Z}^d\Big).$$

Now let E be Jordan measurable and let $\varepsilon > 0$. By Exercise 1.1.5 there are elementary sets $A \subset E \subset B$ with $m(B \setminus A) \leq \varepsilon$. Since $A \subset E$ is elementary, the definition of Jordan inner measure gives $m(A) \leq m_{*,(J)}(E) = m(E)$; since $B \supset E$ is elementary, the definition of Jordan outer measure gives $m(E) = m^{*,(J)}(E) \leq m(B)$. Also, by finite additivity of elementary measure applied to the disjoint

decomposition $B = A \cup (B \setminus A)$,

$$m(B) = m(A) + m(B \setminus A) \le m(A) + \varepsilon.$$

Combining these, $m(E) - \varepsilon \le m(B) - \varepsilon \le m(A)$ and $m(B) \le m(A) + \varepsilon \le m(E) + \varepsilon$.
From $A \subset E \subset B$ we get, for every $N \ge 1$,

$$\frac{1}{N^d}\#\Big(A \cap \frac{1}{N}\mathbb{Z}^d\Big) \le \frac{1}{N^d}\#\Big(E \cap \frac{1}{N}\mathbb{Z}^d\Big) \le \frac{1}{N^d}\#\Big(B \cap \frac{1}{N}\mathbb{Z}^d\Big).$$

Letting $N \to \infty$ and applying the elementary case to A and to B,

$$m(E) - \varepsilon \le m(A) \le \liminf_{N\to\infty} \frac{1}{N^d}\#\Big(E \cap \frac{1}{N}\mathbb{Z}^d\Big) \le \limsup_{N\to\infty} \frac{1}{N^d}\#\Big(E \cap \frac{1}{N}\mathbb{Z}^d\Big) \le m(B) \le m(E) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the limit inferior and limit superior both equal $m(E)$. Hence the limit exists and

$$m(E) = \lim_{N\to\infty} \frac{1}{N^d}\#\Big(E \cap \frac{1}{N}\mathbb{Z}^d\Big),$$

which is (1.1) for the Jordan measurable set E.

Problem 1.1.14. (Metric entropy formulation of Jordan measurability) Define a dyadic cube to be a half-open box of the form

$$\Big[\frac{i_1}{2^n}, \frac{i_1+1}{2^n}\Big) \times \cdots \times \Big[\frac{i_d}{2^n}, \frac{i_d+1}{2^n}\Big)$$

for some integers $n, i_1, \dots, i_d$. Let $E \subset \mathbb{R}^d$ be a bounded set. For each integer n, let $\mathcal{E}_*(E, 2^{-n})$ denote the number of dyadic cubes of sidelength 2^{-n} that are contained in E, and let $\mathcal{E}^*(E, 2^{-n})$ be the number of dyadic cubes of sidelength 2^{-n} that intersect E. Show that E is Jordan measurable if and only if

$$\lim_{n\to\infty} 2^{-dn}\big(\mathcal{E}^*(E, 2^{-n}) - \mathcal{E}_*(E, 2^{-n})\big) = 0,$$

in which case one has

$$m(E) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}_*(E, 2^{-n}) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}^*(E, 2^{-n}).$$

[This quantity $\mathcal{E}^*(E, 2^{-n})$ could be called the (dyadic) metric entropy of E at scale 2^{-n}.]

Solution. Fix n. The dyadic cubes of sidelength 2^{-n}, i.e. the sets $\prod_{j=1}^d [i_j 2^{-n}, (i_j+1)2^{-n})$ with $i_1, \dots, i_d \in \mathbb{Z}$, are pairwise disjoint and cover $\mathbb{R}^d$, since for each coordinate the intervals $[i2^{-n}, (i+1)2^{-n})$, $i \in \mathbb{Z}$, partition $\mathbb{R}$. Each such cube is a box of volume 2^{-dn}.
For a bounded set $S \subset \mathbb{R}^d$ define

$$A_n(S) := \bigcup\{Q : Q \text{ a dyadic cube of sidelength } 2^{-n},\ Q \subset S\}, \quad B_n(S) := \bigcup\{Q : Q \cap S \neq \emptyset\}.$$

Since S is bounded, only finitely many dyadic cubes of sidelength 2^{-n} meet S, so both unions are finite unions of boxes, hence elementary sets, and both counts $\mathcal{E}_*(S, 2^{-n}), \mathcal{E}^*(S, 2^{-n})$ are finite. As the cubes involved are pairwise disjoint of volume 2^{-dn},

$$m(A_n(S)) = 2^{-dn}\mathcal{E}_*(S, 2^{-n}), \qquad m(B_n(S)) = 2^{-dn}\mathcal{E}^*(S, 2^{-n}),$$

and clearly $A_n(S) \subset S \subset B_n(S)$. Write $A_n := A_n(E)$, $B_n := B_n(E)$. By the definitions of Jordan inner and outer measure,

$$m(A_n) \le m_{*,(J)}(E) \le m^{*,(J)}(E) \le m(B_n)$$

for every n; call this the sandwich bound.
The "if" direction. Suppose $2^{-dn}(\mathcal{E}^* - \mathcal{E}_*) = m(B_n) - m(A_n) \to 0$. By the sandwich bound,

$$0 \le m^{*,(J)}(E) - m_{*,(J)}(E) \le m(B_n) - m(A_n) \to 0,$$

so $m^{*,(J)}(E) = m_{*,(J)}(E)$ and E is Jordan measurable. Moreover the sandwich bound reads $m(A_n) \le m(E) \le m(B_n)$, and since $m(B_n) - m(A_n) \to 0$ we get $m(A_n) \to m(E)$ and $m(B_n) \to m(E)$, i.e.

$$m(E) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}_*(E, 2^{-n}) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}^*(E, 2^{-n}).$$

The "only if" direction. We first prove two claims about boxes.
Claim A. For a box $Q = I_1 \times \cdots \times I_d$, with I_j having endpoints $a_j \le b_j$, one has $m(A_n(Q)) \ge \prod_{j=1}^d \max\big(0,\ (b_j - a_j) - 2^{1-n}\big)$.
Indeed, fix j and let $K_j^{(n)}$ be the set of integers k with $a_j 2^n < k$ and $k+1 < b_j 2^n$. For such k we have $[k2^{-n}, (k+1)2^{-n}) \subset (a_j, b_j) \subset I_j$. The integers k satisfying these two strict inequalities are exactly those in the open interval $(a_j 2^n,\ b_j 2^n - 1)$, which contains at least $(b_j - a_j)2^n - 2$ integers (an open interval of length L contains at least $L-1$ integers, and here $L = (b_j - a_j)2^n - 1$). Hence the union $J_j^{(n)} := \bigcup_{k\in K_j^{(n)}} [k2^{-n}, (k+1)2^{-n})$ is elementary of measure at least $\max(0, (b_j - a_j) - 2^{1-n})$. Every product $\prod_j [k_j 2^{-n}, (k_j+1)2^{-n})$ with $k_j \in K_j^{(n)}$ is a dyadic cube of sidelength 2^{-n} contained in Q, and these products are pairwise disjoint with union $\prod_j J_j^{(n)}$. So $A_n(Q) \supset \prod_j J_j^{(n)}$ and, by Exercise 1.1.4 and monotonicity, $m(A_n(Q)) \ge \prod_j m(J_j^{(n)})$, which gives the claim.
Claim B. With Q as above, $m(B_n(Q)) \le \prod_{j=1}^d \big((b_j - a_j) + 2^{1-n}\big)$.
Indeed, if a dyadic cube $\prod_j [k_j 2^{-n}, (k_j+1)2^{-n})$ meets Q then for each j the interval $[k_j 2^{-n}, (k_j+1)2^{-n})$ meets I_j, say at the point x. Then $k_j 2^{-n} \le x \le b_j$ and $(k_j+1)2^{-n} > x \ge a_j$, so k_j lies in the half-open interval $(a_j 2^n - 1,\ b_j 2^n]$. (The first inequality need not be strict: if I_j contains its right endpoint, a dyadic interval may begin exactly at b_j and still meet I_j.) That interval has length $(b_j - a_j)2^n + 1$, and any interval of length L contains at most $L+1$ integers, since the integers it contains are consecutive and span a distance at most L; hence it contains at most $(b_j - a_j)2^n + 2$ integers. Letting $\widetilde{J}_j^{(n)}$ be the union of the corresponding dyadic intervals, we get $m(\widetilde{J}_j^{(n)}) \le (b_j - a_j) + 2^{1-n}$ and $B_n(Q) \subset \prod_j \widetilde{J}_j^{(n)}$, whence the claim by Exercise 1.1.4 and monotonicity.
Consequently, for a fixed box Q both $m(A_n(Q))$ and $m(B_n(Q))$ converge to $|Q|$ as $n \to \infty$ (both bounding products tend to $\prod_j (b_j - a_j) = |Q|$, and $m(A_n(Q)) \le |Q| \le m(B_n(Q))$).
Now let F be an arbitrary elementary set, and write $F = Q_1 \cup \cdots \cup Q_k$ as a finite union of disjoint boxes (Lemma 1.1.2(i)). Since $Q_i \subset F$ we have $A_n(Q_i) \subset A_n(F)$, and the sets $A_n(Q_1), \dots, A_n(Q_k)$ are disjoint, so by finite additivity and monotonicity of elementary measure

$$m(A_n(F)) \ge \sum_{i=1}^k m(A_n(Q_i)) \xrightarrow[n\to\infty]{} \sum_{i=1}^k |Q_i| = m(F).$$

Similarly every dyadic cube meeting F meets some Q_i, so $B_n(F) \subset \bigcup_i B_n(Q_i)$ and, by finite subadditivity,

$$m(B_n(F)) \le \sum_{i=1}^k m(B_n(Q_i)) \xrightarrow[n\to\infty]{} \sum_{i=1}^k |Q_i| = m(F).$$

Since also $m(A_n(F)) \le m(F) \le m(B_n(F))$, we conclude

$$\lim_{n\to\infty} m(A_n(F)) = \lim_{n\to\infty} m(B_n(F)) = m(F) \qquad \text{for every elementary } F.$$

We refer to this as the elementary approximation property.
Finally suppose E is Jordan measurable and let $\varepsilon > 0$. By Exercise 1.1.5 choose elementary sets $A \subset E \subset B$ with $m(B \setminus A) \le \varepsilon$; then $m(A) \le m(E) \le m(B) \le m(A) + \varepsilon$, so $m(A) \ge m(E) - \varepsilon$ and $m(B) \le m(E) + \varepsilon$.
Since $A \subset E$, every dyadic cube contained in A is contained in E, so $A_n(A) \subset A_n$ and $m(A_n) \ge m(A_n(A))$. Since $E \subset B$, every dyadic cube meeting E meets B, so $B_n \subset B_n(B)$ and $m(B_n) \le m(B_n(B))$. By the elementary approximation property there is n_0 such that for all $n \ge n_0$,

$$m(A_n(A)) \ge m(A) - \varepsilon \quad \text{and} \quad m(B_n(B)) \le m(B) + \varepsilon.$$

Hence for $n \geq n_0$,

$$m(A_n) \geq m(A) - \varepsilon \geq m(E) - 2\varepsilon, \qquad m(B_n) \leq m(B) + \varepsilon \leq m(E) + 2\varepsilon,$$

and therefore

$$0 \leq 2^{-dn}\big(\mathcal{E}^*(E, 2^{-n}) - \mathcal{E}_*(E, 2^{-n})\big) = m(B_n) - m(A_n) \leq 4\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $\lim_{n\to\infty} 2^{-dn}(\mathcal{E}^*(E, 2^{-n}) - \mathcal{E}_*(E, 2^{-n})) = 0$. The displayed bounds also give $m(E) - 2\varepsilon \leq m(A_n) \leq m(E) \leq m(B_n) \leq m(E) + 2\varepsilon$ for $n \geq n_0$, so again

$$m(E) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}_*(E, 2^{-n}) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}^*(E, 2^{-n}),$$

which completes the proof of the equivalence and of the measure formula.

Problem 1.1.15. (Uniqueness of Jordan measure) Let $d \geq 1$. Let $m' : \mathcal{J}(\mathbb{R}^d) \to \mathbb{R}^+$ be a map from the collection $\mathcal{J}(\mathbb{R}^d)$ of Jordan-measurable subsets of $\mathbb{R}^d$ to the nonnegative reals that obeys the non-negativity, finite additivity, and translation invariance properties. Show that there exists a constant $c \in \mathbb{R}^+$ such that $m'(E) = c\,m(E)$ for all Jordan measurable sets E. In particular, if we impose the additional normalisation $m'([0,1)^d) = 1$, then $m' \equiv m$.

Solution. Every elementary set is Jordan measurable, and Jordan measure agrees with elementary measure on such sets (this is the corollary of Exercise 1.1.5 recorded immediately after that exercise). Moreover, by Exercise 1.1.1 the collection $\mathcal{E}(\mathbb{R}^d)$ of elementary sets is closed under the Boolean operations and under translations. Hence the restriction $m'|_{\mathcal{E}(\mathbb{R}^d)}$ is a map from $\mathcal{E}(\mathbb{R}^d)$ to $\mathbb{R}^+$ which is non-negative, finitely additive and translation invariant, so Exercise 1.1.3 applies to it: there is a constant $c \in \mathbb{R}^+$ (indeed $c = m'([0,1)^d)$, as the hint to that exercise indicates) with

$$m'(A) = c\,m(A) \qquad \text{for every elementary } A \subset \mathbb{R}^d.$$

It remains to upgrade this from elementary sets to Jordan measurable sets.
First we record that m' is monotone on $\mathcal{J}(\mathbb{R}^d)$. Indeed, if $A \subset B$ are Jordan measurable then $B \setminus A$ is Jordan measurable by Exercise 1.1.6(1), and B is the disjoint union of A and $B \setminus A$; finite additivity of m' and non-negativity of m' give

$$m'(B) = m'(A) + m'(B \setminus A) \geq m'(A).$$

Now let E be an arbitrary Jordan measurable set and let $\varepsilon > 0$. By Exercise 1.1.5 there exist elementary sets $A \subset E \subset B$ with $m(B\setminus A) \leq \varepsilon$. Since $A \subset B$ are elementary, finite additivity of elementary measure gives $m(B) - m(A) = m(B \setminus A) \leq \varepsilon$.
By monotonicity of m' and the identity already established on elementary sets,

$$c\,m(A) = m'(A) \leq m'(E) \leq m'(B) = c\,m(B).$$

By monotonicity of Jordan measure (Exercise 1.1.6(4)) we also have

$$c\,m(A) \leq c\,m(E) \leq c\,m(B).$$

Thus both $m'(E)$ and $c\,m(E)$ lie in the interval $[c\,m(A), c\,m(B)]$, whose length is $c(m(B)-m(A)) \leq c\varepsilon$. Consequently

$$|m'(E) - c\,m(E)| \leq c\varepsilon.$$

As $\varepsilon > 0$ was arbitrary and c is a fixed finite constant, we conclude $m'(E) = c\,m(E)$ for every Jordan measurable E.
Finally, suppose the normalisation $m'([0,1)^d) = 1$ is imposed. The set $[0,1)^d$ is a box of volume 1, so $m([0,1)^d) = 1$, and the identity just proved gives $1 = m'([0,1)^d) = c \cdot 1 = c$. Hence $c = 1$ and $m' \equiv m$ on $\mathcal{J}(\mathbb{R}^d)$.

Problem 1.1.16. Let $d_1, d_2 \geq 1$, and let $E_1 \subset \mathbb{R}^{d_1}$, $E_2 \subset \mathbb{R}^{d_2}$ be Jordan measurable sets. Show that $E_1 \times E_2 \subset \mathbb{R}^{d_1+d_2}$ is Jordan measurable, and $m_{d_1+d_2}(E_1 \times E_2) = m_{d_1}(E_1) \times m_{d_2}(E_2)$.

Solution. Throughout we identify $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with $\mathbb{R}^{d_1+d_2}$ in the obvious way. Since E_1 and E_2 are Jordan measurable they are in particular bounded, so $E_1 \times E_2$ is a bounded subset of $\mathbb{R}^{d_1+d_2}$ and it makes sense to ask for its Jordan measurability.
Let $0 < \varepsilon \leq 1$. By Exercise 1.1.5 there are elementary sets

$$A_1 \subset E_1 \subset B_1 \subset \mathbb{R}^{d_1}, \qquad A_2 \subset E_2 \subset B_2 \subset \mathbb{R}^{d_2}$$

with $m(B_1 \setminus A_1) \leq \varepsilon$ and $m(B_2 \setminus A_2) \leq \varepsilon$. Since $A_i \subset B_i$ are elementary, finite additivity of elementary measure gives

$$m(B_i) - m(A_i) = m(B_i \setminus A_i) \leq \varepsilon \qquad (i = 1, 2).$$

Also, by monotonicity of Jordan measure, $m(A_i) \leq m_{d_i}(E_i) \leq m(B_i)$, and hence $m(B_i) \leq m(A_i) + \varepsilon \leq m_{d_i}(E_i) + 1$.
By Exercise 1.1.4 the products $A_1 \times A_2$ and $B_1 \times B_2$ are elementary subsets of $\mathbb{R}^{d_1+d_2}$, with

$$m(A_1 \times A_2) = m(A_1)m(A_2), \qquad m(B_1 \times B_2) = m(B_1)m(B_2).$$

Clearly

$$A_1 \times A_2 \subset E_1 \times E_2 \subset B_1 \times B_2,$$

and $A_1 \times A_2 \subset B_1 \times B_2$, so by finite additivity

$$\begin{aligned} m\big((B_1 \times B_2) \setminus (A_1 \times A_2)\big) &= m(B_1)m(B_2) - m(A_1)m(A_2) \\ &= m(B_1)\big(m(B_2) - m(A_2)\big) + \big(m(B_1) - m(A_1)\big)m(A_2) \\ &\leq \big(m_{d_1}(E_1) + 1\big)\varepsilon + \varepsilon\, m_{d_2}(E_2) = C\varepsilon, \end{aligned}$$

where $C := m_{d_1}(E_1) + m_{d_2}(E_2) + 1$ is a constant depending only on E_1, E_2.
Note that $C \geq 1 > 0$ and that C does not depend on ε. Hence, given any $\eta > 0$, running the above with $\varepsilon := \min(1, \eta/C)$ produces elementary sets $A_1 \times A_2 \subset E_1 \times E_2 \subset B_1 \times B_2$ with $m\big((B_1 \times B_2) \setminus (A_1 \times A_2)\big) \leq C\varepsilon \leq \eta$. Criterion (2) of Exercise 1.1.5 is therefore satisfied, so $E_1 \times E_2$ is Jordan measurable.
For the value of the measure, note that by monotonicity of Jordan measure applied to the inclusions above,

$$m(A_1)m(A_2) \leq m_{d_1+d_2}(E_1 \times E_2) \leq m(B_1)m(B_2),$$

while trivially

$$m(A_1)m(A_2) \leq m_{d_1}(E_1)\, m_{d_2}(E_2) \leq m(B_1)m(B_2)$$

because $m(A_i) \leq m_{d_i}(E_i) \leq m(B_i)$ and all quantities are non-negative. Both numbers $m_{d_1+d_2}(E_1 \times E_2)$ and $m_{d_1}(E_1)m_{d_2}(E_2)$ therefore lie in an interval of length at most $C\varepsilon$, so

$$\big|m_{d_1+d_2}(E_1 \times E_2) - m_{d_1}(E_1)m_{d_2}(E_2)\big| \leq C\varepsilon.$$

Letting $\varepsilon \to 0$ gives $m_{d_1+d_2}(E_1 \times E_2) = m_{d_1}(E_1)\, m_{d_2}(E_2)$, as required.

Problem 1.1.17. Let P, Q be two polytopes in $\mathbb{R}^d$. Suppose that P can be partitioned into finitely many sub-polytopes which, after being rotated and translated, form a cover of Q, with any two of the sub-polytopes in Q intersecting only at their boundaries. Conclude that P and Q have the same Jordan measure. The converse statement is true in one and two dimensions $d = 1, 2$ (this is the Bolyai-Gerwien theorem), but false in higher dimensions (this was Dehn's negative answer [De1901] to Hilbert's third problem).

Solution. Here "polytope" means compact convex polytope in the sense of footnote 6, i.e. a bounded intersection of finitely many closed half-spaces $\{x \in \mathbb{R}^d : x \cdot v \le c\}$. By Exercise 1.1.9 every such set is Jordan measurable, so P, Q and all the pieces below have well-defined Jordan measures.

Step 1: the boundary of a polytope is a Jordan null set.

We first check that if $H = \{x \in \mathbb{R}^d : x \cdot v = c\}$ is a hyperplane (so $v \neq 0$) and $S \subset H$ is bounded, then S is a Jordan null set. Let L be an orthogonal linear map with $L(v) = |v|e_d$. For $x \in H$ we have $L(x) \cdot L(v) = x \cdot v = c$, i.e. the last coordinate of $L(x)$ equals $c/|v|$. Hence $L(S)$ is a bounded subset of the hyperplane $\{y_d = c/|v|\}$, so it is contained in a degenerate box $[-M, M]^{d-1} \times \{c/|v|\}$ of volume 0 for some M; therefore $m^{*,(J)}(L(S)) = 0$ and $L(S)$ is Jordan measurable with measure zero. Applying Exercise 1.1.11(2) to the linear map L^{-1} we get that $S = L^{-1}(L(S))$ is Jordan measurable with $m(S) = |\det L^{-1}|\, m(L(S)) = 0$.

Now let $R = \bigcap_{k=1}^{K} \{x : x \cdot v_k \le c_k\}$ be a polytope. If some $v_k = 0$ the corresponding half-space is either all of $\mathbb{R}^d$ (if $c_k \ge 0$), and may be discarded, or empty (if $c_k < 0$), in which case $R = \emptyset$ and there is nothing to prove. So assume every $v_k \neq 0$, and put $H_k := \{x : x \cdot v_k = c_k\}$. If $x \in R$ satisfies $x \cdot v_k < c_k$ for every k, then by continuity of the maps $x \mapsto x \cdot v_k$ the same strict inequalities hold on a neighbourhood of x, so x lies in the interior of R. Since R is closed, $\partial R \subset R$, and therefore

$$\partial R \subset \bigcup_{k=1}^{K} (R \cap H_k).$$

Each $R \cap H_k$ is a bounded subset of a hyperplane, hence a Jordan null set by the previous paragraph; a finite union of Jordan null sets is Jordan measurable with measure zero by Exercise 1.1.6, and ∂R, being a subset of that union, is a Jordan null set by Exercise 1.1.12.

Step 2: measure is additive across a dissection.

Suppose R is a Jordan measurable set which is the union of polytopes $R_1, \dots, R_n$ such that $R_i \cap R_j \subset \partial R_i$ whenever $i \neq j$ (this is exactly the hypothesis "any two pieces intersect only at their boundaries"). We claim $m(R) = \sum_{i=1}^{n} m(R_i)$.

Put $R_i' := R_i \setminus (R_1 \cup \cdots \cup R_{i-1})$. By Exercise 1.1.6(1) each R_i' is Jordan measurable, the R_i' are pairwise disjoint, and their union is R; so by finite additivity (Exercise 1.1.6(3), iterated) $m(R) = \sum_i m(R_i')$. Moreover

$$R_i \setminus R_i' = R_i \cap (R_1 \cup \cdots \cup R_{i-1}) = \bigcup_{j<i} (R_i \cap R_j) \subset \partial R_i,$$

which is a Jordan null set by Step 1; since $R_i \setminus R_i'$ is Jordan measurable, Exercise 1.1.12 gives $m(R_i \setminus R_i') = 0$. Finite additivity applied to the disjoint decomposition $R_i = R_i' \cup (R_i \setminus R_i')$ then gives $m(R_i') = m(R_i)$. Hence $m(R) = \sum_i m(R_i)$, as claimed.

Step 3: conclusion.

A word on the reading of the hypothesis. A nonempty polytope P is convex, hence connected, so it cannot be written as a union of two or more disjoint nonempty closed sets; since sub-polytopes are closed, a literal partition of P into disjoint sub-polytopes is only possible with $n = 1$. Accordingly, "partitioned into finitely many sub-polytopes" is to be read in the classical dissection sense: $P = P_1 \cup \cdots \cup P_n$ with $P_i \cap P_j \subset \partial P_i$ for $i \neq j$. Nothing is lost by this: a genuinely disjoint decomposition satisfies $P_i \cap P_j = \emptyset \subset \partial P_i$, so Step 2 covers both readings at once.

Write $P = P_1 \cup \cdots \cup P_n$ for this dissection of P, and let $Q_i := \rho_i(P_i) + w_i$, where ρ_i is a rotation of $\mathbb{R}^d$ and $w_i \in \mathbb{R}^d$, be the resulting pieces. The hypothesis that the Q_i "form a cover of Q, with any two intersecting only at their boundaries" is the statement that the Q_i dissect Q: $Q = Q_1 \cup \cdots \cup Q_n$ and $Q_i \cap Q_j \subset \partial Q_i$ for $i \neq j$. (Here "cover" must be read as an exact cover, i.e. the union of the Q_i is Q and not merely a superset of Q. Some such reading is forced if the statement is to be true: with $d = 2$, cut the unit square $P = [0,1]^2$ into the two halves $P_1 = [0, \frac{1}{2}] \times [0,1]$ and $P_2 = [\frac{1}{2}, 1] \times [0,1]$ and leave them where they are, so $Q_i = P_i$; these two pieces meet only along the segment $\{\frac{1}{2}\} \times [0,1]$, which lies in the boundary of each, and they certainly cover the small square $Q = [0, \frac{1}{2}]^2$. But $m(P) = 1$ while $m(Q) = \frac{1}{4}$. Overhang must therefore be disallowed, which is exactly the requirement that the reassembled pieces have union equal to Q.) (A rotation is a linear isometry, so each Q_i is again a polytope: it is the image of an intersection of closed half-spaces under an invertible affine map, hence again a bounded intersection of closed half-spaces.)

By Step 2 applied to P and to Q,

$$m(P) = \sum_{i=1}^{n} m(P_i), \qquad m(Q) = \sum_{i=1}^{n} m(Q_i).$$

By translation invariance of Jordan measure (Exercise 1.1.6(6)) and Exercise 1.1.11(2)(3),

$$m(Q_i) = m(\rho_i(P_i) + w_i) = m(\rho_i(P_i)) = |\det \rho_i|\, m(P_i) = m(P_i),$$

since a rotation is orthogonal and therefore has determinant of absolute value 1.
Summing over i gives $m(Q) = \sum_i m(P_i) = m(P)$, i.e. P and Q have the same Jordan measure.

Problem 1.1.18. Let $E \subset \mathbb{R}^d$ be a bounded set.

1. Show that E and the closure $\overline{E}$ of E have the same Jordan outer measure.
2. Show that E and the interior E° of E have the same Jordan inner measure.
3. Show that E is Jordan measurable if and only if the topological boundary ∂E of E has Jordan outer measure zero.
4. Show that the bullet-riddled square $[0,1]^2 \backslash \mathbb{Q}^2$, and set of bullets $[0,1]^2 \cap \mathbb{Q}^2$, both have Jordan inner measure zero and Jordan outer measure one. In particular, both sets are not Jordan measurable.

Solution. We first record three elementary facts that will be used repeatedly.
Fact A (monotonicity and subadditivity of Jordan outer measure). If $S \subset T$ are bounded then every elementary $B \supset T$ also contains S, so $m^{*,(J)}(S) \leq m^{*,(J)}(T)$. If S, T are bounded and $\varepsilon > 0$, pick elementary $B_1 \supset S$, $B_2 \supset T$ with $m(B_1) \leq m^{*,(J)}(S) + \varepsilon$ and $m(B_2) \leq m^{*,(J)}(T) + \varepsilon$; then $B_1 \cup B_2$ is elementary (Exercise 1.1.1), contains $S \cup T$, and has $m(B_1 \cup B_2) \leq m(B_1) + m(B_2)$ by finite subadditivity of elementary measure. Letting $\varepsilon \to 0$,

$$m^{*,(J)}(S \cup T) \leq m^{*,(J)}(S) + m^{*,(J)}(T).$$

Similarly $m_{*,(J)}$ is monotone, and one always has $m_{*,(J)}(S) \leq m^{*,(J)}(S)$, since any elementary $A \subset S \subset B$ satisfy $A \subset B$ and hence $m(A) \leq m(B)$.
Fact B (closing up an elementary set does not change its measure). Let B be elementary. By Lemma 1.1.2(i) write $B = B_1 \cup \cdots \cup B_k$ as a finite union of disjoint boxes, discarding any empty ones, so that $m(B) = \sum_j |B_j|$. Then $\overline{B} = \overline{B_1} \cup \cdots \cup \overline{B_k}$, and for a nonempty box $B_j = I_{j,1} \times \cdots \times I_{j,d}$ we have $\overline{B_j} = \overline{I_{j,1}} \times \cdots \times \overline{I_{j,d}}$, which is again a box with $|\overline{B_j}| = |B_j|$ (closing an interval does not change its length). Hence $\overline{B}$ is elementary, and by finite subadditivity and monotonicity of elementary measure

$$m(B) \leq m(\overline{B}) \leq \sum_j |\overline{B_j}| = \sum_j |B_j| = m(B),$$

so $m(\overline{B}) = m(B)$.
Fact C (opening up an elementary set does not change its measure). Let A be elementary, written as above as a disjoint union of boxes $B_1, \ldots, B_k$. Set $A^\flat := B_1^\circ \cup \cdots \cup B_k^\circ$. The sets B_j° are open boxes (possibly empty), pairwise disjoint, and $|B_j^\circ| = |B_j|$: if every side of B_j has positive length this is clear, and if some side has zero length then both volumes are 0. Hence $A^\flat$ is an open elementary set with $A^\flat \subset A$ and

$$m(A^\flat) = \sum_j |B_j^\circ| = \sum_j |B_j| = m(A).$$

Part (1). Since $E \subset \overline{E}$, Fact A gives $m^{*,(J)}(E) \leq m^{*,(J)}(\overline{E})$. Conversely, let $B \supset E$ be elementary. Then $\overline{E} \subset \overline{B}$, and $\overline{B}$ is elementary with $m(\overline{B}) = m(B)$ by Fact B; hence $m^{*,(J)}(\overline{E}) \leq m(B)$. Taking the infimum over all elementary $B \supset E$ yields $m^{*,(J)}(\overline{E}) \leq m^{*,(J)}(E)$. The two inequalities give equality.
Part (2). Since $E^\circ \subset E$, monotonicity gives $m_{*,(J)}(E^\circ) \leq m_{*,(J)}(E)$. Conversely, let $A \subset E$ be elementary and let $A^\flat$ be the open elementary set produced in Fact C, so $A^\flat \subset A \subset E$ and $m(A^\flat) =$

$m(A)$. As $A^\flat$ is open and contained in E, we have $A^\flat \subset E^\circ$, whence $m_{*,(J)}(E^\circ) \geq m(A^\flat) = m(A)$. Taking the supremum over all elementary $A \subset E$ gives $m_{*,(J)}(E^\circ) \geq m_{*,(J)}(E)$, and equality follows.
Part (3). Suppose first that E is Jordan measurable, and let $\varepsilon > 0$. By Exercise 1.1.5 there are elementary sets $A \subset E \subset B$ with $m(B \setminus A) \leq \varepsilon$. Let $A^\flat \subset A$ be the open elementary set of Fact C and $\overline{B}$ the closed elementary set of Fact B, so $m(A^\flat) = m(A)$ and $m(\overline{B}) = m(B)$. Since $A^\flat$ is open and contained in E, $A^\flat \subset E^\circ$; since $\overline{B}$ is closed and contains E, $\overline{E} \subset \overline{B}$. Therefore

$$\partial E = \overline{E} \setminus E^\circ \subset \overline{B} \setminus A^\flat,$$

and $\overline{B} \setminus A^\flat$ is elementary (Exercise 1.1.1) with measure

$$m(\overline{B} \setminus A^\flat) = m(\overline{B}) - m(A^\flat) = m(B) - m(A) = m(B \setminus A) \leq \varepsilon,$$

using finite additivity of elementary measure and $A^\flat \subset \overline{B}$, $A \subset B$. Hence $m^{*,(J)}(\partial E) \leq \varepsilon$ for every $\varepsilon > 0$, i.e. $m^{*,(J)}(\partial E) = 0$.
Conversely, suppose $m^{*,(J)}(\partial E) = 0$, and let $\varepsilon > 0$. Choose an elementary $C \supset \partial E$ with $m(C) \leq \varepsilon$. We may enlarge C to an open elementary set. Write C as a disjoint union of boxes $B_1, \ldots, B_k$ using Lemma 1.1.2(i), adjoining the empty box if necessary so that $k \geq 1$; thus $m(C) = \sum_{j=1}^k |B_j|$. Writing $B_j = I_{j,1} \times \cdots \times I_{j,d}$ with $I_{j,i}$ having endpoints $a_{j,i} \leq b_{j,i}$, put

$$B_j'(\delta) := (a_{j,1} - \delta, b_{j,1} + \delta) \times \cdots \times (a_{j,d} - \delta, b_{j,d} + \delta)$$

for $\delta > 0$. This is an open box containing $\overline{B_j}$, with volume $\prod_{i=1}^d (b_{j,i} - a_{j,i} + 2\delta)$, which tends to $\prod_{i=1}^d (b_{j,i} - a_{j,i}) = |B_j|$ as $\delta \to 0$; so we may fix δ small enough that $|B_j'| \leq |B_j| + \varepsilon/k$ for every j, where $B_j' := B_j'(\delta)$. Set $U := B_1' \cup \cdots \cup B_k'$. Then U is an open elementary set with $\partial E \subset C \subset U$ and, by finite subadditivity of elementary measure, $m(U) \leq \sum_j |B_j'| \leq m(C) + \varepsilon \leq 2\varepsilon$.
Let $K := \overline{E} \setminus U$. As E is bounded, $\overline{E}$ is compact, and K is a closed subset of it, hence compact. Since $\overline{E} = E^\circ \cup \partial E$ and $\partial E \subset U$, we have $K \subset E^\circ$. Now for each $x \in K$ choose an open box Q_x with $x \in Q_x \subset E^\circ$ (possible as E° is open); by compactness finitely many of these, $Q_{x_1}, \ldots, Q_{x_r}$, cover K. The set $A := Q_{x_1} \cup \cdots \cup Q_{x_r}$ is elementary with $K \subset A \subset E^\circ \subset E$, so

$$m^{*,(J)}(K) \leq m(A) \leq m_{*,(J)}(E).$$

Since $\overline{E} \subset K \cup U$, Fact A and part (1) give

$$m^{*,(J)}(E) = m^{*,(J)}(\overline{E}) \leq m^{*,(J)}(K) + m^{*,(J)}(U) \leq m_{*,(J)}(E) + m(U) \leq m_{*,(J)}(E) + 2\varepsilon.$$

Letting $\varepsilon \to 0$ gives $m^{*,(J)}(E) \leq m_{*,(J)}(E)$; the reverse inequality always holds (Fact A), so the inner and outer Jordan measures of E coincide and E is Jordan measurable.
Part (4). Write $S := [0,1]^2 \cap \mathbb{Q}^2$ (the bullets) and $T := [0,1]^2 \setminus \mathbb{Q}^2$ (the bullet-riddled square). Both are bounded.
Closures. Every point of $[0,1]^2$ is a limit of points of $\mathbb{Q}^2 \cap [0,1]^2$, since $\mathbb{Q}$ is dense in $[0,1]$; hence $\overline{S} = [0,1]^2$ (the closure is contained in the closed set $[0,1]^2$ and contains it). For T: given $(x,y) \in [0,1]^2$ and $\delta > 0$, choose an irrational $x' \in [0,1]$ with $|x - x'| < \delta$ (the irrationals are dense in $[0,1]$); then $(x', y) \in T$ because its first coordinate is irrational, so $(x', y) \notin \mathbb{Q}^2$. Hence $\overline{T} = [0,1]^2$ as well.
Interiors. If U is a nonempty open subset of $\mathbb{R}^2$ then U contains a point of $\mathbb{Q}^2$ (density of $\mathbb{Q}^2$) and also a point with an irrational coordinate. Consequently no nonempty open set is contained in S, and none is contained in T; that is, $S^\circ = T^\circ = \emptyset$.
Now $[0,1]^2$ is a box of volume 1, hence elementary with $m([0,1]^2) = 1$, so $m^{*,(J)}([0,1]^2) = 1$. By part (1),

$$m^{*,(J)}(S) = m^{*,(J)}(\overline{S}) = m^{*,(J)}([0,1]^2) = 1, \qquad m^{*,(J)}(T) = m^{*,(J)}(\overline{T}) = 1.$$

By part (2), and since the only elementary subset of $\emptyset$ is $\emptyset$ itself,

$$m_{*,(J)}(S) = m_{*,(J)}(S^\circ) = m_{*,(J)}(\emptyset) = 0, \qquad m_{*,(J)}(T) = m_{*,(J)}(T^\circ) = 0.$$

For each of S and T the inner measure is 0 and the outer measure is 1; these differ, so neither set is Jordan measurable. (Consistently with part (3), the boundary of each is $\partial S = \partial T = [0,1]^2$, of outer measure $1 \neq 0$.)

Problem 1.1.19. (Caratheodory type property) Let $E \subset \mathbb{R}^d$ be a bounded set, and $F \subset \mathbb{R}^d$ be an elementary set. Show that

$$m^{*,(J)}(E) = m^{*,(J)}(E \cap F) + m^{*,(J)}(E \setminus F).$$

Solution. Both $E \cap F$ and $E \setminus F$ are subsets of the bounded set E, hence bounded, so all three outer measures are finite.
The inequality $\leq$. This is just finite subadditivity of Jordan outer measure. Let $\varepsilon > 0$ and choose elementary sets $B_1 \supset E \cap F$ and $B_2 \supset E \setminus F$ with

$$m(B_1) \leq m^{*,(J)}(E \cap F) + \varepsilon, \qquad m(B_2) \leq m^{*,(J)}(E \setminus F) + \varepsilon.$$

By Exercise 1.1.1 the set $B_1 \cup B_2$ is elementary, and it contains $(E \cap F) \cup (E \setminus F) = E$. By finite subadditivity of elementary measure,

$$m^{*,(J)}(E) \leq m(B_1 \cup B_2) \leq m(B_1) + m(B_2) \leq m^{*,(J)}(E \cap F) + m^{*,(J)}(E \setminus F) + 2\varepsilon.$$

Letting $\varepsilon \to 0$ gives the desired inequality.
The inequality $\geq$. This is where the elementarity of F is used. Let $B \supset E$ be any elementary set. By Exercise 1.1.1 the sets $B \cap F$ and $B \setminus F$ are elementary; they are disjoint and their union is B. Moreover

$$E \cap F \subset B \cap F, \qquad E \setminus F \subset B \setminus F,$$

the first because $E \subset B$, and the second because $x \in E \setminus F$ implies $x \in B$ and $x \notin F$. Hence, by the definition of Jordan outer measure as an infimum over elementary supersets,

$$m^{*,(J)}(E \cap F) \leq m(B \cap F), \qquad m^{*,(J)}(E \setminus F) \leq m(B \setminus F).$$

Adding these and using finite additivity of elementary measure on the disjoint decomposition $B = (B \cap F) \cup (B \setminus F)$,

$$m^{*,(J)}(E \cap F) + m^{*,(J)}(E \setminus F) \leq m(B \cap F) + m(B \setminus F) = m(B).$$

This holds for every elementary $B \supset E$, so taking the infimum over such B gives

$$m^{*,(J)}(E \cap F) + m^{*,(J)}(E \setminus F) \leq m^{*,(J)}(E).$$

Combining the two inequalities yields $m^{*,(J)}(E) = m^{*,(J)}(E \cap F) + m^{*,(J)}(E \setminus F)$, as required.

Problem 1.1.20. (Piecewise constant functions) Let $[a, b]$ be an interval. A piecewise constant function $f : [a, b] \to \mathbb{R}$ is a function for which there exists a partition of $[a, b]$ into finitely many intervals $I_1, \ldots, I_n$, such that f is equal to a constant c_i on each of the intervals I_i. If f is piecewise constant, show that the expression

$$\sum_{i=1}^{n} c_i |I_i|$$

is independent of the choice of partition used to demonstrate the piecewise constant nature of f. We will denote this quantity by p.c. $\int_a^b f(x)\,dx$, and refer to it as the piecewise constant integral of f on $[a, b]$.

Solution. Suppose $I_1, \ldots, I_n$ is a partition of $[a, b]$ into intervals with $f \equiv c_i$ on I_i, and $I'_1, \ldots, I'_{n'}$ is a second such partition with $f \equiv c'_j$ on I'_j. (Here "partition" means the intervals are pairwise disjoint and their union is $[a, b]$; degenerate intervals, of length zero, and the empty interval are allowed, and

on an empty I_i the constant c_i may be chosen arbitrarily.) We must show

$$\sum_{i=1}^{n} c_i|I_i| = \sum_{j=1}^{n'} c'_j|I'_j|.$$

Consider the common refinement, consisting of the sets $I_i \cap I'_j$ for $1 \le i \le n$, $1 \le j \le n'$. We first check that the intersection of two intervals is again an interval in the sense of Definition 1.1.1. It suffices to observe that every bounded convex subset $C \subset \mathbb{R}$ is such an interval. Indeed, if $C = \emptyset$ we may write $C = (0,0)$, a degenerate interval of length 0. Otherwise put $\alpha := \inf C$ and $\beta := \sup C$, which are finite since C is bounded and nonempty, and satisfy $\alpha \le \beta$. Every x with $\alpha < x < \beta$ lies strictly between two points of C and hence lies in C by convexity, so $(\alpha,\beta) \subset C \subset [\alpha,\beta]$; therefore C is one of $[\alpha,\beta]$, $[\alpha,\beta)$, $(\alpha,\beta]$, (α,β), i.e. an interval, of length $\beta - \alpha$. Since I_i and I'_j are convex and bounded, so is $I_i \cap I'_j$, which is therefore an interval; in particular it is a one-dimensional box, so Lemma 1.1.2 applies to it.
Fix i. The sets $I_i \cap I'_j$, $1 \le j \le n'$, are pairwise disjoint (since the I'_j are) and their union is $I_i \cap \bigcup_j I'_j = I_i \cap [a,b] = I_i$. Thus I_i is partitioned into the finitely many disjoint intervals (i.e. one-dimensional boxes) $I_i \cap I'_j$, and Lemma 1.1.2(ii) gives

$$|I_i| = m(I_i) = \sum_{j=1}^{n'} |I_i \cap I'_j|.$$

Symmetrically, $|I'_j| = \sum_{i=1}^{n} |I_i \cap I'_j|$ for each j.
Consequently

$$\sum_{i=1}^{n} c_i|I_i| = \sum_{i=1}^{n}\sum_{j=1}^{n'} c_i|I_i \cap I'_j|, \qquad \sum_{j=1}^{n'} c'_j|I'_j| = \sum_{i=1}^{n}\sum_{j=1}^{n'} c'_j|I_i \cap I'_j|.$$

It therefore suffices to check that, for each pair (i,j),

$$c_i|I_i \cap I'_j| = c'_j|I_i \cap I'_j|.$$

If $I_i \cap I'_j = \emptyset$ then $|I_i \cap I'_j| = 0$ and both sides vanish. If $I_i \cap I'_j \neq \emptyset$, pick a point x in it; then $x \in I_i$ gives $f(x) = c_i$, and $x \in I'_j$ gives $f(x) = c'_j$, so $c_i = c'_j$ and the two sides agree.
Summing over all pairs (i,j) gives $\sum_i c_i|I_i| = \sum_j c'_j|I'_j|$, which is the required independence of the partition. The common value may therefore be written unambiguously as p.c. $\int_a^b f(x)\,dx$.

Problem 1.1.21. (Basic properties of the piecewise constant integral) Let $[a,b]$ be an interval, and let $f, g : [a,b] \to \mathbb{R}$ be piecewise constant functions. Establish the following statements:

1. (Linearity) For any real number c, cf and $f+g$ are piecewise constant, with p.c. $\int_a^b cf(x)\,dx = c$ p.c. $\int_a^b f(x)\,dx$ and p.c. $\int_a^b f(x)+g(x)\,dx =$ p.c. $\int_a^b f(x)\,dx +$ p.c. $\int_a^b g(x)\,dx$.
2. (Monotonicity) If $f \le g$ pointwise (i.e. $f(x) \le g(x)$ for all $x \in [a,b]$) then p.c. $\int_a^b f(x)\,dx \le$ p.c. $\int_a^b g(x)\,dx$.
3. (Indicator) If E is an elementary subset of $[a,b]$, then the indicator function $1_E : [a,b] \to \mathbb{R}$ (defined by setting $1_E(x) := 1$ when $x \in E$ and $1_E(x) := 0$ otherwise) is piecewise constant, and p.c. $\int_a^b 1_E(x)\,dx = m(E)$.

Solution. A preliminary reduction. Let f be piecewise constant with respect to a partition $I_1, \ldots, I_n$ of $[a,b]$ into intervals (values c_i), and g piecewise constant with respect to a partition $I'_1, \ldots, I'_{n'}$ (values c'_j). As in the solution to Exercise 1.1.20, the sets $J_{ij} := I_i \cap I'_j$ are intervals, are pairwise disjoint, and their union is $[a,b]$; so, relabelling them as $J_1, \ldots, J_m$, we obtain a single partition of $[a,b]$ into finitely many intervals on which both f and g are constant, say $f \equiv \alpha_k$ and $g \equiv \beta_k$ on J_k. By Exercise 1.1.20

the piecewise constant integrals may be computed with this common partition:

$$\text{p.c.}\int_a^b f(x)\,dx = \sum_{k=1}^m \alpha_k |J_k|, \qquad \text{p.c.}\int_a^b g(x)\,dx = \sum_{k=1}^m \beta_k |J_k|.$$

(As before, on an empty J_k the "value" of a function is vacuous and we may assign any constant; such a J_k has $|J_k| = 0$ and contributes nothing.)
Part (1) (Linearity). Fix $c \in \mathbb{R}$. The function cf is constant, equal to $c\alpha_k$, on each J_k, so cf is piecewise constant with respect to the partition $J_1, \ldots, J_m$. By Exercise 1.1.20,

$$\text{p.c.}\int_a^b cf(x)\,dx = \sum_{k=1}^m (c\alpha_k)|J_k| = c\sum_{k=1}^m \alpha_k |J_k| = c\,\text{p.c.}\int_a^b f(x)\,dx.$$

Similarly, $f+g$ is constant, equal to $\alpha_k + \beta_k$, on each J_k, hence piecewise constant, and

$$\text{p.c.}\int_a^b f(x)+g(x)\,dx = \sum_{k=1}^m (\alpha_k+\beta_k)|J_k| = \sum_{k=1}^m \alpha_k|J_k| + \sum_{k=1}^m \beta_k|J_k|,$$

which is p.c. $\int_a^b f(x)\,dx$ + p.c. $\int_a^b g(x)\,dx$.
Part (2) (Monotonicity). Suppose $f(x) \le g(x)$ for all $x \in [a,b]$, and use the common partition $J_1, \ldots, J_m$ as above. Fix k. If $J_k \neq \emptyset$, pick $x \in J_k$; then $\alpha_k = f(x) \le g(x) = \beta_k$, and since $|J_k| \ge 0$ we get $\alpha_k|J_k| \le \beta_k|J_k|$. If $J_k = \emptyset$ then $|J_k| = 0$ and both terms are 0. Summing over k,

$$\text{p.c.}\int_a^b f(x)\,dx = \sum_{k=1}^m \alpha_k|J_k| \le \sum_{k=1}^m \beta_k|J_k| = \text{p.c.}\int_a^b g(x)\,dx.$$

Part (3) (Indicator). Let $E \subset [a,b]$ be elementary. Since $d=1$ here, E is a finite union of intervals, and by Lemma 1.1.2(i) we may write it as a finite union of pairwise disjoint intervals

$$E = J_1 \cup \cdots \cup J_r, \qquad m(E) = \sum_{k=1}^r |J_k|$$

(the second identity being Lemma 1.1.2(ii)). The set $[a,b] \setminus E$ is elementary by Exercise 1.1.1, since $[a,b]$ is an interval and hence elementary; applying Lemma 1.1.2(i) again, write $[a,b]\setminus E = J_{r+1} \cup \cdots \cup J_s$ with $J_{r+1}, \ldots, J_s$ pairwise disjoint intervals.
Then $J_1, \ldots, J_s$ are pairwise disjoint intervals whose union is $E \cup ([a,b] \setminus E) = [a,b]$, i.e. they form a partition of $[a,b]$ into finitely many intervals. On each J_k with $k \le r$ we have $J_k \subset E$, so $1_E \equiv 1$ there; on each J_k with $k > r$ we have $J_k \cap E = \emptyset$, so $1_E \equiv 0$ there. Hence 1_E is piecewise constant, and by Exercise 1.1.20 its piecewise constant integral may be computed with this partition:

$$\text{p.c.}\int_a^b 1_E(x)\,dx = \sum_{k=1}^r 1\cdot|J_k| + \sum_{k=r+1}^s 0\cdot|J_k| = \sum_{k=1}^r |J_k| = m(E),$$

as required.

Problem 1.1.22. Let $[a,b]$ be an interval, and $f : [a,b] \to \mathbb{R}$ be a bounded function. Show that f is Riemann integrable if and only if it is Darboux integrable, in which case the Riemann integral and Darboux integrals are equal.

Solution. If $[a,b]$ has zero length both notions are trivial (every function is Riemann integrable with integral 0 by the convention in Definition 1.1.5, and every piecewise constant function has piecewise constant integral 0, so both Darboux integrals vanish). So assume $a < b$, and fix M with $|f| \le M$ on $[a,b]$.

Everything rests on the following approximation lemma, which says that Riemann sums of a piecewise constant function converge to its piecewise constant integral.

Lemma. Let $g : [a,b] \to \mathbb{R}$ be piecewise constant with $|g| \le M$, and let $I_1, \dots, I_k$ be a partition of $[a,b]$ into intervals with $g \equiv c_j$ on I_j. Then for every tagged partition $\mathcal{P} = ((x_0, \dots, x_n), (x_1^*, \dots, x_n^*))$ of $[a,b]$,

$$\left| \mathcal{R}(g, \mathcal{P}) - \text{p.c.} \int_a^b g(x)\, dx \right| \le 2Mk\, \Delta(\mathcal{P}).$$

Proof of Lemma. Write $\delta := \Delta(\mathcal{P})$. Every tag x_i^* lies in exactly one I_j, and then $g(x_i^*) = c_j$. Grouping the Riemann sum according to which I_j contains the tag,

$$\mathcal{R}(g, \mathcal{P}) = \sum_{j=1}^{k} c_j S_j, \qquad S_j := \sum_{i\,:\, x_i^* \in I_j} \delta x_i,$$

while p.c. $\int_a^b g = \sum_{j=1}^k c_j |I_j|$ by Exercise 1.1.20. So it suffices to prove $|S_j - |I_j|| \le 2\delta$ for each j.

For the upper bound: if $x_i^* \in I_j$ then $[x_{i-1}, x_i]$ contains a point of I_j and has length at most δ, so $[x_{i-1}, x_i]$ is contained in the interval $I_j^{+\delta}$ obtained by extending I_j by δ at each end, which has length $|I_j| + 2\delta$. The intervals $[x_{i-1}, x_i]$ have pairwise disjoint interiors, so the sum of the lengths of those occurring in S_j is at most $|I_j| + 2\delta$; that is, $S_j \le |I_j| + 2\delta$.

For the lower bound: let $I_j^{-\delta} := \{x \in I_j : \text{dist}(x, \mathbb{R} \setminus I_j) > \delta\}$, an open interval of length at least $|I_j| - 2\delta$ (possibly empty, in which case there is nothing to prove). If $[x_{i-1}, x_i]$ meets $I_j^{-\delta}$, then since $\delta x_i \le \delta$ every point of $[x_{i-1}, x_i]$ is within distance δ of a point of $I_j^{-\delta}$, whence $[x_{i-1}, x_i] \subset I_j$; in particular $x_i^* \in I_j$ and i contributes to S_j. The intervals $[x_{i-1}, x_i]$ meeting $I_j^{-\delta}$ cover $I_j^{-\delta}$, so the sum of their lengths is at least $|I_j^{-\delta}| \ge |I_j| - 2\delta$. Hence $S_j \ge |I_j| - 2\delta$, and the Lemma follows since $|c_j| \le M$ (as $c_j = g(x)$ for any $x \in I_j$; empty I_j contribute nothing). $\square$

Darboux integrable implies Riemann integrable. Suppose $\underline{\int_a^b} f = \overline{\int_a^b} f =: I$. Let $\varepsilon > 0$. By definition of the lower and upper Darboux integrals as a supremum and an infimum, there are piecewise constant $g \le f \le h$ on $[a,b]$ with

$$\text{p.c.} \int_a^b g > I - \varepsilon, \qquad \text{p.c.} \int_a^b h < I + \varepsilon.$$

Both g and h are bounded (a piecewise constant function takes finitely many values), so by the Lemma there is $\delta > 0$ such that every tagged partition $\mathcal{P}$ with $\Delta(\mathcal{P}) \le \delta$ satisfies

$$|\mathcal{R}(g, \mathcal{P}) - \text{p.c.} \int_a^b g| \le \varepsilon \quad \text{and} \quad |\mathcal{R}(h, \mathcal{P}) - \text{p.c.} \int_a^b h| \le \varepsilon.$$

Since $g \le f \le h$ pointwise and all $\delta x_i \ge 0$, we have $\mathcal{R}(g, \mathcal{P}) \le \mathcal{R}(f, \mathcal{P}) \le \mathcal{R}(h, \mathcal{P})$ for every tagged partition. Combining, for every $\mathcal{P}$ with $\Delta(\mathcal{P}) \le \delta$,

$$I - 2\varepsilon < \mathcal{R}(f, \mathcal{P}) < I + 2\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $\mathcal{R}(f, \mathcal{P}) \to I$ as $\Delta(\mathcal{P}) \to 0$, i.e. f is Riemann integrable with $\int_a^b f(x)\, dx = I$.

Riemann integrable implies Darboux integrable. Suppose f is Riemann integrable with integral I, and let $\varepsilon > 0$. Choose $\delta > 0$ so that $|\mathcal{R}(f, \mathcal{P}) - I| \le \varepsilon$ for every tagged partition with $\Delta(\mathcal{P}) \le \delta$. Fix n with $(b-a)/n \le \delta$ and let $x_i := a + i(b-a)/n$ be the uniform partition points; write $\delta x_i = (b-a)/n$. Put

$$M_i := \sup_{x \in [x_{i-1}, x_i]} f(x), \qquad m_i := \inf_{x \in [x_{i-1}, x_i]} f(x),$$

which are finite since f is bounded.

Choose tags $x_i^* \in [x_{i-1}, x_i]$ with $f(x_i^*) \ge M_i - \varepsilon/(b-a)$. Then

$$\sum_{i=1}^{n} M_i\, \delta x_i \le \mathcal{R}(f, \mathcal{P}) + \frac{\varepsilon}{b-a} \sum_{i=1}^{n} \delta x_i = \mathcal{R}(f, \mathcal{P}) + \varepsilon \le I + 2\varepsilon.$$

Now let h be the function equal to M_i on $(x_{i-1}, x_i]$ for $i = 1, \dots, n$ and equal to M_1 at the single point a. These $n+1$ intervals partition $[a, b]$, so h is piecewise constant, and $h \geq f$ pointwise by construction. Its piecewise constant integral is $\sum_i M_i\, \delta x_i$ (the singleton $\{a\}$ has length 0). Hence

$$\overline{\int_a^b} f(x)\, dx \leq \sum_{i=1}^{n} M_i\, \delta x_i \leq I + 2\varepsilon.$$

Symmetrically, choosing tags with $f(x_i^*) \leq m_i + \varepsilon/(b-a)$ and letting g equal m_i on $(x_{i-1}, x_i]$ and m_1 at a, we get $\underline{\int_a^b} f \geq \sum_i m_i \delta x_i \geq I - 2\varepsilon$.

Since $\varepsilon > 0$ was arbitrary and $\underline{\int_a^b} f \leq \overline{\int_a^b} f$ always, we conclude

$$I \leq \underline{\int_a^b} f(x)\, dx \leq \overline{\int_a^b} f(x)\, dx \leq I,$$

so f is Darboux integrable with Darboux integral I. Together with the previous direction, the two notions coincide and the two integrals agree.

Problem 1.1.23. Show that any continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable. More generally, show that any bounded, piecewise continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable. [A function $f : [a, b] \to \mathbb{R}$ is piecewise continuous if one can partition $[a, b]$ into finitely many intervals, such that f is continuous on each interval.]

Solution. By Exercise 1.1.22 it suffices in each case to verify Darboux integrability, i.e. that

$$\overline{\int_a^b} f(x)\, dx - \underline{\int_a^b} f(x)\, dx \leq \varepsilon$$

for every $\varepsilon > 0$. Throughout we use the following construction: given a partition of an interval J into finitely many intervals $J_1, \dots, J_n$ and a bounded f, the functions $\sum_i (\sup_{J_i} f) 1_{J_i}$ and $\sum_i (\inf_{J_i} f) 1_{J_i}$ are piecewise constant on J, sandwich f, and their piecewise constant integrals differ by $\sum_i (\sup_{J_i} f - \inf_{J_i} f)|J_i|$, by Exercise 1.1.20 and Exercise 1.1.21.

The continuous case. If $a = b$ there is nothing to prove, so assume $a < b$. A continuous function on the compact interval $[a, b]$ is bounded and uniformly continuous. Let $\varepsilon > 0$ and pick $\eta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever $x, y \in [a, b]$ with $|x - y| \leq \eta$. Choose n with $(b-a)/n \leq \eta$ and let $x_i := a + i(b-a)/n$. Partition $[a, b]$ into the intervals $\{a\}, (x_0, x_1], \dots, (x_{n-1}, x_n]$. On each of these, any two points are within η of each other, so

$$\sup_{J_i} f - \inf_{J_i} f \leq \varepsilon.$$

Hence with g, h the two piecewise constant functions above we have $g \leq f \leq h$ and

$$\text{p.c.}\int_a^b h - \text{p.c.}\int_a^b g \leq \varepsilon \sum_i |J_i| = \varepsilon(b-a).$$

Therefore $\overline{\int_a^b} f - \underline{\int_a^b} f \leq \varepsilon(b-a)$ for every $\varepsilon > 0$, so the upper and lower Darboux integrals coincide, and f is Riemann integrable by Exercise 1.1.22.

The piecewise continuous case. Let f be bounded, say $|f| \leq M$, and let $I_1, \dots, I_k$ be a partition of $[a, b]$ into intervals with $f|_{I_j}$ continuous for each j. Fix $\varepsilon > 0$.

Let I_j have endpoints $\alpha_j \leq \beta_j$. For $\eta > 0$ set

$$K_j := [\alpha_j + \eta,\ \beta_j - \eta] \cap I_j,$$

which is a compact interval contained in I_j (empty if $\beta_j - \alpha_j < 2\eta$; note $[\alpha_j + \eta, \beta_j - \eta] \subset (\alpha_j, \beta_j) \subset I_j$ automatically), and satisfies $|I_j| - |K_j| \leq 2\eta$. The restriction $f|_{K_j}$ is continuous on a compact interval,

hence uniformly continuous; so the continuous case above, applied on K_j with ε there replaced by $\varepsilon/(k(b-a)+1)$, produces piecewise constant functions $g_j \le f|_{K_j} \le h_j$ on K_j with

$$\text{p.c.}\int_{K_j} h_j - \text{p.c.}\int_{K_j} g_j \le \varepsilon/k.$$

The set $R := [a,b] \setminus \bigcup_{j=1}^k K_j$ is a finite union of intervals of total length at most $\sum_j (|I_j| - |K_j|) \le 2k\eta$: indeed the I_j partition $[a,b]$, so $R = \bigcup_j (I_j \setminus K_j)$, and each $I_j \setminus K_j$ is a union of at most two intervals of total length at most 2η.
Define g on $[a,b]$ to equal g_j on K_j and $-M$ on R, and h to equal h_j on K_j and $+M$ on R. Each of g, h is constant on each interval of a common refinement of the finitely many intervals involved, so both are piecewise constant on $[a,b]$; and $g \le f \le h$ pointwise, using $|f| \le M$ on R. By Exercise 1.1.20 the piecewise constant integral is additive over the pieces, so

$$\text{p.c.}\int_a^b h - \text{p.c.}\int_a^b g = \sum_{j=1}^k \Big(\text{p.c.}\int_{K_j} h_j - \text{p.c.}\int_{K_j} g_j\Big) + 2M\,|R| \le \varepsilon + 4Mk\eta.$$

Choosing $\eta \le \varepsilon/(4Mk+1)$ makes the right-hand side at most 2ε. As $\varepsilon > 0$ was arbitrary, the upper and lower Darboux integrals of f agree, so f is Darboux integrable, hence Riemann integrable by Exercise 1.1.22.

Problem 1.1.24. (Basic properties of the Riemann integral) Let $[a,b]$ be an interval, and let $f, g : [a,b] \to \mathbb{R}$ be Riemann integrable. Establish the following statements:

1. (Linearity) For any real number c, cf and $f+g$ are Riemann integrable, with $\int_a^b cf(x)\,dx = c \cdot \int_a^b f(x)\,dx$ and $\int_a^b f(x)+g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$.
2. (Monotonicity) If $f \le g$ pointwise (i.e. $f(x) \le g(x)$ for all $x \in [a,b]$) then $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$.
3. (Indicator) If E is a Jordan measurable subset of $[a,b]$, then the indicator function $1_E : [a,b] \to \mathbb{R}$ (defined by setting $1_E(x) := 1$ when $x \in E$ and $1_E(x) := 0$ otherwise) is Riemann integrable, and $\int_a^b 1_E(x)\,dx = m(E)$.

Finally, show that these properties uniquely define the Riemann integral, in the sense that the functional $f \mapsto \int_a^b f(x)\,dx$ is the only map from the space of Riemann integrable functions on $[a,b]$ to $\mathbb{R}$ which obeys all three of the above properties.

Solution. Throughout we work with the Darboux formulation, which is legitimate by Exercise 1.1.22: a bounded f is Riemann integrable exactly when $\underline{\int_a^b} f = \overline{\int_a^b} f$, and the common value is $\int_a^b f$. Recall also that Riemann integrable functions are bounded (Definition 1.1.5 and the remark following it), so all the Darboux quantities below are finite. We freely use Exercise 1.1.21 for linearity and monotonicity of the piecewise constant integral, and the fact that sums, scalar multiples, and pointwise maxima of piecewise constant functions are piecewise constant (pass to a common refinement of the two partitions).
Part 1: linearity.
First take $c > 0$. If $g_0 \le f$ is piecewise constant then $cg_0 \le cf$ is piecewise constant with $\text{p.c.}\int cg_0 = c\,\text{p.c.}\int g_0$; conversely if $g_1 \le cf$ is piecewise constant then $g_1/c \le f$ is. So the two sets of piecewise constant lower integrals correspond under multiplication by c, giving $\underline{\int} cf = c\underline{\int} f$, and symmetrically $\overline{\int} cf = c\overline{\int} f$. As the two Darboux integrals of f agree, so do those of cf, and $\int cf = c\int f$.
For $c = 0$ the function cf is identically 0, which is piecewise constant with integral 0; so $\int cf = 0 = c\int f$.
For $c = -1$ use the reflection identity recorded in the text: since $h \ge -f$ if and only if $-h \le f$, and $\text{p.c.}\int h = -\text{p.c.}\int(-h)$,

$$\overline{\int_a^b}(-f) = \inf_{h \ge -f} \text{p.c.}\int h = \inf_{-h \le f} \Big(-\text{p.c.}\int(-h)\Big) = -\sup_{g \le f} \text{p.c.}\int g = -\underline{\int_a^b} f,$$

and likewise $\underline{\int}(-f) = -\overline{\int} f$. Since $\underline{\int} f = \overline{\int} f$, the two Darboux integrals of $-f$ agree and $\int(-f) = -\int f$. A general $c < 0$ is handled by writing $cf = |c| \cdot (-f)$ and combining the two cases.

Now additivity. Let $g_1 \le f$ and $g_2 \le g$ be piecewise constant. Then $g_1 + g_2 \le f + g$ is piecewise constant with p.c. $\int (g_1 + g_2) = \text{p.c.} \int g_1 + \text{p.c.} \int g_2$. Taking suprema over g_1 and then g_2,

$$\underline{\int_a^b}(f+g) \ge \underline{\int_a^b} f + \underline{\int_a^b} g.$$

Symmetrically, using piecewise constant $h_1 \ge f$, $h_2 \ge g$,

$$\overline{\int_a^b}(f+g) \le \overline{\int_a^b} f + \overline{\int_a^b} g.$$

Since f and g are Darboux integrable, the right-hand sides both equal $\int f + \int g$, and therefore

$$\int_a^b f + \int_a^b g \le \underline{\int_a^b}(f+g) \le \overline{\int_a^b}(f+g) \le \int_a^b f + \int_a^b g.$$

So all four quantities coincide; $f+g$ (which is bounded) is Darboux integrable with $\int(f+g) = \int f + \int g$.

Part 2: monotonicity. Suppose $f \le g$ pointwise. Every piecewise constant $h \ge g$ also satisfies $h \ge f$, so the infimum defining $\overline{\int} f$ is over a larger set than that defining $\overline{\int} g$, giving $\overline{\int} f \le \overline{\int} g$. Since both functions are Darboux integrable this reads $\int_a^b f \le \int_a^b g$.

Part 3: indicators. Let $E \subset [a,b]$ be Jordan measurable and let $\varepsilon > 0$. By Exercise 1.1.5 there are elementary sets $A \subset E \subset B$ with $m(B \setminus A) \le \varepsilon$. Replacing B by $B \cap [a,b]$ (still elementary, still containing E, and only shrinking $B \setminus A$) we may assume $A \subset B \subset [a,b]$.

An elementary subset of $\mathbb{R}$ contained in $[a,b]$ is a finite union of intervals, and by splitting at the finitely many endpoints involved we may partition $[a,b]$ into finitely many intervals on each of which such an indicator is constant; so 1_A and 1_B are piecewise constant on $[a,b]$, with p.c. $\int_a^b 1_A = m(A)$ and p.c. $\int_a^b 1_B = m(B)$ by Exercise 1.1.21(3). Since $A \subset E \subset B$ we have $1_A \le 1_E \le 1_B$, so

$$m(A) \le \underline{\int_a^b} 1_E \le \overline{\int_a^b} 1_E \le m(B),$$

while also $m(A) \le m(E) \le m(B)$ by monotonicity of Jordan measure (Exercise 1.1.6). By finite additivity, $m(B) - m(A) = m(B \setminus A) \le \varepsilon$. Hence the upper and lower Darboux integrals of 1_E differ by at most ε, and each differs from $m(E)$ by at most ε. Letting $\varepsilon \to 0$, 1_E is Darboux integrable with $\int_a^b 1_E(x)\,dx = m(E)$, and so Riemann integrable with that value by Exercise 1.1.22.

Uniqueness. Let L be any map from the Riemann integrable functions on $[a,b]$ to $\mathbb{R}$ obeying properties (1), (2), (3). Note first that every piecewise constant function is Riemann integrable (it is bounded and piecewise continuous, so Exercise 1.1.23 applies), so L is defined on such functions.

Let g be piecewise constant, say $g \equiv c_j$ on the intervals $I_1, \dots, I_n$ partitioning $[a,b]$. Then $g = \sum_{j=1}^n c_j 1_{I_j}$ pointwise. Each I_j is an interval, hence an elementary and therefore Jordan measurable subset of $[a,b]$ with $m(I_j) = |I_j|$. By property (1) applied repeatedly, and then property (3),

$$L(g) = \sum_{j=1}^n c_j L(1_{I_j}) = \sum_{j=1}^n c_j |I_j| = \text{p.c.} \int_a^b g(x)\,dx.$$

Now let f be Riemann integrable and $\varepsilon > 0$. By Darboux integrability there are piecewise constant $g \le f \le h$ with

$$\text{p.c.} \int_a^b g \ge \int_a^b f - \varepsilon, \qquad \text{p.c.} \int_a^b h \le \int_a^b f + \varepsilon.$$

By property (2) applied twice, $L(g) \le L(f) \le L(h)$, i.e.

$$\int_a^b f(x)\,dx - \varepsilon \le L(f) \le \int_a^b f(x)\,dx + \varepsilon.$$

Letting $\varepsilon \to 0$ gives $L(f) = \int_a^b f(x)\,dx$, as required.

Problem 1.1.25. (Area interpretation of the Riemann integral) Let $[a,b]$ be an interval, and let $f : [a,b] \to \mathbb{R}$ be a bounded function. Show that f is Riemann integrable if and only if the sets $E_+ := \{(x,t) : x \in [a,b];\ 0 \leq t \leq f(x)\}$ and $E_- := \{(x,t) : x \in [a,b];\ f(x) \leq t \leq 0\}$ are both Jordan measurable in $\mathbb{R}^2$, in which case one has

$$\int_a^b f(x)\,dx = m^2(E_+) - m^2(E_-),$$

where m^2 denotes two-dimensional Jordan measure. [Hint: First establish this in the case when f is non-negative.]

Solution. The case $a = b$ is trivial (both sets are contained in a single vertical segment, which is a degenerate box of measure zero, and the integral is 0 by convention), so assume $a < b$. Fix M with $|f| \leq M$.

Step 0: a column decomposition of elementary sets. Let $A \subset [a,b] \times \mathbb{R}$ be an elementary subset of $\mathbb{R}^2$, say $A = \bigcup_{r=1}^N P_r \times Q_r$ with P_r, Q_r intervals; discarding any empty boxes we may assume every P_r and Q_r is non-empty, so that $P_r \times Q_r \subset A \subset [a,b] \times \mathbb{R}$ forces $P_r \subset [a,b]$. Let D be the finite set consisting of a, b, and all endpoints of the P_r lying in $[a,b]$. The points of D cut $[a,b]$ into finitely many intervals $I_1, \ldots, I_n$ (the singletons $\{d\}$ for $d \in D$, together with the open intervals between consecutive points of D) which partition $[a,b]$ and have the property that each $P_r \cap [a,b]$ is a union of some of them; in particular, for each j and r, either $I_j \subset P_r$ or $I_j \cap P_r = \emptyset$. Consequently

$$A \cap (I_j \times \mathbb{R}) = \bigcup_r (P_r \cap I_j) \times Q_r = I_j \times C_j, \qquad C_j := \bigcup_{r\,:\,I_j \subset P_r} Q_r,$$

and C_j is an elementary subset of $\mathbb{R}$. Note that C_j is exactly the vertical slice $A_x := \{t : (x,t) \in A\}$ for every $x \in I_j$. By finite additivity of elementary measure and Exercise 1.1.4,

$$m^2(A) = \sum_{j=1}^n m^2(I_j \times C_j) = \sum_{j=1}^n |I_j|\, m(C_j).$$

Step 1: the non-negative case. Assume $f \geq 0$. Then $E_- = \{(x,0) : f(x) = 0\}$ is contained in the degenerate box $[a,b] \times \{0\}$, which has elementary measure 0; hence $m^{*,(J)}(E_-) = 0$ and E_- is Jordan measurable with $m^2(E_-) = 0$. So we only have to analyse E_+. We claim

$$m_{*,(J)}(E_+) = \underline{\int_a^b} f(x)\,dx, \qquad m^{*,(J)}(E_+) = \overline{\int_a^b} f(x)\,dx.$$

Inner measure, lower bound. Let $g \leq f$ be piecewise constant; replacing g by $\max(g,0)$, which is still piecewise constant, still bounded above by f (as $f \geq 0$), and has at least as large a piecewise constant integral, we may assume $g \geq 0$. Write $g \equiv d_j \geq 0$ on intervals $I_1, \ldots, I_n$ partitioning $[a,b]$, and set $A := \bigcup_j I_j \times [0, d_j]$. This is elementary; the union is disjoint since the I_j are, so $m^2(A) = \sum_j d_j |I_j| = \text{p.c.} \int_a^b g$ by Exercise 1.1.4. If $x \in I_j$ and $0 \leq t \leq d_j$ then $0 \leq t \leq g(x) \leq f(x)$, so $A \subset E_+$. Hence $m_{*,(J)}(E_+) \geq \text{p.c.} \int g$, and taking the supremum over g, $m_{*,(J)}(E_+) \geq \underline{\int_a^b} f$.

Inner measure, upper bound. Let $A \subset E_+$ be elementary. Since $E_+ \subset [a,b] \times \mathbb{R}$, Step 0 applies: $m^2(A) = \sum_j |I_j|\, m(C_j)$ where $C_j = A_x$ for all $x \in I_j$. Fix j with $I_j \neq \emptyset$ and $x \in I_j$; from $A \subset E_+$ we get $C_j = A_x \subset [0, f(x)]$, so $m(C_j) \leq f(x)$ by monotonicity of elementary measure. Thus $m(C_j) \leq \inf_{I_j} f$, and the piecewise constant function $g := \sum_j m(C_j) 1_{I_j}$ satisfies $g \leq f$ on $[a,b]$ with $\text{p.c.} \int_a^b g = m^2(A)$. Hence $m^2(A) \leq \underline{\int_a^b} f$; taking the supremum over A gives $m_{*,(J)}(E_+) \leq \underline{\int_a^b} f$. Combined with the previous paragraph, $m_{*,(J)}(E_+) = \underline{\int_a^b} f$.

Outer measure, upper bound. Let $h \geq f$ be piecewise constant; then $h \geq f \geq 0$, so writing $h \equiv c_j \geq 0$ on intervals $I_1, \ldots, I_n$ partitioning $[a,b]$, the elementary set $B := \bigcup_j I_j \times [0, c_j]$ contains E_+ (if $x \in I_j$ and $0 \leq t \leq f(x)$ then $t \leq c_j$) and has $m^2(B) = \sum_j c_j |I_j| = \text{p.c.} \int h$. Hence $m^{*,(J)}(E_+) \leq \overline{\int_a^b} f$.

Outer measure, lower bound. Let $B \supset E_+$ be elementary. Since $E_+ \subset [a,b] \times [0,M]$, replacing B by $B \cap ([a,b] \times [0,M])$ keeps it elementary, keeps $E_+ \subset B$, and does not increase $m^2(B)$; so assume $B \subset [a,b] \times \mathbb{R}$. By Step 0, $m^2(B) = \sum_j |I_j| m(C_j)$ with $C_j = B_x$ for all $x \in I_j$. For $x \in I_j$ we have $[0, f(x)] \subset (E_+)_x \subset B_x = C_j$, so $f(x) \leq m(C_j)$. Thus $h := \sum_j m(C_j) 1_{I_j}$ is piecewise constant with $h \geq f$ and p.c. $\int_a^b h = m^2(B)$, giving $\overline{\int_a^b} f \leq m^2(B)$. Taking the infimum over B, $\overline{\int_a^b} f \leq m^{*,(J)}(E_+)$, and hence $m^{*,(J)}(E_+) = \overline{\int_a^b} f$.
Consequently, for $f \geq 0$: E_+ is Jordan measurable if and only if the two Darboux integrals of f agree, i.e. (Exercise 1.1.22) if and only if f is Riemann integrable, in which case $m^2(E_+) = \int_a^b f$. As E_- is always Jordan measurable with measure 0 here, the exercise holds for non-negative f.
Step 2: reduction of the general case. Write $f_+ := \max(f, 0)$ and $f_- := \max(-f, 0)$, so $f = f_+ - f_-$ and both are non-negative and bounded by M.
First, $E_+(f)$ and $E_+(f_+)$ differ only inside $[a,b] \times \{0\}$. Indeed for x with $f(x) \geq 0$ the two vertical slices are both $[0, f(x)]$, while for x with $f(x) < 0$ the slice of $E_+(f)$ is empty and that of $E_+(f_+)$ is $\{0\}$. So $E_+(f) \subset E_+(f_+)$ and $E_+(f_+) \setminus E_+(f) \subset [a,b] \times \{0\} =: Z$, a set with $m^2(Z) = 0$. Every subset N of Z satisfies $0 \leq m_{*,(J)}(N) \leq m^{*,(J)}(N) \leq m^{*,(J)}(Z) = 0$, so is Jordan measurable with measure zero. Hence by Boolean closure and additivity of Jordan measure (Exercise 1.1.6), each of $E_+(f)$, $E_+(f_+)$ is Jordan measurable if and only if the other is (write one as the union of the other with a Jordan null set, or as a difference of Jordan measurable sets), and then $m^2(E_+(f)) = m^2(E_+(f_+))$.
Second, the reflection $\sigma(x,t) := (x, -t)$ maps boxes to boxes of the same volume, hence maps elementary sets to elementary sets of the same measure, and therefore preserves Jordan inner and outer measure; in particular it preserves Jordan measurability and Jordan measure. Now $\sigma(E_-(f)) = \{(x,t) : x \in [a,b];\ 0 \leq t \leq -f(x)\}$, which by the same slice comparison as above differs from $E_+(f_-)$ only inside Z. Hence $E_-(f)$ is Jordan measurable if and only if $E_+(f_-)$ is, with the same measure.
Third, f is Riemann integrable if and only if both f_+ and f_- are. If f_+, f_- are Riemann integrable then so is $f = f_+ - f_-$, by linearity (Exercise 1.1.24(1)), with $\int f = \int f_+ - \int f_-$. Conversely suppose f is Riemann integrable and let $\varepsilon > 0$; by Exercise 1.1.22 choose piecewise constant $g \leq f \leq h$ with p.c. $\int (h - g) \leq \varepsilon$. The functions $g^+ := \max(g, 0)$ and $h^+ := \max(h, 0)$ are piecewise constant with $g^+ \leq f_+ \leq h^+$, and since $t \mapsto \max(t, 0)$ is non-decreasing and 1-Lipschitz we have $0 \leq h^+ - g^+ \leq h - g$ pointwise. By monotonicity and linearity of the piecewise constant integral (Exercise 1.1.21),

$$\text{p.c.} \int_a^b h^+ - \text{p.c.} \int_a^b g^+ \leq \text{p.c.} \int_a^b (h - g) \leq \varepsilon,$$

so $\overline{\int} f_+ - \underline{\int} f_+ \leq \varepsilon$ for all ε, and f_+ is Riemann integrable. Applying this to $-f$ (Riemann integrable by Exercise 1.1.24(1)) and noting $f_- = (-f)_+$ gives the integrability of f_-.
Step 3: conclusion. Combining the three points of Step 2 with Step 1:

f Riemann integrable $\iff$ f_+, f_- Riemann integrable $\iff$ $E_+(f_+), E_+(f_-)$ Jordan measurable $\iff$ E_+, E_-

and in that case

$$m^2(E_+) - m^2(E_-) = m^2(E_+(f_+)) - m^2(E_+(f_-)) = \int_a^b f_+(x)\,dx - \int_a^b f_-(x)\,dx = \int_a^b f(x)\,dx,$$

using Step 1 for the middle equality and linearity (Exercise 1.1.24(1)) for the last.

Problem 1.1.26. Extend the definition of the Riemann and Darboux integrals to higher dimensions, in such a way that analogues of all the previous results hold.

Solution. Fix $d \geq 1$. Throughout, a box means a box $B = \prod_{l=1}^d I_l$ in $\mathbb{R}^d$ as in Section 1.1, with volume $|B| = \prod_l |I_l|$.
Definitions.
A partition of a box B is a finite collection $\mathcal{Q} = \{B_1, \ldots, B_n\}$ of pairwise disjoint boxes with $\bigcup_i B_i =$

B. A tagged partition $\mathcal{P}$ of B is such a $\mathcal{Q}$ together with tags $x_i^* \in \overline{B_i} \cap B$. Its norm is $\Delta(\mathcal{P}) := \max_i \operatorname{diam}(B_i)$, and the Riemann sum of $f : B \to \mathbb{R}$ is $\mathcal{R}(f, \mathcal{P}) := \sum_{i=1}^n f(x_i^*)|B_i|$. We say f is Riemann integrable on B with integral $\int_B f$ if for every $\varepsilon > 0$ there is $\delta > 0$ with $|\mathcal{R}(f, \mathcal{P}) - \int_B f| \le \varepsilon$ for all tagged partitions $\mathcal{P}$ with $\Delta(\mathcal{P}) \le \delta$. If $|B| = 0$ we declare every f Riemann integrable with integral 0. This differs cosmetically from Definition 1.1.5 when $d = 1$, in that the pieces are taken disjoint rather than closed and overlapping at endpoints (a partition $a = x_0 < \cdots < x_n = b$ corresponds to the disjoint intervals $\{a\}, (x_0, x_1], \ldots, (x_{n-1}, x_n]$, with the same tags, now lying in the closures of the pieces). The two conventions define the same notion: the argument below shows that either one is equivalent to Darboux integrability with the same value, and the Darboux integral does not refer to tagged partitions at all.

A function $g : B \to \mathbb{R}$ is piecewise constant if there is a partition $\{B_1, \ldots, B_n\}$ of B with g constant, say $= c_i$, on each B_i; we set p.c. $\int_B g := \sum_i c_i|B_i|$. This is well defined: given two such partitions $\{B_i\}, \{B_j'\}$, the collection $\{B_i \cap B_j'\}$ is again a partition of B into boxes (an intersection of boxes is a box), and by finite additivity of elementary measure $|B_i| = \sum_j |B_i \cap B_j'|$; since $c_i = c_j'$ whenever $B_i \cap B_j' \neq \emptyset$, both $\sum_i c_i|B_i|$ and $\sum_j c_j'|B_j'|$ equal $\sum_{i,j} c_i|B_i \cap B_j'|$. This is the analogue of Exercise 1.1.20, and is in fact easier than in one dimension.

The lower and upper Darboux integrals of a bounded $f : B \to \mathbb{R}$ are

$$\underline{\int_B} f := \sup_{g \le f \text{ p.c.}} \text{p.c.} \int_B g, \qquad \overline{\int_B} f := \inf_{h \ge f \text{ p.c.}} \text{p.c.} \int_B h,$$

and f is Darboux integrable when they agree.

Finally, for a Jordan measurable $\Omega \subset \mathbb{R}^d$ and bounded $f : \Omega \to \mathbb{R}$, define $\int_\Omega f := \int_B \tilde{f}$ where $B \supset \Omega$ is any box and $\tilde{f}$ is f extended by 0; this is independent of B, since enlarging B only adds regions on which $\tilde{f}$ vanishes, and piecewise constant minorants and majorants can be taken to vanish there too.

Analogue of Exercise 1.1.21. Linearity and monotonicity of p.c. $\int_B$ follow exactly as in one dimension, passing to the common refinement $\{B_i \cap B_j'\}$ described above. For the indicator property, let $E \subset B$ be elementary, say a union of boxes $R_1, \ldots, R_N$. In each coordinate l, the endpoints of the l-th sides of $B, R_1, \ldots, R_N$ cut the l-th side of B into finitely many intervals; taking products gives a partition of B into finitely many boxes, each of which is, for each r, either contained in R_r or disjoint from it, and hence either contained in E or disjoint from E. Thus 1_E is piecewise constant, and by finite additivity p.c. $\int_B 1_E = m^d(E)$.

Analogue of Exercise 1.1.22 (Riemann equals Darboux). The key lemma again is that Riemann sums of a piecewise constant g, with $|g| \le M$ and $g \equiv c_j$ on the boxes $B_1, \ldots, B_k$, converge to p.c. $\int_B g$. Grouping by which B_j contains the tag, $\mathcal{R}(g, \mathcal{P}) = \sum_j c_j S_j$ with $S_j := \sum_{i : x_i^* \in B_j} |P_i|$, where $\{P_i\}$ are the boxes of $\mathcal{P}$; so it suffices to show $S_j \to |B_j|$ uniformly as $\delta := \Delta(\mathcal{P}) \to 0$. If $x_i^* \in B_j$ then $\operatorname{diam}(P_i) \le \delta$ forces P_i into the box $B_j^{+\delta}$ obtained by extending B_j by δ in each coordinate direction; the P_i are disjoint, so $S_j \le |B_j^{+\delta}| = \prod_l (s_{j,l} + 2\delta)$, where $s_{j,l}$ are the side lengths of B_j. Conversely let $B_j^{-2\delta} := \{x \in B_j : \operatorname{dist}(x, \mathbb{R}^d \setminus B_j) > 2\delta\}$, an open box of volume at least $\prod_l \max(s_{j,l} - 4\delta, 0)$. If P_i meets $B_j^{-2\delta}$ then, as $\operatorname{diam}(P_i) \le \delta$, we get $\overline{P_i} \subset B_j$ and in particular $x_i^* \in B_j$, so i contributes to S_j; and those P_i cover $B_j^{-2\delta}$, whence $S_j \ge |B_j^{-2\delta}|$. Both bounds tend to $|B_j|$ as $\delta \to 0$, uniformly in $\mathcal{P}$, proving the lemma.

Given the lemma, the two implications go exactly as in Exercise 1.1.22. If f is Darboux integrable with value I, sandwich f between piecewise constant $g \le f \le h$ with integrals within ε of I; then $\mathcal{R}(g, \mathcal{P}) \le \mathcal{R}(f, \mathcal{P}) \le \mathcal{R}(h, \mathcal{P})$ and the lemma pins $\mathcal{R}(f, \mathcal{P})$ within 2ε of I for small $\Delta(\mathcal{P})$. If instead f is Riemann integrable with value I, take the uniform grid partition $\{P_i\}$ of B into congruent boxes of diameter at most δ, put $M_i := \sup_{\overline{P_i} \cap B} f$ and choose tags with $f(x_i^*) \ge M_i - \varepsilon/(|B| + 1)$; then $h := \sum_i M_i 1_{P_i}$ is piecewise constant with $h \ge f$ and p.c. $\int_B h \le I + 2\varepsilon$, so $\overline{\int_B} f \le I + 2\varepsilon$, and symmetrically $\underline{\int_B} f \ge I - 2\varepsilon$.

Analogue of Exercise 1.1.23. If $f : B \to \mathbb{R}$ is continuous on the compact box $\overline{B}$ (or is continuous and bounded on B and extends continuously to $\overline{B}$), it is uniformly continuous there; given $\varepsilon > 0$ take the uniform grid partition of B into boxes of diameter at most the corresponding modulus of continuity, and the associated $g = \sum_i (\inf_{P_i} f) 1_{P_i}$ and $h = \sum_i (\sup_{P_i} f) 1_{P_i}$ satisfy $g \le f \le h$ with p.c. $\int_B (h - g) \le \varepsilon |B|$.

Calling f piecewise continuous when B can be partitioned into finitely many boxes $B_1, \dots, B_k$ with $f|_{B_j}$ continuous, the one-dimensional argument carries over verbatim: shrink each B_j to a compact box $K_j \subset B_j$ with $|B_j| - |K_j|$ small (shrink every side by η; the volumes converge as $\eta \to 0$), apply the continuous case on each K_j, and use the constants $\pm M$ on the leftover elementary set $B \setminus \bigcup_j K_j$, whose measure can be made arbitrarily small. So every bounded piecewise continuous f on B is Riemann integrable.

Analogue of Exercise 1.1.24. Linearity, monotonicity, and the indicator property $\int_B 1_E = m^d(E)$ for Jordan measurable $E \subset B$ all have the same proofs as in Exercise 1.1.24, working with the Darboux formulation: for the indicator one again invokes Exercise 1.1.5 to find elementary $A \subset E \subset B'$ with $m^d(B' \setminus A) \le \varepsilon$ (replacing B' by $B' \cap B$), and uses the indicator property of the piecewise constant integral established above. Uniqueness likewise: any functional obeying the three properties agrees with p.c. $\int_B$ on piecewise constant functions (which are finite linear combinations of indicators of boxes, and boxes are Jordan measurable), hence with $\int_B$ everywhere by squeezing. The same statements then transfer to any Jordan measurable domain Ω by the definition $\int_\Omega f = \int_B \tilde{f}$.

Analogue of Exercise 1.1.25. For bounded $f : B \to \mathbb{R}$, set $E_\pm \subset \mathbb{R}^{d+1}$ by $E_+ := \{(x,t) : x \in B;\ 0 \le t \le f(x)\}$ and $E_- := \{(x,t) : x \in B;\ f(x) \le t \le 0\}$. Then f is Riemann integrable on B if and only if E_+ and E_- are Jordan measurable in $\mathbb{R}^{d+1}$, in which case $\int_B f = m^{d+1}(E_+) - m^{d+1}(E_-)$. The proof of Exercise 1.1.25 goes through with the interval partitions of $[a,b]$ replaced by box partitions of B: Step 0 becomes the statement that any elementary $A \subset B \times \mathbb{R}$ can be written, after refining by the grid generated in the first d coordinates by the finitely many boxes constituting A and by B itself, as a disjoint union $\bigsqcup_j P_j \times C_j$ with $\{P_j\}$ a box partition of B and $C_j \subset \mathbb{R}$ elementary equal to the slice A_x for every $x \in P_j$, so that $m^{d+1}(A) = \sum_j |P_j|\, m^1(C_j)$ by Exercise 1.1.4. The rest of the argument (the four inequalities identifying $m_{*,(J)}(E_+)$ and $m^{*,(J)}(E_+)$ with the lower and upper Darboux integrals for $f \ge 0$; the reduction of general f to $f_\pm$ using that $B \times \{0\}$ is Jordan null and that the reflection $(x,t) \mapsto (x,-t)$ preserves Jordan measure) is unchanged.

Finally, the earlier structural results of Section 1.1 that were used remain available in this generality, since Exercises 1.1.4, 1.1.5 and 1.1.6 are already stated for arbitrary dimension.

2.2 §1.2 — Lebesgue Measure

Problem 1.2.1. Show that the countable union $\bigcup_{n=1}^\infty E_n$ or countable intersection $\bigcap_{n=1}^\infty E_n$ of Jordan measurable sets $E_1, E_2, \dots \subset \mathbb{R}$ need not be Jordan measurable, even when bounded.

Solution. Enumerate the rationals in $[0,1]$ as $\mathbb{Q} \cap [0,1] = \{q_1, q_2, q_3, \dots\}$ (this set is countably infinite). We first record two computations. Let $A \subset \mathbb{R}$ be an elementary set contained in $\mathbb{Q} \cap [0,1]$. By Definition 1.1.1, A is a finite union of intervals $I_1, \dots, I_k$, and each I_j is contained in $A \subset \mathbb{Q}$. A non-degenerate interval contains an irrational number, so every I_j is degenerate (a single point, or empty). Hence A is a finite set of points; discarding repetitions and writing A as the union of the disjoint degenerate intervals $\{x\}$, $x \in A$, each of length 0, Lemma 1.1.2(ii) gives $m(A) = 0$. Taking the supremum over all such A, the Jordan inner measure satisfies

$$m_{*,(J)}(\mathbb{Q} \cap [0,1]) = 0.$$

The same argument applies verbatim to any subset of $[0,1]$ containing no non-degenerate interval, in particular to $[0,1] \setminus \mathbb{Q}$, so $m_{*,(J)}([0,1] \setminus \mathbb{Q}) = 0$ as well. On the other hand both $\mathbb{Q} \cap [0,1]$ and $[0,1] \setminus \mathbb{Q}$ are dense in $[0,1]$, so each has closure $[0,1]$, and by Exercise 1.1.18(1) together with monotonicity of Jordan outer measure,

$$m^{*,(J)}(\mathbb{Q} \cap [0,1]) = m^{*,(J)}([0,1]) = 1, \qquad m^{*,(J)}([0,1] \setminus \mathbb{Q}) = m^{*,(J)}([0,1]) = 1.$$

(Here $m^{*,(J)}([0,1]) = 1$ because $[0,1]$ is elementary of measure 1.) Thus neither set is Jordan measurable, and both are bounded.

Countable unions. Put $E_n := \{q_n\}$. Each E_n is a degenerate interval, hence elementary, hence Jordan

measurable with $m(E_n) = 0$ (Exercise 1.1.5). Their union is

$$\bigcup_{n=1}^{\infty} E_n = \mathbb{Q} \cap [0,1],$$

which is bounded but, by the computation above, has Jordan inner measure 0 and Jordan outer measure 1, so it is not Jordan measurable. Hence a countable union of Jordan measurable sets need not be Jordan measurable.

Countable intersections. Put $E_n := [0,1] \setminus \{q_1, \ldots, q_n\}$. Removing finitely many points from the interval $[0,1]$ leaves a finite union of intervals, so E_n is elementary and therefore Jordan measurable, with $m(E_n) = 1$ (the removed points form an elementary set of measure 0, and elementary measure is finitely additive). Their intersection is

$$\bigcap_{n=1}^{\infty} E_n = [0,1] \setminus (\mathbb{Q} \cap [0,1]) = [0,1] \setminus \mathbb{Q},$$

which is bounded but has Jordan inner measure 0 and Jordan outer measure 1, hence is not Jordan measurable. Hence a countable intersection of Jordan measurable sets need not be Jordan measurable. (One can also obtain the intersection example from the union example by complementation within $[0,1]$: if $\bigcap_n E_n$ were Jordan measurable, then by Exercise 1.1.6(1) so would be $[0,1] \setminus \bigcap_n E_n = \mathbb{Q} \cap [0,1]$, a contradiction.)

Problem 1.2.2. Give an example of a sequence of uniformly bounded, Riemann integrable functions $f_n : [0,1] \to \mathbb{R}$ for $n = 1, 2, \ldots$ that converge pointwise to a bounded function $f : [0,1] \to \mathbb{R}$ that is not Riemann integrable. What happens if we replace pointwise convergence with uniform convergence?

Solution. The example. Enumerate $\mathbb{Q} \cap [0,1] = \{q_1, q_2, q_3, \ldots\}$ and set

$$f_n := 1_{\{q_1, \ldots, q_n\}} : [0,1] \to \mathbb{R},$$

the indicator of the first n rationals. The set $\{q_1, \ldots, q_n\}$ is a finite union of degenerate intervals, hence elementary, so by Exercise 1.1.21(3) each f_n is piecewise constant with p.c. $\int_0^1 f_n = m(\{q_1, \ldots, q_n\}) = 0$. A piecewise constant function is Darboux integrable: taking $g = f_n$ in the lower Darboux integral and $h = f_n$ in the upper one gives $\underline{\int_0^1} f_n \geq \text{p.c.} \int_0^1 f_n \geq \overline{\int_0^1} f_n$, while the reverse inequality $\underline{\int_0^1} f_n \leq \overline{\int_0^1} f_n$ holds for any bounded function by the monotonicity of the piecewise constant integral (Exercise 1.1.21(2)), since $g \leq f_n \leq h$ forces $\text{p.c.} \int_0^1 g \leq \text{p.c.} \int_0^1 h$. Hence f_n is Darboux integrable, so Riemann integrable by Exercise 1.1.22, with $\int_0^1 f_n = 0$. The family is uniformly bounded, since $0 \leq f_n \leq 1$ for every n.

The f_n converge pointwise to $f := 1_{\mathbb{Q} \cap [0,1]}$. Indeed, if $x \in [0,1]$ is irrational then $f_n(x) = 0 = f(x)$ for all n; if x is rational then $x = q_N$ for some N, and $f_n(x) = 1 = f(x)$ for all $n \geq N$. The limit f is bounded ($0 \leq f \leq 1$).

But f is not Riemann integrable. By Exercise 1.1.22 it suffices to show its lower and upper Darboux integrals differ. Let $g \leq f$ be piecewise constant, say $g = c_i$ on the intervals $I_1, \ldots, I_k$ partitioning $[0,1]$. If I_i is non-degenerate it contains an irrational point x, at which $c_i = g(x) \leq f(x) = 0$; if I_i is degenerate then $|I_i| = 0$. Hence $\text{p.c.} \int_0^1 g = \sum_i c_i |I_i| \leq 0$, and taking $g \equiv 0$ shows the bound is attained, so the lower Darboux integral of f is 0. Dually, let $h \geq f$ be piecewise constant, $h = c_i$ on $I_1, \ldots, I_k$. If I_i is non-degenerate it contains a rational point x, at which $c_i = h(x) \geq f(x) = 1$. Hence

$$\text{p.c.} \int_0^1 h = \sum_i c_i |I_i| \geq \sum_{I_i \text{ non-degenerate}} |I_i| = 1,$$

and taking $h \equiv 1$ shows this is attained, so the upper Darboux integral of f is 1. Since $0 \neq 1$, f is not Darboux integrable, hence not Riemann integrable.

Uniform convergence. If instead $f_n \to f$ uniformly on $[0,1]$ with each f_n Riemann integrable, then f is Riemann integrable and $\int_0^1 f = \lim_{n \to \infty} \int_0^1 f_n$. (In particular no counterexample exists, and the uniform boundedness hypothesis is automatic.)

To see this, let $\varepsilon > 0$ and choose N such that $|f_n(x) - f(x)| \le \varepsilon$ for all $x \in [0,1]$ and all $n \ge N$; fix any such n. Since f_n is Riemann integrable it is bounded, hence so is f, and the Darboux integrals of f are defined. If $g \le f_n$ is piecewise constant, then $g - \varepsilon \le f_n - \varepsilon \le f$ and $g - \varepsilon$ is piecewise constant, so by Exercise 1.1.21(1) and Definition 1.1.6,

$$\underline{\int_0^1} f \;\ge\; \text{p.c.}\int_0^1 g - \varepsilon.$$

Taking the supremum over such g, and using that f_n is Darboux integrable with Darboux integral $\int_0^1 f_n$, gives $\underline{\int_0^1} f \ge \int_0^1 f_n - \varepsilon$. Symmetrically, if $h \ge f_n$ is piecewise constant then $h + \varepsilon \ge f$, so $\overline{\int_0^1} f \le \int_0^1 f_n + \varepsilon$. Combining, and using that the lower Darboux integral never exceeds the upper one (Exercise 1.1.21(2)),

$$0 \;\le\; \overline{\int_0^1} f - \underline{\int_0^1} f \;\le\; 2\varepsilon.$$

The bound $\overline{\int_0^1} f - \underline{\int_0^1} f \le 2\varepsilon$ holds for every $\varepsilon > 0$ (the quantity on the left does not depend on ε or n), so the two Darboux integrals of f agree; thus f is Darboux integrable, hence Riemann integrable by Exercise 1.1.22, and we write $\int_0^1 f$ for the common value. Feeding this back into the two displayed inequalities, we get

$$\int_0^1 f_n - \varepsilon \;\le\; \int_0^1 f \;\le\; \int_0^1 f_n + \varepsilon \qquad \text{for every } n \ge N.$$

Since $\varepsilon > 0$ was arbitrary and N was chosen from it, this says exactly that $\int_0^1 f_n \to \int_0^1 f$.

Problem 1.2.3. (The outer measure axioms.)
- (i) (Empty set) $m^*(\emptyset) = 0$.
- (ii) (Monotonicity) If $E \subset F \subset \mathbb{R}^d$, then $m^*(E) \le m^*(F)$.
- (iii) (Countable subadditivity) If $E_1, E_2, \ldots \subset \mathbb{R}^d$ is a countable sequence of sets, then $m^*\left(\bigcup_{n=1}^\infty E_n\right) \le \sum_{n=1}^\infty m^*(E_n)$.

(Hint: Use the axiom of countable choice, Tonelli's theorem for series, and the $\varepsilon/2^n$ trick used previously to show that countable sets had outer measure zero.)

Solution. Throughout, recall the definition

$$m^*(E) := \inf_{\bigcup_{n=1}^\infty B_n \supset E;\ B_1, B_2, \ldots \text{ boxes}} \sum_{n=1}^\infty |B_n|,$$

the infimum of the total volume of a countable cover of E by boxes. Note that boxes are allowed to be degenerate or empty, and that every subset of $\mathbb{R}^d$ admits at least one such cover (cover $\mathbb{R}^d$ by unit cubes), so the infimum is over a non-empty set and m^* takes values in $[0, +\infty]$.

(i) Every term $|B_n|$ is non-negative, so $m^*(\emptyset) \ge 0$. Conversely, take $B_n := \emptyset$ for every n (or $B_n := \{0\}$, a degenerate box). These boxes cover $\emptyset$ and $\sum_{n=1}^\infty |B_n| = 0$, so $m^*(\emptyset) \le 0$. Hence $m^*(\emptyset) = 0$.

(ii) Let $E \subset F$. If $B_1, B_2, \ldots$ is any countable family of boxes with $\bigcup_n B_n \supset F$, then also $\bigcup_n B_n \supset E$. Thus every quantity $\sum_n |B_n|$ admissible in the infimum defining $m^*(F)$ is also admissible in the infimum defining $m^*(E)$; the latter infimum is over a larger set of numbers, so it is no larger. Hence $m^*(E) \le m^*(F)$.

(iii) Write $E := \bigcup_{n=1}^\infty E_n$. If $\sum_{n=1}^\infty m^*(E_n) = +\infty$ there is nothing to prove, so assume the sum is finite; in particular each $m^*(E_n)$ is finite.

Let $\varepsilon > 0$. For each n, since $m^*(E_n)$ is the infimum of a non-empty set of reals and $m^*(E_n) + \varepsilon/2^n > m^*(E_n)$, there exists a countable family of boxes $B_{n,1}, B_{n,2}, \ldots$ with

$$E_n \subset \bigcup_{m=1}^\infty B_{n,m}, \qquad \sum_{m=1}^\infty |B_{n,m}| \le m^*(E_n) + \frac{\varepsilon}{2^n}.$$

(Choosing one such family simultaneously for every n uses the axiom of countable choice.)
The doubly indexed collection $\{B_{n,m}\}_{n,m\geq 1}$ is a countable family of boxes, being a countable union of countable families, so it may be re-enumerated as a single sequence of boxes; and it covers E, since any $x \in E$ lies in some E_n and hence in some $B_{n,m}$. All the volumes $|B_{n,m}|$ are non-negative, so by Tonelli's theorem for series the total sum is independent of the enumeration and may be computed as an iterated sum:

$$\sum_{n,m} |B_{n,m}| \;=\; \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} |B_{n,m}| \;\leq\; \sum_{n=1}^{\infty}\left(m^*(E_n) + \frac{\varepsilon}{2^n}\right) \;=\; \sum_{n=1}^{\infty} m^*(E_n) + \varepsilon,$$

using $\sum_{n=1}^{\infty} \varepsilon/2^n = \varepsilon$. Since this re-enumerated family is an admissible cover of E, the definition of m^* gives

$$m^*(E) \;\leq\; \sum_{n=1}^{\infty} m^*(E_n) + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $m^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m^*(E_n)$, as claimed.
Finally, finite subadditivity follows: given $E_1, \ldots, E_k$, apply (iii) to the sequence $E_1, \ldots, E_k, \emptyset, \emptyset, \ldots$ and use (i).

Problem 1.2.4. Let $E, F \subset \mathbb{R}^d$ be disjoint closed sets, with at least one of E, F being compact. Show that $\operatorname{dist}(E, F) > 0$. Give a counterexample to show that this claim fails when the compactness hypothesis is dropped.

Solution. Recall $\operatorname{dist}(E, F) := \inf\{|x - y| : x \in E, y \in F\}$.
The positive result. If one of E, F is empty the infimum is over the empty set, hence equals $+\infty > 0$, and we are done; so assume both are non-empty. By symmetry we may assume E is compact.
Suppose for contradiction that $\operatorname{dist}(E, F) = 0$. Then for each $n \geq 1$ we may pick $x_n \in E$ and $y_n \in F$ with $|x_n - y_n| < 1/n$, so that $|x_n - y_n| \to 0$. Since E is compact it is bounded (indeed by the Heine-Borel theorem compactness is equivalent to being closed and bounded in $\mathbb{R}^d$), so the sequence (x_n) is bounded, and by the Bolzano-Weierstrass theorem it has a subsequence x_{n_k} converging to some $x \in \mathbb{R}^d$; as E is closed, $x \in E$. Then

$$|y_{n_k} - x| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - x| \longrightarrow 0,$$

so $y_{n_k} \to x$ as well; as F is closed and $y_{n_k} \in F$, we get $x \in F$. Hence $x \in E \cap F$, contradicting the disjointness of E and F. Therefore $\operatorname{dist}(E, F) > 0$.
(Equivalently: the function $x \mapsto \operatorname{dist}(x, F) = \inf_{y \in F} |x - y|$ is continuous, being 1-Lipschitz, and is strictly positive at every point of E because F is closed and disjoint from E; a continuous function on the compact set E attains its infimum, so that infimum, which equals $\operatorname{dist}(E, F)$, is strictly positive.)
Failure without compactness. Two closed sets, neither compact. In $\mathbb{R}$ take

$$E := \{1, 2, 3, \ldots\}, \qquad F := \left\{n + \tfrac{1}{2n} : n = 1, 2, 3, \ldots\right\}.$$

Both sets are closed: each meets any bounded subset of $\mathbb{R}$ in only finitely many points, so neither has a limit point in $\mathbb{R}$, and a set with no limit points is closed. They are disjoint, since $n + \frac{1}{2n}$ is never an integer (as $0 < \frac{1}{2n} < 1$). Neither is compact, being unbounded. Yet

$$\operatorname{dist}(E, F) \leq \inf_{n \geq 1} \left|\left(n + \tfrac{1}{2n}\right) - n\right| = \inf_{n \geq 1} \frac{1}{2n} = 0,$$

so $\operatorname{dist}(E, F) = 0$ even though $E \cap F = \emptyset$. Embedding these sets in the first coordinate axis, $E \times \{0\}^{d-1}$ and $F \times \{0\}^{d-1}$ give the same counterexample in $\mathbb{R}^d$.
For $d \geq 2$ one also has the picturesque example $E := \{(x, 0) : x \in \mathbb{R}\}$ (the horizontal axis) and $F := \{(x, 1/x) : x > 0\}$ (a branch of a hyperbola), both closed and disjoint with $\operatorname{dist}(E, F) \leq \inf_{x>0} 1/x = 0$. That F is closed follows since if $(x_n, 1/x_n) \to (a, b)$ with $x_n > 0$ then $a \geq 0$, and $a = 0$ would force $1/x_n \to +\infty$, contradicting convergence; so $a > 0$ and the limit $(a, 1/a)$ lies in F.

Problem 1.2.5. Show that if a set $E \subset \mathbb{R}^d$ is expressible as the countable union of almost disjoint boxes, then the Lebesgue outer measure of E is equal to the Jordan inner measure: $m^*(E) = m_{*,(J)}(E)$, where we extend the definition of Jordan inner measure to unbounded sets in the obvious manner.

Solution. Here the Jordan inner measure of an arbitrary (possibly unbounded) set $E \subset \mathbb{R}^d$ is

$$m_{*,(J)}(E) := \sup_{A \subset E,\ A \text{ elementary}} m(A) \ \in [0, +\infty],$$

the supremum now being allowed to take the value $+\infty$. Suppose $E = \bigcup_{n=1}^{\infty} B_n$ where $B_1, B_2, \ldots$ are almost disjoint boxes (their interiors are pairwise disjoint). By Lemma 1.2.9,

$$m^*(E) = \sum_{n=1}^{\infty} |B_n|.$$

The inequality $m_{*,(J)}(E) \leq m^*(E)$. This holds for any set E whatsoever. Indeed, let $A \subset E$ be elementary. By Lemma 1.2.6 the Lebesgue outer measure of an elementary set equals its elementary measure, so by monotonicity of Lebesgue outer measure (Exercise 1.2.3(ii)),

$$m(A) = m^*(A) \leq m^*(E).$$

Taking the supremum over all elementary $A \subset E$ gives $m_{*,(J)}(E) \leq m^*(E)$. (For bounded E this is the left-hand inequality of (1.2).)
The inequality $m^*(E) \leq m_{*,(J)}(E)$. Here we use the hypothesis. Fix a natural number N. The set $A_N := B_1 \cup \ldots \cup B_N$ is a finite union of boxes, hence elementary, and $A_N \subset E$. Since the boxes are almost disjoint, the finite additivity property (1.3) gives

$$m(A_N) = |B_1| + \ldots + |B_N|.$$

By definition of the Jordan inner measure as a supremum over elementary subsets of E,

$$m_{*,(J)}(E) \ \geq \ m(A_N) \ = \ \sum_{n=1}^{N} |B_n|.$$

This holds for every N, so letting $N \to \infty$ (the partial sums are non-decreasing, and converge in $[0, +\infty]$ to the full sum) we obtain

$$m_{*,(J)}(E) \ \geq \ \sum_{n=1}^{\infty} |B_n| \ = \ m^*(E).$$

In particular, if $\sum_n |B_n| = +\infty$ then both quantities are $+\infty$.
Combining the two inequalities, $m^*(E) = m_{*,(J)}(E)$, as required.
Combining this with Lemma 1.2.11, we obtain in particular that for every open set $U \subset \mathbb{R}^d$ the Lebesgue outer measure $m^*(U)$ coincides with the Jordan inner measure $m_{*,(J)}(U)$.

Problem 1.2.6. Give an example to show that the reverse statement

$$m^*(E) = \sup_{U \subset E,\ U \text{ open}} m^*(U)$$

is false. (For the corrected version of this statement, see Exercise 1.2.15.)

Solution. Take $d = 1$ and let $E := [0, 1] \setminus \mathbb{Q}$ be the set of irrationals in the unit interval.
First we compute $m^*(E) = 1$. By monotonicity (Exercise 1.2.3(ii)) and Lemma 1.2.6, $m^*(E) \leq$

$m^*([0,1]) = m([0,1]) = 1$. Conversely, $[0,1] = E \cup (\mathbb{Q} \cap [0,1])$, and $\mathbb{Q} \cap [0,1]$ is countable, hence has Lebesgue outer measure zero by Example 1.2.1; so by finite subadditivity (a consequence of Exercise 1.2.3(i),(iii)),

$$1 = m^*([0,1]) \leq m^*(E) + m^*(\mathbb{Q} \cap [0,1]) = m^*(E) + 0.$$

Hence $m^*(E) = 1$.

Now we compute the right-hand side. Let $U \subset E$ be open. If U were non-empty, it would contain an open interval (a,b) with $a < b$; but every non-degenerate interval contains a rational number, which does not lie in E, a contradiction. So the only open subset of E is $U = \emptyset$, and therefore

$$\sup_{U \subset E,\ U \text{ open}} m^*(U) = m^*(\emptyset) = 0$$

by Exercise 1.2.3(i).

Thus $m^*(E) = 1 \neq 0 = \sup_{U \subset E, U \text{ open}} m^*(U)$, so the reverse (inner regularity by open sets) statement fails.

Two remarks. First, the failure is not caused by non-measurability: $E = [0,1] \cap (\mathbb{R} \setminus \mathbb{Q})$ is Lebesgue measurable, being the intersection of a closed set with the complement of the countable, hence null, hence measurable, set $\mathbb{Q}$ (Lemma 1.2.13(ii),(iii),(v),(vii)). So even for measurable sets one cannot approximate from inside by open sets. Second, the essential obstruction is that E has empty interior while having positive outer measure; the correct inner regularity statement replaces open sets by compact sets and is the content of Exercise 1.2.15.

Problem 1.2.7. (Criteria for measurability.) Let $E \subset \mathbb{R}^d$. Show that the following are equivalent:

- (i) E is Lebesgue measurable.
- (ii) (Outer approximation by open) For every $\varepsilon > 0$, one can contain E in an open set U with $m^*(U \setminus E) \leq \varepsilon$.
- (iii) (Almost open) For every $\varepsilon > 0$, one can find an open set U such that $m^*(U \Delta E) \leq \varepsilon$. (In other words, E differs from an open set by a set of outer measure at most ε.)
- (iv) (Inner approximation by closed) For every $\varepsilon > 0$, one can find a closed set F contained in E with $m^*(E \setminus F) \leq \varepsilon$.
- (v) (Almost closed) For every $\varepsilon > 0$, one can find a closed set F such that $m^*(F \Delta E) \leq \varepsilon$. (In other words, E differs from a closed set by a set of outer measure at most ε.)
- (vi) (Almost measurable) For every $\varepsilon > 0$, one can find a Lebesgue measurable set E_ε such that $m^*(E_\varepsilon \Delta E) \leq \varepsilon$. (In other words, E differs from a measurable set by a set of outer measure at most ε.)

(Hint: Some of these deductions are either trivial or very easy. To deduce (i) from (vi), use the $\varepsilon/2^n$ trick to show that E is contained in a Lebesgue measurable set E'_ε with $m^*(E'_\varepsilon \Delta E) \leq \varepsilon$, and then take countable intersections to show that E differs from a Lebesgue measurable set by a null set.)

Solution. We prove the two chains

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(vi)} \Rightarrow \text{(i)}, \qquad \text{(i)} \Rightarrow \text{(iv)} \Rightarrow \text{(v)} \Rightarrow \text{(vi)},$$

which together with (vi) $\Rightarrow$ (i) give the equivalence of all six statements.

(i) $\Leftrightarrow$ (ii). Statement (ii) is verbatim Definition 1.2.2 of Lebesgue measurability, so there is nothing to prove.

(ii) $\Rightarrow$ (iii). Given $\varepsilon > 0$, take U open with $E \subset U$ and $m^*(U \setminus E) \leq \varepsilon$. Since $E \subset U$ we have $E \setminus U = \emptyset$, so $U \Delta E = (U \setminus E) \cup (E \setminus U) = U \setminus E$, and hence $m^*(U \Delta E) \leq \varepsilon$.

(iii) $\Rightarrow$ (vi). Every open set is Lebesgue measurable by Lemma 1.2.13(i), so the set U produced by (iii) may be taken as the set E_ε in (vi).

(i) $\Rightarrow$ (iv). Suppose E is Lebesgue measurable. By Lemma 1.2.13(v), the complement $E^c := \mathbb{R}^d \setminus E$ is also Lebesgue measurable. Let $\varepsilon > 0$. Applying Definition 1.2.2 to E^c, there is an open set $U \supset E^c$ with $m^*(U \setminus E^c) \leq \varepsilon$. Put $F := \mathbb{R}^d \setminus U$, a closed set. Taking complements in $E^c \subset U$ gives $F \subset E$. Moreover

$$E \setminus F = E \cap U = U \setminus E^c,$$

so $m^*(E \setminus F) = m^*(U \setminus E^c) \le \varepsilon$, which is (iv).
(iv) $\Rightarrow$ (v). If $F \subset E$ is closed with $m^*(E \setminus F) \le \varepsilon$, then $F \setminus E = \emptyset$, so $F \Delta E = E \setminus F$ and $m^*(F \Delta E) \le \varepsilon$.
(v) $\Rightarrow$ (vi). Every closed set is Lebesgue measurable by Lemma 1.2.13(ii), so the closed set F from (v) serves as E_ε.
(vi) $\Rightarrow$ (i). This is the substantial implication; we follow the hint in two steps.
Step 1: from (vi) we can always take the approximating measurable set to contain E. We claim that for every $\varepsilon > 0$ there is a Lebesgue measurable set E'_ε with

$$E \subset E'_\varepsilon \quad \text{and} \quad m^*(E'_\varepsilon \setminus E) \le \varepsilon$$

(so that $m^*(E'_\varepsilon \Delta E) = m^*(E'_\varepsilon \setminus E) \le \varepsilon$). Indeed, put $\delta := \varepsilon/3$ and use (vi) to find a Lebesgue measurable A with $m^*(A \Delta E) \le \delta$. In particular, by monotonicity,

$$m^*(E \setminus A) \le \delta \quad \text{and} \quad m^*(A \setminus E) \le \delta,$$

since both $E \setminus A$ and $A \setminus E$ are subsets of $A \Delta E$. As $m^*(E \setminus A) \le \delta$ is finite, outer regularity (Lemma 1.2.12) supplies an open set $V \supset E \setminus A$ with

$$m^*(V) \le m^*(E \setminus A) + \delta \le 2\delta.$$

Set $E'_\varepsilon := A \cup V$. It is Lebesgue measurable, being the union of two measurable sets (Lemma 1.2.13(i) and (vi)). It contains E, because $E = (E \cap A) \cup (E \setminus A) \subset A \cup V$. Finally

$$E'_\varepsilon \setminus E = (A \cup V) \setminus E \subset (A \setminus E) \cup V,$$

so by monotonicity and finite subadditivity of Lebesgue outer measure (Exercise 1.2.3),

$$m^*(E'_\varepsilon \setminus E) \le m^*(A \setminus E) + m^*(V) \le \delta + 2\delta = 3\delta = \varepsilon,$$

proving the claim.
Step 2: intersect. For each natural number $n \ge 1$, apply Step 1 with $\varepsilon = 1/n$ to obtain a Lebesgue measurable set $E'_{1/n} \supset E$ with $m^*(E'_{1/n} \setminus E) \le 1/n$, and set

$$E'' := \bigcap_{n=1}^{\infty} E'_{1/n}.$$

By Lemma 1.2.13(vii), E'' is Lebesgue measurable, and clearly $E \subset E''$ since E is contained in each $E'_{1/n}$. For every n we have $E'' \setminus E \subset E'_{1/n} \setminus E$, so by monotonicity

$$m^*(E'' \setminus E) \le \frac{1}{n} \qquad \text{for every } n \ge 1,$$

and letting $n \to \infty$ gives $m^*(E'' \setminus E) = 0$. Thus $N := E'' \setminus E$ is a null set, hence Lebesgue measurable by Lemma 1.2.13(iii).
Since $E \subset E''$ we have $E = E'' \setminus N = E'' \cap (\mathbb{R}^d \setminus N)$. By Lemma 1.2.13(v) the set $\mathbb{R}^d \setminus N$ is measurable, and by Lemma 1.2.13(vii) the intersection of the two measurable sets E'' and $\mathbb{R}^d \setminus N$ is measurable. Hence E is Lebesgue measurable, which is (i).
This closes both chains, so (i)-(vi) are all equivalent.
(Note where the hypotheses were used: the only non-formal ingredients are outer regularity (Lemma 1.2.12), used in Step 1 to replace the uncontrolled part $E \setminus A$ by an open set of small outer measure, and the closure properties of the class of Lebesgue measurable sets from Lemma 1.2.13.)

Problem 1.2.8. Show that every Jordan measurable set is Lebesgue measurable.

Solution. We first record that every elementary set is Lebesgue measurable. Indeed, let $B \subset \mathbb{R}^d$ be a box, and let B° be its interior, an open box. The difference $B \setminus B^\circ$ is contained in the boundary ∂B, which is a finite union of degenerate boxes (each with one side collapsed to a point), hence an elementary set of elementary measure 0; by Lemma 1.2.6 its Lebesgue outer measure is 0, and so by monotonicity $m^*(B \setminus B^\circ) = 0$. Thus $B = B^\circ \cup (B \setminus B^\circ)$ is the union of an open set and a null set, hence Lebesgue measurable by Lemma 1.2.13(i), (iii), (vi). An elementary set is a finite union of boxes, so it is Lebesgue measurable by Lemma 1.2.13(vi).
Now let $E \subset \mathbb{R}^d$ be Jordan measurable, so that E is bounded and

$$m_{*,(J)}(E) = m^{*,(J)}(E) = m(E) < \infty.$$

Let $\varepsilon > 0$. By the definitions of Jordan inner and outer measure we may find elementary sets A, B with

$$A \subset E \subset B, \qquad m(A) \geq m(E) - \varepsilon, \qquad m(B) \leq m(E) + \varepsilon.$$

Since $A \subset B$ are elementary, $B \setminus A$ is elementary and finite additivity of elementary measure gives

$$m(B \setminus A) = m(B) - m(A) \leq 2\varepsilon.$$

Now A is Lebesgue measurable by the first paragraph, and since $A \subset E \subset B$ we have

$$A \Delta E = E \setminus A \subset B \setminus A,$$

so by monotonicity of Lebesgue outer measure (Exercise 1.2.3) and Lemma 1.2.6,

$$m^*(A \Delta E) \leq m^*(B \setminus A) = m(B \setminus A) \leq 2\varepsilon.$$

Thus for every $\varepsilon > 0$ the set E differs from a Lebesgue measurable set by a set of Lebesgue outer measure at most 2ε. By the criterion (vi) $\implies$ (i) of Exercise 1.2.7, E is Lebesgue measurable.
(Incidentally, since m^* agrees with $m^{*,(J)}$ and $m_{*,(J)}$ on Jordan measurable sets by (1.2), the Lebesgue measure of E equals its Jordan measure, so Lebesgue measure genuinely extends Jordan measure.)

Problem 1.2.9. (Middle thirds Cantor set) Let $I_0 := [0,1]$ be the unit interval, let $I_1 := [0,1/3] \cup [2/3,1]$ be I_0 with the interior of the middle third interval removed, let $I_2 := [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]$ be I_1 with the interior of the middle third of each of the two intervals of I_1 removed, and so forth. More formally, write

$$I_n := \bigcup_{a_1,\ldots,a_n \in \{0,2\}} \left[\sum_{i=1}^n \frac{a_i}{3^i}, \sum_{i=1}^n \frac{a_i}{3^i} + \frac{1}{3^n} \right].$$

Let $C := \bigcap_{n=1}^\infty I_n$ be the intersection of all the elementary sets I_n. Show that C is compact, uncountable, and a null set.

Solution. Throughout, for a string $a = (a_1, \ldots, a_n) \in \{0,2\}^n$ write

$$p_a := \sum_{i=1}^n \frac{a_i}{3^i}, \qquad J_a := \left[p_a,\ p_a + 3^{-n}\right],$$

so that $I_n = \bigcup_{a \in \{0,2\}^n} J_a$ is a union of 2^n closed intervals of length 3^{-n}.
Compactness. Each J_a is closed and bounded, so each I_n, a finite union of closed sets, is closed; and $I_n \subset [0,1]$ is bounded. Hence $C = \bigcap_{n \geq 1} I_n$ is an intersection of closed sets, hence closed, and is bounded (being contained in $I_1 \subset [0,1]$). By the Heine-Borel theorem, C is compact.
Null set. The intervals J_a, $a \in \{0,2\}^n$, are almost disjoint: if $a \neq b$ are distinct strings of length n, let

i be the first index where they differ, say $a_i = 0$ and $b_i = 2$. Then

$$p_a + 3^{-n} \leq \sum_{j<i} \frac{a_j}{3^j} + \sum_{j>i} \frac{2}{3^j} + 3^{-n} = \sum_{j<i} \frac{a_j}{3^j} + 3^{-i} - 3^{-n} + 3^{-n} = \sum_{j<i} \frac{a_j}{3^j} + 3^{-i},$$

while $p_b \geq \sum_{j<i} \frac{a_j}{3^j} + 2 \cdot 3^{-i}$, so J_a lies entirely to the left of J_b (indeed with a gap of length at least 3^{-i}). In particular I_n is an elementary set of elementary measure

$$m(I_n) = 2^n \cdot 3^{-n} = (2/3)^n.$$

Since $C \subset I_n$, monotonicity of Lebesgue outer measure (Exercise 1.2.3) together with Lemma 1.2.6 gives

$$m^*(C) \leq m^*(I_n) = m(I_n) = (2/3)^n$$

for every n. Letting $n \to \infty$ yields $m^*(C) = 0$, i.e. C is a null set (and hence Lebesgue measurable, by Lemma 1.2.13(iii), with $m(C) = 0$).

Uncountability. Define $\Phi : \{0,2\}^{\mathbb{N}} \to [0,1]$ by

$$\Phi\big((a_i)_{i=1}^{\infty}\big) := \sum_{i=1}^{\infty} \frac{a_i}{3^i},$$

the series converging absolutely since $0 \leq a_i/3^i \leq 2 \cdot 3^{-i}$.

First, Φ takes values in C. Indeed, fix n and write $a^{(n)} := (a_1, \ldots, a_n)$. Then

$$p_{a^{(n)}} \leq \Phi(a) = p_{a^{(n)}} + \sum_{i>n} \frac{a_i}{3^i} \leq p_{a^{(n)}} + \sum_{i>n} \frac{2}{3^i} = p_{a^{(n)}} + 3^{-n},$$

so $\Phi(a) \in J_{a^{(n)}} \subset I_n$. As n was arbitrary, $\Phi(a) \in \bigcap_n I_n = C$.

Second, Φ is injective. Let $a \neq b$ in $\{0,2\}^{\mathbb{N}}$, and let i be the least index with $a_i \neq b_i$; say $a_i = 0$, $b_i = 2$. Put $P := \sum_{j<i} a_j 3^{-j} = \sum_{j<i} b_j 3^{-j}$. Then

$$\Phi(a) = P + \sum_{j>i} \frac{a_j}{3^j} \leq P + \sum_{j>i} \frac{2}{3^j} = P + 3^{-i},$$

while

$$\Phi(b) = P + \frac{2}{3^i} + \sum_{j>i} \frac{b_j}{3^j} \geq P + 2 \cdot 3^{-i} > P + 3^{-i} \geq \Phi(a).$$

Hence $\Phi(a) \neq \Phi(b)$.

Finally, $\{0,2\}^{\mathbb{N}}$ is uncountable: given any sequence $a^{(1)}, a^{(2)}, \ldots$ of elements of $\{0,2\}^{\mathbb{N}}$, the element c defined by $c_i := 2 - a_i^{(i)}$ lies in $\{0,2\}^{\mathbb{N}}$ and differs from $a^{(i)}$ in the i-th coordinate for every i, so no sequence exhausts $\{0,2\}^{\mathbb{N}}$. Since Φ is an injection of this uncountable set into C, the set C is uncountable.

(Thus C is an uncountable null set; in particular the converse of the observation that countable sets have Lebesgue outer measure zero is very far from true.)

Problem 1.2.10. (This exercise presumes some familiarity with point-set topology.) Show that the half-open interval $[0,1)$ cannot be expressed as the countable union of disjoint closed intervals. (Hint: It is easy to prevent $[0,1)$ from being expressed as the finite union of disjoint closed intervals. Next, assume for sake of contradiction that $[0,1)$ is the union of infinitely many closed intervals, and conclude that $[0,1)$ is homeomorphic to the middle thirds Cantor set, which is absurd. It is also possible to proceed using the Baire category theorem (Section 1.7 of *An epsilon of room, Vol. I.*) For an additional challenge, show that $[0,1)$ cannot be expressed as the countable union of disjoint closed sets.

Solution. We prove the stronger "additional challenge" statement, which contains the stated one:
Claim. $[0,1)$ cannot be written as the union of a countable family of pairwise disjoint nonempty closed subsets of $\mathbb{R}$, unless the family consists of a single set (and no single closed set equals $[0,1)$, since $[0,1)$ is not closed).
In particular $[0,1)$ is not a countable union of disjoint closed intervals: a nonempty closed interval $[a,b]$ is a closed set, a single one cannot equal $[0,1)$ (as 1 is a limit point of $[0,1)$ not lying in $[0,1)$, so $[0,1)$ is not closed), and empty members of the family may be discarded.
Step 0: reduction to a compact interval. Suppose for contradiction that

$$[0,1) = \bigsqcup_{n\in N} F_n$$

with N countable, the F_n nonempty, closed in $\mathbb{R}$, and pairwise disjoint, and with at least two indices. Pick $p \in F_1$ and $q \in F_2$ for two distinct indices; then $p \neq q$ by disjointness, and setting $c := \min(p,q)$, $d := \max(p,q)$ we get a nondegenerate compact interval $[c,d] \subset [0,1)$. Writing $C_n := F_n \cap [c,d]$, we obtain

$$[c,d] = \bigsqcup_{n\in N} C_n$$

as a countable disjoint union of closed (hence compact) sets, of which at least two are nonempty (namely those containing c and d). Discard the empty ones and re-index, so that all C_n are nonempty, n ranging over an index set of size at least 2. It therefore suffices to prove:
Sierpinski's theorem for an interval. A nondegenerate compact interval $[c,d]$ is not the union of a countable family of at least two pairwise disjoint nonempty closed sets.
Step 1: the finite case. Suppose $[c,d] = C_1 \sqcup \cdots \sqcup C_k$ with $k \geq 2$, the C_j nonempty, closed, pairwise disjoint. Then each C_j is also open in $[c,d]$, being the complement in $[c,d]$ of the closed set $\bigcup_{i\neq j} C_i$. Thus C_1 and $C_2 \cup \cdots \cup C_k$ are two disjoint nonempty relatively open sets covering $[c,d]$, contradicting the connectedness of the interval $[c,d]$.
Step 2: setup for the infinite case. Suppose now $[c,d] = \bigsqcup_{n=1}^{\infty} C_n$ with the C_n nonempty, closed and pairwise disjoint. All topological notions below (interior, boundary) are relative to the space $X := [c,d]$. Put

$$U := \bigcup_{n=1}^{\infty} \operatorname{int}_X(C_n), \qquad B := X \setminus U.$$

Since U is open in X, the set B is closed in X, hence compact, hence a complete metric space.
Step 3: B is nonempty. Suppose $B = \emptyset$, so that the relatively open sets $\operatorname{int}_X(C_n)$ cover the compact set X. Extract a finite subcover $\operatorname{int}_X(C_{n_1}), \ldots, \operatorname{int}_X(C_{n_k})$, discarding any empty members, so all are nonempty. If $k = 1$ then $X \subset C_{n_1}$, forcing all other C_n to be empty, contrary to hypothesis. If $k \geq 2$, then $\operatorname{int}_X(C_{n_1})$ and $\bigcup_{j\geq 2} \operatorname{int}_X(C_{n_j})$ are disjoint (by disjointness of the C_n), nonempty, relatively open, and cover X, contradicting connectedness of X. Hence $B \neq \emptyset$.
Step 4: each $B \cap C_n$ has empty interior in B. Suppose not: there is an index n and an open set $V \subset \mathbb{R}$ with

$$\emptyset \neq B \cap V \subset C_n.$$

Pick $x \in B \cap V$ and $r > 0$ with $(x-r, x+r) \subset V$. Note $x \in C_n$ but $x \notin \operatorname{int}_X(C_n)$ (as $x \in B$). Hence the relative neighbourhood $(x - r/2, x + r/2) \cap X$ is not contained in C_n: there exists

$$y \in (x - r/2, x + r/2) \cap X, \qquad y \notin C_n.$$

Then $y \in C_m$ for some $m \neq n$. Assume $y > x$ (the case $y < x$ is symmetric). Consider

$$z := \inf\{s \in [x,y] : s \in C_m\}.$$

The set $\{s \in [x,y] : s \in C_m\}$ is nonempty (it contains y) and closed, so z belongs to it: $z \in C_m$ and $z \in [x,y]$. Since $x \in C_n$ and $C_n \cap C_m = \emptyset$, we have $z > x$, and by definition of the infimum

$$[x,z) \cap C_m = \emptyset.$$

Consequently $z \notin \mathrm{int}_X(C_m)$: every relative neighbourhood of z in X contains points of $[x,z)$, which are not in C_m (here we use $z > x \geq c$, so points of $[x,z)$ arbitrarily close to z lie in X). Moreover, for $k \neq m$ we have $z \notin C_k$ (as $z \in C_m$ and the sets are disjoint), so certainly $z \notin \mathrm{int}_X(C_k)$. Therefore $z \notin U$, i.e. $z \in B$. But $z \in [x,y] \subset (x-r, x+r) \subset V$, so $z \in B \cap V \subset C_n$, contradicting $z \in C_m$ with $m \neq n$.

Step 5: Baire. By Steps 2 and 3, B is a nonempty compact, hence complete, metric space, and

$$B = \bigcup_{n=1}^{\infty} (B \cap C_n)$$

expresses B as a countable union of sets that are closed in B and, by Step 4, have empty interior in B. This contradicts the Baire category theorem, which asserts that a complete metric space is not the countable union of nowhere dense closed sets.

This contradiction completes Step 2-5, hence Sierpinski's theorem for the interval, hence the Claim, and with it the exercise: $[0,1)$ is not a countable union of disjoint closed intervals, nor even of disjoint closed sets.

Remark (sketch of the route suggested by the hint; the proof above is already complete and does not depend on this). For the special case of closed intervals one can avoid Baire's theorem in the above form and argue directly. If $[0,1) = \bigsqcup_n [a_n, b_n]$ with infinitely many disjoint closed intervals, let F be the set of all endpoints a_n, b_n; then $F = [0,1) \setminus \bigcup_n (a_n, b_n)$ is countable and relatively closed in $[0,1)$. One checks that no b_n has a right neighbourhood free of endpoints (if $(b_n, b_n+\varepsilon)$ contained no endpoint then, since two points of $(b_n, b_n+\varepsilon)$ lying in different intervals would force an endpoint between them, all of $(b_n, b_n+\varepsilon)$ would lie in a single $[a_m, b_m]$ with $a_m \leq b_n$, contradicting $b_n \notin [a_m, b_m]$), and symmetrically no $a_n > 0$ has a left neighbourhood free of endpoints. Choosing two of the intervals, one entirely to the left of the other, gives $b_m < a_k$ with $[b_m, a_k] \subset [0,1)$, and $F \cap [b_m, a_k]$ is then a nonempty, countable, compact set with no isolated points, which is impossible (a nonempty perfect compact metric space is uncountable). This exhibits $F \cap [b_m, a_k]$ as a Cantor-like set, which is the "homeomorphic to the middle thirds Cantor set" phenomenon alluded to in the hint.

Problem 1.2.11. (Monotone convergence theorem for measurable sets)

- (i) (Upward monotone convergence) Let $E_1 \subset E_2 \subset \cdots \subset \mathbb{R}^d$ be a countable non-decreasing sequence of Lebesgue measurable sets. Show that $m\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_n)$. (Hint: Express $\bigcup_{n=1}^{\infty} E_n$ as the countable union of the lacunae $E_n \setminus \bigcup_{n'=1}^{n-1} E_{n'}$.)
- (ii) (Downward monotone convergence) Let $\mathbb{R}^d \supset E_1 \supset E_2 \supset \ldots$ be a countable non-increasing sequence of Lebesgue measurable sets. If at least one of the $m(E_n)$ is finite, show that $m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_n)$.
- (iii) Give a counterexample to show that the hypothesis that at least one of the $m(E_n)$ is finite in the downward monotone convergence theorem cannot be dropped.

Solution. Throughout we use that the Lebesgue measurable sets are closed under countable unions, countable intersections and complements (Lemma 1.2.13), that m is countably additive on them (Lemma 1.2.15), and that m is monotone (this is monotonicity of m^*, Exercise 1.2.3).

(i). Set $E_0 := \emptyset$ and define the lacunae

$$A_n := E_n \setminus \bigcup_{n'=1}^{n-1} E_{n'} = E_n \setminus E_{n-1} \qquad (n \geq 1),$$

the second equality using the nesting $E_1 \subset E_2 \subset \ldots$. Each $A_n = E_n \cap (\mathbb{R}^d \setminus E_{n-1})$ is Lebesgue measurable by Lemma 1.2.13(v), (vii).

The A_n are pairwise disjoint: if $n < n'$ then $A_n \subset E_n \subset E_{n'-1}$ while $A_{n'} \cap E_{n'-1} = \emptyset$.

Moreover, for each $N \geq 1$,

$$\bigcup_{n=1}^{N} A_n = E_N.$$

Indeed "$\subset$" is clear since $A_n \subset E_n \subset E_N$ for $n \leq N$; conversely if $x \in E_N$, let $n \leq N$ be the least index with $x \in E_n$, so $x \notin E_{n-1}$ and hence $x \in A_n$. Letting $N \to \infty$ in the same argument gives $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_n$.
By countable additivity (Lemma 1.2.15(ii)) applied to the disjoint measurable sets A_n, and then finite additivity applied to $A_1, \ldots, A_N$,

$$m\left(\bigcup_{n=1}^{\infty} E_n\right) = m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n) = \lim_{N\to\infty} \sum_{n=1}^{N} m(A_n) = \lim_{N\to\infty} m(E_N),$$

where the third equality is the definition of an infinite sum of non-negative terms (the partial sums are non-decreasing, so the limit exists in $[0, +\infty]$). This is the claim; note that no finiteness hypothesis was needed. In particular the limit $\lim_N m(E_N)$ exists in $[0, +\infty]$, as it is a limit of a non-decreasing sequence (by monotonicity).
(ii). Suppose $m(E_{n_0}) < \infty$ for some n_0. Discarding $E_1, \ldots, E_{n_0-1}$ changes neither $\bigcap_n E_n$ nor $\lim_n m(E_n)$ (both depend only on the tail), so we may assume $m(E_1) < \infty$. By monotonicity, $m(E_n) \leq m(E_1) < \infty$ for all n, and also $m\left(\bigcap_n E_n\right) < \infty$.
Set $F_n := E_1 \setminus E_n$. These are Lebesgue measurable, and they are non-decreasing since the E_n are non-increasing. Their union is

$$\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} (E_1 \setminus E_n) = E_1 \setminus \bigcap_{n=1}^{\infty} E_n,$$

by de Morgan. Applying part (i) to the F_n,

$$m\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_1 \setminus E_n).$$

Now for any measurable $A \subset E_1$ we have the disjoint decomposition $E_1 = A \sqcup (E_1 \setminus A)$, so by finite additivity $m(E_1) = m(A) + m(E_1 \setminus A)$; since $m(E_1) < \infty$, all terms are finite and we may rearrange:

$$m(E_1 \setminus A) = m(E_1) - m(A).$$

Applying this with $A = E_n$ and with $A = \bigcap_n E_n$ (both measurable subsets of E_1), the displayed identity becomes

$$m(E_1) - m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \big(m(E_1) - m(E_n)\big) = m(E_1) - \lim_{n\to\infty} m(E_n),$$

the last step being legitimate because $m(E_1)$ is a finite real number (and $\lim_n m(E_n)$ exists, the sequence $m(E_n)$ being non-increasing and bounded below by 0). Cancelling the finite quantity $m(E_1)$ gives

$$m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_n).$$

(iii). Take $E_n := \{x \in \mathbb{R}^d : |x| \geq n\}$, a closed (hence Lebesgue measurable, by Lemma 1.2.13(ii)) non-increasing sequence. Each E_n contains boxes of arbitrarily large volume, so by monotonicity and Lemma 1.2.6, $m(E_n) = +\infty$ for every n; thus $\lim_{n\to\infty} m(E_n) = +\infty$. On the other hand $\bigcap_{n=1}^{\infty} E_n = \emptyset$, since any $x \in \mathbb{R}^d$ fails to lie in E_n once $n > |x|$, so

$$m\left(\bigcap_{n=1}^{\infty} E_n\right) = m(\emptyset) = 0 \neq +\infty.$$

Hence the finiteness hypothesis in (ii) cannot be dropped. (In dimension $d = 1$ one may take $E_n = [n, +\infty)$.)

Problem 1.2.12. Show that any map $E \mapsto m(E)$ from Lebesgue measurable sets to elements of $[0, +\infty]$ that obeys the above empty set and countable additivity axioms will also obey the monotonicity and countable subadditivity axioms from Exercise 1.2.3, when restricted to Lebesgue measurable sets of course.

Solution. Let μ be any map from the Lebesgue measurable subsets of $\mathbb{R}^d$ to $[0, +\infty]$ satisfying

- (Empty set) $\mu(\emptyset) = 0$;
- (Countable additivity) $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$ whenever $E_1, E_2, \ldots$ are pairwise disjoint Lebesgue measurable sets.

(We write μ rather than m to emphasise that no property of Lebesgue measure beyond these two axioms is being used. Recall from Lemma 1.2.13 that the Lebesgue measurable sets form a σ-algebra, so all the sets constructed below are measurable and μ may be applied to them. Sums of terms in $[0, +\infty]$ are always defined, as the partial sums are non-decreasing.)

Finite additivity. Let $E_1, \ldots, E_k$ be pairwise disjoint Lebesgue measurable sets. Extend the list by setting $E_n := \emptyset$ for $n > k$; the extended sequence is still pairwise disjoint and consists of measurable sets, and has the same union. Countable additivity and the empty set axiom give

$$\mu(E_1 \cup \cdots \cup E_k) = \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{k} \mu(E_n) + \sum_{n>k} \mu(\emptyset) = \sum_{n=1}^{k} \mu(E_n).$$

Monotonicity. Let $E \subset F$ be Lebesgue measurable. Then $F \setminus E = F \cap (\mathbb{R}^d \setminus E)$ is Lebesgue measurable by Lemma 1.2.13(v), (vii), and $F = E \sqcup (F \setminus E)$ is a disjoint union. By finite additivity,

$$\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E),$$

since $\mu(F \setminus E) \geq 0$ (this is valid in $[0, +\infty]$: if either term is $+\infty$ then $\mu(F) = +\infty$ and the inequality is trivial).

Countable subadditivity. Let $E_1, E_2, \ldots$ be arbitrary Lebesgue measurable sets (not assumed disjoint). Define the disjointified sets

$$A_n := E_n \setminus \bigcup_{n'=1}^{n-1} E_{n'} \qquad (n \geq 1),$$

which are Lebesgue measurable by Lemma 1.2.13(v), (vi), (vii). They are pairwise disjoint: if $n < n'$ then $A_n \subset E_n$ while $A_{n'} \cap E_n = \emptyset$. They satisfy $A_n \subset E_n$, and

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_n :$$

the inclusion "$\subset$" is clear, and conversely if $x \in \bigcup_n E_n$, taking n least with $x \in E_n$ gives $x \in A_n$. Hence by countable additivity followed by the monotonicity just proved (applied to $A_n \subset E_n$),

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \leq \sum_{n=1}^{\infty} \mu(E_n),$$

the last inequality holding termwise and hence for the (possibly infinite) sums. This is countable subadditivity, and finite subadditivity follows by padding with empty sets as before.

Problem 1.2.13. We say that a sequence E_n of sets in $\mathbb{R}^d$ converges pointwise to another set E in $\mathbb{R}^d$ if the indicator functions 1_{E_n} converge pointwise to 1_E.

- (i) Show that if the E_n are all Lebesgue measurable, and converge pointwise to E, then E is Lebesgue measurable also. (Hint: use the identity $1_E(x) = \liminf_{n\to\infty} 1_{E_n}(x)$ or $1_E(x) = \limsup_{n\to\infty} 1_{E_n}(x)$ to write E in terms of countable unions and intersections of the E_n.)

- (ii) (Dominated convergence theorem) Suppose that the E_n are all contained in another Lebesgue measurable set F of finite measure. Show that $m(E_n)$ converges to $m(E)$. (Hint: use the upward and downward monotone convergence theorems, Exercise 1.2.11.)
- (iii) Give a counterexample to show that the dominated convergence theorem fails if the E_n are not contained in a set of finite measure, even if we assume that the $m(E_n)$ are all uniformly bounded.

Solution. Introduce the sets

$$A_N := \bigcap_{n\geq N} E_n, \qquad B_N := \bigcup_{n\geq N} E_n \qquad (N \geq 1),$$

which are Lebesgue measurable by Lemma 1.2.13(vi), (vii). Note that A_N is non-decreasing in N, B_N is non-increasing in N, and $A_N \subset E_N \subset B_N$ for every N.
(i). Since the indicator functions take only the values 0 and 1, pointwise convergence $1_{E_n}(x) \to 1_E(x)$ means precisely that the sequence $1_{E_n}(x)$ is eventually constant with eventual value $1_E(x)$. Hence:

$$x \in E \iff 1_{E_n}(x) = 1 \text{ for all sufficiently large } n \iff x \in \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} E_n = \bigcup_{N=1}^{\infty} A_N.$$

(This is the identity $1_E = \liminf_n 1_{E_n}$ in set-theoretic form.) Thus

$$E = \bigcup_{N=1}^{\infty} A_N,$$

which is Lebesgue measurable, being a countable union of countable intersections of Lebesgue measurable sets (Lemma 1.2.13(vi), (vii)).
Symmetrically, $x \notin E$ iff $1_{E_n}(x) = 0$ for all sufficiently large n, i.e. iff x lies in only finitely many E_n; hence $x \in E$ iff x lies in infinitely many E_n, which gives the dual identity

$$E = \bigcap_{N=1}^{\infty} B_N$$

(the identity $1_E = \limsup_n 1_{E_n}$). We record both, as both are used in (ii).
(ii). Assume $E_n \subset F$ for all n with F Lebesgue measurable and $m(F) < \infty$. By (i), E is measurable. The sequence $A_1 \subset A_2 \subset \dots$ is non-decreasing measurable with union E, so by upward monotone convergence (Exercise 1.2.11(i)),

$$\lim_{N\to\infty} m(A_N) = m(E).$$

The sequence $B_1 \supset B_2 \supset \dots$ is non-increasing measurable with intersection E, and $B_1 = \bigcup_{n\geq 1} E_n \subset F$, so by monotonicity $m(B_1) \leq m(F) < \infty$. The finiteness hypothesis of downward monotone convergence (Exercise 1.2.11(ii)) is therefore met, and

$$\lim_{N\to\infty} m(B_N) = m(E).$$

Since $A_N \subset E_N \subset B_N$, monotonicity gives

$$m(A_N) \leq m(E_N) \leq m(B_N)$$

for every N, and both outer terms converge to $m(E)$. By the squeeze theorem, $m(E_N) \to m(E)$.
(iii). Work in $\mathbb{R}^d$ and set

$$E_n := [n, n+1] \times [0,1]^{d-1} \qquad (n = 1, 2, \dots),$$

a closed box, hence Lebesgue measurable, with $m(E_n) = 1$ for every n by Lemma 1.2.6. (For $d = 1$ this is just $E_n = [n, n+1]$.) The $m(E_n)$ are thus uniformly bounded.

The E_n converge pointwise to $E := \emptyset$: given $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, we have $x \notin E_n$ as soon as $n > x_1$, so $1_{E_n}(x) = 0$ for all large n, i.e. $1_{E_n}(x) \to 0 = 1_\emptyset(x)$.
But $m(E) = m(\emptyset) = 0$, while $m(E_n) = 1$ for all n, so $m(E_n)$ does not converge to $m(E)$. Of course the E_n here are not contained in any set of finite measure: any measurable F containing all of them satisfies $m(F) \geq m(E_1 \cup \cdots \cup E_N) = N$ for every N (the E_n being almost disjoint, or genuinely disjoint if one uses $[n, n+1)$ instead), hence $m(F) = +\infty$.

Problem 1.2.14. Let $E \subset \mathbb{R}^d$. Show that E is contained in a Lebesgue measurable set of measure exactly equal to $m^*(E)$.

Solution. We construct a G_δ set $A \supset E$ with $m(A) = m^*(E)$.
Case 1: $m^*(E) = +\infty$. Take $A := \mathbb{R}^d$. This is open, hence Lebesgue measurable by Lemma 1.2.13(i); it contains E; and it contains boxes of arbitrarily large volume, so by monotonicity of m^* (Exercise 1.2.3) and Lemma 1.2.6 we have $m(A) = m^*(\mathbb{R}^d) = +\infty = m^*(E)$.
Case 2: $m^*(E) < \infty$. By outer regularity (Lemma 1.2.12),

$$m^*(E) = \inf_{E \subset U,\ U \text{ open}} m^*(U),$$

so for each $n \geq 1$ we may choose an open set $U_n \supset E$ with

$$m^*(U_n) \leq m^*(E) + \frac{1}{n}.$$

Put

$$A := \bigcap_{n=1}^{\infty} U_n.$$

Each U_n is open, hence Lebesgue measurable (Lemma 1.2.13(i)), so A is Lebesgue measurable by Lemma 1.2.13(vii); and $E \subset U_n$ for every n gives $E \subset A$.
By monotonicity of Lebesgue outer measure (Exercise 1.2.3) applied to $E \subset A \subset U_n$,

$$m^*(E) \leq m(A) = m^*(A) \leq m^*(U_n) \leq m^*(E) + \frac{1}{n}$$

for every $n \geq 1$. Letting $n \to \infty$ (and using that $m^*(E)$ is finite, so that the squeeze is meaningful) we conclude

$$m(A) = m^*(E),$$

as required.
Remark. In both cases the set A produced is a G_δ set containing E; such an A is called a measurable hull of E. It is not in general unique, and one should not conclude that $m^*(A \setminus E) = 0$: that would force E itself to be measurable, since $E = A \cap (\mathbb{R}^d \setminus (A \setminus E))$ would then be an intersection of measurable sets (using Lemma 1.2.13(iii) for the null set $A \setminus E$, then (v) and (vii)), which is false for a non-measurable E. The subtraction $m(A) - m^*(E) = 0$ simply cannot be converted into a statement about $m^*(A \setminus E)$, since m^* is only subadditive, not additive, on non-measurable sets.

Problem 1.2.15. (Inner regularity) Let $E \subset \mathbb{R}^d$ be Lebesgue measurable. Show that

$$m(E) = \sup_{K \subset E,\ K \text{ compact}} m(K).$$

Solution. Write $S := \sup_{K \subset E,\, K \text{ compact}} m(K)$.
One inequality is immediate: every compact $K \subset E$ is closed, hence Lebesgue measurable by Lemma 1.2.13(ii), and monotonicity of Lebesgue outer measure (Exercise 1.2.3) gives $m(K) = m^*(K) \leq m^*(E) = m(E)$. Taking suprema, $S \leq m(E)$.

For the reverse inequality it suffices to show that $S > t$ for every real $t < m(E)$.
Fix such a t. Put $E_n := E \cap \overline{B}(0,n)$, where $\overline{B}(0,n)$ is the closed ball of radius n about the origin. Each E_n is Lebesgue measurable, being the intersection of the measurable set E with a closed (hence measurable) set, by Lemma 1.2.13(ii),(vii). The sets E_n increase and $\bigcup_{n=1}^{\infty} E_n = E$, so by upward monotone convergence (Exercise 1.2.11(i)),

$$m(E) = \lim_{n\to\infty} m(E_n).$$

Since $t < m(E)$, there is an n with $m(E_n) > t$. Moreover E_n is bounded, so it is contained in a box B, and monotonicity together with Lemma 1.2.6 gives $m(E_n) \le |B| < \infty$; thus $m(E_n)$ is finite.
Set $\varepsilon := m(E_n) - t > 0$. Since E_n is Lebesgue measurable, the criterion of inner approximation by closed sets (Exercise 1.2.7(iv)) supplies a closed set $F \subset E_n$ with $m^*(E_n \setminus F) \le \varepsilon/2$. The set F is closed and, being a subset of $\overline{B}(0,n)$, bounded; by the Heine-Borel theorem F is compact. Also $F \subset E_n \subset E$. Finally, $E_n \subset F \cup (E_n \setminus F)$, so finite subadditivity (Exercise 1.2.3) yields

$$m(E_n) \le m(F) + m^*(E_n \setminus F) \le m(F) + \varepsilon/2,$$

whence

$$m(F) \ge m(E_n) - \varepsilon/2 = t + \varepsilon/2 > t.$$

Thus $S \ge m(F) > t$. As $t < m(E)$ was arbitrary, $S \ge m(E)$ (this covers the case $m(E) = +\infty$ as well, since then t may be taken arbitrarily large). Combining the two inequalities gives $m(E) = S$, as claimed.

Problem 1.2.16. (Criteria for finite measure) Let $E \subset \mathbb{R}^d$. Show that the following are equivalent:
(i) E is Lebesgue measurable with finite measure.
(ii) (Outer approximation by open) For every $\varepsilon > 0$, one can contain E in an open set U of finite measure with $m^*(U \setminus E) \le \varepsilon$.
(iii) (Almost open bounded) E differs from a bounded open set by a set of arbitrarily small Lebesgue outer measure. (In other words, for every $\varepsilon > 0$ there exists a bounded open set U such that $m^*(E \Delta U) \le \varepsilon$.)
(iv) (Inner approximation by compact) For every $\varepsilon > 0$, one can find a compact set F contained in E with $m^*(E \setminus F) \le \varepsilon$.
(v) (Almost compact) E differs from a compact set by a set of arbitrarily small Lebesgue outer measure.
(vi) (Almost bounded measurable) E differs from a bounded Lebesgue measurable set by a set of arbitrarily small Lebesgue outer measure.
(vii) (Almost finite measure) E differs from a Lebesgue measurable set with finite measure by a set of arbitrarily small Lebesgue outer measure.
(viii) (Almost elementary) E differs from an elementary set by a set of arbitrarily small Lebesgue outer measure.
(ix) (Almost dyadically elementary) For every $\varepsilon > 0$, there exists an integer n and a finite union F of closed dyadic cubes of sidelength 2^{-n} such that $m^*(E \Delta F) \le \varepsilon$.
[One can interpret the equivalence of (i) and (ix) as asserting that Lebesgue measurable sets are those which look (locally) "pixelated" at sufficiently fine scales.]

Solution. Throughout we use two elementary remarks. First, for any sets A, B, C one has the triangle inequality for symmetric differences,

$$A \Delta C \subset (A \Delta B) \cup (B \Delta C),$$

since a point lying in exactly one of A, C must lie in exactly one of A, B or in exactly one of B, C. Combined with subadditivity (Exercise 1.2.3) this gives $m^*(A\Delta C) \le m^*(A\Delta B) + m^*(B\Delta C)$. Second, a bounded set S is contained in some box B, so $m^*(S) \le |B| < \infty$ by monotonicity and Lemma 1.2.6; hence a bounded measurable set automatically has finite measure.

We prove the two cycles

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(ix)} \Rightarrow \text{(viii)} \Rightarrow \text{(vii)} \Rightarrow \text{(i)}$$

and

$$\text{(i)} \Rightarrow \text{(iv)} \Rightarrow \text{(v)} \Rightarrow \text{(vi)} \Rightarrow \text{(vii)},$$

which together show all nine statements are equivalent.
(i) implies (ii). Let $\varepsilon > 0$. By outer regularity (Lemma 1.2.12) there is an open $U \supset E$ with $m^*(U) \leq m^*(E) + \varepsilon = m(E) + \varepsilon < \infty$. Open sets are measurable (Lemma 1.2.13(i)), so U has finite measure, and $U \setminus E$ is measurable by Lemma 1.2.13(v),(vii). Since U is the disjoint union of E and $U \setminus E$, finite additivity (Lemma 1.2.15) gives $m(U) = m(E) + m(U \setminus E)$; as $m(E)$ is finite we may subtract it and obtain $m^*(U \setminus E) = m(U) - m(E) \leq \varepsilon$.
(ii) implies (iii). Let $\varepsilon > 0$ and pick an open $U \supset E$ of finite measure with $m^*(U \setminus E) \leq \varepsilon/2$. Put $U_R := U \cap B(0,R)$ with $B(0,R)$ the open ball of radius R; each U_R is open and bounded, and $U_1 \subset U_2 \subset \cdots$ with $\bigcup_{R=1}^{\infty} U_R = U$. By upward monotone convergence (Exercise 1.2.11(i)), $m(U_R) \to m(U) < \infty$, so we may fix an integer R with $m(U \setminus U_R) = m(U) - m(U_R) \leq \varepsilon/2$ (the first equality is finite additivity, legitimate since $m(U_R) \leq m(U) < \infty$). Since $E \subset U$ we have $U_R \setminus E \subset U \setminus E$ and $E \setminus U_R \subset U \setminus U_R$, so

$$m^*(E \Delta U_R) \leq m^*(U \setminus E) + m(U \setminus U_R) \leq \varepsilon.$$

Thus U_R is a bounded open set as required.
(iii) implies (ix). Let $\varepsilon > 0$ and choose a bounded open U with $m^*(E \Delta U) \leq \varepsilon/2$. For each integer $n \geq 0$ let F_n be the union of all closed dyadic cubes of sidelength 2^{-n} that are contained in U. Since U is bounded, only finitely many dyadic cubes of a given sidelength meet U, so F_n is a finite union of closed dyadic cubes of sidelength 2^{-n}. If a dyadic cube of sidelength 2^{-n} lies in U then so does each of its 2^d children of sidelength 2^{-n-1}, whence $F_n \subset F_{n+1}$. Moreover $\bigcup_n F_n = U$: as noted in the proof of Lemma 1.2.11, every x in the open set U lies in some closed dyadic cube contained in U, and hence (passing to a descendant containing x) in one of sidelength 2^{-n} for all large n. The F_n are closed, hence measurable, and $m(U) < \infty$ because U is bounded, so upward monotone convergence gives $m(F_n) \to m(U)$ and therefore $m(U \setminus F_n) = m(U) - m(F_n) \leq \varepsilon/2$ for some n. Since $F_n \subset U$ we have $U \Delta F_n = U \setminus F_n$, and the triangle inequality gives

$$m^*(E \Delta F_n) \leq m^*(E \Delta U) + m(U \setminus F_n) \leq \varepsilon.$$

(If one insists that the union defining F in (ix) be non-empty, note that enlarging n only decreases $m(U \setminus F_n)$, and that adjoining to F_n one further closed dyadic cube of sidelength 2^{-n} increases $m^*(E \Delta F_n)$ by at most 2^{-nd}; so choosing n large enough handles the degenerate case $F_n = \emptyset$ as well.)
(ix) implies (viii). A finite union of closed dyadic cubes is a finite union of boxes, hence an elementary set.
(viii) implies (vii). Every elementary set is Jordan measurable, hence Lebesgue measurable by Exercise 1.2.8, and its Lebesgue measure equals its elementary measure by Lemma 1.2.6, which is finite. So an elementary set is a Lebesgue measurable set of finite measure.
(vii) implies (i). For each n pick a Lebesgue measurable S_n of finite measure with $m^*(E \Delta S_n) \leq 1/n$. In particular E differs from a measurable set by a set of arbitrarily small outer measure, so E is Lebesgue measurable by Exercise 1.2.7 ((vi) implies (i)). Furthermore $E \setminus S_1 \subset E \Delta S_1$, so $E \subset S_1 \cup (E \Delta S_1)$ and subadditivity gives

$$m(E) \leq m(S_1) + m^*(E \Delta S_1) \leq m(S_1) + 1 < \infty.$$

(i) implies (iv). Let $\varepsilon > 0$. By Exercise 1.2.7(iv) there is a closed set $F' \subset E$ with $m^*(E \setminus F') \leq \varepsilon/2$. The sets $E \cap \overline{B}(0,R)$ are measurable and increase to E, so by upward monotone convergence and the finiteness of $m(E)$ we may choose R with $m(E \setminus \overline{B}(0,R)) \leq \varepsilon/2$. Put $F := F' \cap \overline{B}(0,R)$; this is closed and bounded, hence compact by Heine-Borel, and $F \subset E$. Since

$$E \setminus F \subset (E \setminus F') \cup (E \setminus \overline{B}(0,R)),$$

subadditivity gives $m^*(E \setminus F) \leq \varepsilon$.
(iv) implies (v). If $F \subset E$ is compact with $m^*(E \setminus F) \leq \varepsilon$, then $E \Delta F = E \setminus F$, so $m^*(E \Delta F) \leq \varepsilon$.
(v) implies (vi). A compact set is closed and bounded, hence a bounded Lebesgue measurable set by Lemma 1.2.13(ii).
(vi) implies (vii). A bounded Lebesgue measurable set has finite measure, by the second remark at the start of the solution.
This completes both cycles, and hence the proof that (i) through (ix) are all equivalent.

Problem 1.2.17. (Caratheodory criterion, one direction) Let $E \subset \mathbb{R}^d$. Show that the following are equivalent:
(i) E is Lebesgue measurable.
(ii) For every elementary set A, one has $m(A) = m^*(A \cap E) + m^*(A \setminus E)$.
(iii) For every box B, one has $|B| = m^*(B \cap E) + m^*(B \setminus E)$.

Solution. We prove (i) implies (ii) implies (iii) implies (i).
(i) implies (ii). Let A be an elementary set. Then A is Lebesgue measurable (Exercise 1.2.8) and $m^*(A) = m(A)$ is finite (Lemma 1.2.6). If E is Lebesgue measurable then $A \cap E$ and $A \setminus E$ are Lebesgue measurable by Lemma 1.2.13(v),(vii); they are disjoint with union A, so finite additivity (Lemma 1.2.15) gives

$$m(A) = m(A \cap E) + m(A \setminus E) = m^*(A \cap E) + m^*(A \setminus E).$$

(ii) implies (iii). A box B is an elementary set whose elementary measure is $|B|$, so this is a special case.
(iii) implies (i). Since $E = \bigcup_{N=1}^{\infty} (E \cap B_N)$ with $B_N := [-N, N]^d$, and countable unions of measurable sets are measurable (Lemma 1.2.13(vi)), it suffices to show that $F := E \cap B_N$ is Lebesgue measurable for each fixed N. Write $B := B_N$ and note that $B \setminus E = B \setminus F$, so the hypothesis applied to the box B reads

$$|B| = m^*(F) + m^*(B \setminus F),$$

and both terms on the right are finite, being at most $|B|$ by monotonicity.
Let $\varepsilon > 0$. By outer regularity (Lemma 1.2.12) choose open sets

$$U \supset F, \qquad V \supset B \setminus F$$

with $m(U) \leq m^*(F) + \varepsilon$ and $m(V) \leq m^*(B \setminus F) + \varepsilon$; both are open, hence measurable (Lemma 1.2.13(i)), and both have finite measure.
Since $B = F \cup (B \setminus F)$, we have $B \subset U \cup V$ and hence $m(U \cup V) \geq m(B) = |B|$ by monotonicity and Lemma 1.2.6. On the other hand, decomposing $U \cup V = U \sqcup (V \setminus U)$ and $V = (V \cap U) \sqcup (V \setminus U)$ and using finite additivity for the measurable sets involved (all of finite measure),

$$m(U \cup V) + m(U \cap V) = m(U) + m(V).$$

Therefore

$$m(U \cap V) = m(U) + m(V) - m(U \cup V) \leq \big(m^*(F) + \varepsilon\big) + \big(m^*(B \setminus F) + \varepsilon\big) - |B| = 2\varepsilon.$$

The open set U contains F, but U itself may be far too large outside B, so we trim it. Choose $\delta > 0$ small enough that the open box $W := (-N - \delta, N + \delta)^d$ satisfies

$$m(W \setminus B) = m(W) - m(B) = (2N + 2\delta)^d - (2N)^d \leq \varepsilon,$$

which is possible by continuity of $\delta \mapsto (2N + 2\delta)^d$. Put $\tilde{U} := U \cap W$, an open set containing F (indeed $F \subset U$ and $F \subset B \subset W$). If $x \in \tilde{U} \setminus F$ then either $x \in B$, in which case $x \in U \cap (B \setminus F) \subset U \cap V$, or $x \notin B$, in which case $x \in W \setminus B$. Hence

$$\tilde{U} \setminus F \subset (U \cap V) \cup (W \setminus B),$$

and subadditivity gives

$$m^*(\tilde{U} \setminus F) \leq 2\varepsilon + \varepsilon = 3\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, F is contained in open sets whose difference from F has arbitrarily small outer measure; by Definition 1.2.2, $F = E \cap B_N$ is Lebesgue measurable. As explained above, E is then Lebesgue measurable.

Problem 1.2.18. (Inner measure) Let $E \subset \mathbb{R}^d$ be a bounded set. Define the Lebesgue inner measure $m_*(E)$ of E by the formula

$$m_*(E) := m(A) - m^*(A \setminus E)$$

for any elementary set A containing E.

(i) Show that this definition is well defined, i.e. that if A, A' are two elementary sets containing E, that $m(A) - m^*(A \setminus E)$ is equal to $m(A') - m^*(A' \setminus E)$.

(ii) Show that $m_*(E) \leq m^*(E)$, and that equality holds if and only if E is Lebesgue measurable.

Solution. Since E is bounded it is contained in some box, so elementary sets $A \supset E$ do exist and the definition is not vacuous. Note also that for an elementary set A we have $m^*(A) = m(A) < \infty$ by Lemma 1.2.6, and that elementary sets are Lebesgue measurable (Exercise 1.2.8).

(i) Suppose first that $A' \subset A$ are elementary sets both containing E. We claim

$$m^*(A \setminus E) = m^*(A' \setminus E) + m(A \setminus A').$$

Note that $A \setminus E$ is the disjoint union of $A' \setminus E$ and $A \setminus A'$ (a point of $A \setminus E$ either lies in A', and then in $A' \setminus E$, or not).

Upper bound: by Exercise 1.2.14 there is a Lebesgue measurable set $S \supset A' \setminus E$ with $m(S) = m^*(A' \setminus E)$; replacing S by $S \cap A'$ (still measurable, still containing $A' \setminus E$, and of measure between $m^*(A' \setminus E)$ and $m(S)$ by monotonicity) we may assume $S \subset A'$. Then $A \setminus E \subset S \cup (A \setminus A')$, so by subadditivity

$$m^*(A \setminus E) \leq m(S) + m(A \setminus A') = m^*(A' \setminus E) + m(A \setminus A').$$

Lower bound: by Exercise 1.2.14 again there is a measurable $T \supset A \setminus E$ with $m(T) = m^*(A \setminus E)$, and replacing T by $T \cap A$ we may assume $T \subset A$, so that $m(T) \leq m(A) < \infty$. Now $A \setminus A' \subset A \setminus E \subset T$, and $A \setminus A'$ is measurable, so finite additivity gives

$$m(T) = m(A \setminus A') + m\big(T \setminus (A \setminus A')\big).$$

The measurable set $T \setminus (A \setminus A')$ contains $A' \setminus E$, because $A' \setminus E \subset T$ and $A' \setminus E$ is disjoint from $A \setminus A'$. Hence by monotonicity

$$m^*(A' \setminus E) \leq m\big(T \setminus (A \setminus A')\big) = m^*(A \setminus E) - m(A \setminus A'),$$

which is the desired lower bound. This proves the claim.

Since A is the disjoint union of the elementary sets A' and $A \setminus A'$, finite additivity of elementary measure gives $m(A) = m(A') + m(A \setminus A')$, and therefore

$$m(A) - m^*(A \setminus E) = m(A') + m(A \setminus A') - m^*(A' \setminus E) - m(A \setminus A') = m(A') - m^*(A' \setminus E),$$

all quantities being finite. This settles the case $A' \subset A$.

For general elementary A, A' containing E, the set $A \cap A'$ is elementary and contains E, and it is contained in each of A and A'. Applying the case just proved twice,

$$m(A) - m^*(A \setminus E) = m(A \cap A') - m^*\big((A \cap A') \setminus E\big) = m(A') - m^*(A' \setminus E),$$

so the definition of $m_*(E)$ does not depend on the choice of A.

(ii) Fix an elementary set $A \supset E$. Since $A = E \cup (A \setminus E)$, subadditivity gives $m(A) = m^*(A) \leq m^*(E) + m^*(A \setminus E)$; as $m^*(A \setminus E) \leq m(A) < \infty$ we may rearrange to get

$$m_*(E) = m(A) - m^*(A \setminus E) \leq m^*(E).$$

Suppose now that E is Lebesgue measurable. Then $A \setminus E$ is measurable and A is the disjoint union of E and $A \setminus E$, so finite additivity (Lemma 1.2.15) gives $m(A) = m(E) + m(A \setminus E)$. All terms are finite (they are at most $m(A)$), so

$$m_*(E) = m(A) - m(A \setminus E) = m(E) = m^*(E),$$

giving equality.
Conversely suppose $m_*(E) = m^*(E)$, i.e.

$$m(A) = m^*(E) + m^*(A \setminus E),$$

with all three quantities finite. We argue as in Exercise 1.2.17. Let $\varepsilon > 0$. By outer regularity (Lemma 1.2.12) choose open sets $U \supset E$ and $V \supset A \setminus E$ with

$$m(U) \le m^*(E) + \varepsilon, \qquad m(V) \le m^*(A \setminus E) + \varepsilon,$$

both of finite measure. Since $A \subset U \cup V$ we have $m(U \cup V) \ge m(A)$, while finite additivity for measurable sets of finite measure gives $m(U \cup V) + m(U \cap V) = m(U) + m(V)$. Hence

$$m(U \cap V) = m(U) + m(V) - m(U \cup V) \le \big(m^*(E) + \varepsilon\big) + \big(m^*(A \setminus E) + \varepsilon\big) - m(A) = 2\varepsilon.$$

Next, again by outer regularity, choose an open set $O \supset A$ with $m(O) \le m(A) + \varepsilon$; since A is measurable of finite measure, $m(O \setminus A) = m(O) - m(A) \le \varepsilon$. Put $\tilde{U} := U \cap O$, an open set containing E (as $E \subset U$ and $E \subset A \subset O$). If $x \in \tilde{U} \setminus E$ then either $x \in A$, so that $x \in U \cap (A \setminus E) \subset U \cap V$, or $x \notin A$, so that $x \in O \setminus A$. Therefore

$$\tilde{U} \setminus E \subset (U \cap V) \cup (O \setminus A), \qquad m^*(\tilde{U} \setminus E) \le 2\varepsilon + \varepsilon = 3\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, E is Lebesgue measurable by Definition 1.2.2.

Problem 1.2.19. [Define a G_δ set to be a countable intersection $\bigcap_{n=1}^{\infty} U_n$ of open sets, and an F_σ set to be a countable union $\bigcup_{n=1}^{\infty} F_n$ of closed sets.]
Let $E \subset \mathbb{R}^d$. Show that the following are equivalent:
(i) E is Lebesgue measurable.
(ii) E is a G_δ set with a null set removed.
(iii) E is the union of an F_σ set and a null set.

Solution. We show (i) implies (ii), (i) implies (iii), and that each of (ii), (iii) implies (i).
(ii) implies (i). Suppose $E = G \setminus N$ where $G = \bigcap_{n=1}^{\infty} U_n$ is a G_δ set and N is null. Each U_n is Lebesgue measurable by Lemma 1.2.13(i), so G is Lebesgue measurable by Lemma 1.2.13(vii). The null set N is Lebesgue measurable by Lemma 1.2.13(iii), so $\mathbb{R}^d \setminus N$ is measurable by Lemma 1.2.13(v), and therefore $E = G \cap (\mathbb{R}^d \setminus N)$ is measurable by Lemma 1.2.13(vii).
(iii) implies (i). Suppose $E = F \cup N$ where $F = \bigcup_{n=1}^{\infty} F_n$ is an F_σ set and N is null. Each F_n is measurable by Lemma 1.2.13(ii), so F is measurable by Lemma 1.2.13(vi); N is measurable by Lemma 1.2.13(iii); and hence so is their union, again by Lemma 1.2.13(vi).
(i) implies (ii). Let E be Lebesgue measurable. By Definition 1.2.2, for each positive integer n there is an open set $U_n \supset E$ with $m^*(U_n \setminus E) \le 1/n$. Put $G := \bigcap_{n=1}^{\infty} U_n$, a G_δ set containing E. For every n we have $G \setminus E \subset U_n \setminus E$, so by monotonicity $m^*(G \setminus E) \le 1/n$; letting $n \to \infty$ gives $m^*(G \setminus E) = 0$, i.e. $N := G \setminus E$ is a null set. Since $E \subset G$ we have $E = G \setminus N$, which is a G_δ set with a null set removed.
(i) implies (iii). Let E be Lebesgue measurable. By the criterion of inner approximation by closed sets (Exercise 1.2.7(iv)), for each positive integer n there is a closed set $F_n \subset E$ with $m^*(E \setminus F_n) \le 1/n$. Put $F := \bigcup_{n=1}^{\infty} F_n$, an F_σ set with $F \subset E$. For every n we have $E \setminus F \subset E \setminus F_n$, so $m^*(E \setminus F) \le 1/n$ for all n, whence $N := E \setminus F$ is null. Since $F \subset E$ we get $E = F \cup N$, the union of an F_σ set and a null set.
(Alternatively, (i) implies (iii) follows from (i) implies (ii) applied to the complement: writing $\mathbb{R}^d \setminus E =$

$G \setminus N$ with G a G_δ and $N \subset G$ null, and taking complements, $E = (\mathbb{R}^d \setminus G) \cup N$, and the complement of a G_δ set is an F_σ set by de Morgan's laws.)

Problem 1.2.20. (Translation invariance) If $E \subset \mathbb{R}^d$ is Lebesgue measurable, show that $E + x$ is Lebesgue measurable for any $x \in \mathbb{R}^d$, and that $m(E + x) = m(E)$.

Solution. Fix $x \in \mathbb{R}^d$ and write $\tau_x(y) := y + x$, so that $E + x = \tau_x(E)$.
Step 1: Lebesgue outer measure is translation invariant, i.e. $m^*(S + x) = m^*(S)$ for every $S \subset \mathbb{R}^d$.
If $B = \prod_{i=1}^d I_i$ is a box, with each I_i an interval with endpoints $a_i \le b_i$, then $B + x = \prod_{i=1}^d (I_i + x_i)$ is again a box, with the i-th side an interval of the same type with endpoints $a_i + x_i \le b_i + x_i$. Hence $|B + x| = \prod_{i=1}^d (b_i - a_i) = |B|$.
Now let $S \subset \bigcup_{n=1}^\infty B_n$ be any cover of S by countably many boxes. Then $S+x \subset \bigcup_{n=1}^\infty (B_n+x)$ is a cover of $S+x$ by countably many boxes of the same total volume $\sum_n |B_n+x| = \sum_n |B_n|$. Taking the infimum over all such covers in the definition of Lebesgue outer measure gives $m^*(S+x) \le m^*(S)$. Applying this inequality to the set $S + x$ and the translation by $-x$ gives $m^*(S) = m^*((S + x) + (-x)) \le m^*(S + x)$, so the two are equal.
Step 2: $E + x$ is Lebesgue measurable.
Let $\varepsilon > 0$. Since E is Lebesgue measurable, Definition 1.2.2 gives an open set $U \supset E$ with $m^*(U \setminus E) \le \varepsilon$. The translation τ_x is a homeomorphism of $\mathbb{R}^d$ (it is a bijection, and both it and its inverse τ_{-x} are continuous), so $U + x$ is open; and clearly $E + x \subset U + x$. Moreover translation commutes with set difference,

$$(U + x) \setminus (E + x) = (U \setminus E) + x,$$

since $y + x$ lies in $U + x$ but not in $E + x$ exactly when $y \in U$ and $y \notin E$. By Step 1,

$$m^*\big((U + x) \setminus (E + x)\big) = m^*(U \setminus E) \le \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $E + x$ is Lebesgue measurable by Definition 1.2.2.
Step 3: the measures agree. Lebesgue measure of a measurable set is by definition its Lebesgue outer measure, so by Step 1,

$$m(E + x) = m^*(E + x) = m^*(E) = m(E). \qquad \square$$

Problem 1.2.21. (Change of variables) If $E \subset \mathbb{R}^d$ is Lebesgue measurable, and $T : \mathbb{R}^d \to \mathbb{R}^d$ is a linear transformation, show that $T(E)$ is Lebesgue measurable, and that $m(T(E)) = |\det T|\, m(E)$. [We caution that if $T : \mathbb{R}^d \to \mathbb{R}^{d'}$ is a linear map to a space $\mathbb{R}^{d'}$ of strictly smaller dimension than $\mathbb{R}^d$, then $T(E)$ need not be Lebesgue measurable; see Exercise 1.2.27.]

Solution. Throughout, $\|T\| := \sup_{|y|\le 1} |Ty|$ denotes the operator norm of T, which is finite, and we use the convention $0 \cdot \infty = 0$ (needed when T is singular and $m(E) = \infty$).
Lemma A (linear maps send null sets to null sets). If $N \subset \mathbb{R}^d$ is null, then $T(N)$ is null.
Proof. We may assume $T \neq 0$. First, if Q is a closed cube of sidelength r, then $\operatorname{diam} Q = r\sqrt{d}$, so $\operatorname{diam} T(Q) \le \|T\| r\sqrt{d}$; a set of diameter δ has each coordinate projection contained in an interval of length δ, hence is contained in a closed cube of sidelength δ. Therefore

$$m^*(T(Q)) \le \big(\|T\|\sqrt{d}\big)^d r^d = C|Q|, \qquad C := \big(\|T\|\sqrt{d}\big)^d.$$

Now let $\varepsilon > 0$. By outer regularity (Lemma 1.2.12) there is an open $U \supset N$ with $m^*(U) \le \varepsilon$. By Lemma 1.2.11, $U = \bigcup_n Q_n$ is a countable union of almost disjoint closed cubes, and by Lemma 1.2.9, $\sum_n |Q_n| = m^*(U) \le \varepsilon$. Since $T(N) \subset T(U) = \bigcup_n T(Q_n)$, countable subadditivity gives $m^*(T(N)) \le \sum_n C|Q_n| \le C\varepsilon$. Letting $\varepsilon \to 0$ gives $m^*(T(N)) = 0$. $\square$
Lemma B (proper subspaces are null). If $V \subsetneq \mathbb{R}^d$ is a proper linear subspace, then $m^*(V) = 0$. Consequently every affine hyperplane is null.
Proof. Choose a unit vector v orthogonal to V, so that $V \subset H := \{y : y \cdot v = 0\}$. Fix $R > 0$ and let

$H_R := H \cap [-R,R]^d$, $c := m^*(H_R) \le (2R)^d < \infty$. For $t \in \mathbb{R}$ the translate $H_R + tv$ lies in the hyperplane $\{y : y \cdot v = t\}$, and $m^*(H_R + tv) = c$ by translation invariance of outer measure (Step 1 of Exercise 1.2.20). If $s \neq t$ and $y \in H_R + sv$, $y' \in H_R + tv$, then by Cauchy-Schwarz $|y - y'| \ge |(y - y') \cdot v| = |t - s|$, so these translates are separated by positive distance. Fix $N \ge 1$ and set $S_k := H_R + (k/N)v$ for $k = 1, \dots, N$. Then $\operatorname{dist}(S_1 \cup \cdots \cup S_k, S_{k+1}) \ge 1/N > 0$, so applying Lemma 1.2.5 repeatedly,

$$m^*(S_1 \cup \cdots \cup S_N) = Nc.$$

On the other hand $S_1 \cup \cdots \cup S_N \subset [-R-1, R+1]^d$, so $Nc \le (2R+2)^d$ for every N, forcing $c = 0$. Since $H = \bigcup_{R=1}^{\infty} H_R$, countable subadditivity gives $m^*(H) = 0$, and monotonicity gives $m^*(V) = 0$. An affine hyperplane is a translate of a linear one, so it is null too, by translation invariance. $\square$

Lemma C (images of measurable sets). If E is Lebesgue measurable, then $T(E)$ is Lebesgue measurable.

Proof. By Exercise 1.2.19 ((i) implies (iii)), $E = F \cup N$ with $F = \bigcup_n F_n$ an F_σ set and N null. Each closed set F_n is the countable union of the compact sets $F_n \cap \overline{B}(0,k)$, so F is a countable union of compact sets K_j. As T is continuous, each $T(K_j)$ is compact, hence Lebesgue measurable by Lemma 1.2.13(ii). By Lemma A, $T(N)$ is null, hence measurable by Lemma 1.2.13(iii). Therefore $T(E) = \big(\bigcup_j T(K_j)\big) \cup T(N)$ is measurable by Lemma 1.2.13(vi). $\square$

Case 1: $\det T = 0$. Then $T(\mathbb{R}^d)$ has dimension at most $d-1$, so it is a proper subspace and hence null by Lemma B. Thus $T(E) \subset T(\mathbb{R}^d)$ is null, hence measurable with $m(T(E)) = 0$; and $|\det T|\, m(E) = 0 \cdot m(E) = 0$ under our convention. The identity holds.

Case 2: $\det T \neq 0$, so T is a linear bijection. Measurability of $T(E)$ is Lemma C, so only the measure identity remains.

Set $c := m\big(T([0,1)^d)\big)$, which is finite because $T([0,1)^d)$ is a bounded measurable set. Two remarks will be used repeatedly. First, for measurable E and $x \in \mathbb{R}^d$ we have $T(E + x) = T(E) + Tx$, so $m(T(E+x)) = m(T(E))$ by Exercise 1.2.20. Second, since T is injective, it carries disjoint sets to disjoint sets and commutes with unions and differences; combined with countable additivity (Lemma 1.2.15) this makes $E \mapsto m(T(E))$ countably additive on measurable sets.

Claim 1: if Q is a half-open dyadic cube of sidelength 2^{-n}, then $m(T(Q)) = c\,|Q|$.

Indeed, $[0,1)^d$ is the disjoint union of the 2^{nd} translates of $Q_0 := [0, 2^{-n})^d$ by vectors whose coordinates lie in $\{0, 2^{-n}, \dots, (2^n - 1)2^{-n}\}$. By the two remarks,

$$c = m\big(T([0,1)^d)\big) = \sum_{\text{translates}} m\big(T(Q_0 + x)\big) = 2^{nd}\, m(T(Q_0)),$$

so $m(T(Q_0)) = c\, 2^{-nd} = c|Q_0|$, and any Q as in the claim is a translate of Q_0.

Claim 1': the same holds for a closed dyadic cube Q of sidelength 2^{-n}. Let Q' denote the half-open dyadic cube of the same sidelength with the same lower corner, so $Q' \subset Q$ and $Q \setminus Q'$ is contained in the union of the $2d$ affine hyperplanes carrying the faces of Q, hence is null by Lemma B and subadditivity; by Lemma A, $T(Q \setminus Q')$ is null too. Since T is injective, $T(Q)$ is the disjoint union of $T(Q')$ and $T(Q \setminus Q')$, so $m(T(Q)) = c|Q'| = c|Q|$.

Claim 2: $m(T(U)) = c\, m(U)$ for every open $U \subset \mathbb{R}^d$.

By Lemma 1.2.11, $U = \bigcup_n Q_n$ with Q_n almost disjoint closed dyadic cubes, and by Lemma 1.2.9, $m(U) = \sum_n |Q_n|$. The interiors Q_n° are pairwise disjoint, and $Z := U \setminus \bigcup_n Q_n^\circ \subset \bigcup_n \partial Q_n$ is null by Lemma B and countable subadditivity. Since ∂Q_n is null we have $m(Q_n^\circ) = |Q_n|$ and, using Lemma A and Claim 1', $m(T(Q_n^\circ)) = m(T(Q_n)) = c|Q_n|$. As T is injective, $T(U)$ is the disjoint union of the sets $T(Q_n^\circ)$ together with $T(Z)$, and $T(Z)$ is null by Lemma A. Countable additivity now gives

$$m(T(U)) = \sum_n m(T(Q_n^\circ)) + m(T(Z)) = c \sum_n |Q_n| = c\, m(U).$$

Claim 3: $c > 0$. Suppose $c = 0$. Now $\mathbb{R}^d$ is the union of the countably many half-open unit cubes $[0,1)^d + k$, $k \in \mathbb{Z}^d$, so by the first remark and Claim 1 each $T([0,1)^d + k)$ has measure $c = 0$. Since T is surjective, $\mathbb{R}^d = T(\mathbb{R}^d) = \bigcup_{k \in \mathbb{Z}^d} T([0,1)^d + k)$, and countable subadditivity would give $m(\mathbb{R}^d) = 0$, contradicting $m([0,1]^d) = 1$ (Lemma 1.2.6) and monotonicity. Hence $c > 0$.

Claim 4: $m(T(E)) = c\, m(E)$ for every measurable E.

Suppose first $m(E) < \infty$. By Exercise 1.2.16 ((i) implies (ii)), for each n there is an open $U_n \supset E$ of finite measure with $m(U_n \setminus E) \le 1/n$; replacing U_n by $U_1 \cap \cdots \cap U_n$ we may assume $U_1 \supset U_2 \supset \cdots$. Let $G := \bigcap_n U_n$, so $E \subset G$ and $m(G \setminus E) \le m(U_n \setminus E) \le 1/n$ for all n, whence $G \setminus E$ is null and $m(G) = m(E)$. By Lemma A, $T(G \setminus E)$ is null, and since $T(G)$ is the disjoint union of $T(E)$ and $T(G \setminus E)$ we get $m(T(G)) = m(T(E))$. The sets $T(U_n)$ decrease with intersection $T(G)$ (using injectivity of T), and $m(T(U_1)) = c\,m(U_1) < \infty$ by Claim 2, so downward monotone convergence (Exercise 1.2.11(ii)) applies to both sequences and gives

$$m(T(E)) = m(T(G)) = \lim_{n\to\infty} m(T(U_n)) = \lim_{n\to\infty} c\,m(U_n) = c\,m(G) = c\,m(E).$$

If $m(E) = \infty$, apply the finite case to $E \cap B(0,k)$ and let $k \to \infty$: the sets $T(E \cap B(0,k))$ increase to $T(E)$, so upward monotone convergence (Exercise 1.2.11(i)) gives $m(T(E)) = \lim_k c\,m(E \cap B(0,k)) = \infty = c\,m(E)$, using $c > 0$.

Identification of the constant. For invertible S, T write $c(T)$ for the constant produced above. Applying Claim 4 twice, $m(ST(E)) = c(S)m(T(E)) = c(S)c(T)m(E)$, while also $m(ST(E)) = c(ST)m(E)$; taking $E = [0,1]^d$, whose measure is 1, gives the multiplicativity

$$c(ST) = c(S)\,c(T).$$

If T is orthogonal, then $|Ty| = |y|$ for all y, so T maps the closed unit ball B into itself, and onto it by surjectivity; thus $T(B) = B$ and Claim 4 gives $m(B) = c(T)m(B)$. Since $[-1/(2\sqrt{d}), 1/(2\sqrt{d})]^d \subset B \subset [-1,1]^d$ we have $0 < d^{-d/2} \le m(B) \le 2^d < \infty$, so $c(T) = 1$. This agrees with $|\det T| = 1$ for orthogonal T.

If $T = \operatorname{diag}(\sigma_1, \dots, \sigma_d)$ with all $\sigma_i > 0$, then $T([0,1)^d) = \prod_{i=1}^d [0, \sigma_i)$ is a box of volume $\prod_i \sigma_i$, so $c(T) = \prod_i \sigma_i = |\det T|$.

Finally, let T be an arbitrary invertible linear map. By the singular value decomposition, $T = U\Sigma V^{\mathsf{T}}$ where U, V are orthogonal and $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_d)$ with $\sigma_i > 0$ (strict positivity because T is invertible). By multiplicativity and the two computations above,

$$c(T) = c(U)\,c(\Sigma)\,c(V^{\mathsf{T}}) = 1 \cdot \prod_{i=1}^{d} \sigma_i \cdot 1 = \prod_{i=1}^{d} \sigma_i,$$

while $|\det T| = |\det U|\,|\det \Sigma|\,|\det V^{\mathsf{T}}| = \prod_{i=1}^d \sigma_i$ as well. Hence $c(T) = |\det T|$, and Claim 4 becomes

$$m(T(E)) = |\det T|\, m(E)$$

for every Lebesgue measurable E. Together with Case 1 this proves the exercise. □

Problem 1.2.22. Let $d, d' \ge 1$ be natural numbers.
(i) If $E \subset \mathbb{R}^d$ and $F \subset \mathbb{R}^{d'}$, show that $(m^{d+d'})^*(E \times F) \le (m^d)^*(E)(m^{d'})^*(F)$, where $(m^d)^*$ denotes d-dimensional Lebesgue measure, etc.
(ii) Let $E \subset \mathbb{R}^d$, $F \subset \mathbb{R}^{d'}$ be Lebesgue measurable sets. Show that $E \times F \subset \mathbb{R}^{d+d'}$ is Lebesgue measurable, with $m^{d+d'}(E \times F) = m^d(E) \cdot m^{d'}(F)$. (Note that we allow E or F to have infinite measure, and so one may have to divide into cases or take advantage of the monotone convergence theorem for Lebesgue measure, Exercise 1.2.11.)

Solution. Throughout we use the convention $0 \cdot \infty = \infty \cdot 0 = 0$, which is the standard convention in measure theory and the only one under which the assertions are true (see the degenerate case treated at the end of (i)).

Part (i). Write $a := (m^d)^*(E)$ and $b := (m^{d'})^*(F)$.

Suppose first that a and b are both finite. Let $\varepsilon > 0$. By the definition of Lebesgue outer measure as an infimum over countable coverings by boxes, we may find boxes $B_1, B_2, \dots \subset \mathbb{R}^d$ with $E \subset \bigcup_n B_n$ and $\sum_{n=1}^\infty |B_n| \le a + \varepsilon$, and boxes $B'_1, B'_2, \dots \subset \mathbb{R}^{d'}$ with $F \subset \bigcup_m B'_m$ and $\sum_{m=1}^\infty |B'_m| \le b + \varepsilon$.

The products $B_n \times B'_m$, for (n,m) ranging over the countable index set $\mathbb{N} \times \mathbb{N}$, are boxes in $\mathbb{R}^{d+d'}$:

indeed if $B_n = I_1 \times \cdots \times I_d$ and $B'_m = J_1 \times \cdots \times J_{d'}$ are products of intervals, then $B_n \times B'_m = I_1 \times \cdots \times I_d \times J_1 \times \cdots \times J_{d'}$ is again a product of intervals, and its volume is

$$|B_n \times B'_m| = \Big(\prod_{i=1}^{d} |I_i|\Big)\Big(\prod_{j=1}^{d'} |J_j|\Big) = |B_n|\,|B'_m|.$$

Moreover these products cover $E \times F$: if $x \in E$ and $y \in F$ then $x \in B_n$ for some n and $y \in B'_m$ for some m, so $(x, y) \in B_n \times B'_m$. Hence, by the definition of $(m^{d+d'})^*$ as an infimum over countable coverings by boxes, and by the fact that a double series of non-negative terms may be summed in any order,

$$(m^{d+d'})^*(E \times F) \le \sum_{n,m} |B_n|\,|B'_m| = \Big(\sum_n |B_n|\Big)\Big(\sum_m |B'_m|\Big) \le (a+\varepsilon)(b+\varepsilon).$$

Letting $\varepsilon \to 0$ gives $(m^{d+d'})^*(E \times F) \le ab$, as claimed.
Now suppose a or b is infinite. If the other quantity is strictly positive then the right-hand side ab is $+\infty$ and there is nothing to prove. The only remaining case is (after swapping the roles of the two factors if necessary) $a = 0$ and $b = +\infty$, where the convention $0 \cdot \infty = 0$ forces us to show that $E \times F$ is null. Since $F \subset \mathbb{R}^{d'} = \bigcup_{k=1}^{\infty} [-k, k]^{d'}$, we have

$$E \times F \subset \bigcup_{k=1}^{\infty} \big(E \times [-k, k]^{d'}\big),$$

and for each k the finite case treated above applies (with $a = 0$ and $(m^{d'})^*([-k, k]^{d'}) = (2k)^{d'} < \infty$) and gives $(m^{d+d'})^*(E \times [-k, k]^{d'}) \le 0 \cdot (2k)^{d'} = 0$. By countable subadditivity of Lebesgue outer measure, $(m^{d+d'})^*(E \times F) = 0 = ab$. This completes (i).
Part (ii), reduction to the bounded case. Set $E_k := E \cap [-k, k]^d$ and $F_k := F \cap [-k, k]^{d'}$ for $k = 1, 2, \ldots$. These are bounded Lebesgue measurable sets (intersections of measurable sets, Lemma 1.2.13), and $E_1 \times F_1 \subset E_2 \times F_2 \subset \cdots$ is a non-decreasing sequence of sets whose union is $E \times F$. So if we prove the claim for bounded measurable sets, then $E \times F$ is measurable by Lemma 1.2.13(vi), and by the upward monotone convergence theorem (Exercise 1.2.11(i)) applied three times,

$$m^{d+d'}(E \times F) = \lim_{k\to\infty} m^{d+d'}(E_k \times F_k) = \lim_{k\to\infty} m^d(E_k)\, m^{d'}(F_k),$$

while $m^d(E_k) \uparrow m^d(E)$ and $m^{d'}(F_k) \uparrow m^{d'}(F)$. It remains to check that if $\alpha_k \uparrow \alpha$ and $\beta_k \uparrow \beta$ in $[0, +\infty]$ then $\alpha_k \beta_k \to \alpha\beta$ with our convention. If $\alpha = 0$ then every $\alpha_k = 0$, so $\alpha_k\beta_k = 0 \to 0 = \alpha\beta$; symmetrically if $\beta = 0$. If $\alpha, \beta > 0$, pick k_0 with $\alpha_{k_0}, \beta_{k_0} > 0$; then for $k \ge k_0$ the products $\alpha_k\beta_k$ increase to $\alpha\beta$ (this is clear when α, β are finite, and when say $\alpha = \infty$ we have $\alpha_k\beta_k \ge \alpha_k\beta_{k_0} \to \infty = \alpha\beta$). So the limit is $m^d(E)m^{d'}(F)$, as required.
Part (ii), measurability in the bounded case. Assume now E, F are measurable and bounded; fix open boxes $W \subset \mathbb{R}^d$ and $W' \subset \mathbb{R}^{d'}$ with $E \subset W$, $F \subset W'$, and put $C := |W| + |W'| < \infty$.
Let $\varepsilon > 0$. Since E is Lebesgue measurable there is an open $U_0 \supset E$ with $m^*(U_0 \setminus E) \le \varepsilon$; replacing U_0 by $U := U_0 \cap W$ we obtain an open set U with $E \subset U \subset W$ and $m^*(U \setminus E) \le \varepsilon$ (as $U \setminus E \subset U_0 \setminus E$). Similarly choose open V with $F \subset V \subset W'$ and $m^*(V \setminus F) \le \varepsilon$.
Then $U \times V$ is an open subset of $\mathbb{R}^{d+d'}$ (a product of open sets is open) containing $E \times F$, and

$$(U \times V) \setminus (E \times F) \subset \big((U \setminus E) \times V\big) \cup \big(U \times (V \setminus F)\big) :$$

indeed if $(x, y) \in U \times V$ does not lie in $E \times F$, then either $x \notin E$, in which case $(x, y) \in (U \setminus E) \times V$, or else $x \in E$ and $y \notin F$, in which case $(x, y) \in U \times (V \setminus F)$. Hence by subadditivity of outer measure and part (i),

$$m^*\big((U \times V) \setminus (E \times F)\big) \le m^*(U \setminus E)\, m^*(V) + m^*(U)\, m^*(V \setminus F) \le \varepsilon\,|W'| + |W|\,\varepsilon = C\varepsilon.$$

As $\varepsilon > 0$ was arbitrary and C does not depend on ε, the set $E \times F$ can be contained in an open set whose difference with it has arbitrarily small outer measure; by the definition of Lebesgue measurability (equivalently Exercise 1.2.7(ii)), $E \times F$ is Lebesgue measurable.

Part (ii), the identity in the bounded case. Let $B \subset \mathbb{R}^d$ and $B' \subset \mathbb{R}^{d'}$ be closed boxes with $E \subset B$ and $F \subset B'$. The sets $B \setminus E$ and $B' \setminus F$ are bounded and measurable, so by the previous paragraph all three of

$$E \times F, \qquad (B \setminus E) \times F, \qquad B \times (B' \setminus F)$$

are measurable subsets of $\mathbb{R}^{d+d'}$. They are pairwise disjoint: the first two are disjoint in the first coordinate, and the third is disjoint from both because its points have second coordinate outside F. Their union is $B \times B'$: given $(x, y) \in B \times B'$, either $y \notin F$, giving a point of the third set, or $y \in F$, in which case $x \in E$ or $x \in B \setminus E$, giving a point of the first or second set. Since $B \times B'$ is a box of volume $|B||B'|$ (Lemma 1.2.6), finite additivity of Lebesgue measure (a special case of Lemma 1.2.15(ii), padding with empty sets) gives

$$|B|\,|B'| = m^{d+d'}(E \times F) + m^{d+d'}\big((B \setminus E) \times F\big) + m^{d+d'}\big(B \times (B' \setminus F)\big).$$

All quantities here are finite, being at most $|B||B'|$. By part (i) and the additivity relations $m^d(B \setminus E) = |B| - m^d(E)$, $m^{d'}(B' \setminus F) = |B'| - m^{d'}(F)$ (finite additivity again),

$$m^{d+d'}\big((B \setminus E) \times F\big) \le (|B| - m^d(E))\, m^{d'}(F), \qquad m^{d+d'}\big(B \times (B' \setminus F)\big) \le |B|\,(|B'| - m^{d'}(F)).$$

Substituting and rearranging,

$$\begin{aligned} m^{d+d'}(E \times F) &\ge |B||B'| - (|B| - m^d(E))m^{d'}(F) - |B|(|B'| - m^{d'}(F)) \\ &= |B||B'| - |B|m^{d'}(F) + m^d(E)m^{d'}(F) - |B||B'| + |B|m^{d'}(F) \\ &= m^d(E)\, m^{d'}(F). \end{aligned}$$

Together with the reverse inequality from part (i), this gives $m^{d+d'}(E \times F) = m^d(E)m^{d'}(F)$ for bounded measurable E, F, and by the reduction above the same identity holds in general.

Problem 1.2.23. (Uniqueness of Lebesgue measure) Show that Lebesgue measure $E \mapsto m(E)$ is the only map from Lebesgue measurable sets to $[0, +\infty]$ that obeys the following axioms:
(i) (Empty set) $m(\emptyset) = 0$.
(ii) (Countable additivity) If $E_1, E_2, \ldots \subset \mathbb{R}^d$ is a countable sequence of disjoint Lebesgue measurable sets, then $m(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n)$.
(iii) (Translation invariance) If E is Lebesgue measurable and $x \in \mathbb{R}^d$, then $m(E + x) = m(E)$.
(iv) (Normalisation) $m([0,1]^d) = 1$.
Hint: First show that m must match elementary measure on elementary sets, then show that m is bounded by outer measure.

Solution. Lebesgue measure itself obeys these axioms: (i) and (ii) are Lemma 1.2.15, (iii) is Exercise 1.2.20, and (iv) holds because $[0,1]^d$ is a box of volume 1 (Lemma 1.2.6). So the content of the exercise is uniqueness. Let μ be any map from the Lebesgue measurable subsets of $\mathbb{R}^d$ to $[0, +\infty]$ obeying (i)-(iv); we must show $\mu(E) = m(E)$ for every measurable E.
Two consequences of the axioms will be used constantly. First, μ is finitely additive: given finitely many disjoint measurable $E_1, \ldots, E_N$, apply (ii) to the sequence $E_1, \ldots, E_N, \emptyset, \emptyset, \ldots$ and use (i). Second, by Exercise 1.2.12, μ is also monotone and countably subadditive on measurable sets.
Step 1: μ is a finite positive multiple of volume on half-open unit cubes. Write $c := \mu([0,1)^d)$. By monotonicity $c \le \mu([0,1]^d) = 1$, so c is finite. For a lower bound, decompose $[0,1]^d$ according to which coordinates equal 1: for $S \subset \{1, \ldots, d\}$ let $P_S := \{x \in [0,1]^d : x_i = 1 \text{ for } i \in S,\ x_i \in [0,1) \text{ for } i \notin S\}$. These 2^d sets cover $[0,1]^d$, and $P_S \subset e_S + [0,1)^d$ where $e_S := \sum_{i \in S} e_i$ (a coordinate equal to 1 lies in $[1,2)$). Hence by subadditivity, monotonicity and translation invariance,

$$1 = \mu([0,1]^d) \le \sum_{S \subset \{1,\ldots,d\}} \mu(e_S + [0,1)^d) = 2^d c,$$

so $c \ge 2^{-d} > 0$.

Step 2: $\mu(B) = c|B|$ for every box B. Fix $n \geq 0$. The cube $[0,1)^d$ is the disjoint union of the 2^{nd} half-open dyadic cubes $\prod_{i=1}^d [j_i 2^{-n}, (j_i+1)2^{-n})$ with $0 \leq j_i < 2^n$, each of which is a translate of $[0, 2^{-n})^d$. By finite additivity and translation invariance, $c = 2^{nd}\mu([0,2^{-n})^d)$, i.e.

$$\mu([0,2^{-n})^d) = c\,2^{-nd} = c\,\big|[0,2^{-n})^d\big|.$$

Now let $B = \prod_{i=1}^d [a_i, b_i)$ be a half-open box all of whose endpoints are integer multiples of 2^{-n}. Then B is the disjoint union of exactly $\prod_i (b_i - a_i)2^n = |B|2^{nd}$ translates of $[0,2^{-n})^d$, so finite additivity and translation invariance give $\mu(B) = |B|2^{nd} \cdot c2^{-nd} = c|B|$.
Next let $B = \prod_{i=1}^d I_i$ be an arbitrary box, where I_i is an interval (open, closed or half-open) with endpoints $a_i \leq b_i$, so that $(a_i, b_i) \subset I_i \subset [a_i, b_i]$ and $|B| = \prod_i (b_i - a_i)$. Let $\delta > 0$ and choose n with $2^{-n} \leq \delta$.
For the upper bound, set $a_i^- := 2^{-n}\lfloor 2^n a_i \rfloor$ and $b_i^+ := 2^{-n}\lceil 2^n b_i \rceil$, which are integer multiples of 2^{-n} with $a_i^- > a_i - 2^{-n}$ and $b_i^+ < b_i + 2^{-n}$. Then $B \subset \prod_{i=1}^d [a_i^-,\, b_i^+ + 2^{-n})$, a half-open box with dyadic endpoints of denominator 2^{-n} whose i-th side has length less than $b_i - a_i + 3 \cdot 2^{-n}$. Monotonicity and the dyadic case give

$$\mu(B) \;\leq\; c \prod_{i=1}^d (b_i - a_i + 3\delta).$$

For the lower bound, suppose first that $b_i - a_i > 4\delta$ for every i, and set $\alpha_i := 2^{-n}\lceil 2^n(a_i + \delta)\rceil$ and $\beta_i := 2^{-n}\lfloor 2^n(b_i - \delta)\rfloor$. Then $a_i < a_i + \delta \leq \alpha_i < a_i + 2\delta$ and $b_i - 2\delta < \beta_i \leq b_i - \delta < b_i$, so $[\alpha_i, \beta_i) \subset (a_i, b_i) \subset I_i$ and $\beta_i - \alpha_i > b_i - a_i - 4\delta > 0$. Hence $\prod_{i=1}^d [\alpha_i, \beta_i)$ is a nonempty half-open dyadic box contained in B, and monotonicity and the dyadic case give

$$\mu(B) \;\geq\; c \prod_{i=1}^d (b_i - a_i - 4\delta).$$

If $b_i > a_i$ for every i, both bounds are available for all small δ, and letting $\delta \to 0$ yields $\mu(B) = c\prod_i (b_i - a_i) = c|B|$. If instead $b_i = a_i$ for some i, then $|B| = 0$, and the upper bound alone gives $\mu(B) \leq c\prod_i (b_i - a_i + 3\delta) \to 0$ as $\delta \to 0$, so $\mu(B) = 0 = c|B|$ in that case as well.
Step 3: $c = 1$, and μ agrees with elementary measure on elementary sets. By Lemma 1.1.2, every elementary set A is a finite disjoint union of boxes $B_1, \ldots, B_N$, whence by finite additivity and Step 2, $\mu(A) = c\sum_j |B_j| = c\,m(A)$. Taking $A = [0,1]^d$ and using the normalisation axiom gives $1 = c \cdot 1$, so $c = 1$ and $\mu(A) = m(A)$ for all elementary A; in particular $\mu(B) = |B|$ for every box B.
Step 4: $\mu(E) \leq m(E)$ for every measurable E. Let $B_1, B_2, \ldots$ be any countable family of boxes covering E. Then $E \subset \bigcup_n B_n$, so by monotonicity and countable subadditivity of μ together with Step 3,

$$\mu(E) \leq \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} |B_n|.$$

Taking the infimum over all such coverings, and recalling that Lebesgue outer measure is by definition this infimum, we get $\mu(E) \leq m^*(E) = m(E)$.
Step 5: equality for bounded sets. Suppose E is measurable and contained in some box B; we may take B to be a half-open box, so $m(B) = |B| < \infty$. The set $B \setminus E$ is measurable (Lemma 1.2.13), and E and $B \setminus E$ are disjoint with union B. Finite additivity for μ and for m (Lemma 1.2.15) gives

$$\mu(E) + \mu(B \setminus E) = \mu(B) = |B| = m(B) = m(E) + m(B \setminus E).$$

All four terms are finite (each is at most $|B|$). By Step 4, $\mu(E) \leq m(E)$ and $\mu(B \setminus E) \leq m(B \setminus E)$; if either inequality were strict, the left-hand sum would be strictly smaller than the right-hand sum, contradicting the displayed identity. Hence $\mu(E) = m(E)$.
Step 6: the general case. Partition $\mathbb{R}^d$ into the countably many disjoint half-open unit cubes $Q_k := k + [0,1)^d$, $k \in \mathbb{Z}^d$. For measurable E, the sets $E \cap Q_k$ are disjoint measurable sets with union E, each contained in a box, so by countable additivity of μ and of m, and by Step 5,

$$\mu(E) = \sum_{k \in \mathbb{Z}^d} \mu(E \cap Q_k) = \sum_{k \in \mathbb{Z}^d} m(E \cap Q_k) = m(E).$$

(The sums are over a countable index set of non-negative terms, so the order of summation is immaterial.) Thus $\mu = m$, and Lebesgue measure is the unique such map.

Problem 1.2.24. (Lebesgue measure as the completion of elementary measure) The purpose of the following exercise is to indicate how Lebesgue measure can be viewed as a metric completion of elementary measure in some sense. To avoid some technicalities we will not work in all of $\mathbb{R}^d$, but in some fixed elementary set A (e.g. $A = [0,1]^d$).
(i) Let $2^A := \{E : E \subset A\}$ be the power set of A. We say that two sets $E, F \in 2^A$ are equivalent if $E \Delta F$ is a null set. Show that this is an equivalence relation.
(ii) Let $2^A/\sim$ be the set of equivalence classes $[E] := \{F \in 2^A : E \sim F\}$ of 2^A with respect to the above equivalence relation. Define a distance $d : 2^A/\sim \times 2^A/\sim \to \mathbb{R}^+$ between two equivalence classes $[E], [E']$ by defining $d([E],[E']) := m^*(E\Delta E')$. Show that this distance is well-defined (in the sense that $m(E\Delta E') = m(F\Delta F')$ whenever $[E] = [F]$ and $[E'] = [F']$) and gives $2^A/\sim$ the structure of a complete metric space.
(iii) Let $\mathcal{E} \subset 2^A$ be the elementary subsets of A, and let $\mathcal{L} \subset 2^A$ be the Lebesgue measurable subsets of A. Show that $\mathcal{L}/\sim$ is the closure of $\mathcal{E}/\sim$ with respect to the metric defined above. In particular, $\mathcal{L}/\sim$ is a complete metric space that contains $\mathcal{E}/\sim$ as a dense subset; in other words, $\mathcal{L}/\sim$ is a metric completion of $\mathcal{E}/\sim$.
(iv) Show that Lebesgue measure $m : \mathcal{L} \to \mathbb{R}^+$ descends to a continuous function $m : \mathcal{L}/\sim \to \mathbb{R}^+$, which by abuse of notation we shall still call m. Show that $m : \mathcal{L}/\sim \to \mathbb{R}^+$ is the unique continuous extension of the analogous elementary measure function $m : \mathcal{E}/\sim \to \mathbb{R}^+$ to $\mathcal{L}/\sim$.

Solution. Throughout we use two elementary facts about null sets, both immediate from monotonicity and countable subadditivity of Lebesgue outer measure: a subset of a null set is null, and a countable union of null sets is null. We also use repeatedly the algebraic identity that symmetric difference makes 2^A into an abelian group with $1_{E\Delta F} = 1_E + 1_F \bmod 2$; in particular Δ is associative and commutative. (In (ii) the parenthetical remark should of course read $m^*(E\Delta E') = m^*(F\Delta F')$, since the sets involved need not be measurable.)
Part (i). Reflexivity: $E\Delta E = \emptyset$, which is null. Symmetry: $E\Delta F = F\Delta E$. Transitivity: if $E\Delta F$ and $F\Delta G$ are null, then from

$$E\Delta G = (E\Delta F)\Delta(F\Delta G) \subset (E\Delta F) \cup (F\Delta G)$$

we see $E\Delta G$ is a subset of a union of two null sets, hence null. So $\sim$ is an equivalence relation.
Part (ii), well-definedness. First note the general inequality

$$m^*(X) \leq m^*(Y) + m^*(X\Delta Y) \qquad (X, Y \subset \mathbb{R}^d),$$

which follows from $X \subset Y \cup (X\Delta Y)$ and subadditivity. Now suppose $[E] = [F]$ and $[E'] = [F']$, i.e. $E\Delta F$ and $E'\Delta F'$ are null. Then

$$(E\Delta E')\Delta(F\Delta F') = (E\Delta F)\Delta(E'\Delta F') \subset (E\Delta F) \cup (E'\Delta F')$$

is null, so applying the general inequality with $X = E\Delta E'$, $Y = F\Delta F'$ and then with the roles reversed gives $m^*(E\Delta E') = m^*(F\Delta F')$. Thus d is well-defined.
Part (ii), metric axioms. Since $E, E' \subset A$ and A is elementary, hence bounded, $d([E],[E']) = m^*(E\Delta E') \leq m^*(A) = m(A) < \infty$, so d indeed takes values in $\mathbb{R}^+$. It vanishes exactly when $E\Delta E'$ is null, i.e. exactly when $[E] = [E']$. It is symmetric since $E\Delta E' = E'\Delta E$. Finally, $E\Delta E'' \subset (E\Delta E') \cup (E'\Delta E'')$, so subadditivity of outer measure gives the triangle inequality

$$d([E],[E'']) \leq d([E],[E']) + d([E'],[E'']).$$

Part (ii), completeness. Let $([E_n])_{n\geq 1}$ be a Cauchy sequence in $2^A/\sim$. It suffices to produce a convergent subsequence, since a Cauchy sequence with a convergent subsequence converges. Passing to a subsequence, we may assume $d([E_n],[E_{n+1}]) \leq 2^{-n}$ for all n.

Put $F_n := E_n \Delta E_{n+1}$, so $m^*(F_n) \leq 2^{-n}$, and set $G_N := \bigcup_{n \geq N} F_n$ and $G := \bigcap_{N \geq 1} G_N$. By countable subadditivity, $m^*(G_N) \leq \sum_{n \geq N} 2^{-n} = 2^{-N+1}$, and hence by monotonicity $m^*(G) = 0$: G is null.
If $x \in A \setminus G$, then $x \notin G_N$ for some N, i.e. $x \notin F_n$ for all $n \geq N$, which means $1_{E_n}(x)$ is constant for $n \geq N$. So the limit $\lim_{n \to \infty} 1_{E_n}(x)$ exists for every $x \in A \setminus G$, and we may define

$$E := \{x \in A \setminus G : x \in E_n \text{ for all sufficiently large } n\} \subset A.$$

Fix n and let $x \in E \Delta E_n$ with $x \notin G$. If x belonged to no F_k with $k \geq n$, then $1_{E_k}(x) = 1_{E_n}(x)$ for all $k \geq n$, and hence $1_E(x) = 1_{E_n}(x)$, contradicting $x \in E \Delta E_n$. Therefore $x \in G_n$. This shows $E \Delta E_n \subset G \cup G_n$, whence

$$d([E_n], [E]) = m^*(E \Delta E_n) \leq m^*(G) + m^*(G_n) \leq 2^{-n+1} \to 0.$$

So $[E_n] \to [E]$, and $2^A / \sim$ is complete.
Part (iii). First observe that $\mathcal{L}$ is a union of equivalence classes, so that $\mathcal{L}/\sim$ makes sense as a subset of $2^A/\sim$: if E is Lebesgue measurable and $E \sim F$, then $F = E \Delta (E \Delta F)$ differs from E by a null set, and null sets are Lebesgue measurable (Lemma 1.2.13(iii)), so F is measurable. (For $\mathcal{E}$ we simply take $\mathcal{E}/\sim$ to mean the set of classes $[F]$ with F elementary.)
Density of $\mathcal{E}/\sim$ in $\mathcal{L}/\sim$: let $E \in \mathcal{L}$ and $\varepsilon > 0$. Since $E \subset A$ and A is bounded, $m(E) \leq m(A) < \infty$, so E is a Lebesgue measurable set of finite measure; by the equivalence of (i) and (viii) in Exercise 1.2.16 there is an elementary set F with $m^*(E \Delta F) \leq \varepsilon$. Replacing F by $F \cap A$ (still elementary, and now a subset of A) only improves matters: since $E \subset A$ we have $E \setminus (F \cap A) = E \setminus F$ and $(F \cap A) \setminus E \subset F \setminus E$, so $E \Delta (F \cap A) \subset E \Delta F$ and $d([E], [F \cap A]) \leq \varepsilon$. Hence every point of $\mathcal{L}/\sim$ is a limit of points of $\mathcal{E}/\sim$.
$\mathcal{L}/\sim$ is closed in $2^A/\sim$: suppose $[E_n] \in \mathcal{L}/\sim$ and $[E_n] \to [E]$ for some $E \in 2^A$. Then for every $\varepsilon > 0$ there is n with $m^*(E \Delta E_n) \leq \varepsilon$, i.e. E differs from the measurable set E_n by a set of outer measure at most ε. By the "almost measurable" criterion, Exercise 1.2.7(vi), E is Lebesgue measurable, so $[E] \in \mathcal{L}/\sim$.
Combining: every elementary set is a finite union of boxes (Lemma 1.1.2) is in particular Jordan measurable, hence Lebesgue measurable by Exercise 1.2.8, so $\mathcal{E}/\sim \subset \mathcal{L}/\sim$; since $\mathcal{L}/\sim$ is closed, so the closure of $\mathcal{E}/\sim$ is contained in $\mathcal{L}/\sim$; and by density the closure contains $\mathcal{L}/\sim$. Hence $\overline{\mathcal{E}/\sim} = \mathcal{L}/\sim$. Being a closed subset of the complete metric space $2^A/\sim$, the space $\mathcal{L}/\sim$ is itself complete, and it contains $\mathcal{E}/\sim$ as a dense subset; that is, it is a metric completion of $\mathcal{E}/\sim$.
Part (iv). If $E, F \in \mathcal{L}$ with $E \sim F$, then $E \subset F \cup (E \Delta F)$ gives $m(E) \leq m(F) + m^*(E \Delta F) = m(F)$, and symmetrically $m(F) \leq m(E)$; so $m(E) = m(F)$ and m descends to a well-defined function on $\mathcal{L}/\sim$, with values in $\mathbb{R}^+$ since $m(E) \leq m(A) < \infty$. The same computation applied to arbitrary $E, F \in \mathcal{L}$ gives $m(E) \leq m(F) + m^*(E \Delta F)$ and $m(F) \leq m(E) + m^*(E \Delta F)$, i.e.

$$|m([E]) - m([F])| \leq d([E], [F]),$$

all quantities being finite. So m is 1-Lipschitz on $\mathcal{L}/\sim$, in particular continuous, and it clearly restricts on $\mathcal{E}/\sim$ to elementary measure (Lebesgue measure agrees with elementary measure on elementary sets).
Uniqueness of the continuous extension: if $f, g : \mathcal{L}/\sim \to \mathbb{R}^+$ are continuous and agree on the dense subset $\mathcal{E}/\sim$, then for any $[E] \in \mathcal{L}/\sim$ choose elementary $F_n \subset A$ with $[F_n] \to [E]$ (possible by part (iii)); then $f([E]) = \lim_n f([F_n]) = \lim_n g([F_n]) = g([E])$. Hence m is the unique continuous extension of elementary measure from $\mathcal{E}/\sim$ to $\mathcal{L}/\sim$.

Problem 1.2.25. Define a continuously differentiable curve in $\mathbb{R}^d$ to be a set of the form $\{\gamma(t) : a \leq t \leq b\}$ where $[a, b]$ is a closed interval and $\gamma : [a, b] \to \mathbb{R}^d$ is a continuously differentiable function.
(i) If $d \geq 2$, show that every continuously differentiable curve has Lebesgue measure zero. (Why is the condition $d \geq 2$ necessary?)
(ii) Conclude that if $d \geq 2$, then the unit cube $[0, 1]^d$ cannot be covered by countably many continuously differentiable curves.
We remark that if the curve is only assumed to be continuous, rather than continuously differentiable,

then these claims fail, thanks to the existence of space-filling curves.

Solution. Part (i). Let $\Gamma := \{\gamma(t) : a \le t \le b\}$ with $\gamma : [a,b] \to \mathbb{R}^d$ continuously differentiable, and assume $d \ge 2$. We may assume $a < b$, since otherwise Γ is a single point and hence null.
Since γ' is continuous on the compact interval $[a,b]$, it is bounded there: put

$$M := \sup_{t\in[a,b]} |\gamma'(t)| < \infty,$$

where $|\cdot|$ is the Euclidean norm. The curve is then Lipschitz with constant M: for $a \le s \le t \le b$, applying the fundamental theorem of calculus to each coordinate gives $\gamma(t) - \gamma(s) = \int_s^t \gamma'(u)\,du$, and hence

$$|\gamma(t) - \gamma(s)| \le \int_s^t |\gamma'(u)|\,du \le M(t-s).$$

Now fix a natural number $n \ge 1$ and partition $[a,b]$ into the n intervals

$$I_k := \Big[a + \tfrac{(k-1)(b-a)}{n},\ a + \tfrac{k(b-a)}{n}\Big], \qquad k = 1, \dots, n,$$

each of length $(b-a)/n$. Let t_k be the left endpoint of I_k. By the Lipschitz bound, every point $\gamma(t)$ with $t \in I_k$ satisfies $|\gamma(t) - \gamma(t_k)| \le M(b-a)/n$, so $\gamma(I_k)$ is contained in the closed cube Q_k centred at $\gamma(t_k)$ of sidelength $2M(b-a)/n$ (a Euclidean ball of radius r is contained in the cube of sidelength $2r$ with the same centre). Since $\Gamma = \bigcup_{k=1}^n \gamma(I_k)$, the cubes $Q_1, \dots, Q_n$ cover Γ, and by monotonicity and finite subadditivity of Lebesgue outer measure together with the formula for the measure of a box (Lemma 1.2.6),

$$m^*(\Gamma) \le \sum_{k=1}^n |Q_k| = n\left(\frac{2M(b-a)}{n}\right)^d = \big(2M(b-a)\big)^d\, n^{1-d}.$$

Because $d \ge 2$, the exponent $1-d$ is negative, so letting $n \to \infty$ gives $m^*(\Gamma) = 0$. Thus Γ is a null set, hence Lebesgue measurable with $m(\Gamma) = 0$ by Lemma 1.2.13(iii). (Alternatively, Γ is compact, being the continuous image of a compact set, so its measurability is automatic from Lemma 1.2.13(ii).)
The hypothesis $d \ge 2$ is exactly what makes the exponent n^{1-d} decay: for $d = 1$ the bound reads $m^*(\Gamma) \le 2M(b-a)$, with no gain as $n \to \infty$. And indeed the conclusion is false for $d = 1$: the curve $\gamma(t) := t$ on $[0,1]$ has image $[0,1]$, of Lebesgue measure 1. The geometric point is that a curve is a one-dimensional object; only when it sits inside a space of dimension at least 2 does it fail to fill up positive volume. The gain in the estimate comes from covering the curve by n cubes of sidelength $O(1/n)$, whose total volume is $O(n^{1-d})$, and this is small precisely when $d > 1$.
Part (ii). Suppose for contradiction that $d \ge 2$ and that $[0,1]^d \subset \bigcup_{n=1}^\infty \Gamma_n$ where each Γ_n is a continuously differentiable curve. By part (i), $m^*(\Gamma_n) = 0$ for every n. By monotonicity and countable subadditivity of Lebesgue outer measure,

$$1 = \big|[0,1]^d\big| = m^*\big([0,1]^d\big) \le \sum_{n=1}^\infty m^*(\Gamma_n) = 0,$$

a contradiction. Hence $[0,1]^d$ cannot be covered by countably many continuously differentiable curves.

Problem 1.2.26. (Outer measure is not finitely additive) Show that there exist disjoint bounded subsets E, F of the real line such that $m^*(E \cup F) \ne m^*(E) + m^*(F)$. (Hint: Show that the set constructed in the proof of the above proposition has positive outer measure.)

Solution. Let $V \subset [0,1]$ be the set constructed in the proof of Proposition 1.2.18: choosing (by the axiom of choice) one representative $x_C \in C \cap [0,1]$ from each coset $C \in \mathbb{R}/\mathbb{Q}$, we set $V := \{x_C : C \in \mathbb{R}/\mathbb{Q}\}$.

Recall from that proof the two inclusions

$$[0,1] \subset \bigcup_{q\in\mathbb{Q}\cap[-1,1]} (V+q), \qquad \bigcup_{q\in\mathbb{Q}\cap[-1,1]} (V+q) \subset [-1,2],$$

which are the inclusions (1.4) and (1.5) of that proof, and the fact that the translates $V+q$, $q\in\mathbb{Q}$, are pairwise disjoint: if $x+q = y+q'$ with $x,y\in V$ and $q\neq q'$ rational, then $x-y = q'-q$ is a nonzero rational, so x and y are two distinct elements of V lying in the same coset of $\mathbb{Q}$, contrary to the construction.

We first record that Lebesgue outer measure is translation invariant on arbitrary sets: if $A\subset\mathbb{R}^d$ and $x\in\mathbb{R}^d$, then $B\mapsto B+x$ is a bijection from countable box-coverings of A onto countable box-coverings of $A+x$ preserving each $|B|$ (a translate of a box is a box of the same volume), so the two infima agree and $m^*(A+x) = m^*(A)$.

Step 1: V has positive outer measure. The rationals in $[-1,1]$ form a countable set, so by countable subadditivity of outer measure, monotonicity, the inclusion (1.4), and translation invariance,

$$1 = m^*([0,1]) \le \sum_{q\in\mathbb{Q}\cap[-1,1]} m^*(V+q) = \sum_{q\in\mathbb{Q}\cap[-1,1]} m^*(V).$$

If $m^*(V)$ were 0 the right-hand side would be 0, which is absurd. Hence $c := m^*(V) > 0$ (and $c\le 1$ by monotonicity, since $V\subset[0,1]$).

Step 2: finite additivity of m^* on disjoint bounded sets fails. Suppose for contradiction that the additivity property

$$m^*(A\cup B) = m^*(A) + m^*(B) \qquad \text{whenever } A,B\subset\mathbb{R} \text{ are disjoint and bounded.}$$

holds. We claim this implies, by induction on N, that $m^*(A_1\cup\cdots\cup A_N) = \sum_{n=1}^N m^*(A_n)$ for any pairwise disjoint bounded sets $A_1,\dots,A_N\subset\mathbb{R}$. The case $N=1$ is trivial; and if the claim holds for N, then $A_1\cup\cdots\cup A_N$ and A_{N+1} are disjoint bounded sets, so the assumed additivity property and the inductive hypothesis give

$$m^*(A_1\cup\cdots\cup A_{N+1}) = m^*(A_1\cup\cdots\cup A_N) + m^*(A_{N+1}) = \sum_{n=1}^{N+1} m^*(A_n).$$

Now pick N distinct rationals $q_1,\dots,q_N\in\mathbb{Q}\cap[-1,1]$, where N is any integer with $N > 3/c$ (possible since $c>0$, and since $\mathbb{Q}\cap[-1,1]$ is infinite). The sets $A_n := V+q_n$ are bounded, pairwise disjoint (as noted above), and each has $m^*(A_n) = c$ by translation invariance. By the claim and by (1.5) together with monotonicity,

$$Nc = \sum_{n=1}^N m^*(A_n) = m^*\Big(\bigcup_{n=1}^N (V+q_n)\Big) \le m^*([-1,2]) = 3,$$

contradicting $N > 3/c$.

Therefore the assumed additivity property is false: there exist disjoint bounded subsets $E,F\subset\mathbb{R}$ with $m^*(E\cup F) \neq m^*(E)+m^*(F)$. (By subadditivity, the failure is necessarily in the direction $m^*(E\cup F) < m^*(E)+m^*(F)$.)

Explicitly, one may extract such a pair from the argument as follows. With $A_1,\dots,A_N$ as above, let k be the largest index in $\{1,\dots,N\}$ for which $m^*(A_1\cup\cdots\cup A_k) = \sum_{n\le k} m^*(A_n)$; this is well defined ($k=1$ always works) and $k<N$, since the displayed contradiction shows the identity fails for N. Taking $E := A_1\cup\cdots\cup A_k$ and $F := A_{k+1}$, these are disjoint bounded sets with $m^*(E)+m^*(F) = \sum_{n\le k+1} m^*(A_n) \neq m^*(E\cup F)$, by the maximality of k.

Problem 1.2.27. (Projections of measurable sets need not be measurable) Let $\pi : \mathbb{R}^2 \to \mathbb{R}$ be the coordinate projection $\pi(x, y) := x$. Show that there exists a measurable subset E of $\mathbb{R}^2$ such that $\pi(E)$ is not measurable.

Solution. By Proposition 1.2.18 there is a set $V \subset [0, 1]$ which is not Lebesgue measurable (as a subset of $\mathbb{R}$). Define

$$E := V \times \{0\} \subset \mathbb{R}^2.$$

E is Lebesgue measurable in $\mathbb{R}^2$. Indeed, E is contained in the segment $[0, 1] \times \{0\}$, which is a degenerate box in $\mathbb{R}^2$ of volume $1 \cdot 0 = 0$; by Lemma 1.2.6 its outer measure equals its volume, namely 0. Hence by monotonicity $m^*(E) = 0$, so E is a null set and is therefore Lebesgue measurable, with $m(E) = 0$, by Lemma 1.2.13(iii). (Equivalently, one may quote Exercise 1.2.22(i): $(m^2)^*(V \times \{0\}) \leq (m^1)^*(V)\,(m^1)^*(\{0\}) = (m^1)^*(V) \cdot 0 = 0$, the product being legitimate since $(m^1)^*(V) \leq 1 < \infty$.)
On the other hand,

$$\pi(E) = \{x : (x, 0) \in V \times \{0\}\} = V,$$

which is not Lebesgue measurable. So E is a measurable (indeed null) subset of $\mathbb{R}^2$ whose projection $\pi(E)$ is non-measurable, as required.
This example also justifies the caution recorded in Exercise 1.2.21: for a linear map $T : \mathbb{R}^d \to \mathbb{R}^d$ of a space to itself one has that $T(E)$ is measurable whenever E is, but for a linear map $T : \mathbb{R}^d \to \mathbb{R}^{d'}$ into a space of strictly smaller dimension this can fail, since such a map can carry a null set onto a set of positive outer measure (here the null segment $V \times \{0\}$ is carried onto the non-measurable set V, which by Step 1 of Exercise 1.2.26 has positive outer measure).

2.3 §1.3 — The Lebesgue Integral

Problem 1.3.1. (Basic properties of the simple unsigned integral) Let $f, g : \mathbb{R}^d \to [0, +\infty]$ be simple unsigned functions.

- (i) (Unsigned linearity) We have

$$\text{Simp} \int_{\mathbb{R}^d} f(x) + g(x)\ dx = \text{Simp} \int_{\mathbb{R}^d} f(x)\ dx + \text{Simp} \int_{\mathbb{R}^d} g(x)\ dx$$

and

$$\text{Simp} \int_{\mathbb{R}^d} cf(x)\ dx = c \times \text{Simp} \int_{\mathbb{R}^d} f(x)\ dx$$

for all $c \in [0, +\infty]$.

- (ii) (Finiteness) We have $\text{Simp} \int_{\mathbb{R}^d} f(x)\ dx < \infty$ if and only if f is finite almost everywhere, and its support has finite measure.
- (iii) (Vanishing) We have $\text{Simp} \int_{\mathbb{R}^d} f(x)\ dx = 0$ if and only if f is zero almost everywhere.
- (iv) (Equivalence) If f and g agree almost everywhere, then $\text{Simp} \int_{\mathbb{R}^d} f(x)\ dx = \text{Simp} \int_{\mathbb{R}^d} g(x)\ dx$.
- (v) (Monotonicity) If $f(x) \leq g(x)$ for almost every $x \in \mathbb{R}^d$, then $\text{Simp} \int_{\mathbb{R}^d} f(x)\ dx \leq \text{Simp} \int_{\mathbb{R}^d} g(x)\ dx$.
- (vi) (Compatibility with Lebesgue measure) For any Lebesgue measurable E, one has $\text{Simp} \int_{\mathbb{R}^d} 1_E(x)\ dx = m(E)$.

Furthermore, show that the simple unsigned integral $f \mapsto \text{Simp} \int_{\mathbb{R}^d} f(x)\ dx$ is the only map from the space $\text{Simp}^+(\mathbb{R}^d)$ of unsigned simple functions to $[0, +\infty]$ that obeys all of the above properties.

Solution. Throughout we use the arithmetic of $[0, +\infty]$ with the convention $0 \cdot \infty = \infty \cdot 0 = 0$; with this convention addition and multiplication on $[0, +\infty]$ are commutative, associative and distributive, and both operations are monotone in each variable.
Canonical form. We first record a normal form that will be used repeatedly. Let $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ be unsigned simple. As observed in the text preceding Definition 1.3.3, the k sets $E_1, \ldots, E_k$ partition $\mathbb{R}^d$ into at most 2^k disjoint measurable atoms, each atom being an intersection $\bigcap_{i \in S} E_i \cap \bigcap_{i \notin S} (\mathbb{R}^d \setminus E_i)$ for some $S \subset \{1, \ldots, k\}$; and f is constant on each atom (equal to $\sum_{i \in S} c_i$ there). Consequently f

takes at most 2^k distinct values $a_1, \dots, a_m \in [0, +\infty]$, each level set

$$A_p := f^{-1}(\{a_p\}) = \{x \in \mathbb{R}^d : f(x) = a_p\}$$

is a finite union of atoms and hence Lebesgue measurable, the A_p are disjoint, and $\bigcup_{p=1}^m A_p = \mathbb{R}^d$. Thus

$$f = a_1 1_{A_1} + \cdots + a_m 1_{A_m},$$

and by Lemma 1.3.4 (well-definedness) the two representations give the same integral, so

$$\text{Simp} \int_{\mathbb{R}^d} f(x)\, dx = \sum_{p=1}^m a_p m(A_p).$$

We call this the canonical form of f. More generally, given finitely many simple functions $f, g, \dots$ we may run the same atom construction on the union of all the sets appearing in their representations; this produces a single finite partition $\mathbb{R}^d = A_1 \cup \cdots \cup A_m$ into disjoint measurable sets on which all of $f, g, \dots$ are simultaneously constant, and (again by Lemma 1.3.4) the integral of each of them is computed by summing (its constant value on A_p) times $m(A_p)$. We call this a common refinement.
(i) Linearity. Write $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ and $g = d_1 1_{F_1} + \cdots + d_l 1_{F_l}$. Then

$$f + g = c_1 1_{E_1} + \cdots + c_k 1_{E_k} + d_1 1_{F_1} + \cdots + d_l 1_{F_l}$$

is itself a representation of $f + g$ as a finite unsigned combination of indicators of measurable sets, so $f + g$ is simple and, directly from Definition 1.3.3,

$$\text{Simp} \int_{\mathbb{R}^d} f + g = \sum_{i=1}^k c_i m(E_i) + \sum_{j=1}^l d_j m(F_j) = \text{Simp} \int_{\mathbb{R}^d} f + \text{Simp} \int_{\mathbb{R}^d} g.$$

Similarly, for $c \in [0, +\infty]$ we have $cf = (cc_1)1_{E_1} + \cdots + (cc_k)1_{E_k}$, so cf is simple and

$$\text{Simp} \int_{\mathbb{R}^d} cf = \sum_{i=1}^k (cc_i) m(E_i) = c \sum_{i=1}^k c_i m(E_i) = c\, \text{Simp} \int_{\mathbb{R}^d} f,$$

using distributivity in $[0, +\infty]$. (Both computations are legitimate because, by Lemma 1.3.4, the value of the integral does not depend on which representation we use.)
(ii) Finiteness. Take the canonical form $f = \sum_{p=1}^m a_p 1_{A_p}$, so that $\text{Simp} \int f = \sum_p a_p m(A_p)$. A finite sum in $[0, +\infty]$ is finite if and only if every summand is finite, so

$$\text{Simp} \int_{\mathbb{R}^d} f < \infty \iff a_p m(A_p) < \infty \text{ for every } p.$$

Now examine the possible a_p. If $a_p = 0$ the term is 0 always. If $0 < a_p < \infty$ then $a_p m(A_p) < \infty$ iff $m(A_p) < \infty$. If $a_p = +\infty$ then $a_p m(A_p) < \infty$ iff $m(A_p) = 0$ (since $\infty \cdot t = \infty$ for $t > 0$ and $\infty \cdot 0 = 0$). Note also that the support $\{f \neq 0\}$ is exactly $\bigcup_{p: a_p > 0} A_p$, a disjoint union, so by finite additivity of Lebesgue measure

$$m(\{f \neq 0\}) = \sum_{p: a_p > 0} m(A_p),$$

and $\{f = +\infty\} = \bigcup_{p: a_p = \infty} A_p$.
Suppose first the integral is finite. Then for every p with $a_p = \infty$ we get $m(A_p) = 0$, whence $m(\{f = \infty\}) = 0$, i.e. f is finite a.e.; and for every p with $a_p > 0$ we get $m(A_p) < \infty$ (this covers both $0 < a_p < \infty$ and $a_p = \infty$), so $m(\{f \neq 0\})$ is a finite sum of finite numbers, hence finite.
Conversely, suppose f is finite a.e. and $m(\{f \neq 0\}) < \infty$. If $a_p = \infty$ then $A_p \subset \{f = \infty\}$ is null, so $a_p m(A_p) = \infty \cdot 0 = 0$. If $0 < a_p < \infty$ then $A_p \subset \{f \neq 0\}$ gives $m(A_p) \leq m(\{f \neq 0\}) < \infty$ by monotonicity of measure, so $a_p m(A_p) < \infty$. If $a_p = 0$ the term vanishes. Hence every summand is finite and so is the integral.

(iii) Vanishing. Again with the canonical form, $\sum_p a_p m(A_p) = 0$ iff $a_p m(A_p) = 0$ for every p, i.e. iff for every p with $a_p > 0$ we have $m(A_p) = 0$. By the identity $m(\{f \neq 0\}) = \sum_{p:a_p>0} m(A_p)$ from (ii), this holds iff $m(\{f \neq 0\}) = 0$, i.e. iff $f = 0$ almost everywhere.
(iv) Equivalence. Let $N := \{x : f(x) \neq g(x)\}$, a null set by hypothesis. Note first that N is measurable (every null set is Lebesgue measurable). Take the canonical form $f = \sum_p a_p 1_{A_p}$. Then $f1_{\mathbb{R}^d \backslash N} = \sum_p a_p 1_{A_p \backslash N}$ is simple, and since $m(A_p \cap N) = 0$ we get $m(A_p) = m(A_p \setminus N) + m(A_p \cap N) = m(A_p \setminus N)$ by finite additivity. Hence

$$\text{Simp} \int_{\mathbb{R}^d} f = \sum_p a_p m(A_p) = \sum_p a_p m(A_p \setminus N) = \text{Simp} \int_{\mathbb{R}^d} f1_{\mathbb{R}^d \backslash N}.$$

The same computation applied to g gives $\text{Simp} \int g = \text{Simp} \int g1_{\mathbb{R}^d \backslash N}$. But $f1_{\mathbb{R}^d \backslash N}$ and $g1_{\mathbb{R}^d \backslash N}$ are literally the same function (they agree off N by definition of N, and both vanish on N), so the two integrals coincide.
(v) Monotonicity. Let $N := \{x : f(x) > g(x)\}$, which is null by hypothesis. By the computation in (iv), replacing f by $f1_{\mathbb{R}^d \backslash N}$ and g by $g1_{\mathbb{R}^d \backslash N}$ changes neither integral, and after this replacement we have $f \leq g$ at every point of $\mathbb{R}^d$. So we may assume $f \leq g$ everywhere. Now take a common refinement $\mathbb{R}^d = A_1 \cup \cdots \cup A_m$ as above, with $f \equiv \alpha_p$ and $g \equiv \beta_p$ on A_p. Discarding empty A_p (which contribute 0 to both sides), we have $\alpha_p \leq \beta_p$ for each remaining p, since we may evaluate the pointwise inequality at any point of A_p. Therefore, by monotonicity of multiplication and addition in $[0, +\infty]$,

$$\text{Simp} \int_{\mathbb{R}^d} f = \sum_p \alpha_p m(A_p) \leq \sum_p \beta_p m(A_p) = \text{Simp} \int_{\mathbb{R}^d} g.$$

(vi) Compatibility. The function 1_E is the unsigned simple function with representation $1 \cdot 1_E$ (i.e. $k = 1$, $c_1 = 1$, $E_1 = E$), so Definition 1.3.3 gives $\text{Simp} \int_{\mathbb{R}^d} 1_E(x)\, dx = 1 \cdot m(E) = m(E)$.
Uniqueness. Let $I : \text{Simp}^+(\mathbb{R}^d) \to [0, +\infty]$ be any map obeying (i)–(vi). In fact only (i) and (vi) are needed. Let $f \in \text{Simp}^+(\mathbb{R}^d)$ and choose any representation $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ with $c_i \in [0, +\infty]$ and E_i measurable. Applying the additivity in (i) repeatedly ($k-1$ times, by induction on k) and then the homogeneity in (i) to each summand,

$$I(f) = \sum_{i=1}^k I(c_i 1_{E_i}) = \sum_{i=1}^k c_i I(1_{E_i}),$$

and by (vi), $I(1_{E_i}) = m(E_i)$. Hence

$$I(f) = \sum_{i=1}^k c_i m(E_i) = \text{Simp} \int_{\mathbb{R}^d} f(x)\, dx$$

by Definition 1.3.3. (In the degenerate case $k = 0$, i.e. f is the zero function 0, apply the homogeneity in (i) with scalar $c = 0$ to the zero function: since $0 \cdot 0 = 0$ as functions, we get $I(0) = 0 \cdot I(0) = 0$, the last step because $0 \cdot t = 0$ for every $t \in [0, +\infty]$ under the convention $0 \cdot \infty = 0$; and likewise $\text{Simp} \int 0 = 0$ from the empty representation.) So I agrees with the simple unsigned integral on all of $\text{Simp}^+(\mathbb{R}^d)$, proving uniqueness.

Problem 1.3.2. (Basic properties of the complex-valued simple integral) Let $f, g : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable simple functions.

- (i) ($*$-linearity) We have

$$\text{Simp} \int_{\mathbb{R}^d} f(x) + g(x)\, dx = \text{Simp} \int_{\mathbb{R}^d} f(x)\, dx + \text{Simp} \int_{\mathbb{R}^d} g(x)\, dx$$

and

$$\text{Simp} \int_{\mathbb{R}^d} cf(x)\ dx = c \times \text{Simp} \int_{\mathbb{R}^d} f(x)\ dx \qquad (1.11)$$

for all $c \in \mathbb{C}$. Also we have

$$\text{Simp} \int_{\mathbb{R}^d} \overline{f(x)}\ dx = \overline{\text{Simp} \int_{\mathbb{R}^d} f(x)\ dx}.$$

- (ii) (Equivalence) If f and g agree almost everywhere, then $\text{Simp} \int_{\mathbb{R}^d} f(x)\ dx = \text{Simp} \int_{\mathbb{R}^d} g(x)\ dx$.
- (iii) (Compatibility with Lebesgue measure) For any Lebesgue measurable E, one has $\text{Simp} \int_{\mathbb{R}^d} 1_E(x)\ dx = m(E)$.

(Hints: Work out the real-valued counterpart of the linearity property first. To establish (1.11), treat the cases $c > 0$, $c = 0$, $c = -1$ separately. To deal with the additivity for real functions f, g, start with the identity

$$f + g = (f + g)_+ - (f + g)_- = (f_+ - f_-) + (g_+ - g_-)$$

and rearrange the second inequality so that no subtraction appears.)
Furthermore, show that the complex-valued simple integral $f \mapsto \text{Simp} \int_{\mathbb{R}^d} f(x)\ dx$ is the only map from the space $\text{Simp}_{\text{abs}}(\mathbb{R}^d)$ of absolutely integrable simple functions to $\mathbb{C}$ that obeys all of the above properties.

Solution. Throughout, "simple" means complex-valued simple in the sense of Definition 1.3.2, and the integrals are those of Definition 1.3.6. We repeatedly use Exercise 1.3.1.
Preliminary remarks. If h is a complex-valued simple function then $|h|$, $\operatorname{Re} h$, $\operatorname{Im} h$ and (for real-valued h) $h_+ = \max(h, 0)$, $h_- = \max(-h, 0)$ are all simple. Indeed, running the atom construction from the text preceding Lemma 1.3.4 on a representation $h = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ shows that h is constant on each of the finitely many disjoint measurable atoms determined by $E_1, \ldots, E_k$; hence h takes only finitely many values $a_1, \ldots, a_m$, each level set $A_p := h^{-1}(\{a_p\})$ is measurable, and $h = \sum_p a_p 1_{A_p}$. Applying any function $\Phi : \mathbb{C} \to \mathbb{C}$ (or $\Phi : \mathbb{C} \to [0, +\infty]$) with $\Phi(0) = 0$ pointwise yields $\Phi \circ h = \sum_p \Phi(a_p) 1_{A_p}$, again simple. Taking $\Phi = |\cdot|, \operatorname{Re}, \operatorname{Im}, (\cdot)_+, (\cdot)_-$ gives the claim.
Moreover if f is absolutely integrable then so are $\operatorname{Re} f$, $\operatorname{Im} f$, $(\operatorname{Re} f)_\pm$, $(\operatorname{Im} f)_\pm$, because each of these is pointwise dominated in magnitude by $|f|$, so Exercise 1.3.1(v) gives that their absolute values have integral at most $\text{Simp} \int |f| < \infty$. In particular all four unsigned quantities

$$\text{Simp} \int (\operatorname{Re} f)_+, \quad \text{Simp} \int (\operatorname{Re} f)_-, \quad \text{Simp} \int (\operatorname{Im} f)_+, \quad \text{Simp} \int (\operatorname{Im} f)_-$$

are finite real numbers, so the expressions in Definition 1.3.6 are well-defined and $\text{Simp} \int f \in \mathbb{C}$, with $\text{Simp} \int f \in \mathbb{R}$ when f is real-valued.
Key lemma (subtraction-free evaluation). Let $f : \mathbb{R}^d \to \mathbb{R}$ be an absolutely integrable simple function, and let h_1, h_2 be unsigned simple functions with $\text{Simp} \int h_1 < \infty$ and $\text{Simp} \int h_2 < \infty$ (so that, by Exercise 1.3.1(ii), each h_i is finite outside a null set). Suppose that $f(x) = h_1(x) - h_2(x)$ at every x at which $h_1(x)$ and $h_2(x)$ are both finite, so that the subtraction makes sense. Then

$$\text{Simp} \int_{\mathbb{R}^d} f = \text{Simp} \int_{\mathbb{R}^d} h_1 - \text{Simp} \int_{\mathbb{R}^d} h_2.$$

Proof: let N be the null set off which both h_1 and h_2 are finite. For $x \notin N$ all four numbers $f_+(x), f_-(x), h_1(x), h_2(x)$ are finite reals and $f_+(x) - f_-(x) = h_1(x) - h_2(x)$, so rearranging so that no subtraction appears,

$$f_+(x) + h_2(x) = h_1(x) + f_-(x).$$

Thus the two unsigned simple functions $f_+ + h_2$ and $h_1 + f_-$ agree almost everywhere, so by Exercise 1.3.1(iv) they have the same unsigned integral, and by Exercise 1.3.1(i) this reads

$$\text{Simp} \int f_+ + \text{Simp} \int h_2 = \text{Simp} \int h_1 + \text{Simp} \int f_-.$$

All four quantities are finite real numbers, so we may subtract and obtain $\operatorname{Simp}\int f_+ - \operatorname{Simp}\int f_- = \operatorname{Simp}\int h_1 - \operatorname{Simp}\int h_2$, which is the claim by Definition 1.3.6. $\square$

(i) Additivity, real case. Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be absolutely integrable simple. Then $f+g$ is simple, and $|f+g| \le |f|+|g|$ pointwise, so by Exercise 1.3.1(v),(i) we get $\operatorname{Simp}\int |f+g| \le \operatorname{Simp}\int |f| + \operatorname{Simp}\int |g| < \infty$; thus $f+g$ is absolutely integrable. Following the hint,

$$f+g = (f_+ - f_-) + (g_+ - g_-) = (f_+ + g_+) - (f_- + g_-),$$

so we may apply the key lemma with $h_1 := f_+ + g_+$ and $h_2 := f_- + g_-$; these are unsigned simple, and by Exercise 1.3.1(i) their integrals are $\operatorname{Simp}\int f_+ + \operatorname{Simp}\int g_+ < \infty$ and $\operatorname{Simp}\int f_- + \operatorname{Simp}\int g_- < \infty$ respectively. The lemma gives

$$\operatorname{Simp}\int (f+g) = \left(\operatorname{Simp}\int f_+ + \operatorname{Simp}\int g_+\right) - \left(\operatorname{Simp}\int f_- + \operatorname{Simp}\int g_-\right),$$

and regrouping these finite reals gives $\operatorname{Simp}\int f + \operatorname{Simp}\int g$.

(i) Homogeneity, real case. Let f be real absolutely integrable simple and $c \in \mathbb{R}$. Then cf is simple, and $|cf| = |c|\,|f|$ pointwise, so $\operatorname{Simp}\int |cf| = |c| \operatorname{Simp}\int |f| < \infty$ by Exercise 1.3.1(i); thus cf is absolutely integrable and its integral is defined. If $c > 0$ then $(cf)_+ = cf_+$ and $(cf)_- = cf_-$ pointwise, so by Exercise 1.3.1(i),

$$\operatorname{Simp}\int cf = c\operatorname{Simp}\int f_+ - c\operatorname{Simp}\int f_- = c\operatorname{Simp}\int f.$$

If $c = 0$ then $cf \equiv 0$, whose positive and negative parts vanish identically, so $\operatorname{Simp}\int cf = 0 - 0 = 0 = 0 \cdot \operatorname{Simp}\int f$ (the last product being legitimate since $\operatorname{Simp}\int f$ is a finite real). If $c = -1$ then $(-f)_+ = f_-$ and $(-f)_- = f_+$, so

$$\operatorname{Simp}\int (-f) = \operatorname{Simp}\int f_- - \operatorname{Simp}\int f_+ = -\operatorname{Simp}\int f.$$

For general $c < 0$ write $c = (-1)\cdot|c|$ and compose the cases $c > 0$ and $c = -1$.

(i) Complex case. Let f, g be absolutely integrable complex simple functions. As in the real case, $|f+g| \le |f| + |g|$ and $|cf| = |c|\,|f|$ pointwise, so $f+g$ and cf ($c \in \mathbb{C}$) are again absolutely integrable simple functions by Exercise 1.3.1(i),(v), and all the integrals below are defined. Since $\operatorname{Re}(f+g) = \operatorname{Re} f + \operatorname{Re} g$ and $\operatorname{Im}(f+g) = \operatorname{Im} f + \operatorname{Im} g$, and all six of these functions are real absolutely integrable simple functions, Definition 1.3.6 and the real additivity just proved give

$$\begin{aligned}
\operatorname{Simp}\int (f+g) &= \operatorname{Simp}\int (\operatorname{Re} f + \operatorname{Re} g) + i\operatorname{Simp}\int (\operatorname{Im} f + \operatorname{Im} g) \\
&= \left(\operatorname{Simp}\int \operatorname{Re} f + i\operatorname{Simp}\int \operatorname{Im} f\right) + \left(\operatorname{Simp}\int \operatorname{Re} g + i\operatorname{Simp}\int \operatorname{Im} g\right) \\
&= \operatorname{Simp}\int f + \operatorname{Simp}\int g.
\end{aligned}$$

For (1.11), first let $c \in \mathbb{R}$. Then $\operatorname{Re}(cf) = c\operatorname{Re} f$ and $\operatorname{Im}(cf) = c\operatorname{Im} f$, so by real homogeneity

$$\operatorname{Simp}\int cf = c\operatorname{Simp}\int \operatorname{Re} f + ic\operatorname{Simp}\int \operatorname{Im} f = c\operatorname{Simp}\int f.$$

Next let $c = i$. Then $\operatorname{Re}(if) = -\operatorname{Im} f$ and $\operatorname{Im}(if) = \operatorname{Re} f$, so using the case $c = -1$ of real homogeneity,

$$\operatorname{Simp}\int if = -\operatorname{Simp}\int \operatorname{Im} f + i\operatorname{Simp}\int \operatorname{Re} f = i\left(\operatorname{Simp}\int \operatorname{Re} f + i\operatorname{Simp}\int \operatorname{Im} f\right) = i\operatorname{Simp}\int f.$$

For general $c = a + bi$ with $a, b \in \mathbb{R}$, write $cf = af + b(if)$ and combine complex additivity with the two cases just treated:

$$\operatorname{Simp}\int cf = a\operatorname{Simp}\int f + b\,i\operatorname{Simp}\int f = (a+bi)\operatorname{Simp}\int f = c\operatorname{Simp}\int f.$$

Finally, $\mathrm{Re}(\overline{f}) = \mathrm{Re}\, f$ and $\mathrm{Im}(\overline{f}) = -\mathrm{Im}\, f$, and $\overline{f}$ is again simple and absolutely integrable (as $|\overline{f}| = |f|$), so

$$\mathrm{Simp} \int \overline{f} = \mathrm{Simp} \int \mathrm{Re}\, f + i\, \mathrm{Simp} \int (-\mathrm{Im}\, f) = \mathrm{Simp} \int \mathrm{Re}\, f - i\, \mathrm{Simp} \int \mathrm{Im}\, f = \overline{\mathrm{Simp} \int f},$$

where the last step uses that $\mathrm{Simp} \int \mathrm{Re}\, f$ and $\mathrm{Simp} \int \mathrm{Im}\, f$ are real numbers.
(ii) Equivalence. Suppose $f = g$ almost everywhere. Then also $\mathrm{Re}\, f = \mathrm{Re}\, g$ and $\mathrm{Im}\, f = \mathrm{Im}\, g$ almost everywhere, and hence, taking positive and negative parts pointwise (an operation that respects pointwise equality), the four pairs $(\mathrm{Re}\, f)_\pm$ and $(\mathrm{Re}\, g)_\pm$, $(\mathrm{Im}\, f)_\pm$ and $(\mathrm{Im}\, g)_\pm$ each agree almost everywhere. By Exercise 1.3.1(iv) the corresponding unsigned simple integrals agree, and since $\mathrm{Simp} \int f$ is by Definition 1.3.6 built from exactly these four numbers by the same formula as $\mathrm{Simp} \int g$, we conclude $\mathrm{Simp} \int f = \mathrm{Simp} \int g$. (Note also that $|f| = |g|$ a.e., so f is absolutely integrable iff g is.)
(iii) Compatibility. Note first that $|1_E| = 1_E$, so by Exercise 1.3.1(vi) the unsigned integral of $|1_E|$ is $m(E)$; hence 1_E belongs to $\mathrm{Simp}_{\mathrm{abs}}(\mathbb{R}^d)$ precisely when $m(E) < \infty$, and the assertion is to be read for such E (otherwise the left-hand side is not defined). For such E the function 1_E is real-valued and non-negative, so $\mathrm{Re}\, 1_E = 1_E$, $\mathrm{Im}\, 1_E = 0$, $(1_E)_+ = 1_E$ and $(1_E)_- = 0$. Definition 1.3.6 therefore gives

$$\mathrm{Simp} \int_{\mathbb{R}^d} 1_E(x)\, dx = \left(\mathrm{Simp} \int_{\mathbb{R}^d} 1_E(x)\, dx - \mathrm{Simp} \int_{\mathbb{R}^d} 0\, dx \right) + i \cdot 0,$$

where the two integrals in the bracket are unsigned ones; by Exercise 1.3.1(vi) the first is $m(E)$ and the second is 0, so the total is $m(E)$.
Uniqueness. Let $I : \mathrm{Simp}_{\mathrm{abs}}(\mathbb{R}^d) \to \mathbb{C}$ obey (i), (ii), (iii); we show $I(f) = \mathrm{Simp} \int f$ for all $f \in \mathrm{Simp}_{\mathrm{abs}}(\mathbb{R}^d)$. Only (i) and (iii) are needed. Fix such an f. By the preliminary remarks, f takes finitely many values and, writing $a_1, \dots, a_m$ for its non-zero values and $A_p := f^{-1}(\{a_p\})$, each A_p is measurable, the A_p are disjoint, and

$$f = a_1 1_{A_1} + \cdots + a_m 1_{A_m}.$$

Since f is absolutely integrable, and since $|f| \geq |a_p| 1_{A_p}$ pointwise for each fixed p, Exercise 1.3.1(v),(vi),(i) give $|a_p| m(A_p) \leq \mathrm{Simp} \int |f| < \infty$; as $|a_p| > 0$, this forces $m(A_p) < \infty$. Hence each 1_{A_p} lies in $\mathrm{Simp}_{\mathrm{abs}}(\mathbb{R}^d)$. Applying the additivity of (i) $m - 1$ times and then (1.11) to each summand,

$$I(f) = \sum_{p=1}^m I(a_p 1_{A_p}) = \sum_{p=1}^m a_p I(1_{A_p}) = \sum_{p=1}^m a_p m(A_p),$$

the last step by (iii). The same computation applies to the genuine simple integral (using the properties (i), (iii) already proved for it), giving $\mathrm{Simp} \int f = \sum_p a_p m(A_p)$. Hence $I(f) = \mathrm{Simp} \int f$. In the degenerate case $f \equiv 0$ (i.e. $m = 0$), (1.11) with $c = 0$ gives $I(0) = I(0 \cdot 0) = 0 \cdot I(0) = 0 = \mathrm{Simp} \int 0$. This proves uniqueness.

Problem 1.3.3.

- (i) Show that every continuous function $f : \mathbb{R}^d \to [0, +\infty]$ is measurable.
- (ii) Show that every unsigned simple function is measurable.
- (iii) Show that the supremum, infimum, limit superior, or limit inferior of unsigned measurable functions is unsigned measurable.
- (iv) Show that an unsigned function that is equal almost everywhere to an unsigned measurable function, is itself measurable.
- (v) Show that if a sequence f_n of unsigned measurable functions converges pointwise almost everywhere to an unsigned limit f, then f is also measurable.
- (vi) If $f : \mathbb{R}^d \to [0, +\infty]$ is measurable and $\varphi : [0, +\infty] \to [0, +\infty]$ is continuous, show that $\varphi \circ f : \mathbb{R}^d \to [0, +\infty]$ is measurable.

- (vii) If f, g are unsigned measurable functions, show that $f+g$ and fg are measurable.

Solution. We give $[0,+\infty]$ its usual topology: a set $W \subset [0,+\infty]$ is open iff $W \cap [0,+\infty)$ is open in $[0,+\infty)$ and, in the case $+\infty \in W$, W contains a ray $(\lambda,+\infty]$. In particular $[0,+\infty)$ is itself open, so any relatively open subset of $[0,+\infty)$ is open in $[0,+\infty]$. Throughout we freely use the equivalences of Lemma 1.3.9, referring to its clauses by number. In "supremum, infimum, limit superior, limit inferior" we understand, as usual, an at most countable family $f_1, f_2, \dots$ of functions.
A preliminary claim. If $f : \mathbb{R}^d \to [0,+\infty]$ is measurable and $W \subset [0,+\infty]$ is open, then $f^{-1}(W)$ is Lebesgue measurable. Indeed, $W \cap [0,+\infty)$ is relatively open in $[0,+\infty)$, so $f^{-1}(W \cap [0,+\infty))$ is measurable by Lemma 1.3.9(xi); and $f^{-1}(\{+\infty\}) = \{x \in \mathbb{R}^d : f(x) \geq +\infty\}$ is measurable by Lemma 1.3.9(vi) applied with $\lambda = +\infty$. Since $f^{-1}(W)$ equals $f^{-1}(W \cap [0,+\infty))$ if $+\infty \notin W$, and equals $f^{-1}(W \cap [0,+\infty)) \cup f^{-1}(\{+\infty\})$ if $+\infty \in W$, it is measurable in either case.
(i) Let $f : \mathbb{R}^d \to [0,+\infty]$ be continuous. We verify Lemma 1.3.9(v). Fix $\lambda \in [0,+\infty]$. If $\lambda = +\infty$ then $\{x : f(x) > \lambda\} = \emptyset$, which is measurable. If $\lambda < +\infty$, then $(\lambda,+\infty]$ is an open subset of $[0,+\infty]$, so by continuity

$$\{x \in \mathbb{R}^d : f(x) > \lambda\} = f^{-1}((\lambda,+\infty])$$

is an open subset of $\mathbb{R}^d$, and every open subset of $\mathbb{R}^d$ is Lebesgue measurable. Hence (v) holds and f is measurable.
(ii) Let f be an unsigned simple function. Set $f_n := f$ for every n. This is a sequence of unsigned simple functions with $f_n(x) \to f(x)$ for every $x \in \mathbb{R}^d$ (the sequence is constant). By Definition 1.3.8, f is measurable.
(iii) Let $f_1, f_2, \cdots : \mathbb{R}^d \to [0,+\infty]$ be measurable.
Supremum: put $F := \sup_n f_n$. For any $\lambda \in [0,+\infty]$ we have $F(x) > \lambda$ if and only if $f_n(x) > \lambda$ for at least one n (if $f_n(x) \leq \lambda$ for all n then λ is an upper bound, so $F(x) \leq \lambda$; conversely $f_n(x) > \lambda$ forces $F(x) \geq f_n(x) > \lambda$). Hence

$$\{x : F(x) > \lambda\} = \bigcup_{n=1}^{\infty} \{x : f_n(x) > \lambda\},$$

a countable union of Lebesgue measurable sets (Lemma 1.3.9(v) for each f_n), hence Lebesgue measurable. By Lemma 1.3.9(v), F is measurable.
Infimum: put $G := \inf_n f_n$. Symmetrically, $G(x) < \lambda$ if and only if $f_n(x) < \lambda$ for some n, so

$$\{x : G(x) < \lambda\} = \bigcup_{n=1}^{\infty} \{x : f_n(x) < \lambda\}$$

is measurable for every λ by Lemma 1.3.9(vii) applied to each f_n; hence G is measurable by Lemma 1.3.9(vii).
Limit superior and limit inferior: by definition

$$\limsup_{n\to\infty} f_n = \inf_{N\geq 1} \left(\sup_{n\geq N} f_n \right), \qquad \liminf_{n\to\infty} f_n = \sup_{N\geq 1} \left(\inf_{n\geq N} f_n \right).$$

By the supremum case, each $\sup_{n\geq N} f_n$ is measurable; by the infimum case, the countable infimum of these is measurable. Hence $\limsup_n f_n$ is measurable, and symmetrically so is $\liminf_n f_n$.
(iv) Let $g : \mathbb{R}^d \to [0,+\infty]$ be measurable and let $f : \mathbb{R}^d \to [0,+\infty]$ agree with g outside a null set N. Fix $\lambda \in [0,+\infty]$. Then

$$\{f > \lambda\} = (\{g > \lambda\} \setminus N) \cup (\{f > \lambda\} \cap N),$$

since off N the conditions $f > \lambda$ and $g > \lambda$ are the same. The first set is measurable, being the difference of the measurable set $\{g > \lambda\}$ (Lemma 1.3.9(v)) and the null set N (null sets are Lebesgue measurable). The second set is a subset of the null set N, hence is itself null and therefore measurable. So $\{f > \lambda\}$ is measurable for every λ, and f is measurable by Lemma 1.3.9(v).
(v) Suppose $f_n \to f$ pointwise almost everywhere, with each f_n measurable. Put $g := \limsup_{n\to\infty} f_n$, which is measurable by (iii). Let E be the co-null set of x for which $f_n(x) \to f(x)$. For $x \in E$ the

limit exists, so it equals the limit superior, i.e. $f(x) = g(x)$. Hence $f = g$ almost everywhere, and f is measurable by (iv).
(vi) Let f be measurable and $\varphi : [0, +\infty] \to [0, +\infty]$ continuous. We verify Lemma 1.3.9(xi) for $\varphi \circ f$. Let $U \subset [0, +\infty)$ be relatively open; as noted above, U is then open in $[0, +\infty]$, so by continuity of φ the set $W := \varphi^{-1}(U)$ is open in $[0, +\infty]$. Therefore

$$(\varphi \circ f)^{-1}(U) = f^{-1}(\varphi^{-1}(U)) = f^{-1}(W)$$

is Lebesgue measurable by the preliminary claim. By Lemma 1.3.9(xi), $\varphi \circ f$ is measurable.
(vii) Let $f, g : \mathbb{R}^d \to [0, +\infty]$ be measurable. By Lemma 1.3.9(iv) there are increasing sequences $0 \leq f_1 \leq f_2 \leq \dots$ and $0 \leq g_1 \leq g_2 \leq \dots$ of unsigned simple functions with $f = \sup_n f_n$ and $g = \sup_n g_n$; since the sequences are increasing, $f_n(x) \to f(x)$ and $g_n(x) \to g(x)$ for every x.
Sums. Each $f_n + g_n$ is an unsigned simple function (the sum of two finite unsigned combinations of indicators of measurable sets is again one, as noted after Definition 1.3.2). For each fixed x, the sequences $f_n(x)$ and $g_n(x)$ increase to $f(x)$ and $g(x)$ in $[0, +\infty]$, and for increasing sequences in $[0, +\infty]$ one has $\lim_n(a_n + b_n) = \lim_n a_n + \lim_n b_n$: if both limits are finite this is the usual limit law, and if (say) $\lim_n a_n = +\infty$ then $a_n + b_n \geq a_n \to +\infty$, matching the right-hand side. Hence $f_n + g_n \to f + g$ pointwise everywhere, and $f + g$ is measurable by Definition 1.3.8.
Products. First, the product of two unsigned simple functions is unsigned simple: writing them in canonical form $f_n = \sum_p a_p 1_{A_p}$, $g_n = \sum_q b_q 1_{B_q}$ with $\{A_p\}$ and $\{B_q\}$ finite measurable partitions of $\mathbb{R}^d$ (see the atom construction in the text before Lemma 1.3.4), we get

$$f_n g_n = \sum_{p,q} a_p b_q 1_{A_p \cap B_q},$$

a finite unsigned combination of indicators of measurable sets. Next we check that $f_n(x)g_n(x) \to f(x)g(x)$ for every x, using the convention $0 \cdot \infty = 0$. Write $a_n := f_n(x) \uparrow a := f(x)$ and $b_n := g_n(x) \uparrow b := g(x)$. If $a = 0$ then $a_n = 0$ for all n (the a_n are non-negative and increase to 0), so $a_n b_n = 0 \to 0 = ab$; symmetrically if $b = 0$. If $a, b > 0$ and both are finite, this is the usual product limit law. If $a, b > 0$ and, say, $a = +\infty$, then choose N with $b_N > 0$ (possible since $b_n \uparrow b > 0$) and set $\beta := b_N > 0$; for $n \geq N$ we have $a_n b_n \geq \beta a_n \to +\infty$, so $a_n b_n \to +\infty = ab$; the case $b = +\infty$ is symmetric. Hence $f_n g_n \to fg$ pointwise everywhere, and fg is measurable by Definition 1.3.8.

Problem 1.3.4. Let $f : \mathbb{R}^d \to [0, +\infty]$. Show that f is a bounded unsigned measurable function if and only if f is the uniform limit of bounded simple functions.

Solution. Here "bounded" means that there is a finite M with $0 \leq f(x) \leq M$ for all x (in particular a bounded function is finite-valued), and "uniform limit" means that there are unsigned simple functions f_n with $\sup_{x \in \mathbb{R}^d} |f_n(x) - f(x)| \to 0$ as $n \to \infty$, where we interpret $|a - b| = +\infty$ whenever exactly one of a, b equals $+\infty$.
Necessity. Suppose f is bounded and measurable, say $0 \leq f \leq M < \infty$. For each $n \geq 1$ define

$$f_n(x) := \frac{1}{n} \lfloor n f(x) \rfloor,$$

i.e. $f_n(x)$ is the largest integer multiple of $1/n$ not exceeding $f(x)$. Explicitly, f_n takes only the finitely many values k/n with $0 \leq k \leq \lfloor nM \rfloor$, and

$$f_n^{-1}\left(\left\{\tfrac{k}{n}\right\}\right) = f^{-1}\left(\left[\tfrac{k}{n}, \tfrac{k+1}{n}\right)\right),$$

which is Lebesgue measurable by Lemma 1.3.9(x), since $[k/n, (k+1)/n)$ is an interval in $[0, +\infty)$. Hence

$$f_n = \sum_{k=0}^{\lfloor nM \rfloor} \frac{k}{n} 1_{f^{-1}([k/n, (k+1)/n))}$$

is a finite unsigned combination of indicators of measurable sets, i.e. an unsigned simple function by Definition 1.3.2; and $0 \leq f_n \leq M$, so it is bounded. Finally, by construction $0 \leq f(x) - f_n(x) < 1/n$ for every x, so

$$\sup_{x \in \mathbb{R}^d} |f_n(x) - f(x)| \leq \frac{1}{n} \to 0,$$

i.e. $f_n \to f$ uniformly. Thus f is the uniform limit of bounded simple functions.
Sufficiency. Conversely, suppose f_n are bounded unsigned simple functions with $f_n \to f$ uniformly. First, f is finite everywhere: if $f(x_0) = +\infty$ for some x_0, then since each f_n is bounded (hence finite) we would have $|f_n(x_0) - f(x_0)| = +\infty$ for every n, contradicting uniform convergence. Next, f is bounded: choose n_0 with $\sup_x |f_{n_0}(x) - f(x)| \leq 1$, and let $M_{n_0} < \infty$ bound f_{n_0}; then $f(x) \leq f_{n_0}(x) + 1 \leq M_{n_0} + 1$ for every x. Finally, uniform convergence implies pointwise convergence, so f is the pointwise limit of the unsigned simple functions f_n, and is therefore measurable by Definition 1.3.8.

Problem 1.3.5. Show that an unsigned function $f : \mathbb{R}^d \to [0, +\infty]$ is a simple function if and only if it is measurable and takes on at most finitely many values.

Solution. Necessity. Suppose f is an unsigned simple function, say $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ with $c_i \in [0, +\infty]$ and $E_i \subset \mathbb{R}^d$ Lebesgue measurable. By Exercise 1.3.3(ii), f is measurable. Moreover, as observed in the text preceding Lemma 1.3.4, the sets $E_1, \ldots, E_k$ partition $\mathbb{R}^d$ into at most 2^k disjoint atoms $\bigcap_{i \in S} E_i \cap \bigcap_{i \notin S} (\mathbb{R}^d \setminus E_i)$, indexed by $S \subset \{1, \ldots, k\}$, and on the atom indexed by S the function f is constant, with value $\sum_{i \in S} c_i$. Since these atoms cover $\mathbb{R}^d$, the range of f is contained in the finite set $\{\sum_{i \in S} c_i : S \subset \{1, \ldots, k\}\}$, so f takes at most 2^k values.
Sufficiency. Conversely, suppose f is measurable and its range is a finite set $\{a_1, \ldots, a_m\} \subset [0, +\infty]$, with the a_p distinct. Put $A_p := f^{-1}(\{a_p\})$. Each A_p is Lebesgue measurable: indeed

$$A_p = \{x : f(x) \geq a_p\} \cap \{x : f(x) \leq a_p\},$$

where $\{f \geq a_p\}$ is measurable by Lemma 1.3.9(vi) and $\{f \leq a_p\}$ is the complement of the measurable set $\{f > a_p\}$ (Lemma 1.3.9(v)), hence measurable; and the intersection of two measurable sets is measurable. The A_p are pairwise disjoint and cover $\mathbb{R}^d$ (every x has $f(x)$ equal to exactly one a_p), so

$$f = a_1 1_{A_1} + \cdots + a_m 1_{A_m}$$

holds pointwise: at a point $x \in A_p$ the right-hand side equals $a_p = f(x)$, all other terms vanishing. This exhibits f as a finite unsigned combination of indicator functions of Lebesgue measurable sets, so f is an unsigned simple function by Definition 1.3.2.

Problem 1.3.6. Let $f : \mathbb{R}^d \to [0, +\infty]$ be an unsigned measurable function. Show that the region $\{(x, t) \in \mathbb{R}^d \times \mathbb{R} : 0 \leq t \leq f(x)\}$ is a measurable subset of $\mathbb{R}^{d+1}$. (There is a converse to this statement, but we will wait until Exercise 1.7.24 to prove it, once we have the Fubini-Tonelli theorem (Corollary 1.7.23) available to us.)

Solution. For an unsigned function $g : \mathbb{R}^d \to [0, +\infty]$ write

$$R(g) := \{(x, t) \in \mathbb{R}^d \times \mathbb{R} : 0 \leq t < g(x)\}$$

for the "open" region under the graph of g, and let

$$E := \{(x, t) \in \mathbb{R}^d \times \mathbb{R} : 0 \leq t \leq f(x)\}$$

be the region we must show is measurable.
Step 1: $R(g)$ is measurable whenever g is an unsigned simple function. By the atom construction preceding Lemma 1.3.4 we may write g in canonical form $g = \sum_{p=1}^m a_p 1_{A_p}$, where $a_1, \ldots, a_m \in [0, +\infty]$ are the distinct values of g and $A_p := g^{-1}(\{a_p\})$ are disjoint Lebesgue measurable sets covering $\mathbb{R}^d$.

For $x \in A_p$, the slice $\{t : (x,t) \in R(g)\}$ is exactly $[0, a_p)$, which is the empty set if $a_p = 0$, the interval $[0, a_p)$ if $0 < a_p < \infty$, and the ray $[0, +\infty)$ if $a_p = +\infty$. Hence

$$R(g) = \bigcup_{p=1}^{m} A_p \times [0, a_p).$$

Each A_p is Lebesgue measurable in $\mathbb{R}^d$ and each $[0, a_p)$ is Lebesgue measurable in $\mathbb{R}$ (it is either empty, or a bounded interval and hence an elementary set, or the ray $[0, +\infty) = \bigcup_{k \geq 1} [0, k)$, a countable union of such), so by Exercise 1.2.22(ii) each product $A_p \times [0, a_p) \subset \mathbb{R}^{d+1}$ is Lebesgue measurable. A finite union of measurable sets is measurable, so $R(g)$ is measurable.
Step 2: $R(g)$ is measurable whenever g is unsigned measurable. By Lemma 1.3.9(iv) there is an increasing sequence $0 \leq g_1 \leq g_2 \leq \ldots$ of unsigned simple functions with $g = \sup_n g_n$. I claim

$$R(g) = \bigcup_{n=1}^{\infty} R(g_n).$$

Indeed, if $0 \leq t < g_n(x)$ for some n then $t < g_n(x) \leq g(x)$, giving the inclusion $\supset$. Conversely, if $0 \leq t < g(x) = \sup_n g_n(x)$, then t is not an upper bound for the $g_n(x)$, so $t < g_n(x)$ for some n, giving $\subset$. Each $R(g_n)$ is measurable by Step 1, and countable unions of Lebesgue measurable sets are Lebesgue measurable, so $R(g)$ is measurable.
Step 3: expressing E as a countable intersection. For each $n \geq 1$ let $f + \frac{1}{n}$ denote the function $x \mapsto f(x) + \frac{1}{n}$ (with $+\infty + \frac{1}{n} = +\infty$). The constant function $\frac{1}{n}$ is unsigned simple, hence measurable by Exercise 1.3.3(ii), so $f + \frac{1}{n}$ is unsigned measurable by Exercise 1.3.3(vii). I claim

$$E = \bigcap_{n=1}^{\infty} R\big(f + \tfrac{1}{n}\big).$$

Suppose first $(x,t) \in E$, so $t \geq 0$ and $t \leq f(x)$. Then for every n we have $t \leq f(x) < f(x) + \frac{1}{n}$ if $f(x) < \infty$, and $t < +\infty = f(x) + \frac{1}{n}$ if $f(x) = +\infty$ (recall t is a real number). Either way $0 \leq t < f(x) + \frac{1}{n}$, so $(x,t) \in R(f + \frac{1}{n})$ for every n.
Conversely, suppose $0 \leq t < f(x) + \frac{1}{n}$ for every $n \geq 1$. If $f(x) = +\infty$ then certainly $t \leq f(x)$. If $f(x) < \infty$, then $t - \frac{1}{n} < f(x)$ for all n, and letting $n \to \infty$ gives $t \leq f(x)$. In both cases $(x,t) \in E$.
By Step 2 each $R(f + \frac{1}{n})$ is a Lebesgue measurable subset of $\mathbb{R}^{d+1}$, and countable intersections of Lebesgue measurable sets are Lebesgue measurable. Hence E is a measurable subset of $\mathbb{R}^{d+1}$, as required.

Problem 1.3.7. Let $f : \mathbb{R}^d \to \mathbb{C}$ be an almost everywhere defined complex-valued function. Then the following are equivalent:

- (i) f is measurable.
- (ii) f is the pointwise almost everywhere limit of complex-valued simple functions.
- (iii) The (magnitudes of the) positive and negative parts of $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are unsigned measurable functions.
- (iv) $f^{-1}(U)$ is Lebesgue measurable for every open set $U \subset \mathbb{C}$.
- (v) $f^{-1}(K)$ is Lebesgue measurable for every closed set $K \subset \mathbb{C}$.

Solution. Reduction to everywhere-defined functions. Say f is defined on $\mathbb{R}^d \setminus N_0$ with N_0 null, and extend it to all of $\mathbb{R}^d$ by setting $f := 0$ on N_0. This is harmless for every one of the five statements. Statements (i) and (ii) are unaffected because the notion of almost everywhere convergence ignores null sets. Statement (iii) is unaffected because, as remarked after Exercise 1.3.3(iv), unsigned measurability of an almost everywhere defined function means measurability of an arbitrary extension, and by Exercise 1.3.3(iv) any two extensions are simultaneously measurable. Statements (iv) and (v) are unaffected because changing f on the null set N_0 changes each preimage $f^{-1}(S)$ only within N_0, i.e. by

a subset of a null set, and such a modification takes Lebesgue measurable sets to Lebesgue measurable sets. So assume from now on that f is defined on all of $\mathbb{R}^d$, and write

$$u := (\operatorname{Re} f)_+, \quad v := (\operatorname{Re} f)_-, \quad p := (\operatorname{Im} f)_+, \quad q := (\operatorname{Im} f)_-,$$

where $s_+ := \max(s, 0)$ and $s_- := \max(-s, 0)$ for $s \in \mathbb{R}$. These are functions $\mathbb{R}^d \to [0, +\infty)$, and $\operatorname{Re} f = u - v$, $\operatorname{Im} f = p - q$.
Statements (i) and (ii) are literally the same statement, by Definition 1.3.11. We prove (ii) $\Rightarrow$ (iii) $\Rightarrow$ (ii), (iii) $\Rightarrow$ (iv) $\Rightarrow$ (iii), and (iv) $\Leftrightarrow$ (v).
(ii) $\Rightarrow$ (iii). Let f_n be complex-valued simple functions with $f_n \to f$ pointwise almost everywhere, say on the co-null set E. As noted in the solution to Exercise 1.3.2, if h is a complex simple function then so is $\Phi \circ h$ for any Φ with $\Phi(0) = 0$; taking $\Phi(z) = (\operatorname{Re} z)_+$ we see that $u_n := (\operatorname{Re} f_n)_+$ is an unsigned simple function. For $x \in E$ we have $f_n(x) \to f(x)$ in $\mathbb{C}$, hence $\operatorname{Re} f_n(x) \to \operatorname{Re} f(x)$ in $\mathbb{R}$, hence $u_n(x) \to u(x)$, because $s \mapsto s_+$ is continuous on $\mathbb{R}$. So u is the pointwise almost everywhere limit of unsigned simple functions, and is therefore unsigned measurable by Lemma 1.3.9(iii). The same argument with $\Phi(z) = (\operatorname{Re} z)_-, (\operatorname{Im} z)_+, (\operatorname{Im} z)_-$ shows that v, p, q are unsigned measurable.
(iii) $\Rightarrow$ (ii). Suppose u, v, p, q are unsigned measurable. By Lemma 1.3.9(iv) applied to u, there is an increasing sequence of unsigned simple functions u_n, each of them bounded, with $u = \sup_n u_n$; since the sequence is increasing, $u_n(x) \to u(x)$ for every x. Being bounded, each u_n is finite-valued; rewriting it in canonical form $u_n = \sum_p a_p 1_{A_p}$ over its (finitely many, now finite) distinct values, as in the atom construction preceding Lemma 1.3.4, exhibits it as a finite linear combination of indicators of measurable sets with coefficients in $\mathbb{C}$, so $u_n \in \operatorname{Simp}(\mathbb{R}^d)$, the complex vector space of complex-valued simple functions of Definition 1.3.2. Choose likewise bounded unsigned simple $v_n \to v$, $p_n \to p$, $q_n \to q$ pointwise. Then

$$f_n := (u_n - v_n) + i(p_n - q_n)$$

is a complex-valued simple function, since $\operatorname{Simp}(\mathbb{R}^d)$ is a complex vector space. For every x, all eight quantities involved are finite reals, so

$$f_n(x) \to (u(x) - v(x)) + i(p(x) - q(x)) = \operatorname{Re} f(x) + i \operatorname{Im} f(x) = f(x).$$

Thus f is the pointwise (indeed everywhere) limit of complex-valued simple functions, giving (ii).
(iii) $\Rightarrow$ (iv). Assume u, v, p, q are unsigned measurable. We first show that $\{x : \operatorname{Re} f(x) > \lambda\}$ and $\{x : \operatorname{Re} f(x) < \lambda\}$ are Lebesgue measurable for every $\lambda \in \mathbb{R}$.
If $\lambda \geq 0$ then $\operatorname{Re} f(x) > \lambda$ holds if and only if $u(x) > \lambda$: for if $\operatorname{Re} f(x) > \lambda \geq 0$ then $u(x) = \operatorname{Re} f(x) > \lambda$, while if $u(x) > \lambda \geq 0$ then $u(x) \neq 0$, so $u(x) = \operatorname{Re} f(x)$. Hence $\{\operatorname{Re} f > \lambda\} = \{u > \lambda\}$ is measurable by Lemma 1.3.9(v). If $\lambda < 0$ then $\operatorname{Re} f(x) > \lambda$ holds if and only if $v(x) < -\lambda$: if $\operatorname{Re} f(x) > \lambda$ then $-\operatorname{Re} f(x) < -\lambda$ and $0 < -\lambda$, so $v(x) = \max(-\operatorname{Re} f(x), 0) < -\lambda$; conversely $-\operatorname{Re} f(x) \leq v(x) < -\lambda$ gives $\operatorname{Re} f(x) > \lambda$. Hence $\{\operatorname{Re} f > \lambda\} = \{v < -\lambda\}$ is measurable by Lemma 1.3.9(vii).
Symmetrically, if $\lambda \leq 0$ then $\operatorname{Re} f(x) < \lambda$ if and only if $v(x) > -\lambda$, and if $\lambda > 0$ then $\operatorname{Re} f(x) < \lambda$ if and only if $u(x) < \lambda$; in either case the set is measurable by Lemma 1.3.9(v) or (vii). The same statements hold for $\operatorname{Im} f$ in place of $\operatorname{Re} f$, using p, q instead of u, v.
Consequently, for any real numbers $a < b$ and $c < d$, the preimage of the open rectangle $Q = \{z \in \mathbb{C} : a < \operatorname{Re} z < b,\ c < \operatorname{Im} z < d\}$ is

$$f^{-1}(Q) = \{\operatorname{Re} f > a\} \cap \{\operatorname{Re} f < b\} \cap \{\operatorname{Im} f > c\} \cap \{\operatorname{Im} f < d\},$$

a finite intersection of measurable sets, hence measurable. Finally, let $U \subset \mathbb{C}$ be open. Every point of U lies in an open rectangle with rational corners contained in U, so U is the union of the (countably many) rational-cornered open rectangles it contains; writing $U = \bigcup_{j=1}^{\infty} Q_j$ accordingly, we get

$$f^{-1}(U) = \bigcup_{j=1}^{\infty} f^{-1}(Q_j),$$

a countable union of measurable sets, hence Lebesgue measurable. This is (iv).

(iv) ⇒ (iii). Assume $f^{-1}(U)$ is measurable for every open $U \subset \mathbb{C}$. We verify Lemma 1.3.9(v) for $u = (\operatorname{Re} f)_+$. Let $\lambda \in [0, +\infty]$. If $\lambda = +\infty$ then $\{u > \lambda\} = \emptyset$ (as u is finite-valued), which is measurable. If $\lambda \in [0, +\infty)$ then, as computed above, $\{u > \lambda\} = \{\operatorname{Re} f > \lambda\} = f^{-1}(U_\lambda)$ where $U_\lambda := \{z \in \mathbb{C} : \operatorname{Re} z > \lambda\}$ is an open half-plane; so $\{u > \lambda\}$ is measurable by hypothesis. Hence u is unsigned measurable. Similarly, for $\lambda \in [0, +\infty)$ we have $\{v > \lambda\} = \{\operatorname{Re} f < -\lambda\} = f^{-1}(\{z : \operatorname{Re} z < -\lambda\})$, an open half-plane preimage, so v is unsigned measurable; and the same argument applied to the half-planes $\{\operatorname{Im} z > \lambda\}$ and $\{\operatorname{Im} z < -\lambda\}$ shows that p and q are unsigned measurable. This is (iii).
(iv) ⇔ (v). For any $S \subset \mathbb{C}$ we have $f^{-1}(\mathbb{C} \setminus S) = \mathbb{R}^d \setminus f^{-1}(S)$, and the Lebesgue measurable sets form a σ-algebra, so are closed under complementation. Thus if (iv) holds and $K \subset \mathbb{C}$ is closed, then $U := \mathbb{C} \setminus K$ is open and $f^{-1}(K) = \mathbb{R}^d \setminus f^{-1}(U)$ is measurable, giving (v). Conversely, if (v) holds and $U \subset \mathbb{C}$ is open, then $K := \mathbb{C} \setminus U$ is closed and $f^{-1}(U) = \mathbb{R}^d \setminus f^{-1}(K)$ is measurable, giving (iv).
This closes all the implications, so (i)–(v) are equivalent.

Problem 1.3.8. (i) Show that every continuous function $f : \mathbb{R}^d \to \mathbb{C}$ is measurable.
(ii) Show that a function $f : \mathbb{R}^d \to \mathbb{C}$ is simple if and only if it is measurable and takes on at most finitely many values.
(iii) Show that a complex-valued function that is equal almost everywhere to a measurable function, is itself measurable.
(iv) Show that if a sequence f_n of complex-valued measurable functions converges pointwise almost everywhere to a complex-valued limit f, then f is also measurable.
(v) If $f : \mathbb{R}^d \to \mathbb{C}$ is measurable and $\varphi : \mathbb{C} \to \mathbb{C}$ is continuous, show that $\varphi \circ f : \mathbb{R}^d \to \mathbb{C}$ is measurable.
(vi) If f, g are measurable functions, show that $f + g$ and fg are measurable.

Solution. Throughout we use the equivalences of Exercise 1.3.7: an almost everywhere defined $f : \mathbb{R}^d \to \mathbb{C}$ is measurable if and only if it is a pointwise a.e. limit of complex-valued simple functions, if and only if the magnitudes of the positive and negative parts of $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$ are unsigned measurable, if and only if $f^{-1}(U)$ is Lebesgue measurable for every open $U \subset \mathbb{C}$, if and only if $f^{-1}(K)$ is Lebesgue measurable for every closed $K \subset \mathbb{C}$.
(i) Let f be continuous and let $U \subset \mathbb{C}$ be open. By continuity $f^{-1}(U)$ is an open subset of $\mathbb{R}^d$, and every open set is Lebesgue measurable by Lemma 1.2.13(i). Thus $f^{-1}(U)$ is Lebesgue measurable for every open U, and f is measurable by Exercise 1.3.7(iv).
(ii) Suppose first that f is simple, say $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ with $E_1, \ldots, E_k \subset \mathbb{R}^d$ Lebesgue measurable and $c_1, \ldots, c_k \in \mathbb{C}$. The k sets $E_1, \ldots, E_k$ partition $\mathbb{R}^d$ into at most 2^k disjoint (measurable) atoms, on each of which every 1_{E_i}, and hence f, is constant; so f takes at most 2^k distinct values, in particular finitely many. Moreover f is the pointwise limit of the constant sequence $f, f, f, \ldots$ of simple functions, so f is measurable by Definition 1.3.11.
Conversely, suppose f is measurable and its range is the finite set $\{c_1, \ldots, c_k\}$ of distinct complex numbers. For each i the singleton $\{c_i\}$ is a closed subset of $\mathbb{C}$, so by Exercise 1.3.7(v) the set $E_i := f^{-1}(\{c_i\})$ is Lebesgue measurable. These sets are disjoint and cover $\mathbb{R}^d$, and f is equal to c_i on E_i; hence

$$f = c_1 1_{E_1} + \cdots + c_k 1_{E_k},$$

which is a finite linear combination of indicators of measurable sets, i.e. a simple function in the sense of Definition 1.3.2.
(iii) Let g be measurable and $f = g$ outside a null set N. By Definition 1.3.11 there are simple g_n with $g_n \to g$ pointwise outside a null set N'. Then $g_n \to f$ pointwise outside $N \cup N'$, which is null (a finite union of null sets is null). So f is a pointwise a.e. limit of simple functions, hence measurable.
(iv) Write $u := \operatorname{Re}(f)$ and $u_n := \operatorname{Re}(f_n)$, and likewise v, v_n for the imaginary parts. Off a null set (the union of the countably many null sets on which the f_n are undefined, together with the null set where convergence fails, which is again null) we have $u_n \to u$ and $v_n \to v$ pointwise. The maps $t \mapsto \max(t, 0)$ and $t \mapsto \max(-t, 0)$ are continuous on $\mathbb{R}$, so $(u_n)_+ \to u_+$, $(u_n)_- \to u_-$, $(v_n)_+ \to v_+$ and $(v_n)_- \to v_-$ pointwise almost everywhere. By Exercise 1.3.7(iii) applied to each measurable f_n, the four sequences consist of unsigned measurable functions; by Exercise 1.3.3(v) (a pointwise a.e. limit

of unsigned measurable functions is unsigned measurable) the four limits u_+, u_-, v_+, v_- are unsigned measurable. By Exercise 1.3.7(iii) again, now in the reverse direction, f is measurable.
(v) Let $U \subset \mathbb{C}$ be open. Since φ is continuous, $\varphi^{-1}(U)$ is open in $\mathbb{C}$, and

$$(\varphi \circ f)^{-1}(U) = f^{-1}(\varphi^{-1}(U))$$

is Lebesgue measurable by Exercise 1.3.7(iv) applied to f. (If f is only defined almost everywhere, all preimages are understood as subsets of the co-null set where f is defined, and adjoining or deleting a null set preserves measurability.) Hence $\varphi \circ f$ is measurable, by Exercise 1.3.7(iv) again.
(vi) By Definition 1.3.11 choose simple functions f_n, g_n with $f_n \to f$ and $g_n \to g$ pointwise almost everywhere; the exceptional set is the union of two null sets and hence null. The space $\mathrm{Simp}(\mathbb{R}^d)$ of complex-valued simple functions is a commutative $*$-algebra, so $f_n + g_n$ and $f_n g_n$ are again simple. Since addition and multiplication are continuous on $\mathbb{C}$, we get $f_n + g_n \to f + g$ and $f_n g_n \to fg$ pointwise almost everywhere. Thus $f + g$ and fg are pointwise a.e. limits of simple functions, hence measurable.

Problem 1.3.9. Let $f : [a, b] \to \mathbb{R}$ be a Riemann integrable function. Show that if one extends f to all of $\mathbb{R}$ by defining $f(x) = 0$ for $x \notin [a, b]$, then f is measurable.

Solution. Recall that a Riemann integrable function is by definition bounded, and that by Exercise 1.1.22 Riemann integrability is equivalent to Darboux integrability: the lower and upper Darboux integrals of Definition 1.1.6,

$$\underline{\int_a^b} f = \sup_{g \le f \text{ p.c.}} \text{p.c.} \int_a^b g(x)\,dx, \qquad \overline{\int_a^b} f = \inf_{h \ge f \text{ p.c.}} \text{p.c.} \int_a^b h(x)\,dx,$$

are equal. Hence for each $n \ge 1$ we may pick piecewise constant functions $g_n, h_n : [a, b] \to \mathbb{R}$ with $g_n \le f \le h_n$ on $[a, b]$ and

$$\text{p.c.} \int_a^b h_n(x)\,dx - \text{p.c.} \int_a^b g_n(x)\,dx \le \frac{1}{n}.$$

Extend g_n, h_n to all of $\mathbb{R}$ by zero, and keep the extension of f by zero. Since f vanishes outside $[a, b]$, we still have $g_n \le f \le h_n$ on all of $\mathbb{R}$.
Each extended g_n (and h_n) is a simple function: by definition a piecewise constant function is constant on each interval of a partition of $[a, b]$ into finitely many intervals $I_1, \ldots, I_m$, so the extension is $\sum_{j=1}^m c_j 1_{I_j}$, a finite linear combination of indicator functions of intervals, and intervals are Lebesgue measurable. Thus g_n, h_n are simple in the sense of Definition 1.3.2.
Now set

$$G_n := \max(g_1, \ldots, g_n), \qquad H_n := \min(h_1, \ldots, h_n).$$

Pointwise maxima and minima of finitely many piecewise constant functions are piecewise constant (refine the partitions), so G_n, H_n are again simple functions vanishing outside $[a, b]$, and $G_n \le f \le H_n$ everywhere. By construction $G_1 \le G_2 \le \ldots$ and $H_1 \ge H_2 \ge \ldots$, and all of them are bounded (between $\inf f$ and $\sup f$ on $[a, b]$, and 0 off $[a, b]$; more crudely, $G_n \ge g_1$ and $H_n \le h_1$, both bounded). Being monotone and bounded, the limits

$$G(x) := \lim_{n\to\infty} G_n(x), \qquad H(x) := \lim_{n\to\infty} H_n(x)$$

exist for every $x \in \mathbb{R}$, and $G \le f \le H$ pointwise. As pointwise limits of simple functions, G and H are measurable (Definition 1.3.11).
It remains to show $G = H$ almost everywhere; then $G = f = H$ off a null set, so f agrees almost everywhere with the measurable function G, and f is measurable by Exercise 1.3.8(iii).
Fix $\varepsilon > 0$ and put $\phi_n := H_n - G_n$, a non-negative piecewise constant function on $[a, b]$ (vanishing off $[a, b]$). Since $G_n \ge g_n$ and $H_n \le h_n$, monotonicity of the piecewise constant integral (Exercise 1.1.21) gives

$$\text{p.c.} \int_a^b \phi_n(x)\,dx \le \text{p.c.} \int_a^b h_n(x)\,dx - \text{p.c.} \int_a^b g_n(x)\,dx \le \frac{1}{n}.$$

The set $A_{n,\varepsilon} := \{x \in [a,b] : \phi_n(x) \geq \varepsilon\}$ is a finite union of intervals, hence elementary, and $\varepsilon 1_{A_{n,\varepsilon}} \leq \phi_n$ pointwise on $[a,b]$. By monotonicity, linearity and the indicator property of the piecewise constant integral (Exercise 1.1.21),

$$\varepsilon\, m(A_{n,\varepsilon}) = \text{p.c.}\int_a^b \varepsilon 1_{A_{n,\varepsilon}}(x)\,dx \leq \text{p.c.}\int_a^b \phi_n(x)\,dx \leq \frac{1}{n},$$

so $m(A_{n,\varepsilon}) \leq \frac{1}{\varepsilon n}$.
Since $H \leq H_n$ and $G \geq G_n$ for every n, we have $0 \leq H - G \leq \phi_n$ pointwise, whence

$$\{x \in \mathbb{R} : H(x) - G(x) \geq \varepsilon\} \subset A_{n,\varepsilon}$$

for every n (the left-hand set is contained in $[a,b]$ because $H = G = 0$ outside). By monotonicity of Lebesgue outer measure, together with the fact that the Lebesgue outer measure of an elementary set equals its elementary measure (Lemma 1.2.6),

$$m^*(\{H - G \geq \varepsilon\}) \leq m(A_{n,\varepsilon}) \leq \frac{1}{\varepsilon n}$$

for all n, and letting $n \to \infty$ we conclude $m^*(\{H - G \geq \varepsilon\}) = 0$; that is, $\{H - G \geq \varepsilon\}$ is a null set. Finally

$$\{x : H(x) > G(x)\} = \bigcup_{k=1}^{\infty} \{x : H(x) - G(x) \geq 1/k\}$$

is a countable union of null sets, hence null. Therefore $H = G$ almost everywhere, and by the sandwich $G \leq f \leq H$ we get $f = G$ almost everywhere. Since G is measurable, so is f, by Exercise 1.3.8(iii).

Problem 1.3.10. (Basic properties of the lower Lebesgue integral) Let $f, g : \mathbb{R}^d \to [0, +\infty]$ be unsigned functions (not necessarily measurable).
(i) (Compatibility with the simple integral) If f is simple, then $\underline{\int_{\mathbb{R}^d}} f(x)\,dx = \overline{\int_{\mathbb{R}^d}} f(x)\,dx = \text{Simp}\int_{\mathbb{R}^d} f(x)\,dx$.
(ii) (Monotonicity) If $f \leq g$ pointwise almost everywhere, then $\underline{\int_{\mathbb{R}^d}} f(x)\,dx \leq \underline{\int_{\mathbb{R}^d}} g(x)\,dx$ and $\overline{\int_{\mathbb{R}^d}} f(x)\,dx \leq \overline{\int_{\mathbb{R}^d}} g(x)\,dx$.
(iii) (Homogeneity) If $c \in [0, +\infty)$, then $\underline{\int_{\mathbb{R}^d}} cf(x)\,dx = c\underline{\int_{\mathbb{R}^d}} f(x)\,dx$. (The claim unfortunately fails for $c = +\infty$, but this is somewhat tricky to show.)
(iv) (Equivalence) If f, g agree almost everywhere, then $\underline{\int_{\mathbb{R}^d}} f(x)\,dx = \underline{\int_{\mathbb{R}^d}} g(x)\,dx$ and $\overline{\int_{\mathbb{R}^d}} f(x)\,dx = \overline{\int_{\mathbb{R}^d}} g(x)\,dx$.
(v) (Superadditivity) $\underline{\int_{\mathbb{R}^d}} f(x) + g(x)\,dx \geq \underline{\int_{\mathbb{R}^d}} f(x)\,dx + \underline{\int_{\mathbb{R}^d}} g(x)\,dx$.
(vi) (Subadditivity of upper integral) $\overline{\int_{\mathbb{R}^d}} f(x) + g(x)\,dx \leq \overline{\int_{\mathbb{R}^d}} f(x)\,dx + \overline{\int_{\mathbb{R}^d}} g(x)\,dx$.
(vii) (Divisibility) For any measurable set E, one has $\underline{\int_{\mathbb{R}^d}} f(x)\,dx = \underline{\int_{\mathbb{R}^d}} f(x)1_E(x)\,dx + \underline{\int_{\mathbb{R}^d}} f(x)1_{\mathbb{R}^d \backslash E}(x)\,dx$.
(viii) (Horizontal truncation) As $n \to \infty$, $\underline{\int_{\mathbb{R}^d}} \min(f(x), n)\,dx$ converges to $\underline{\int_{\mathbb{R}^d}} f(x)\,dx$.
(ix) (Vertical truncation) As $n \to \infty$, $\underline{\int_{\mathbb{R}^d}} f(x)1_{|x| \leq n}\,dx$ converges to $\underline{\int_{\mathbb{R}^d}} f(x)\,dx$. [Hint: From Exercise 1.2.11 one has $m(E \cap \{x : |x| \leq n\}) \to m(E)$ for any measurable set E.]
(x) (Reflection) If $f + g$ is a simple function that is bounded with finite measure support (i.e. it is absolutely integrable), then $\text{Simp}\int_{\mathbb{R}^d} f(x) + g(x)\,dx = \underline{\int_{\mathbb{R}^d}} f(x)\,dx + \overline{\int_{\mathbb{R}^d}} g(x)\,dx$.

Solution. Throughout, "simple" means unsigned simple in the sense of Definition 1.3.2, and we use Definition 1.3.12:

$$\underline{\int_{\mathbb{R}^d}} f = \sup_{0 \leq \varphi \leq f,\ \varphi \text{ simple}} \text{Simp}\int_{\mathbb{R}^d} \varphi, \qquad \overline{\int_{\mathbb{R}^d}} f = \inf_{h \geq f,\ h \text{ simple}} \text{Simp}\int_{\mathbb{R}^d} h.$$

We also record two standard facts used repeatedly. First, by the Venn diagram argument preceding

Lemma 1.3.4, any simple φ may be written as $\varphi = \sum_{i=1}^k c_i 1_{A_i}$ with $A_1, \dots, A_k$ disjoint measurable sets partitioning $\mathbb{R}^d$ and $c_i \in [0, +\infty]$, and then $\mathrm{Simp} \int \varphi = \sum_{i=1}^k c_i m(A_i)$ (with the convention $0 \cdot \infty = 0$). Second, if E is measurable and $\varphi = \sum_i c_i 1_{A_i}$ is simple, then $\varphi 1_E = \sum_i c_i 1_{A_i \cap E}$ is simple; likewise, writing two simple functions over a common such partition, their pointwise minimum and maximum are simple, and if $u \leq v$ are simple and everywhere finite then $v - u$ is an unsigned simple function (over a common partition $u = \sum_i b_i 1_{A_i}$, $v = \sum_i a_i 1_{A_i}$ with $b_i \leq a_i$ finite, one has $v - u = \sum_i (a_i - b_i) 1_{A_i}$).
(i) Let f be simple. Since f is admissible as a competitor in both suprema/infima ($f \leq f$), we get $\underline{\int} f \geq \mathrm{Simp} \int f$ and $\overline{\int} f \leq \mathrm{Simp} \int f$. Conversely, if $\varphi \leq f$ is simple then $\mathrm{Simp} \int \varphi \leq \mathrm{Simp} \int f$ by monotonicity of the simple integral (Exercise 1.3.1(v)), so $\underline{\int} f \leq \mathrm{Simp} \int f$; and if $h \geq f$ is simple then $\mathrm{Simp} \int h \geq \mathrm{Simp} \int f$, so $\overline{\int} f \geq \mathrm{Simp} \int f$. All three quantities agree.
(ii) By Definition 1.3.5, the hypothesis that $f \leq g$ pointwise almost everywhere says exactly that $N := \{x : f(x) > g(x)\}$ has Lebesgue measure zero. By Lemma 1.2.13(iii),(v), N and its complement are Lebesgue measurable, so indicator functions of N and $\mathbb{R}^d \setminus N$ are simple.
Lower integrals: let $\varphi \leq f$ be simple, and set $\tilde{\varphi} := \varphi 1_{\mathbb{R}^d \setminus N}$, again simple. Off N we have $\tilde{\varphi} = \varphi \leq f \leq g$, and on N we have $\tilde{\varphi} = 0 \leq g$; so $\tilde{\varphi} \leq g$ everywhere. Since $\tilde{\varphi} = \varphi$ outside the null set N, Exercise 1.3.1(iv) gives $\mathrm{Simp} \int \varphi = \mathrm{Simp} \int \tilde{\varphi} \leq \underline{\int} g$. Taking the supremum over φ yields $\underline{\int} f \leq \underline{\int} g$.
Upper integrals: let $h \geq g$ be simple, and set $\tilde{h} := h + \infty \cdot 1_N$, again simple. Off N, $\tilde{h} = h \geq g \geq f$; on N, $\tilde{h} = +\infty \geq f$. So $\tilde{h} \geq f$ everywhere, and $\tilde{h} = h$ outside the null set N, so $\mathrm{Simp} \int \tilde{h} = \mathrm{Simp} \int h$ by Exercise 1.3.1(iv). Hence $\overline{\int} f \leq \mathrm{Simp} \int h$; taking the infimum over h gives $\overline{\int} f \leq \overline{\int} g$.
(iii) If $c = 0$ then cf is the zero function, whose only simple minorant is the zero function, so $\underline{\int} cf = 0 = 0 \cdot \underline{\int} f$ (using $0 \cdot \infty = 0$). If $c > 0$, then $\varphi \mapsto c\varphi$ is a bijection from $\{\varphi \text{ simple} : 0 \leq \varphi \leq f\}$ onto $\{\psi \text{ simple} : 0 \leq \psi \leq cf\}$, with inverse $\psi \mapsto c^{-1}\psi$, and $\mathrm{Simp} \int c\varphi = c\, \mathrm{Simp} \int \varphi$ by Exercise 1.3.1(i). Since $t \mapsto ct$ is an increasing bijection of $[0, +\infty]$, it commutes with suprema, so

$$\underline{\int} cf = \sup_{0 \leq \varphi \leq f} c\, \mathrm{Simp} \int \varphi = c \sup_{0 \leq \varphi \leq f} \mathrm{Simp} \int \varphi = c \underline{\int} f.$$

(iv) If $f = g$ almost everywhere then $f \leq g$ a.e. and $g \leq f$ a.e., so both inequalities of (ii) hold in both directions, giving equality of the lower integrals and of the upper integrals.
(v) Let $\varphi \leq f$ and $\psi \leq g$ be simple. Then $\varphi + \psi$ is simple and $\varphi + \psi \leq f + g$ pointwise, so by unsigned linearity of the simple integral (Exercise 1.3.1(i)),

$$\underline{\int} (f + g) \geq \mathrm{Simp} \int (\varphi + \psi) = \mathrm{Simp} \int \varphi + \mathrm{Simp} \int \psi.$$

Taking the supremum over φ and then over ψ (using that $\sup_{a \in A, b \in B}(a + b) = \sup A + \sup B$ for non-empty $A, B \subset [0, +\infty]$) gives $\underline{\int}(f + g) \geq \underline{\int} f + \underline{\int} g$.
(vi) Dually, let $h_1 \geq f$ and $h_2 \geq g$ be simple. Then $h_1 + h_2$ is simple and $\geq f + g$, so

$$\overline{\int} (f + g) \leq \mathrm{Simp} \int (h_1 + h_2) = \mathrm{Simp} \int h_1 + \mathrm{Simp} \int h_2.$$

Taking the infimum over h_1 and h_2 (using $\inf_{a \in A, b \in B}(a + b) = \inf A + \inf B$ in $[0, +\infty]$) gives $\overline{\int}(f + g) \leq \overline{\int} f + \overline{\int} g$.
(vii) Since $f = f1_E + f1_{\mathbb{R}^d \setminus E}$ pointwise, part (v) gives the inequality $\underline{\int} f \geq \underline{\int} f1_E + \underline{\int} f1_{\mathbb{R}^d \setminus E}$.
For the reverse inequality, let $\varphi \leq f$ be simple. Then $\varphi 1_E$ and $\varphi 1_{\mathbb{R}^d \setminus E}$ are simple (here we use that E is measurable), and they are bounded by $f1_E$ and $f1_{\mathbb{R}^d \setminus E}$ respectively. Since $\varphi = \varphi 1_E + \varphi 1_{\mathbb{R}^d \setminus E}$, Exercise 1.3.1(i) gives

$$\mathrm{Simp} \int \varphi = \mathrm{Simp} \int \varphi 1_E + \mathrm{Simp} \int \varphi 1_{\mathbb{R}^d \setminus E} \leq \underline{\int} f1_E + \underline{\int} f1_{\mathbb{R}^d \setminus E}.$$

Taking the supremum over φ completes the proof.
(viii) The functions $\min(f, n)$ increase in n and are bounded by f, so by (ii) the quantities $\underline{\int} \min(f, n)$ form a non-decreasing sequence bounded above by $\underline{\int} f$; let $L \in [0, +\infty]$ be its limit, so $L \leq \underline{\int} f$.

For the converse, let $\varphi \le f$ be simple, written as $\varphi = \sum_{i=1}^k c_i 1_{A_i}$ with the A_i disjoint and measurable. Then $\min(\varphi, n) = \sum_{i=1}^k \min(c_i, n) 1_{A_i}$ is simple and satisfies $\min(\varphi, n) \le \min(f, n)$, so

$$\underline{\int} \min(f, n) \ge \text{Simp} \int \min(\varphi, n) = \sum_{i=1}^k \min(c_i, n)\, m(A_i).$$

For each fixed i, $\min(c_i, n) m(A_i) \nearrow c_i m(A_i)$ as $n \to \infty$: this is clear if $c_i < \infty$ (the terms are eventually constant, equal to $c_i m(A_i)$, once $n \ge c_i$), while if $c_i = \infty$ then either $m(A_i) = 0$ and all terms are $0 = c_i m(A_i)$, or $m(A_i) > 0$ and $n\, m(A_i) \to \infty = c_i m(A_i)$. Summing the finitely many terms, $\text{Simp} \int \min(\varphi, n) \to \text{Simp} \int \varphi$. Hence $L \ge \text{Simp} \int \varphi$, and taking the supremum over φ gives $L \ge \underline{\int} f$. Thus $L = \underline{\int} f$.

(ix) Write $B_n := \{x \in \mathbb{R}^d : |x| \le n\}$, a measurable (indeed closed) set. The functions $f 1_{B_n}$ increase in n and are bounded by f, so as in (viii) the numbers $\underline{\int} f 1_{B_n}$ increase to some $L \le \underline{\int} f$.

Conversely let $\varphi \le f$ be simple, written $\varphi = \sum_{i=1}^k c_i 1_{A_i}$ with the A_i disjoint measurable. Then $\varphi 1_{B_n} = \sum_i c_i 1_{A_i \cap B_n}$ is simple and $\le f 1_{B_n}$, so

$$\underline{\int} f 1_{B_n} \ge \text{Simp} \int \varphi 1_{B_n} = \sum_{i=1}^k c_i\, m(A_i \cap B_n).$$

By the hint, i.e. upward monotone convergence for measurable sets (Exercise 1.2.11(i)) applied to the increasing sequence $A_i \cap B_n$ with union A_i, we have $m(A_i \cap B_n) \nearrow m(A_i)$. Consequently $c_i m(A_i \cap B_n) \nearrow c_i m(A_i)$: this is immediate for $0 < c_i < \infty$ and trivial for $c_i = 0$; and for $c_i = \infty$, either $m(A_i) = 0$, in which case every $m(A_i \cap B_n) = 0$ and all terms vanish, or $m(A_i) > 0$, in which case $m(A_i \cap B_n) > 0$ for large n and the terms are eventually $+\infty$. Summing the finitely many terms gives $\text{Simp} \int \varphi 1_{B_n} \to \text{Simp} \int \varphi$, whence $L \ge \text{Simp} \int \varphi$. Taking the supremum over φ gives $L \ge \underline{\int} f$, so $L = \underline{\int} f$.

(x) Write $s := f + g$, a simple function which is bounded, say $0 \le s \le M < \infty$, and supported in a measurable set S with $m(S) < \infty$. Since $f, g \ge 0$ and $f + g = s$, we have $f \le s$ and $g \le s$; in particular f and g are bounded by M and vanish outside S. Any simple function ψ with $0 \le \psi \le M$ that vanishes outside S satisfies $\psi \le M 1_S$, so by monotonicity of the simple integral (Exercise 1.3.1(v)) it has $\text{Simp} \int \psi \le M\, m(S) < \infty$; so all quantities appearing below are finite and subtraction is legitimate.

First we show $\underline{\int} f + \overline{\int} g \le \text{Simp} \int s$. Let $\varphi \le f$ be simple. Then $\varphi \le f \le s \le M$ and φ vanishes off S, so φ is a bounded, finite-measure-support simple function; hence $h := s - \varphi$ is a well-defined simple function (a difference of finite-valued simple functions), and it is unsigned since $\varphi \le s$. Moreover $h = f + g - \varphi \ge g$ because $\varphi \le f$. From $h + \varphi = s$ and Exercise 1.3.1(i),

$$\text{Simp} \int h + \text{Simp} \int \varphi = \text{Simp} \int s,$$

so, $\text{Simp} \int \varphi$ being finite, $\text{Simp} \int h = \text{Simp} \int s - \text{Simp} \int \varphi$. Since $h \ge g$ is simple, $\overline{\int} g \le \text{Simp} \int s - \text{Simp} \int \varphi$, i.e. $\text{Simp} \int \varphi + \overline{\int} g \le \text{Simp} \int s$. Taking the supremum over φ gives $\underline{\int} f + \overline{\int} g \le \text{Simp} \int s$.

For the reverse inequality, let $h \ge g$ be simple. Replacing h by $\min(h, s)$ – which is still simple, still dominates g (as $g \le h$ and $g \le s$), and has $\text{Simp} \int \min(h, s) \le \text{Simp} \int h$ by Exercise 1.3.1(v) – we may assume $g \le h \le s$. Then h is bounded by M and supported in S, so $\text{Simp} \int h < \infty$, and $\varphi := s - h$ is a well-defined unsigned simple function with $\varphi = f + g - h \le f$. As before, $\text{Simp} \int \varphi + \text{Simp} \int h = \text{Simp} \int s$, so

$$\underline{\int} f + \text{Simp} \int h \ge \text{Simp} \int \varphi + \text{Simp} \int h = \text{Simp} \int s.$$

Taking the infimum over all such h – which computes $\overline{\int} g$, since the reduction above shows that restricting to majorants $h \le s$ does not change the infimum – gives $\underline{\int} f + \overline{\int} g \ge \text{Simp} \int s$. Combining the two inequalities completes the proof.

Problem 1.3.11. Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable, bounded, and vanishing outside of a set of finite measure. Show that the lower and upper Lebesgue integrals of f agree. (Hint: use Exercise 1.3.4.) There is a converse to this statement, but we will defer it to later notes. What happens if f is allowed to be unbounded, or is not supported inside a set of finite measure?

Solution. Let $M < \infty$ be a bound for f, and let E be a measurable set with $m(E) < \infty$ such that f vanishes outside E.

Since f is a bounded unsigned measurable function, Exercise 1.3.4 provides bounded simple functions f_n converging to f uniformly. Put

$$\varepsilon_n := \sup_{x \in \mathbb{R}^d} |f_n(x) - f(x)|,$$

so $\varepsilon_n \to 0$ and we may assume each $\varepsilon_n < \infty$. Define

$$g_n := \max(f_n - \varepsilon_n, 0)\, 1_E, \qquad h_n := (f_n + \varepsilon_n)\, 1_E.$$

Both are simple: $f_n \pm \varepsilon_n$ is a (real, finite-valued) simple function, taking a maximum with 0 and multiplying by the indicator of the measurable set E keeps it simple, and both are unsigned.

We claim $g_n \le f \le h_n$ pointwise. Outside E we have $f = 0 = g_n = h_n$, so this is clear there. On E: from $|f_n - f| \le \varepsilon_n$ we get $f \ge f_n - \varepsilon_n$, and also $f \ge 0$, so $f \ge \max(f_n - \varepsilon_n, 0) = g_n$; and $f \le f_n + \varepsilon_n = h_n$.

Moreover $0 \le h_n - g_n \le 2\varepsilon_n 1_E$ pointwise, since on E we have $(f_n + \varepsilon_n) - \max(f_n - \varepsilon_n, 0) \le (f_n + \varepsilon_n) - (f_n - \varepsilon_n) = 2\varepsilon_n$ when $f_n - \varepsilon_n \ge 0$, and $\le f_n + \varepsilon_n \le 2\varepsilon_n$ when $f_n - \varepsilon_n < 0$. Hence, writing $h_n = g_n + (h_n - g_n)$ and using unsigned linearity and monotonicity of the simple integral (Exercise 1.3.1(i),(v)),

$$\mathrm{Simp} \int_{\mathbb{R}^d} h_n = \mathrm{Simp} \int_{\mathbb{R}^d} g_n + \mathrm{Simp} \int_{\mathbb{R}^d} (h_n - g_n) \le \mathrm{Simp} \int_{\mathbb{R}^d} g_n + 2\varepsilon_n m(E).$$

Since $g_n \le f$ is simple and $h_n \ge f$ is simple, Definition 1.3.12 gives

$$\overline{\int_{\mathbb{R}^d}} f \le \mathrm{Simp} \int_{\mathbb{R}^d} h_n \le \mathrm{Simp} \int_{\mathbb{R}^d} g_n + 2\varepsilon_n m(E) \le \underline{\int_{\mathbb{R}^d}} f + 2\varepsilon_n m(E).$$

All quantities here are finite: indeed $\underline{\int} f \le \overline{\int} f \le \mathrm{Simp} \int M 1_E = M\, m(E) < \infty$, and $m(E) < \infty$. Letting $n \to \infty$, so that $\varepsilon_n m(E) \to 0$, we conclude $\overline{\int} f \le \underline{\int} f$. The reverse inequality always holds (Definition 1.3.12), so the two integrals agree.

We now address the two closing questions: without either hypothesis the conclusion can fail.

Unbounded f, finite measure support. Work in $d = 1$ and choose disjoint intervals $I_1, I_2, \dots \subset (0, 1]$ with $m(I_k) = 4^{-k}$, and set $f := \sum_{k=1}^{\infty} 2^k 1_{I_k}$ (so $f = 2^k$ on I_k and $f = 0$ elsewhere). This f is measurable (it is the pointwise limit of the simple functions $\sum_{k \le n} 2^k 1_{I_k}$), is supported in $(0, 1]$, and is unbounded.

Its upper integral is infinite. Indeed, let $h \ge f$ be simple. Being simple, h takes at most finitely many values; let $C < \infty$ be the largest finite value among them (and set $C := 0$ if h takes no finite value at all). Choose k with $2^k > C$. On I_k we have $h \ge f = 2^k > C$, so h cannot take any of its finite values there; hence $h = +\infty$ on all of I_k. Since $m(I_k) = 4^{-k} > 0$, monotonicity of the simple integral (Exercise 1.3.1(v)) and Definition 1.3.3 give

$$\mathrm{Simp} \int_{\mathbb{R}} h \;\ge\; \mathrm{Simp} \int_{\mathbb{R}} (+\infty) 1_{I_k} = (+\infty) \cdot m(I_k) = +\infty.$$

As h was an arbitrary simple majorant, $\overline{\int} f = +\infty$.

Its lower integral is finite: we claim $\mathrm{Simp} \int \varphi \le 3$ for every simple φ with $0 \le \varphi \le f$. Such a φ is finite-valued (as f is), so it is bounded by some $C < \infty$, and $\varphi \le \min(f, C)$. If $C < 2$, then $\varphi \le C 1_{(0,1]}$ and $\mathrm{Simp} \int \varphi \le 2$. Otherwise pick $K \ge 1$ with $2^K \le C < 2^{K+1}$; then $\min(f, C) \le 2^k$ on I_k for $k \le K$,

$\min(f, C) = C$ on I_k for $k > K$, and $\min(f, C) = 0$ off $\bigcup_k I_k$, so with $U_K := \bigcup_{k>K} I_k$ (measurable, of measure $\sum_{k>K} 4^{-k} = 4^{-K}/3$ by countable additivity),

$$\varphi \le \sum_{k=1}^{K} 2^k 1_{I_k} + C 1_{U_K},$$

whence by monotonicity of the simple integral

$$\text{Simp} \int \varphi \le \sum_{k=1}^{K} 2^k 4^{-k} + C \cdot \frac{4^{-K}}{3} \le \sum_{k=1}^{\infty} 2^{-k} + 2^{K+1} \frac{4^{-K}}{3} \le 1 + \frac{2}{3} \le 3.$$

So $\underline{\int} f \le 3 < +\infty = \overline{\int} f$, and the two integrals differ.

Bounded f, support of infinite measure. Again in $d = 1$, let $f := \sum_{n=1}^{\infty} 2^{-n} 1_{[n,n+1)}$, which is measurable, bounded by $1/2$, and supported on $[1, +\infty)$, a set of infinite measure.

For the upper integral, let $h \ge f$ be simple. Then $h > 0$ everywhere on $[1, +\infty)$. As h takes finitely many values and $[1, +\infty)$ is partitioned by the finitely many level sets of h intersected with it, at least one such level set has infinite measure, and the corresponding value is positive; hence $\text{Simp} \int h = +\infty$, and $\overline{\int} f = +\infty$.

For the lower integral, let $\varphi \le f$ be simple, and let $c > 0$ be the smallest strictly positive value that φ takes (if φ is identically zero there is nothing to prove). On $[n, n+1)$ we have $\varphi \le f = 2^{-n}$, so once $2^{-n} < c$ the function φ must vanish on $[n, n+1)$. Hence φ is supported in $[1, N+1)$ for some finite N, and there $\varphi \le \sum_{n=1}^{N} 2^{-n} 1_{[n,n+1)}$, so

$$\text{Simp} \int \varphi \le \sum_{n=1}^{N} 2^{-n} \le 1.$$

Therefore $\underline{\int} f \le 1 < +\infty = \overline{\int} f$, and again the two integrals differ.

Problem 1.3.12. (Upper Lebesgue integral and outer Lebesgue measure) Show that for any set $E \subset \mathbb{R}^d$, $\overline{\int_{\mathbb{R}^d}} 1_E(x)\, dx = m^*(E)$. Conclude that the upper and lower Lebesgue integrals are not necessarily additive if no measurability hypotheses are assumed.

Solution. First we prove $\overline{\int} 1_E = m^*(E)$.

Upper bound. By Exercise 1.2.14, E is contained in a Lebesgue measurable set A with $m(A) = m^*(E)$. Then 1_A is a simple function with $1_A \ge 1_E$, so by Definition 1.3.12 and compatibility of the simple integral with Lebesgue measure (Exercise 1.3.1(vi)),

$$\overline{\int_{\mathbb{R}^d}} 1_E \le \text{Simp} \int_{\mathbb{R}^d} 1_A = m(A) = m^*(E).$$

Lower bound. Let $h \ge 1_E$ be any simple function, written as $h = \sum_{i=1}^{k} c_i 1_{A_i}$ with $A_1, \ldots, A_k$ disjoint measurable sets partitioning $\mathbb{R}^d$. The set $A := \{x : h(x) \ge 1\} = \bigcup_{i:c_i \ge 1} A_i$ is then measurable, and it contains E (since $h \ge 1_E = 1$ on E). As $h \ge 1_A$ pointwise, monotonicity of the simple integral (Exercise 1.3.1(v)) and monotonicity of outer measure give

$$\text{Simp} \int_{\mathbb{R}^d} h \ge \text{Simp} \int_{\mathbb{R}^d} 1_A = m(A) \ge m^*(E).$$

Taking the infimum over h gives $\overline{\int} 1_E \ge m^*(E)$, and hence equality.

It is convenient to record the companion fact for the lower integral: for any $E \subset \mathbb{R}^d$,

$$\underline{\int_{\mathbb{R}^d}} 1_E = m_*(E) := \sup\{m(A) : A \subset E,\ A \text{ Lebesgue measurable}\},$$

the inner measure of E. Indeed, for measurable $A \subset E$ the function 1_A is a simple minorant of 1_E with Simp $\int 1_A = m(A)$, giving $\underline{\int} 1_E \geq m_*(E)$. Conversely, if $\varphi \leq 1_E$ is simple, write $\varphi = \sum_{i=1}^k c_i 1_{A_i}$ with the A_i disjoint, measurable, non-empty, and partitioning $\mathbb{R}^d$ (the Venn diagram refinement in the proof of Lemma 1.3.4 discards the empty atoms). For each i with $c_i > 0$ pick $x \in A_i$; then $1_E(x) \geq \varphi(x) = c_i > 0$, forcing $1_E(x) = 1$ and $c_i \leq 1$, and since φ is constant equal to $c_i > 0$ on A_i the same argument at every point of A_i gives $A_i \subset E$. So with $A := \bigcup_{i:c_i>0} A_i \subset E$ (measurable),

$$\text{Simp} \int \varphi = \sum_i c_i m(A_i) \leq \sum_{i:c_i>0} m(A_i) = m(A) \leq m_*(E),$$

using the disjointness of the A_i and finite additivity. Hence $\underline{\int} 1_E = m_*(E)$.
Failure of additivity. Let $E \subset [0,1]$ be a non-measurable set, as constructed in Proposition 1.2.18, and let $F := [0,1] \setminus E$; since $[0,1]$ is measurable, F is non-measurable too (otherwise $E = [0,1] \setminus F$ would be measurable). Then $1_E + 1_F = 1_{[0,1]}$ is simple, so by Exercise 1.3.10(i),

$$\underline{\int}(1_E + 1_F) = \overline{\int}(1_E + 1_F) = m([0,1]) = 1.$$

We claim that $m^*(E) + m^*(F) > 1$ and $m_*(E) + m_*(F) < 1$; granting this, the displayed identity shows that neither the upper nor the lower Lebesgue integral is additive on the pair $1_E, 1_F$.
To see the claim, first note the duality $m_*(E) = 1 - m^*(F)$ (and symmetrically $m_*(F) = 1 - m^*(E)$) for complementary subsets of $[0,1]$. Indeed, by Exercise 1.2.14 choose a measurable $B \supset F$ with $m(B) = m^*(F)$; replacing B by $B \cap [0,1]$ changes neither of these properties (monotonicity of outer measure forces $m(B \cap [0,1]) = m^*(F)$, as $F \subset [0,1]$). Then $[0,1] \setminus B$ is a measurable subset of E, so $m_*(E) \geq 1 - m^*(F)$. Conversely, if $A \subset E$ is measurable, then $[0,1] \setminus A \supset F$ is measurable, so $m^*(F) \leq 1 - m(A)$, i.e. $m(A) \leq 1 - m^*(F)$; taking the supremum gives $m_*(E) \leq 1 - m^*(F)$.
Now suppose, for contradiction, that $m^*(E) + m^*(F) = 1$ (the inequality $m^*(E) + m^*(F) \geq 1$ always holds, by finite subadditivity of outer measure applied to $[0,1] = E \cup F$). Choose measurable hulls $A \supset E$ and $B \supset F$ inside $[0,1]$ with $m(A) = m^*(E)$ and $m(B) = m^*(F)$, as above. Then

$$[0,1] \setminus B \subset E \subset A,$$

both extremes measurable, and

$$m(A) - m([0,1] \setminus B) = m^*(E) - (1 - m^*(F)) = 0.$$

Hence $A \setminus ([0,1] \setminus B)$ is a measurable set of measure zero, and $E \setminus ([0,1] \setminus B)$ is a subset of it, hence a null set and therefore measurable. Consequently $E = ([0,1] \setminus B) \cup (E \setminus ([0,1] \setminus B))$ is measurable, contradicting the choice of E. Therefore $m^*(E) + m^*(F) > 1$, and by the duality

$$m_*(E) + m_*(F) = (1 - m^*(F)) + (1 - m^*(E)) = 2 - (m^*(E) + m^*(F)) < 1.$$

Combining, we obtain the strict inequalities

$$\overline{\int} 1_E + \overline{\int} 1_F = m^*(E) + m^*(F) > 1 = \overline{\int}(1_E + 1_F)$$

and

$$\underline{\int} 1_E + \underline{\int} 1_F = m_*(E) + m_*(F) < 1 = \underline{\int}(1_E + 1_F).$$

So the upper integral is strictly subadditive and the lower integral strictly superadditive here; in particular, without measurability hypotheses neither is additive. (This is consistent with Exercise 1.3.10(v),(vi), which only assert the two inequalities, and shows that the measurability hypothesis in Corollary 1.3.14 cannot be dropped.)

Problem 1.3.13. (Area interpretation of integral) If $f : \mathbb{R}^d \to [0, +\infty]$ is measurable, show that $\int_{\mathbb{R}^d} f(x)\,dx$ is equal to the $d+1$-dimensional Lebesgue measure of the region $\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : 0 \leq t \leq f(x)\}$. (This can be used as an alternate, and more geometrically intuitive, definition of the unsigned Lebesgue integral; it is a more convenient formulation for establishing the basic convergence theorems, but not quite as convenient for establishing basic properties such as additivity.) (Hint: use Exercise 1.2.22.)

Solution. Write $\Omega_f := \{(x,t) \in \mathbb{R}^d \times \mathbb{R} : 0 \leq t \leq f(x)\}$. By Exercise 1.3.6, Ω_f is a Lebesgue measurable subset of $\mathbb{R}^{d+1}$, so $m^{d+1}(\Omega_f)$ is defined. We must show $m^{d+1}(\Omega_f) = \int_{\mathbb{R}^d} f$, where by Definition 1.3.13 the right-hand side is the lower Lebesgue integral $\underline{\int} f$.

Step 1: the case of a simple function. Let φ be an unsigned simple function, written (after the usual Venn diagram refinement) as $\varphi = \sum_{i=1}^k c_i 1_{A_i}$ with $A_1, \ldots, A_k$ disjoint measurable sets partitioning $\mathbb{R}^d$ and $c_i \in [0, +\infty]$. Setting $J_i := [0, c_i]$ when $c_i < \infty$ and $J_i := [0, +\infty)$ when $c_i = +\infty$, we have the disjoint decomposition

$$\Omega_\varphi = \bigcup_{i=1}^k A_i \times J_i,$$

since the A_i partition $\mathbb{R}^d$ and the fibre of Ω_φ above $x \in A_i$ is exactly J_i. By Exercise 1.2.22(ii) each $A_i \times J_i$ is Lebesgue measurable in $\mathbb{R}^{d+1}$ with $m^{d+1}(A_i \times J_i) = m^d(A_i)\, m^1(J_i) = c_i\, m(A_i)$ (note $m^1(J_i) = c_i$ in both cases, and the convention $0 \cdot \infty = 0$ is the same one used in Definition 1.3.3). By finite additivity of Lebesgue measure,

$$m^{d+1}(\Omega_\varphi) = \sum_{i=1}^k c_i\, m(A_i) = \mathrm{Simp} \int_{\mathbb{R}^d} \varphi(x)\,dx.$$

Step 2: $m^{d+1}(\Omega_f) \geq \int_{\mathbb{R}^d} f$. If φ is simple with $0 \leq \varphi \leq f$, then $\Omega_\varphi \subset \Omega_f$, so by monotonicity of Lebesgue measure and Step 1,

$$m^{d+1}(\Omega_f) \geq m^{d+1}(\Omega_\varphi) = \mathrm{Simp} \int_{\mathbb{R}^d} \varphi.$$

Taking the supremum over all such φ and using Definition 1.3.12 gives $m^{d+1}(\Omega_f) \geq \underline{\int} f = \int f$.

Step 3: the graph of f is null. Let

$$G := \{(x, f(x)) : x \in \mathbb{R}^d,\ f(x) < \infty\} = \Omega_f \setminus \Omega_f^<, \qquad \Omega_f^< := \{(x,t) : 0 \leq t < f(x)\}.$$

(The identification is immediate: for x with $f(x) < \infty$ the fibres of Ω_f and $\Omega_f^<$ are $[0, f(x)]$ and $[0, f(x))$, differing by the point $f(x)$; for x with $f(x) = +\infty$ both fibres equal $[0, +\infty)$.)

We claim $m^{d+1}(G) = 0$. Fix $R \in \mathbb{N}$ and let $B_R := \{x \in \mathbb{R}^d : |x| \leq R\}$, a set of finite measure. For each $n \geq 1$ and each integer $j \geq 0$ put

$$E_{j,n} := \{x \in \mathbb{R}^d : j2^{-n} \leq f(x) < (j+1)2^{-n}\},$$

which is Lebesgue measurable by Lemma 1.3.9, and note that the $E_{j,n}$ ($j \geq 0$) are disjoint with union $\{f < \infty\}$. If $f(x) < \infty$ and $x \in E_{j,n}$, then $f(x) \in [j2^{-n}, (j+1)2^{-n}]$, so

$$G \cap (B_R \times \mathbb{R}) \subset \bigcup_{j=0}^\infty \big((E_{j,n} \cap B_R) \times [j2^{-n}, (j+1)2^{-n}]\big).$$

By countable subadditivity of m^{d+1} and Exercise 1.2.22(ii),

$$m^{d+1}\big(G \cap (B_R \times \mathbb{R})\big) \leq \sum_{j=0}^\infty m^d(E_{j,n} \cap B_R)\, 2^{-n} = 2^{-n}\, m\big(\{f < \infty\} \cap B_R\big) \leq 2^{-n} m(B_R),$$

where the middle equality uses countable additivity of m^d over the disjoint sets $E_{j,n} \cap B_R$. Since $m(B_R) < \infty$ and n is arbitrary, $m^{d+1}(G \cap (B_R \times \mathbb{R})) = 0$. Finally $G = \bigcup_{R \in \mathbb{N}} G \cap (B_R \times \mathbb{R})$ is a countable union of null sets, hence null.

Step 4: $m^{d+1}(\Omega_f) \leq \int_{\mathbb{R}^d} f$. By Lemma 1.3.9(iv) there is an increasing sequence $0 \leq f_1 \leq f_2 \leq \dots$ of unsigned simple functions with $f = \sup_n f_n$. Each Ω_{f_n} is measurable (Step 1) and $\Omega_{f_1} \subset \Omega_{f_2} \subset \dots \subset \Omega_f$.

We claim $\Omega_f^< \subset \bigcup_n \Omega_{f_n}$. Indeed, if $0 \leq t < f(x)$, then since $\sup_n f_n(x) = f(x) > t$ there is some n with $f_n(x) > t$, so $(x,t) \in \Omega_{f_n}$.

Therefore, using subadditivity, Step 3, the claim, upward monotone convergence for measurable sets (Exercise 1.2.11(i)), Step 1, and finally $f_n \leq f$ with Definition 1.3.12:

$$\begin{aligned} m^{d+1}(\Omega_f) &\leq m^{d+1}(\Omega_f^<) + m^{d+1}(G) = m^{d+1}(\Omega_f^<) \\ &\leq m^{d+1}\Big(\bigcup_n \Omega_{f_n}\Big) = \lim_{n\to\infty} m^{d+1}(\Omega_{f_n}) = \lim_{n\to\infty} \mathrm{Simp} \int_{\mathbb{R}^d} f_n \leq \underline{\int_{\mathbb{R}^d}} f = \int_{\mathbb{R}^d} f. \end{aligned}$$

Combining Steps 2 and 4 gives $m^{d+1}(\Omega_f) = \int_{\mathbb{R}^d} f(x)\,dx$, as required.

Problem 1.3.14. (Uniqueness of the Lebesgue integral) Show that the Lebesgue integral $f \mapsto \int_{\mathbb{R}^d} f(x)\,dx$ is the only map from measurable unsigned functions $f : \mathbb{R}^d \to [0,+\infty]$ to $[0,+\infty]$ that obeys the following properties for measurable $f, g : \mathbb{R}^d \to [0,+\infty]$:
(i) (Compatibility with the simple integral) If f is simple, then $\int_{\mathbb{R}^d} f(x)\,dx = \mathrm{Simp} \int_{\mathbb{R}^d} f(x)\,dx$.
(ii) (Finite additivity) $\int_{\mathbb{R}^d} f(x) + g(x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx + \int_{\mathbb{R}^d} g(x)\,dx$.
(iii) (Horizontal truncation) As $n \to \infty$, $\int_{\mathbb{R}^d} \min(f(x), n)\,dx$ converges to $\int_{\mathbb{R}^d} f(x)\,dx$.
(iv) (Vertical truncation) As $n \to \infty$, $\int_{\mathbb{R}^d} f(x) 1_{|x| \leq n}\,dx$ converges to $\int_{\mathbb{R}^d} f(x)\,dx$.

Solution. The Lebesgue integral does obey these properties: (i) is Definition 1.3.13 together with Exercise 1.3.10(i), (ii) is Corollary 1.3.14, and (iii), (iv) are Exercise 1.3.10(viii), (ix). So the content of the exercise is uniqueness. Let I be any map from unsigned measurable functions to $[0,+\infty]$ obeying (i)-(iv); we show $I(f) = \int_{\mathbb{R}^d} f$ for every unsigned measurable f.

Throughout, $B_n := \{x \in \mathbb{R}^d : |x| \leq n\}$, a measurable set of finite measure; note that $\min(f,n)$ is measurable by Exercise 1.3.3(vi) (composing f with the continuous map $t \mapsto \min(t,n)$ on $[0,+\infty]$), and $f 1_{B_n}$ is measurable by Exercise 1.3.3(vii).

A difference lemma. If $u \leq v$ are unsigned measurable functions that are everywhere finite, then $w := v - u$ is an unsigned measurable function with $u + w = v$. Indeed, u and v, being finite-valued unsigned measurable functions, are also measurable as complex-valued functions (as remarked after Exercise 1.3.7, the two notions of measurability agree on functions taking values in $[0,+\infty)$). By Exercise 1.3.8(v) (applied with the continuous map $z \mapsto -z$) and Exercise 1.3.8(vi), the function $v - u$ is complex-valued measurable; since it takes values in $[0,+\infty)$, it is unsigned measurable by Exercise 1.3.7(iii). The identity $u + w = v$ is clear.

Monotonicity of I on finite-valued functions. If $u \leq v$ are unsigned measurable and everywhere finite, then by the lemma and (ii),

$$I(v) = I(u + (v-u)) = I(u) + I(v-u) \geq I(u).$$

Step 1: reduction to bounded functions supported in a ball. Suppose we have already shown that $I(g) = \int g$ for every unsigned measurable g that is bounded and vanishes outside some B_m. Let f be an arbitrary unsigned measurable function. For each fixed n and each m, the function $\min(f,n)1_{B_m}$ is measurable, bounded by n, and vanishes outside B_m; so $I(\min(f,n)1_{B_m}) = \int \min(f,n)1_{B_m}$. Letting $m \to \infty$, property (iv) for I (applied to $\min(f,n)$) and Exercise 1.3.10(ix) for the Lebesgue integral give

$$I(\min(f,n)) = \lim_{m\to\infty} I(\min(f,n)1_{B_m}) = \lim_{m\to\infty} \int_{\mathbb{R}^d} \min(f,n)1_{B_m} = \int_{\mathbb{R}^d} \min(f,n).$$

Now letting $n \to \infty$, property (iii) for I and Exercise 1.3.10(viii) for the Lebesgue integral give

$$I(f) = \lim_{n\to\infty} I(\min(f,n)) = \lim_{n\to\infty} \int_{\mathbb{R}^d} \min(f,n) = \int_{\mathbb{R}^d} f.$$

So it suffices to treat the bounded, ball-supported case.
Step 2: the bounded, ball-supported case. Let f be unsigned measurable with $0 \le f \le M < \infty$ everywhere and $f = 0$ outside $E := B_m$, so $m(E) < \infty$. Exactly as in the solution to Exercise 1.3.11, Exercise 1.3.4 furnishes bounded simple functions $f_k \to f$ uniformly, and setting $\varepsilon_k := \sup_x |f_k(x) - f(x)| \to 0$ and

$$g_k := \max(f_k - \varepsilon_k, 0)1_E, \qquad h_k := (f_k + \varepsilon_k)1_E,$$

we obtain simple functions, bounded and supported in E, with

$$g_k \le f \le h_k, \qquad 0 \le h_k - g_k \le 2\varepsilon_k 1_E.$$

We first bound $I(f)$. Since $g_k \le f$ are finite-valued unsigned measurable functions, the difference lemma and (ii) give

$$I(f) = I(g_k) + I(f - g_k) = \mathrm{Simp} \int_{\mathbb{R}^d} g_k + I(f - g_k),$$

using (i) for the simple function g_k. Now $0 \le f - g_k \le h_k - g_k \le 2\varepsilon_k 1_E$, all of these being finite-valued measurable functions, so by monotonicity of I and (i) again,

$$0 \le I(f - g_k) \le I(2\varepsilon_k 1_E) = \mathrm{Simp} \int_{\mathbb{R}^d} 2\varepsilon_k 1_E = 2\varepsilon_k m(E).$$

Hence

$$\mathrm{Simp} \int_{\mathbb{R}^d} g_k \;\le\; I(f) \;\le\; \mathrm{Simp} \int_{\mathbb{R}^d} g_k + 2\varepsilon_k m(E).$$

Next we bound $\int f$ in the same way. Since $g_k \le f$ is simple, Definition 1.3.12 gives $\mathrm{Simp} \int g_k \le \underline{\int} f = \int f$; and since $h_k \ge f$ is simple, $\int f = \underline{\int} f \le \overline{\int} f \le \mathrm{Simp} \int h_k$. As $h_k = g_k + (h_k - g_k)$ with $h_k - g_k \le 2\varepsilon_k 1_E$ simple, unsigned linearity and monotonicity of the simple integral (Exercise 1.3.1(i),(v)) give $\mathrm{Simp} \int h_k \le \mathrm{Simp} \int g_k + 2\varepsilon_k m(E)$. Hence

$$\mathrm{Simp} \int_{\mathbb{R}^d} g_k \;\le\; \int_{\mathbb{R}^d} f \;\le\; \mathrm{Simp} \int_{\mathbb{R}^d} g_k + 2\varepsilon_k m(E).$$

All the quantities here are finite: $\mathrm{Simp} \int g_k \le \mathrm{Simp} \int M 1_E = M\, m(E) < \infty$ and $\varepsilon_k m(E) < \infty$. Thus $I(f)$ and $\int f$ both lie in the interval $[\mathrm{Simp} \int g_k,\ \mathrm{Simp} \int g_k + 2\varepsilon_k m(E)]$, whose length $2\varepsilon_k m(E)$ tends to 0 as $k \to \infty$. Therefore

$$\left| I(f) - \int_{\mathbb{R}^d} f \right| \le 2\varepsilon_k m(E) \xrightarrow[k\to\infty]{} 0,$$

so $I(f) = \int_{\mathbb{R}^d} f$.
By Step 1, I agrees with the Lebesgue integral on all unsigned measurable functions, establishing uniqueness.

Problem 1.3.15. (Translation invariance). Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable. Show that $\int_{\mathbb{R}^d} f(x+y)\,dx = \int_{\mathbb{R}^d} f(x)\,dx$ for any $y \in \mathbb{R}^d$.

Solution. Fix $y \in \mathbb{R}^d$ and write $\tau_y g$ for the translated function $\tau_y g(x) := g(x + y)$, defined for any unsigned $g : \mathbb{R}^d \to [0, +\infty]$. Note that τ_y is a bijection of the space of unsigned functions onto itself, with inverse τ_{-y}, and that it preserves pointwise order: $g \le h$ pointwise if and only if $\tau_y g \le \tau_y h$ pointwise.

Step 1: translation preserves measurability. For any $\lambda \in [0, +\infty]$ we have

$$\{x \in \mathbb{R}^d : \tau_y g(x) > \lambda\} = \{x \in \mathbb{R}^d : x + y \in E_\lambda\} = E_\lambda - y, \qquad E_\lambda := \{x : g(x) > \lambda\}.$$

If g is measurable then each E_λ is Lebesgue measurable by Lemma 1.3.9(v), and hence $E_\lambda - y$ is Lebesgue measurable by Exercise 1.2.20 (translation invariance of Lebesgue measurability). Applying Lemma 1.3.9 in the reverse direction, $\tau_y g$ is measurable. In particular $\tau_y f$ is measurable, so the left-hand integral is defined.
Step 2: translation preserves simplicity and the simple integral. Let g be an unsigned simple function, say $g = \sum_{i=1}^k c_i 1_{E_i}$ with $c_i \in [0, +\infty]$ and E_i Lebesgue measurable. Since $1_E(x+y) = 1_{E-y}(x)$, we get

$$\tau_y g = \sum_{i=1}^k c_i 1_{E_i - y},$$

which is again an unsigned simple function (each $E_i - y$ is measurable by Exercise 1.2.20). Moreover, by the definition of the simple unsigned integral and Exercise 1.2.20 again,

$$\mathrm{Simp}\int_{\mathbb{R}^d} \tau_y g(x)\, dx = \sum_{i=1}^k c_i\, m(E_i - y) = \sum_{i=1}^k c_i\, m(E_i) = \mathrm{Simp}\int_{\mathbb{R}^d} g(x)\, dx.$$

(The simple integral is well defined independently of the chosen representation, by Lemma 1.3.4, so this computation is legitimate.)
Step 3: comparison of the two suprema. By Definition 1.3.13 the unsigned Lebesgue integral is the lower unsigned Lebesgue integral of Definition 1.3.12, so

$$\int_{\mathbb{R}^d} \tau_y f(x)\, dx = \sup_{0 \le h \le \tau_y f;\ h \text{ simple}} \mathrm{Simp}\int_{\mathbb{R}^d} h(x)\, dx.$$

By Steps 1 and 2, the map $g \mapsto \tau_y g$ sends the class $\mathcal{G} := \{g \text{ simple} : 0 \le g \le f\}$ into the class $\mathcal{H} := \{h \text{ simple} : 0 \le h \le \tau_y f\}$; and the map $h \mapsto \tau_{-y} h$ sends $\mathcal{H}$ into $\mathcal{G}$ for the same reasons (applying Steps 1 and 2 with $-y$ in place of y). These two maps are mutually inverse, so $g \mapsto \tau_y g$ is a bijection from $\mathcal{G}$ onto $\mathcal{H}$, and by Step 2 it preserves the value of the simple integral. Two sets of numbers in bijection through a value-preserving map have the same supremum, whence

$$\int_{\mathbb{R}^d} f(x+y)\, dx = \sup_{h \in \mathcal{H}} \mathrm{Simp}\int_{\mathbb{R}^d} h = \sup_{g \in \mathcal{G}} \mathrm{Simp}\int_{\mathbb{R}^d} g = \int_{\mathbb{R}^d} f(x)\, dx,$$

as required. □

Problem 1.3.16. (Linear change of variables). Let $f : \mathbb{R}^d \to [0, +\infty]$ be measurable, and let $T : \mathbb{R}^d \to \mathbb{R}^d$ be an invertible linear transformation. Show that $\int_{\mathbb{R}^d} f(T^{-1}(x))\, dx = |\det T| \int_{\mathbb{R}^d} f(x)\, dx$, or equivalently $\int_{\mathbb{R}^d} f(Tx)\, dx = \frac{1}{|\det T|} \int_{\mathbb{R}^d} f(x)\, dx$.

Solution. Since T is invertible, $D := |\det T|$ is a real number with $0 < D < \infty$. For an unsigned function $g : \mathbb{R}^d \to [0, +\infty]$ write $g \circ T^{-1}$ for the function $x \mapsto g(T^{-1}(x))$. As in Exercise 1.3.15 we argue in three steps.
Step 1: composition with T^{-1} preserves measurability. For $\lambda \in [0, +\infty]$,

$$\{x : g(T^{-1}(x)) > \lambda\} = \{x : T^{-1}(x) \in E_\lambda\} = T(E_\lambda), \qquad E_\lambda := \{x : g(x) > \lambda\},$$

using that T is a bijection. If g is measurable then E_λ is Lebesgue measurable by Lemma 1.3.9(v), hence $T(E_\lambda)$ is Lebesgue measurable by Exercise 1.2.21 (change of variables for Lebesgue measure). By Lemma 1.3.9(v) again, $g \circ T^{-1}$ is measurable. In particular $f \circ T^{-1}$ is measurable, so the left-hand integral is defined.
Step 2: the effect on simple functions. Since T is a bijection, $1_E \circ T^{-1} = 1_{T(E)}$ for every set E. Hence

if $g = \sum_{i=1}^k c_i 1_{E_i}$ is an unsigned simple function (with $c_i \in [0,+\infty]$ and E_i measurable), then

$$g \circ T^{-1} = \sum_{i=1}^k c_i 1_{T(E_i)},$$

which is again an unsigned simple function by Exercise 1.2.21. By Definition 1.3.3, Lemma 1.3.4 (well-definedness) and Exercise 1.2.21,

$$\operatorname{Simp}\int_{\mathbb{R}^d} g(T^{-1}(x))\,dx = \sum_{i=1}^k c_i\, m(T(E_i)) = \sum_{i=1}^k c_i\, D\, m(E_i) = D\sum_{i=1}^k c_i\, m(E_i) = D\operatorname{Simp}\int_{\mathbb{R}^d} g(x)\,dx.$$

(All manipulations are legitimate in $[0,+\infty]$ because $0 < D < \infty$.)
Step 3: comparison of the suprema. Let $\mathcal{G} := \{g \text{ simple} : 0 \le g \le f\}$ and $\mathcal{H} := \{h \text{ simple} : 0 \le h \le f \circ T^{-1}\}$. Since T^{-1} is a bijection of $\mathbb{R}^d$, we have $g \le f$ pointwise if and only if $g \circ T^{-1} \le f \circ T^{-1}$ pointwise; combined with Step 2 this shows $g \mapsto g \circ T^{-1}$ maps $\mathcal{G}$ into $\mathcal{H}$, and $h \mapsto h \circ T$ maps $\mathcal{H}$ into $\mathcal{G}$ (apply the same reasoning to the invertible linear map T^{-1}). The two maps are mutually inverse, so $g \mapsto g \circ T^{-1}$ is a bijection of $\mathcal{G}$ onto $\mathcal{H}$ which multiplies simple integrals by the constant D. Since multiplication by a constant $D \in (0,\infty)$ commutes with suprema in $[0,+\infty]$, Definitions 1.3.12 and 1.3.13 give

$$\int_{\mathbb{R}^d} f(T^{-1}(x))\,dx = \sup_{h\in\mathcal{H}} \operatorname{Simp}\int_{\mathbb{R}^d} h = \sup_{g\in\mathcal{G}} D\operatorname{Simp}\int_{\mathbb{R}^d} g = D\int_{\mathbb{R}^d} f(x)\,dx,$$

which is the first identity.
The equivalent form. Apply the identity just proved with the invertible linear map $S := T^{-1}$ in place of T: since $S^{-1} = T$ and $|\det S| = 1/|\det T|$, it reads

$$\int_{\mathbb{R}^d} f(Tx)\,dx = \int_{\mathbb{R}^d} f(S^{-1}(x))\,dx = |\det S|\int_{\mathbb{R}^d} f(x)\,dx = \frac{1}{|\det T|}\int_{\mathbb{R}^d} f(x)\,dx.$$

□

Problem 1.3.17. (Compatibility with the Riemann integral). Let $f : [a,b] \to [0,+\infty]$ be Riemann integrable. If we extend f to $\mathbb{R}$ by declaring f to equal zero outside of $[a,b]$, show that $\int_{\mathbb{R}} f(x)\,dx = \int_a^b f(x)\,dx$.

Solution. We may assume $[a,b]$ has positive length, since otherwise both sides vanish (a single point is a null set). Since unbounded functions are not Riemann integrable, f is bounded, and in particular takes values in $[0,M]$ for some finite M; so the extension $f : \mathbb{R} \to [0,+\infty)$ is a genuine unsigned function. By Exercise 1.3.9 this extension is measurable, so $\int_{\mathbb{R}} f(x)\,dx$ is defined and, by Definition 1.3.13, equals the lower unsigned Lebesgue integral of Definition 1.3.12. By Exercise 1.1.22, f is also Darboux integrable and its Riemann integral $\int_a^b f(x)\,dx$ coincides with both its lower and its upper Darboux integral.
Piecewise constant functions become simple functions. Let $g : [a,b] \to \mathbb{R}$ be piecewise constant and unsigned, so that $g = \sum_{i=1}^n c_i 1_{I_i}$ on $[a,b]$ for some partition of $[a,b]$ into intervals $I_1,\dots,I_n$ and constants $c_i \in [0,+\infty)$. Let $\tilde{g} : \mathbb{R} \to [0,+\infty)$ denote its extension by zero. Each I_i is an interval, hence an elementary set, hence Lebesgue measurable with $m(I_i) = |I_i|$ by Lemma 1.2.6, so $\tilde{g} = \sum_{i=1}^n c_i 1_{I_i}$ is an unsigned simple function on $\mathbb{R}$ and, by Definition 1.3.3 and Lemma 1.3.4,

$$\operatorname{Simp}\int_{\mathbb{R}} \tilde{g}(x)\,dx = \sum_{i=1}^n c_i\, m(I_i) = \sum_{i=1}^n c_i|I_i| = \text{p.c.}\int_a^b g(x)\,dx.$$

Lower bound. Let $g : [a,b] \to \mathbb{R}$ be piecewise constant with $g \le f$ pointwise on $[a,b]$. Then $g_+ := \max(g,0)$ is again piecewise constant (on each interval I_i of a partition adapted to g the function g is constant, equal to c_i say, so g_+ is constant there too, equal to $\max(c_i,0)$; the same partition therefore works for g_+), is unsigned, still satisfies $g_+ \le f$ pointwise (because $f \ge 0$, so $\max(g,0) \le \max(f,0) =$

f), and p.c.$\int_a^b g_+ \geq$ p.c.$\int_a^b g$ by monotonicity (Exercise 1.1.21(2)). Hence the lower Darboux integral is unchanged if we restrict the supremum to unsigned piecewise constant minorants. For such a g, its zero-extension $\tilde{g}$ satisfies $0 \leq \tilde{g} \leq f$ pointwise on all of $\mathbb{R}$ (outside $[a,b]$ both sides are zero), so $\tilde{g}$ is one of the competitors in the supremum defining the lower unsigned Lebesgue integral, giving

$$\text{p.c.}\int_a^b g(x)\,dx = \text{Simp}\int_{\mathbb{R}} \tilde{g}(x)\,dx \leq \int_{\mathbb{R}} f(x)\,dx.$$

Taking the supremum over all such g yields $\int_a^b f(x)\,dx = \underline{\int_a^b} f \leq \int_{\mathbb{R}} f(x)\,dx$.
Upper bound. Let $h : [a,b] \to \mathbb{R}$ be piecewise constant with $h \geq f$ pointwise on $[a,b]$; since $f \geq 0$ we automatically have $h \geq 0$. Its zero-extension $\tilde{h}$ is an unsigned simple function on $\mathbb{R}$ with $\tilde{h} \geq f$ pointwise on $\mathbb{R}$. By monotonicity of the lower Lebesgue integral (Exercise 1.3.10(ii)) and its compatibility with the simple integral (Exercise 1.3.10(i)),

$$\int_{\mathbb{R}} f(x)\,dx \leq \int_{\mathbb{R}} \tilde{h}(x)\,dx = \text{Simp}\int_{\mathbb{R}} \tilde{h}(x)\,dx = \text{p.c.}\int_a^b h(x)\,dx.$$

Taking the infimum over all such h yields $\int_{\mathbb{R}} f(x)\,dx \leq \overline{\int_a^b} f = \int_a^b f(x)\,dx$.
Combining the two bounds gives $\int_{\mathbb{R}} f(x)\,dx = \int_a^b f(x)\,dx$. $\square$

Problem 1.3.18. Let $f : \mathbb{R}^d \to [0,+\infty]$ be measurable.
- (i) Show that if $\int_{\mathbb{R}^d} f(x)\,dx < \infty$, then f is finite almost everywhere. Give a counterexample to show that the converse statement is false.
- (ii) Show that $\int_{\mathbb{R}^d} f(x)\,dx = 0$ if and only if f is zero almost everywhere.

Solution. Throughout, write $I := \int_{\mathbb{R}^d} f(x)\,dx$. By Lemma 1.3.9(vi) the sets $\{x : f(x) \geq \lambda\}$ are Lebesgue measurable for every $\lambda \in [0,+\infty]$, so all the measures below are defined, and Markov's inequality (Lemma 1.3.15) applies: for every $0 < \lambda < \infty$,

$$m(\{x \in \mathbb{R}^d : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} I.$$

(i). Suppose $I < \infty$. Let $A := \{x : f(x) = +\infty\}$, which is measurable (it is $\bigcap_{n\geq 1}\{f \geq n\}$, a countable intersection of sets measurable by Lemma 1.3.9(vi)). For every positive integer n we have $A \subset \{f \geq n\}$, so by monotonicity of Lebesgue measure and Markov's inequality,

$$m(A) \leq m(\{x : f(x) \geq n\}) \leq \frac{I}{n}.$$

Since $I < \infty$, letting $n \to \infty$ forces $m(A) = 0$. Thus f is finite outside a null set, i.e. almost everywhere. The converse fails: take $f \equiv 1$ on $\mathbb{R}^d$. This function is simple (it is $1 \cdot 1_{\mathbb{R}^d}$) and finite everywhere, yet by Exercise 1.3.10(i) and Exercise 1.3.1(vi),

$$\int_{\mathbb{R}^d} 1\,dx = \text{Simp}\int_{\mathbb{R}^d} 1_{\mathbb{R}^d}(x)\,dx = m(\mathbb{R}^d) = +\infty,$$

so the integral is infinite. (A finite-support example is also available: $f(x) := 1/|x|^d$ for $0 < |x| \leq 1$ and $f(x) := 0$ elsewhere, with $f(0) := 0$, is finite everywhere but has infinite integral, since $f \geq 2^{kd} 1_{\{2^{-k-1} < |x| \leq 2^{-k}\}}$ for each k and these annuli have measure comparable to 2^{-kd}.)
(ii). Suppose first that $f = 0$ almost everywhere. The zero function is simple with $\text{Simp}\int_{\mathbb{R}^d} 0\,dx = 0$, so by the equivalence property of the lower Lebesgue integral (Exercise 1.3.10(iv)) together with its compatibility with the simple integral (Exercise 1.3.10(i)),

$$I = \int_{\mathbb{R}^d} f(x)\,dx = \int_{\mathbb{R}^d} 0\,dx = \text{Simp}\int_{\mathbb{R}^d} 0\,dx = 0.$$

Conversely, suppose $I = 0$. For each positive integer n, Markov's inequality with $\lambda := 1/n$ gives

$$m(\{x : f(x) \geq 1/n\}) \leq nI = 0,$$

so each set $E_n := \{x : f(x) \geq 1/n\}$ is null. Since f is unsigned,

$$\{x : f(x) \neq 0\} = \{x : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n,$$

because $f(x) > 0$ implies $f(x) \geq 1/n$ for some n. A countable union of null sets is null (by countable subadditivity of Lebesgue outer measure), so the support of f is null, i.e. $f = 0$ almost everywhere. □

Problem 1.3.19. (Integration is linear). Show that integration $f \mapsto \int_{\mathbb{R}^d} f(x)\,dx$ is a (complex) linear operation from $L^1(\mathbb{R}^d)$ to $\mathbb{C}$. In other words, show that

$$\int_{\mathbb{R}^d} f(x) + g(x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx + \int_{\mathbb{R}^d} g(x)\,dx$$

and

$$\int_{\mathbb{R}^d} cf(x)\,dx = c\int_{\mathbb{R}^d} f(x)\,dx$$

for all absolutely integrable $f, g : \mathbb{R}^d \to \mathbb{C}$ and complex numbers c. Also establish the identity

$$\int_{\mathbb{R}^d} \overline{f(x)}\,dx = \overline{\int_{\mathbb{R}^d} f(x)\,dx},$$

which makes integration not just a linear operation, but a $*$-linear operation.

Solution. All functions here are only assumed to be defined almost everywhere; since the lower Lebesgue integral is unchanged when a function is modified on a null set (Exercise 1.3.10(iv)), all the pointwise identities used below need only hold almost everywhere, and the countable-union rule for null sets lets us intersect the (finitely many) full-measure sets involved.

Step 0: the operations stay inside L^1. If $f, g \in L^1(\mathbb{R}^d \to \mathbb{C})$ then $f+g$ and cf are measurable by Exercise 1.3.8(vi) (for cf, apply the product part of that exercise with the constant function $x \mapsto c$, which is measurable by Exercise 1.3.8(i) since it is continuous), and $\overline{f} = \varphi \circ f$ with $\varphi(z) := \overline{z}$ continuous is measurable by Exercise 1.3.8(v). From the pointwise triangle inequality $|f+g| \leq |f|+|g|$, monotonicity of the unsigned integral (Exercise 1.3.10(ii)) and finite additivity (Corollary 1.3.14),

$$\int_{\mathbb{R}^d} |f(x) + g(x)|\,dx \leq \int_{\mathbb{R}^d} |f(x)| + |g(x)|\,dx = \|f\|_{L^1} + \|g\|_{L^1} < \infty,$$

and by homogeneity (Exercise 1.3.10(iii)) $\int |cf| = |c| \int |f| < \infty$, while $|\overline{f}| = |f|$. So $f + g$, cf, $\overline{f}$ all lie in L^1. Moreover, for real-valued absolutely integrable f the two unsigned integrals $\int f_\pm$ appearing in (1.12) are finite (as $f_\pm \leq |f|$ pointwise, use Exercise 1.3.10(ii)), so all the differences below are differences of finite real numbers and may be rearranged freely.

Step 1: additivity for real-valued f, g. Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be absolutely integrable and set $h := f + g$. From $h_+ - h_- = h = (f_+ - f_-) + (g_+ - g_-)$ and the finiteness of all six functions almost everywhere, we may rearrange pointwise (almost everywhere) into an identity between unsigned functions:

$$h_+ + f_- + g_- = h_- + f_+ + g_+.$$

Applying Corollary 1.3.14 twice on each side (all six functions are unsigned and measurable, by Exercise 1.3.7(iii)),

$$\int_{\mathbb{R}^d} h_+ + \int_{\mathbb{R}^d} f_- + \int_{\mathbb{R}^d} g_- = \int_{\mathbb{R}^d} h_- + \int_{\mathbb{R}^d} f_+ + \int_{\mathbb{R}^d} g_+.$$

Every term is a finite real number by Step 0, so we may rearrange to obtain

$$\int_{\mathbb{R}^d} h_+ - \int_{\mathbb{R}^d} h_- = \Big(\int_{\mathbb{R}^d} f_+ - \int_{\mathbb{R}^d} f_-\Big) + \Big(\int_{\mathbb{R}^d} g_+ - \int_{\mathbb{R}^d} g_-\Big),$$

which by (1.12) is exactly $\int_{\mathbb{R}^d} f + g = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$.
Step 2: homogeneity for real scalars and real-valued f. Let f be real-valued and absolutely integrable. If $c \geq 0$ then $(cf)_+ = cf_+$ and $(cf)_- = cf_-$ pointwise, so by homogeneity of the unsigned integral (Exercise 1.3.10(iii), valid for $c \in [0, +\infty)$) and (1.12),

$$\int_{\mathbb{R}^d} cf = c\int_{\mathbb{R}^d} f_+ - c\int_{\mathbb{R}^d} f_- = c\int_{\mathbb{R}^d} f.$$

If $c = -1$ then $(-f)_+ = f_-$ and $(-f)_- = f_+$, so $\int(-f) = \int f_- - \int f_+ = -\int f$. A general real $c < 0$ is handled by writing $cf = (-1)\big(|c|f\big)$ and combining the two cases.
Step 3: additivity for complex-valued f, g. If $f \in L^1(\mathbb{R}^d \to \mathbb{C})$ then $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable (Exercise 1.3.7(iii)) and absolutely integrable, since $|\operatorname{Re} f|, |\operatorname{Im} f| \leq |f|$ pointwise. Because $\operatorname{Re}(f+g) = \operatorname{Re} f + \operatorname{Re} g$ and $\operatorname{Im}(f+g) = \operatorname{Im} f + \operatorname{Im} g$, Definition 1.3.17 and Step 1 give

$$\begin{aligned}\int_{\mathbb{R}^d} (f+g) &= \int_{\mathbb{R}^d} \operatorname{Re} f + \operatorname{Re} g \;+\; i\int_{\mathbb{R}^d} \operatorname{Im} f + \operatorname{Im} g \\ &= \Big(\int_{\mathbb{R}^d} \operatorname{Re} f + i\int_{\mathbb{R}^d} \operatorname{Im} f\Big) + \Big(\int_{\mathbb{R}^d} \operatorname{Re} g + i\int_{\mathbb{R}^d} \operatorname{Im} g\Big) = \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g.\end{aligned}$$

Step 4: homogeneity for complex scalars. Write $c = a + ib$ with $a, b \in \mathbb{R}$, and put $X := \int_{\mathbb{R}^d} \operatorname{Re} f$, $Y := \int_{\mathbb{R}^d} \operatorname{Im} f$, so that $\int_{\mathbb{R}^d} f = X + iY$. Pointwise,

$$\operatorname{Re}(cf) = a\operatorname{Re} f - b\operatorname{Im} f, \qquad \operatorname{Im}(cf) = a\operatorname{Im} f + b\operatorname{Re} f.$$

By Steps 1 and 2 (real linearity of the real-valued integral),

$$\int_{\mathbb{R}^d} \operatorname{Re}(cf) = aX - bY, \qquad \int_{\mathbb{R}^d} \operatorname{Im}(cf) = aY + bX,$$

so by Definition 1.3.17,

$$\int_{\mathbb{R}^d} cf = (aX - bY) + i(aY + bX) = (a+ib)(X+iY) = c\int_{\mathbb{R}^d} f.$$

Together with Step 3 this proves that $f \mapsto \int_{\mathbb{R}^d} f$ is a complex-linear map from $L^1(\mathbb{R}^d)$ to $\mathbb{C}$ (it is $\mathbb{C}$-valued by Step 0).
Step 5: conjugation. Since $\operatorname{Re}\overline{f} = \operatorname{Re} f$ and $\operatorname{Im}\overline{f} = -\operatorname{Im} f$, Definition 1.3.17 and Step 2 (with $c = -1$) give

$$\int_{\mathbb{R}^d} \overline{f(x)}\,dx = \int_{\mathbb{R}^d} \operatorname{Re} f + i\int_{\mathbb{R}^d} (-\operatorname{Im} f) = X - iY = \overline{X + iY} = \overline{\int_{\mathbb{R}^d} f(x)\,dx}. \qquad \square$$

Problem 1.3.20. Show that Exercises 1.3.15, 1.3.16, and 1.3.17 also hold for complex-valued, absolutely integrable functions rather than for unsigned measurable functions.

Solution. The common mechanism is that each of the three operations in question commutes with the decomposition

$$f = (\operatorname{Re} f)_+ - (\operatorname{Re} f)_- + i(\operatorname{Im} f)_+ - i(\operatorname{Im} f)_-,$$

which is the decomposition used in (1.12) and Definition 1.3.17 to define the absolutely convergent integral. Indeed, if Φ is any of the substitutions $\Phi f := f(\cdot + y)$, $\Phi f := f \circ T^{-1}$, or $\Phi f :=$ (extension of f by zero), then Φ acts by precomposition with a map (or by extension by the value 0), and therefore commutes with every pointwise operation fixing 0: $\operatorname{Re}(\Phi f) = \Phi(\operatorname{Re} f)$, $(\Phi u)_\pm = \Phi(u_\pm)$, $|\Phi f| = \Phi|f|$.

We use this repeatedly without further comment. Throughout write $u_1 := (\operatorname{Re} f)_+$, $u_2 := (\operatorname{Re} f)_-$, $u_3 := (\operatorname{Im} f)_+$, $u_4 := (\operatorname{Im} f)_-$; these are unsigned measurable by Exercise 1.3.7(iii), each is bounded pointwise by $|f|$, and hence each has finite unsigned integral, and

$$\int_{\mathbb{R}^d} f = \Big(\int u_1 - \int u_2\Big) + i\Big(\int u_3 - \int u_4\Big).$$

Exercise 1.3.15 for $f \in L^1(\mathbb{R}^d \to \mathbb{C})$. Fix $y \in \mathbb{R}^d$ and write $\tau_y f := f(\cdot + y)$. If f is defined on a set E of full measure, then $\tau_y f$ is defined on $E - y$, which again has full measure by Exercise 1.2.20; so $\tau_y f$ is again defined almost everywhere. Each $\tau_y u_j$ is unsigned measurable by Step 1 of Exercise 1.3.15, and since $\operatorname{Re}(\tau_y f) = \tau_y(\operatorname{Re} f)$ etc., the four functions $\tau_y u_j$ are exactly the four unsigned parts of $\tau_y f$; hence $\tau_y f$ is measurable by Exercise 1.3.7(iii). It is absolutely integrable because $|\tau_y f| = \tau_y |f|$ and so $\int |\tau_y f| = \int |f| < \infty$ by Exercise 1.3.15. Applying Exercise 1.3.15 to each u_j and then Definition 1.3.17,

$$\int_{\mathbb{R}^d} f(x+y)\,dx = \Big(\int \tau_y u_1 - \int \tau_y u_2\Big) + i\Big(\int \tau_y u_3 - \int \tau_y u_4\Big) = \Big(\int u_1 - \int u_2\Big) + i\Big(\int u_3 - \int u_4\Big) = \int_{\mathbb{R}^d} f(x)\,dx.$$

(All eight integrals are finite, so the rearrangement is legitimate.)

Exercise 1.3.16 for $f \in L^1(\mathbb{R}^d \to \mathbb{C})$. Let T be an invertible linear map and $D := |\det T| \in (0, \infty)$. Exactly as above, $f \circ T^{-1}$ is defined almost everywhere (the image $T(E)$ of a full-measure set has full measure, since its complement is $T(\mathbb{R}^d \setminus E)$, a null set by Exercise 1.2.21), is measurable by Step 1 of Exercise 1.3.16 together with Exercise 1.3.7(iii), and satisfies $\int |f \circ T^{-1}| = \int |f| \circ T^{-1} = D \int |f| < \infty$ by Exercise 1.3.16, so it is absolutely integrable. Applying Exercise 1.3.16 to each u_j,

$$\int_{\mathbb{R}^d} f(T^{-1}(x))\,dx = \Big(D\int u_1 - D\int u_2\Big) + i\Big(D\int u_3 - D\int u_4\Big) = D\int_{\mathbb{R}^d} f(x)\,dx,$$

since D is a finite positive real and all the integrals involved are finite. Replacing T by T^{-1} as in Exercise 1.3.16 gives the equivalent form $\int f(Tx)\,dx = \frac{1}{|\det T|}\int f(x)\,dx$.

Exercise 1.3.17 for Riemann integrable $f : [a,b] \to \mathbb{C}$. Here f Riemann integrable means that $\operatorname{Re} f$ and $\operatorname{Im} f$ are Riemann integrable, with $\int_a^b f := \int_a^b \operatorname{Re} f + i\int_a^b \operatorname{Im} f$. It therefore suffices to treat a real-valued Riemann integrable $u : [a,b] \to \mathbb{R}$ and show $\int_{\mathbb{R}} \tilde{u} = \int_a^b u$, where the tilde denotes extension by zero; the complex case then follows from Definition 1.3.17, because $\operatorname{Re}\tilde{f} = \widetilde{\operatorname{Re} f}$ and $\operatorname{Im}\tilde{f} = \widetilde{\operatorname{Im} f}$.

So let $u : [a,b] \to \mathbb{R}$ be Riemann integrable; it is bounded, and $\tilde{u}$ is measurable by Exercise 1.3.9, with $|\tilde{u}| \le M 1_{[a,b]}$ for some finite M, so $\int_{\mathbb{R}} |\tilde{u}| \le M(b-a) < \infty$ by Exercise 1.3.10(i),(ii) and $\tilde{u}$ is absolutely integrable. We claim $u_+ := \max(u, 0)$ and $u_- := \max(-u, 0)$ are Riemann integrable on $[a,b]$. Indeed, by Exercise 1.1.22 and Definition 1.1.6, for any $\varepsilon > 0$ there are piecewise constant $g \le u \le h$ on $[a,b]$ with $\text{p.c.}\int_a^b (h-g) \le \varepsilon$. Then $g_+ \le u_+ \le h_+$, these are again piecewise constant, and since $t \mapsto \max(t, 0)$ is 1-Lipschitz and $h \ge g$ we have $0 \le h_+ - g_+ \le h - g$ pointwise, whence $\text{p.c.}\int_a^b (h_+ - g_+) \le \varepsilon$ by Exercise 1.1.21. Letting $\varepsilon \to 0$ shows that the lower and upper Darboux integrals of u_+ agree, so u_+ is Darboux integrable, hence Riemann integrable by Exercise 1.1.22. The same argument applies to $u_- = (-u)_+$, using that $-u$ is Riemann integrable by Exercise 1.1.24(1).

Now $u_\pm : [a,b] \to [0, +\infty)$ are unsigned and Riemann integrable, so the unsigned Exercise 1.3.17 gives $\int_{\mathbb{R}} \widetilde{u_\pm} = \int_a^b u_\pm$, and both quantities are finite. Since extension by zero commutes with taking positive and negative parts, $\widetilde{u_\pm} = (\tilde{u})_\pm$, so by (1.12) and the linearity of the Riemann integral (Exercise 1.1.24(1), applied to $u = u_+ - u_-$),

$$\int_{\mathbb{R}} \tilde{u}(x)\,dx = \int_{\mathbb{R}} (\tilde{u})_+ - \int_{\mathbb{R}} (\tilde{u})_- = \int_a^b u_+(x)\,dx - \int_a^b u_-(x)\,dx = \int_a^b u(x)\,dx,$$

as required. □

Problem 1.3.21. (Absolute summability is a special case of absolute integrability). Let $(c_n)_{n\in\mathbb{Z}}$ be a doubly infinite sequence of complex numbers, and let $f : \mathbb{R} \to \mathbb{C}$ be the function

$$f(x) := \sum_{n\in\mathbb{Z}} c_n 1_{[n,n+1)}(x) = c_{\lfloor x\rfloor}$$

where $\lfloor x\rfloor$ is the greatest integer less than x. Show that f is absolutely integrable if and only if the series $\sum_{n\in\mathbb{Z}} c_n$ is absolutely convergent, in which case one has $\int_{\mathbb{R}} f(x)\,dx = \sum_{n\in\mathbb{Z}} c_n$.

Solution. Note first that the defining sum is not really an infinite sum: the intervals $[n, n+1)$, $n \in \mathbb{Z}$, partition $\mathbb{R}$, so for each x exactly one term is non-zero and $f(x) = c_{\lfloor x\rfloor}$ is unambiguously defined.
Step 1: f is measurable. For each positive integer N set

$$f_N := \sum_{|n|\le N} c_n 1_{[n,n+1)} = f \cdot 1_{[-N,\,N+1)}.$$

Each f_N is a finite linear combination of indicators of Lebesgue measurable sets, hence a complex-valued simple function (Definition 1.3.2), and $f_N(x) \to f(x)$ for every $x \in \mathbb{R}$ (for a given x we have $f_N(x) = f(x)$ as soon as $N \ge |\lfloor x\rfloor|$). By Definition 1.3.11 (or Exercise 1.3.8(iv)), f is measurable.
Step 2: the key computation $\int_{\mathbb{R}} |f| = \sum_{n\in\mathbb{Z}} |c_n|$. Here the right-hand side is interpreted in $[0, +\infty]$ as $\sup_N \sum_{|n|\le N} |c_n|$, which is legitimate since the terms are unsigned. Note $|f| = \sum_n |c_n| 1_{[n,n+1)}$ pointwise, and $|f|$ is unsigned measurable by Step 1 and Exercise 1.3.8(v) (or Exercise 1.3.7(iii)).
For the lower bound, put $g_N := \sum_{|n|\le N} |c_n| 1_{[n,n+1)}$. This is an unsigned simple function with $0 \le g_N \le |f|$ pointwise, and by Definition 1.3.3 together with Lemma 1.2.6 ($m([n, n+1)) = 1$),

$$\mathrm{Simp}\int_{\mathbb{R}} g_N = \sum_{|n|\le N} |c_n|.$$

By Definition 1.3.12, $\int_{\mathbb{R}} |f| \ge \mathrm{Simp}\int_{\mathbb{R}} g_N$ for every N, so $\int_{\mathbb{R}} |f| \ge \sum_{n\in\mathbb{Z}} |c_n|$.
For the upper bound, observe that if $|x| \le N$ then $\lfloor x\rfloor \in \{-N, \dots, N\}$, so $|f| 1_{\{|x|\le N\}} \le g_N$ pointwise. By monotonicity (Exercise 1.3.10(ii)) and compatibility with the simple integral (Exercise 1.3.10(i)),

$$\int_{\mathbb{R}} |f(x)| 1_{\{|x|\le N\}}\,dx \le \mathrm{Simp}\int_{\mathbb{R}} g_N = \sum_{|n|\le N} |c_n| \le \sum_{n\in\mathbb{Z}} |c_n|.$$

By the vertical truncation property (Exercise 1.3.10(ix)) the left-hand side converges to $\int_{\mathbb{R}} |f|$ as $N \to \infty$, so $\int_{\mathbb{R}} |f| \le \sum_{n\in\mathbb{Z}} |c_n|$. This proves Step 2.
Step 3: the equivalence. By Step 1, f is measurable, so by Definition 1.3.17 it is absolutely integrable precisely when $\int_{\mathbb{R}} |f| < \infty$, which by Step 2 happens precisely when $\sum_{n\in\mathbb{Z}} |c_n| < \infty$, i.e. precisely when $\sum_{n\in\mathbb{Z}} c_n$ converges absolutely.
Step 4: the value of the integral. Assume now $\sum_n |c_n| < \infty$. First we record an elementary triangle inequality, proved without the later Lemma 1.3.19: if $u \in L^1(\mathbb{R}^d \to \mathbb{R})$ is real-valued then, since $|u| = u_+ + u_-$, Corollary 1.3.14 and (1.12) give

$$\Big|\int u\Big| = \Big|\int u_+ - \int u_-\Big| \le \int u_+ + \int u_- = \int |u|,$$

and hence for complex-valued $w \in L^1$, using $|\mathrm{Re}\, w|, |\mathrm{Im}\, w| \le |w|$ and monotonicity,

$$\Big|\int w\Big| \le \Big|\int \mathrm{Re}\, w\Big| + \Big|\int \mathrm{Im}\, w\Big| \le \int |\mathrm{Re}\, w| + \int |\mathrm{Im}\, w| \le 2\int |w|.$$

(The crude factor of 2 is harmless here.)
Next, f_N from Step 1 is absolutely integrable (it is bounded and supported in a set of finite measure), and by linearity of the absolutely convergent integral (Exercise 1.3.19) together with $\int_{\mathbb{R}} 1_{[n,n+1)} = m([n, n+1)) = 1$ (Exercise 1.3.10(i) and Exercise 1.3.1(vi)),

$$\int_{\mathbb{R}} f_N(x)\,dx = \sum_{|n|\le N} c_n.$$

Finally set $h_N := f - f_N = \sum_{|n|>N} c_n 1_{[n,n+1)}$, which is absolutely integrable (a difference of two elements of L^1, by Step 3 and Exercise 1.3.19). The function h_N has exactly the same form as f, with the sequence $(c_n 1_{|n|>N})_{n\in\mathbb{Z}}$ in place of (c_n), so Step 2 applied to it gives

$$\int_{\mathbb{R}} |h_N| = \sum_{|n|>N} |c_n|.$$

Combining, by linearity (Exercise 1.3.19) and the inequality above,

$$\Big| \int_{\mathbb{R}} f(x)\,dx - \sum_{|n|\le N} c_n \Big| = \Big| \int_{\mathbb{R}} h_N(x)\,dx \Big| \le 2 \sum_{|n|>N} |c_n|.$$

Since $\sum_n |c_n| < \infty$, the tail $\sum_{|n|>N} |c_n|$ tends to 0 as $N \to \infty$. Hence the symmetric partial sums $\sum_{|n|\le N} c_n$ converge to $\int_{\mathbb{R}} f(x)\,dx$. As the series is absolutely convergent, its sum $\sum_{n\in\mathbb{Z}} c_n$ is well defined independently of any ordering and equals the limit of these symmetric partial sums; therefore

$$\int_{\mathbb{R}} f(x)\,dx = \sum_{n\in\mathbb{Z}} c_n. \qquad \square$$

Problem 1.3.22. If E, F are disjoint measurable subsets of $\mathbb{R}^d$, and $f : E \cup F \to \mathbb{C}$ is absolutely integrable, show that

$$\int_E f(x)\,dx = \int_{E\cup F} f(x) 1_E(x)\,dx$$

and

$$\int_E f(x)\,dx + \int_F f(x)\,dx = \int_{E\cup F} f(x)\,dx.$$

Solution. Recall the definition of the localised integral given just before the exercise: if $A \subset \mathbb{R}^d$ is measurable and $h : A \to \mathbb{C}$ is a function, one writes $\tilde{h} : \mathbb{R}^d \to \mathbb{C}$ for the extension of h by zero, i.e. $\tilde{h}(x) := h(x)$ for $x \in A$ and $\tilde{h}(x) := 0$ for $x \notin A$; one declares h measurable (resp. absolutely integrable) precisely when $\tilde{h}$ is, and one sets $\int_A h(x)\,dx := \int_{\mathbb{R}^d} \tilde{h}(x)\,dx$.
Write $A := E \cup F$, which is measurable since E and F are, and let

$$h : \mathbb{R}^d \to \mathbb{C}, \qquad h(x) := \begin{cases} f(x), & x \in A \\ 0, & x \notin A \end{cases}$$

be the extension of f by zero. By hypothesis h is measurable and $\int_{\mathbb{R}^d} |h(x)|\,dx < \infty$, and by definition

$$\int_{E\cup F} f(x)\,dx = \int_{\mathbb{R}^d} h(x)\,dx.$$

Step 1: the auxiliary functions $h1_E$ and $h1_F$ are absolutely integrable. The indicator 1_E is measurable, since E is a Lebesgue measurable set; hence $h1_E$ is measurable, being a product of measurable functions (Exercise 1.3.8(vi)). Moreover $|h1_E| \le |h|$ pointwise, so by monotonicity of the lower Lebesgue integral (Exercise 1.3.10(ii))

$$\int_{\mathbb{R}^d} |h(x)1_E(x)|\,dx \le \int_{\mathbb{R}^d} |h(x)|\,dx < \infty,$$

so $h1_E$ lies in $L^1(\mathbb{R}^d \to \mathbb{C})$. The same argument applies to $h1_F$.
Step 2: the first identity. Consider the two functions being integrated on the two sides.
On the left, the integrand is the restriction $f|_E : E \to \mathbb{C}$. Its extension by zero to $\mathbb{R}^d$ is the function which equals $f(x)$ for $x \in E$ and 0 for $x \notin E$. Since $E \subset A$, we have $h(x) = f(x)$ for $x \in E$, so this extension is exactly $h1_E$. By Step 1 it is absolutely integrable, so $f|_E$ is absolutely integrable on E and, by definition of the localised integral,

$$\int_E f(x)\,dx = \int_{\mathbb{R}^d} h(x)1_E(x)\,dx.$$

On the right, the integrand is the function $x \mapsto f(x)1_E(x)$ defined on $A = E \cup F$. Its extension by zero to $\mathbb{R}^d$ equals $f(x)1_E(x)$ for $x \in A$ and 0 off A; since h vanishes off A and agrees with f on A, this extension is again exactly $h1_E$. Hence $f1_E$ is absolutely integrable on $E \cup F$ and

$$\int_{E\cup F} f(x)1_E(x)\,dx = \int_{\mathbb{R}^d} h(x)1_E(x)\,dx.$$

The two right-hand sides coincide, which is the first identity. (Both sides are, so to speak, two different names for the same extension-by-zero function $h1_E$.)
Step 3: the second identity. We claim the pointwise identity

$$h(x) = h(x)1_E(x) + h(x)1_F(x) \qquad \text{for all } x \in \mathbb{R}^d.$$

Indeed, if $x \in E$ then, since E and F are disjoint, $1_E(x) = 1$ and $1_F(x) = 0$, and the right-hand side is $h(x)$; symmetrically if $x \in F$; and if $x \notin E \cup F$ then $h(x) = 0$ and both terms on the right vanish. This proves the claim.
All three functions h, $h1_E$, $h1_F$ are absolutely integrable (hypothesis, and Step 1), so linearity of the absolutely convergent Lebesgue integral (Exercise 1.3.19) applies and gives

$$\int_{\mathbb{R}^d} h(x)\,dx = \int_{\mathbb{R}^d} h(x)1_E(x)\,dx + \int_{\mathbb{R}^d} h(x)1_F(x)\,dx.$$

By Step 2 the first term on the right is $\int_E f(x)\,dx$, and by the same argument applied with F in place of E the second term is $\int_F f(x)\,dx$; the left-hand side is $\int_{E\cup F} f(x)\,dx$ by definition. Rearranging,

$$\int_E f(x)\,dx + \int_F f(x)\,dx = \int_{E\cup F} f(x)\,dx,$$

as required.

Problem 1.3.23. Show that the hypothesis that f is absolutely integrable in Lusin's theorem can be relaxed to being locally absolutely integrable (i.e. absolutely integrable on every bounded set), and then relaxed further to that of being measurable (but still finite everywhere or almost everywhere). (To achieve the latter goal, one can replace f locally with a horizontal truncation $f1_{|f|\leq n}$; alternatively, one can replace f with a bounded variant, such as $\frac{f}{(1+|f|^2)^{1/2}}$.)

Solution. Throughout, "the restriction of f to $\mathbb{R}^d \setminus E$ is continuous" means continuity with respect to the subspace topology on $\mathbb{R}^d \setminus E$. We record the local nature of this notion, which is what makes the patching below work.
Observation (locality). Let $E \subset \mathbb{R}^d$, let $x \in \mathbb{R}^d \setminus E$, and suppose V is an open neighbourhood of x. If $g : \mathbb{R}^d \to \mathbb{C}$ agrees with f on $V \setminus E$, and if for some set E' with $x \notin E'$ and $E' \cap V \subset E$ the restriction of g to $\mathbb{R}^d \setminus E'$ is continuous at x, then the restriction of f to $\mathbb{R}^d \setminus E$ is continuous at x. Indeed, given $\delta > 0$, continuity of $g|_{\mathbb{R}^d\setminus E'}$ at x produces $\rho > 0$ with $B(x,\rho) \subset V$ and $|g(y) - g(x)| \leq \delta$ for all $y \in B(x,\rho) \setminus E'$; but every $y \in B(x,\rho) \setminus E$ lies in V and is not in E' (as $E' \cap V \subset E$), so $|f(y) - f(x)| = |g(y) - g(x)| \leq \delta$ for all such y.
Part 1: f locally absolutely integrable.
Let $f : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable on every bounded measurable set, and let $\varepsilon > 0$. Chop $\mathbb{R}^d$ into open cubical shells: put $Q_0 := \emptyset$ and $Q_n := [-n,n]^d$ for $n \geq 1$, and set

$$U_n := (-n,n)^d \setminus Q_{n-1} \qquad (n \geq 1).$$

Each U_n is open (an open cube minus a closed cube), bounded, and the U_n are pairwise disjoint. Moreover

$$S := \mathbb{R}^d \setminus \bigcup_{n\geq 1} U_n \subset \bigcup_{n\geq 1} \partial Q_n$$

is a null set: if $x \notin \bigcup_{n\geq 1} U_n$, let $n \geq 1$ be least with $x \in Q_n$ (such n exists as $|x|$ is finite, and $Q_0 = \emptyset$); then $x \notin Q_{n-1}$, so $x \notin U_n$ forces $x \notin (-n, n)^d$, i.e. $x \in \partial Q_n$. Each ∂Q_n is a finite union of faces of a cube, each face being a box with one degenerate side and hence of measure zero, so ∂Q_n is null; countable subadditivity then makes S null.
Since U_n is bounded and measurable, $f1_{U_n}$ is absolutely integrable on $\mathbb{R}^d$ by hypothesis. Apply Lusin's theorem (Theorem 1.3.28) to $f1_{U_n}$ with parameter $\varepsilon/2^n$: there is a measurable set $E_n \subset \mathbb{R}^d$ with $m(E_n) \leq \varepsilon/2^n$ such that the restriction of $f1_{U_n}$ to $\mathbb{R}^d \setminus E_n$ is continuous. Set

$$E := S \cup \bigcup_{n\geq 1} \big(E_n \cap U_n\big),$$

a measurable set with $m(E) \leq 0 + \sum_{n\geq 1} \varepsilon/2^n = \varepsilon$ by countable subadditivity.
We claim the restriction of f to $\mathbb{R}^d \setminus E$ is continuous. Let $x \in \mathbb{R}^d \setminus E$. Since $x \notin S$, there is a (unique) n with $x \in U_n$, and U_n is open, so $V := U_n$ is an open neighbourhood of x. On V the function $g := f1_{U_n}$ agrees with f; and $E_n \cap V = E_n \cap U_n \subset E$. Also $x \notin E$ and $x \in U_n$ give $x \notin E_n$, so $x \in \mathbb{R}^d \setminus E_n$ and $g|_{\mathbb{R}^d\setminus E_n}$ is continuous at x. The Observation (with $E' := E_n$) shows that $f|_{\mathbb{R}^d\setminus E}$ is continuous at x. As x was arbitrary, the restriction of f to $\mathbb{R}^d \setminus E$ is continuous, and Lusin's theorem holds for locally absolutely integrable f.
Part 2: f merely measurable.
First suppose $f : \mathbb{R}^d \to \mathbb{C}$ is measurable (hence finite everywhere). Let $\Phi : \mathbb{C} \to \mathbb{C}$ be

$$\Phi(z) := \frac{z}{(1 + |z|^2)^{1/2}}.$$

This is continuous on $\mathbb{C}$, so $g := \Phi \circ f$ is measurable by Exercise 1.3.8(v). Writing $z = re^{i\theta}$ with $r \geq 0$ we get $\Phi(z) = \frac{r}{(1+r^2)^{1/2}} e^{i\theta}$, and $r \mapsto r/(1 + r^2)^{1/2}$ is a strictly increasing bijection from $[0, \infty)$ onto $[0, 1)$ with inverse $s \mapsto s/(1 - s^2)^{1/2}$. Hence Φ is a bijection from $\mathbb{C}$ onto the open unit disc $D := \{w \in \mathbb{C} : |w| < 1\}$, with inverse

$$\Psi : D \to \mathbb{C}, \qquad \Psi(w) = \frac{w}{(1 - |w|^2)^{1/2}},$$

which is continuous on D (the denominator is continuous and nonvanishing there).
Now $|g| < 1$ everywhere, so for any bounded measurable set B we have $|g|1_B \leq 1_B$ pointwise and therefore, by monotonicity (Exercise 1.3.10(ii)) together with compatibility of the lower Lebesgue integral with the simple integral (Exercise 1.3.10(i), applied to the simple function 1_B),

$$\int_{\mathbb{R}^d} |g(x)|1_B(x)\, dx \leq \int_{\mathbb{R}^d} 1_B(x)\, dx = m(B) < \infty,$$

the finiteness holding because a bounded measurable set has finite measure. Hence $g1_B$ is absolutely integrable for every bounded measurable B (it is measurable, being a product of measurable functions), i.e. g is locally absolutely integrable. By Part 1, given $\varepsilon > 0$ there is a measurable E with $m(E) \leq \varepsilon$ such that $g|_{\mathbb{R}^d\setminus E}$ is continuous.
Finally, g takes values in D and $f = \Psi \circ g$ pointwise (since Ψ inverts Φ). Therefore

$$f|_{\mathbb{R}^d\setminus E} = \Psi \circ \big(g|_{\mathbb{R}^d\setminus E}\big)$$

is a composition of the continuous map $g|_{\mathbb{R}^d\setminus E} : \mathbb{R}^d \setminus E \to D$ with the continuous map $\Psi : D \to \mathbb{C}$, hence continuous. This is Lusin's theorem for arbitrary measurable f.
(The truncation route indicated in the hint works equally well: on each shell U_n one replaces f by $f1_{|f|\leq k}$, notes that this truncation is bounded and supported in a bounded set, hence absolutely integrable, and uses that by Exercise 1.2.11(ii) the sets $U_n \cap \{|f| > k\}$ have measure tending to 0 as $k \to \infty$, since f is finite; deleting one such set of measure at most $\varepsilon/2^{n+1}$ and applying Lusin to the truncation reduces to Part 1.)
Part 3: f finite only almost everywhere.

Suppose finally that f is measurable and finite almost everywhere (for instance $f : \mathbb{R}^d \to [0, +\infty]$ measurable and a.e. finite, or f defined only almost everywhere). Let Z be the null set on which f is undefined or infinite, and let $f_0 := f1_{\mathbb{R}^d \setminus Z}$, which agrees with f outside a null set, is defined and finite everywhere, and is measurable by Exercise 1.3.8(iii). By Part 2 there is a measurable E_0 with $m(E_0) \leq \varepsilon$ such that $f_0|_{\mathbb{R}^d \setminus E_0}$ is continuous. Put $E := E_0 \cup Z$; then $m(E) \leq \varepsilon + 0 = \varepsilon$, and on $\mathbb{R}^d \setminus E$ we have $f = f_0$, so $f|_{\mathbb{R}^d \setminus E}$ is the restriction of the continuous function $f_0|_{\mathbb{R}^d \setminus E_0}$ to the smaller set $\mathbb{R}^d \setminus E$ and is therefore continuous. This is exactly the manoeuvre described in Remark 1.3.29.

Problem 1.3.24. Show that a function $f : \mathbb{R}^d \to \mathbb{C}$ is measurable if and only if it is the pointwise almost everywhere limit of continuous functions $f_n : \mathbb{R}^d \to \mathbb{C}$. (Hint: if $f : \mathbb{R}^d \to \mathbb{C}$ is measurable and $n \geq 1$, show that there exists a continuous function $f_n : \mathbb{R}^d \to \mathbb{C}$ for which the set $\{x \in B(0, n) : |f(x) - f_n(x)| \geq 1/n\}$ has measure at most $\frac{1}{2^n}$. You may find Exercise 1.3.25 below to be useful for this.) Use this (and Egorov's theorem, Theorem 1.3.26) to give an alternate proof of Lusin's theorem for arbitrary measurable functions.

Solution. The easy direction. Suppose $f_n : \mathbb{R}^d \to \mathbb{C}$ are continuous and $f_n \to f$ pointwise almost everywhere. Each f_n is measurable by Exercise 1.3.8(i), and a pointwise almost everywhere limit of measurable functions is measurable by Exercise 1.3.8(iv). Hence f is measurable.
The main direction. Suppose $f : \mathbb{R}^d \to \mathbb{C}$ is measurable; note that f is finite everywhere, since it is complex-valued.
Step 1 (the quantitative claim of the hint). Fix $n \geq 1$. Put

$$h := f1_{B(0,n)}.$$

Then h is measurable (product of measurable functions, Exercise 1.3.8(vi)) and supported on the ball $B(0, n)$, which is measurable (being open) and of finite measure (being contained in a box). By Exercise 1.3.25(ii) applied to h with parameter 2^{-n-1}, there is a measurable set $E \subset \mathbb{R}^d$ with $m(E) \leq 2^{-n-1}$ such that h is locally bounded off E; in particular there is $M < \infty$ with $|h(x)| \leq M$ for all $x \in B(0, n) \setminus E$. Let $h' := h1_{\mathbb{R}^d \setminus E}$. Then h' is measurable, vanishes outside $B(0, n)$, and satisfies $|h'| \leq M$ everywhere: indeed h' vanishes off $B(0, n) \setminus E$, and on $B(0, n) \setminus E$ it equals h, which is bounded by M there. Consequently $|h'| \leq M1_{B(0,n)}$ pointwise, so by monotonicity and homogeneity of the lower Lebesgue integral (Exercise 1.3.10(ii),(iii)) together with its compatibility with the simple integral (Exercise 1.3.10(i), applied to $1_{B(0,n)}$),

$$\int_{\mathbb{R}^d} |h'(x)|\, dx \leq M\, m(B(0, n)) < \infty,$$

so h' is absolutely integrable. By Theorem 1.3.20(iii) there is a continuous, compactly supported function $f_n : \mathbb{R}^d \to \mathbb{C}$ with

$$\|h' - f_n\|_{L^1(\mathbb{R}^d)} \leq \frac{1}{n2^{n+1}}.$$

Applying Markov's inequality (Lemma 1.3.15) to the unsigned measurable function $|h' - f_n|$ with $\lambda = 1/n$,

$$m\big(\{x \in \mathbb{R}^d : |h'(x) - f_n(x)| \geq 1/n\}\big) \leq n\|h' - f_n\|_{L^1(\mathbb{R}^d)} \leq 2^{-n-1}.$$

Now set

$$F_n := \{x \in B(0, n) : |f(x) - f_n(x)| \geq 1/n\}.$$

If $x \in F_n$ and $x \notin E$, then $x \in B(0, n)$ gives $f(x) = h(x)$, and $x \notin E$ gives $h(x) = h'(x)$; hence $|h'(x) - f_n(x)| = |f(x) - f_n(x)| \geq 1/n$. Therefore

$$F_n \subset E \cup \{x : |h'(x) - f_n(x)| \geq 1/n\},$$

and by subadditivity $m(F_n) \leq 2^{-n-1} + 2^{-n-1} = 2^{-n}$, which is the claim of the hint. (Each F_n is measurable, being the intersection of $B(0, n)$ with a superlevel set of the measurable function $|f - f_n|$.)

Step 2 (almost everywhere convergence). Let $G_N := \bigcup_{n \geq N} F_n$ and $G := \bigcap_{N \geq 1} G_N$. By countable subadditivity,

$$m(G_N) \leq \sum_{n \geq N} 2^{-n} = 2^{-N+1},$$

and since $G \subset G_N$ for every N, monotonicity gives $m(G) \leq 2^{-N+1}$ for every N, whence $m(G) = 0$.
Let $x \in \mathbb{R}^d \backslash G$. Then there is N with $x \notin G_N$, i.e. $x \notin F_n$ for all $n \geq N$. Choose $N' := \max(N, \lceil |x| \rceil + 1)$. For every $n \geq N'$ we have $|x| < n$, so $x \in B(0,n)$, and $x \notin F_n$; by the definition of F_n this forces

$$|f(x) - f_n(x)| < 1/n.$$

Letting $n \to \infty$ we conclude $f_n(x) \to f(x)$. Thus $f_n \to f$ pointwise outside the null set G, i.e. pointwise almost everywhere, with each f_n continuous. This completes the equivalence.
Step 3 (alternate proof of Lusin's theorem for arbitrary measurable f). Let $f : \mathbb{R}^d \to \mathbb{C}$ be measurable and let $\varepsilon > 0$. By Steps 1 and 2 there are continuous $f_n : \mathbb{R}^d \to \mathbb{C}$ with $f_n \to f$ pointwise almost everywhere. By Egorov's theorem (Theorem 1.3.26) there is a Lebesgue measurable set A with $m(A) \leq \varepsilon$ such that $f_n \to f$ locally uniformly on $\mathbb{R}^d \setminus A$.
We claim the restriction of f to $\mathbb{R}^d \setminus A$ is continuous, which is precisely the conclusion of Lusin's theorem (Theorem 1.3.28) with $E := A$. Let $x_0 \in \mathbb{R}^d \setminus A$ and put

$$B := B(x_0, 1) \cap (\mathbb{R}^d \setminus A),$$

a bounded subset of $\mathbb{R}^d \setminus A$. By Definition 1.3.21 (local uniform convergence in the bounded-set sense), $f_n \to f$ uniformly on B. Each restriction $f_n|_B$ is continuous, being the restriction of a continuous function on $\mathbb{R}^d$ to a subset; a uniform limit of continuous functions on a metric space is continuous; hence $f|_B$ is continuous, in particular at x_0.
Finally, B is a relatively open neighbourhood of x_0 in $\mathbb{R}^d \setminus A$, since $B(x_0, 1)$ is open in $\mathbb{R}^d$. Continuity of a restriction at a point only depends on the behaviour of the function on a relative neighbourhood of that point, so $f|_{\mathbb{R}^d \setminus A}$ is continuous at x_0. As $x_0 \in \mathbb{R}^d \setminus A$ was arbitrary, $f|_{\mathbb{R}^d \setminus A}$ is continuous, and Lusin's theorem follows.
(Note that this route proves Lusin's theorem for every measurable f, not merely for absolutely integrable f, so it reproves Exercise 1.3.23 as well; the absolute integrability of f was used nowhere above, only the finiteness of f, which enters through Exercise 1.3.25(ii).)

Problem 1.3.25. (Littlewood-like principles) The following facts are not, strictly speaking, instances of any of Littlewood's three principles, but are in a similar spirit.

- (i) (Absolutely integrable functions almost have bounded support) Let $f : \mathbb{R}^d \to \mathbb{C}$ be an absolutely integrable function, and let $\varepsilon > 0$. Show that there exists a ball $B(0,R)$ outside of which f has an L^1 norm of at most ε, or in other words that $\int_{\mathbb{R}^d \setminus B(0,R)} |f(x)|\, dx \leq \varepsilon$.
- (ii) (Measurable functions are almost locally bounded) Let $f : \mathbb{R}^d \to \mathbb{C}$ be a measurable function supported on a set of finite measure, and let $\varepsilon > 0$. Show that there exists a measurable set $E \subset \mathbb{R}^d$ of measure at most ε outside of which f is locally bounded, or in other words that for every $R > 0$ there exists $M < \infty$ such that $|f(x)| \leq M$ for all $x \in B(0,R) \setminus E$.

Solution. Part (i). The function $|f| : \mathbb{R}^d \to [0, +\infty)$ is unsigned and measurable (as $z \mapsto |z|$ is continuous, Exercise 1.3.8(v)), and $\int_{\mathbb{R}^d} |f(x)|\, dx < \infty$ by absolute integrability. Write $A_n := \{x \in \mathbb{R}^d : |x| \leq n\}$.
By the vertical truncation property of the lower Lebesgue integral (Exercise 1.3.10(ix)), which for measurable functions is a statement about the unsigned Lebesgue integral (Definition 1.3.13),

$$\int_{\mathbb{R}^d} |f(x)| 1_{A_n}(x)\, dx \longrightarrow \int_{\mathbb{R}^d} |f(x)|\, dx \qquad \text{as } n \to \infty.$$

On the other hand, the divisibility property (Exercise 1.3.10(vii)) applied to the measurable set A_n

gives

$$\int_{\mathbb{R}^d} |f(x)|\,dx = \int_{\mathbb{R}^d} |f(x)|1_{A_n}(x)\,dx + \int_{\mathbb{R}^d} |f(x)|1_{\mathbb{R}^d\setminus A_n}(x)\,dx.$$

All three quantities here are finite, because each is bounded above by $\int_{\mathbb{R}^d}|f| < \infty$ (monotonicity, Exercise 1.3.10(ii)). We may therefore subtract, obtaining

$$\int_{\mathbb{R}^d} |f(x)|1_{\mathbb{R}^d\setminus A_n}(x)\,dx = \int_{\mathbb{R}^d} |f(x)|\,dx - \int_{\mathbb{R}^d} |f(x)|1_{A_n}(x)\,dx \longrightarrow 0$$

as $n\to\infty$. Choose n so large that $\int_{\mathbb{R}^d}|f(x)|1_{\mathbb{R}^d\setminus A_n}(x)\,dx \le \varepsilon$, and set $R := n+1$. Since

$$\mathbb{R}^d \setminus B(0,R) = \{x : |x| \ge n+1\} \subset \{x : |x| > n\} = \mathbb{R}^d \setminus A_n,$$

we have $|f|1_{\mathbb{R}^d\setminus B(0,R)} \le |f|1_{\mathbb{R}^d\setminus A_n}$ pointwise, and hence, by monotonicity again,

$$\int_{\mathbb{R}^d\setminus B(0,R)} |f(x)|\,dx = \int_{\mathbb{R}^d} |f(x)|1_{\mathbb{R}^d\setminus B(0,R)}(x)\,dx \le \int_{\mathbb{R}^d} |f(x)|1_{\mathbb{R}^d\setminus A_n}(x)\,dx \le \varepsilon,$$

where the first equality is the definition of the localised integral. This is the required conclusion.
Part (ii). Let $S\subset\mathbb{R}^d$ be a set of finite measure outside of which f vanishes. Since f is complex-valued, $|f(x)|<\infty$ for every x. Define, for $n\ge 1$,

$$S_n := \{x\in\mathbb{R}^d : |f(x)| > n\}.$$

Each S_n is Lebesgue measurable: writing $U_n := \{z\in\mathbb{C} : |z|>n\}$, which is an open subset of $\mathbb{C}$, we have $S_n = f^{-1}(U_n)$, and $f^{-1}(U)$ is Lebesgue measurable for every open $U\subset\mathbb{C}$ by Exercise 1.3.7(iv). The sets S_n are decreasing in n, since $U_{n+1}\subset U_n$. Moreover $S_1 \subset \{x : f(x)\neq 0\} = f^{-1}(\mathbb{C}\setminus\{0\}) \subset S$, and the middle set is measurable by Exercise 1.3.7(iv) again, so by monotonicity of Lebesgue outer measure $m(S_1)\le m^*(S) < \infty$. Finally,

$$\bigcap_{n\ge 1} S_n = \{x : |f(x)| > n \text{ for all } n\ge 1\} = \emptyset,$$

precisely because $|f(x)|$ is finite for every x.
We may therefore apply downward monotone convergence for Lebesgue measure (Exercise 1.2.11(ii)), whose hypothesis that at least one of the sets have finite measure is met by $m(S_1)<\infty$:

$$\lim_{n\to\infty} m(S_n) = m\Big(\bigcap_{n\ge 1} S_n\Big) = m(\emptyset) = 0.$$

Hence there exists $n_0\ge 1$ with $m(S_{n_0})\le\varepsilon$. Put

$$E := S_{n_0} = \{x : |f(x)| > n_0\},$$

a measurable set of measure at most ε. For every $x\in\mathbb{R}^d\setminus E$ we have $|f(x)|\le n_0$ by the very definition of E. In particular, for any $R>0$ the choice $M := n_0$ works: $|f(x)|\le M$ for all $x\in B(0,R)\setminus E$. So f is locally bounded outside E, as required.
Two remarks. First, the argument gives more than local boundedness: outside the exceptional set E the function f is bounded by the single constant n_0, uniformly on all of $\mathbb{R}^d\setminus E$. This stronger form is what makes the exercise usable in Exercise 1.3.24.
Second, the hypothesis that f be finite is essential, and is exactly the point flagged in the discussion following the exercise: it is what forces $\bigcap_n S_n = \emptyset$. If instead f were unsigned and took the value $+\infty$ on a set P of positive measure, then $P\subset S_n$ for every n, so $m(S_n)\ge m(P)>0$ for all n; and indeed no set E with $m(E)<m(P)$ can work, since $|f| = +\infty$ somewhere on $B(0,R)\setminus E$ for R large enough that $m(P\cap B(0,R)) > m(E)$ (such R exists by upward monotone convergence, Exercise 1.2.11(i)).
The finite-measure support hypothesis, by contrast, is only needed for the uniform bound: for a general finite measurable f one recovers the stated local boundedness by applying the above to $f1_{B(0,k)}$ with parameter $\varepsilon/2^k$ for each $k\ge 1$ and taking E to be the union of the resulting exceptional sets.

2.4 §1.4 — Abstract Measure Spaces

Problem 1.4.1. (Elementary algebra) Let $\mathcal{E}[\mathbb{R}^d]$ be the collection of those sets $E \subset \mathbb{R}^d$ that are either elementary sets, or co-elementary sets (i.e. the complement of an elementary set). Show that $\mathcal{E}[\mathbb{R}^d]$ is a Boolean algebra. We will call this algebra the elementary Boolean algebra of $\mathbb{R}^d$.

Solution. Throughout we use Exercise 1.1.1 (Boolean closure): if $A, B \subset \mathbb{R}^d$ are elementary sets then so are $A \cup B$, $A \cap B$ and $A \setminus B$. Recall also from Definition 1.1.1 that an elementary set is a finite union of boxes, and write $\mathcal{E}[\mathbb{R}^d] = \{E : E \text{ elementary}\} \cup \{E : E^c \text{ elementary}\}$.
We verify the three axioms of Definition 1.4.1.
(i) Empty set. The empty set is the union of an empty (finite) collection of boxes, hence elementary; equivalently it is the single box $[0,0) \times \cdots \times [0,0)$, since $[0,0) = \emptyset$ is a legitimate (degenerate) interval in Definition 1.1.1. So $\emptyset \in \mathcal{E}[\mathbb{R}^d]$.
(ii) Complement. Let $E \in \mathcal{E}[\mathbb{R}^d]$. If E is elementary, then E^c is by definition the complement of an elementary set, i.e. co-elementary, and so $E^c \in \mathcal{E}[\mathbb{R}^d]$. If instead E is co-elementary, say $E = F^c$ with F elementary, then $E^c = (F^c)^c = F$ is elementary, so again $E^c \in \mathcal{E}[\mathbb{R}^d]$.
(iii) Finite unions. Let $E, F \in \mathcal{E}[\mathbb{R}^d]$. There are three cases.
Case 1: both E and F are elementary. Then $E \cup F$ is elementary by Exercise 1.1.1 (or directly: a union of two finite unions of boxes is a finite union of boxes). Hence $E \cup F \in \mathcal{E}[\mathbb{R}^d]$.
Case 2: exactly one of them, say E, is elementary, and $F = G^c$ is co-elementary with G elementary. Then

$$E \cup F = E \cup G^c = (E^c \cap G)^c = (G \setminus E)^c.$$

By Exercise 1.1.1 the set $G \setminus E$ is elementary, so $E \cup F$ is co-elementary and therefore lies in $\mathcal{E}[\mathbb{R}^d]$.
Case 3: both are co-elementary, say $E = G^c$ and $F = H^c$ with G, H elementary. By de Morgan,

$$E \cup F = G^c \cup H^c = (G \cap H)^c,$$

and $G \cap H$ is elementary by Exercise 1.1.1, so $E \cup F$ is co-elementary and lies in $\mathcal{E}[\mathbb{R}^d]$.
In every case $E \cup F \in \mathcal{E}[\mathbb{R}^d]$. Thus $\mathcal{E}[\mathbb{R}^d]$ satisfies (i), (ii), (iii) of Definition 1.4.1 and is a Boolean algebra on $\mathbb{R}^d$.
Two remarks. First, the collection of elementary sets on its own is not a Boolean algebra on $\mathbb{R}^d$: every elementary set is bounded (each box is bounded, and a finite union of bounded sets is bounded), so $\mathbb{R}^d$ itself is not elementary, and complements of nonempty elementary sets are never elementary. Adjoining the co-elementary sets is exactly what is needed to repair closure under complementation, and Case 2 and Case 3 above show that no further sets need to be added. Second, no set is both elementary and co-elementary: if E and E^c were both elementary then $\mathbb{R}^d = E \cup E^c$ would be elementary, hence bounded, which is false.

Problem 1.4.2. (Restriction) Let $\mathcal{B}$ be a Boolean algebra on a set X, and let Y be a subset of X (not necessarily $\mathcal{B}$-measurable). Show that the restriction $\mathcal{B} \restriction_Y := \{E \cap Y : E \in \mathcal{B}\}$ of $\mathcal{B}$ to Y is a Boolean algebra on Y. If Y is $\mathcal{B}$-measurable, show that

$$\mathcal{B} \restriction_Y = \mathcal{B} \cap 2^Y = \{E \subset Y : E \in \mathcal{B}\}.$$

Solution. Part 1: $\mathcal{B} \restriction_Y$ is a Boolean algebra on Y. Every member of $\mathcal{B} \restriction_Y$ is of the form $E \cap Y$ and hence is a subset of Y, so $\mathcal{B} \restriction_Y$ is a collection of subsets of Y, as required by Definition 1.4.1 with ambient set Y. We check the three axioms, remembering that complements are now taken inside Y.
(i) Empty set. Since $\emptyset \in \mathcal{B}$, we have $\emptyset = \emptyset \cap Y \in \mathcal{B} \restriction_Y$.
(ii) Complement. Let $A \in \mathcal{B} \restriction_Y$, say $A = E \cap Y$ with $E \in \mathcal{B}$. Then, using the distributive law,

$$Y \setminus A = Y \setminus (E \cap Y) = Y \cap (E \cap Y)^c = Y \cap (E^c \cup Y^c) = (Y \cap E^c) \cup \emptyset = E^c \cap Y,$$

where the complements $E^c := X \setminus E$ and $Y^c := X \setminus Y$ are taken in X. Since $\mathcal{B}$ is a Boolean algebra, $E^c \in \mathcal{B}$, and therefore $Y \setminus A = E^c \cap Y \in \mathcal{B} \restriction_Y$.

(iii) Finite unions. Let $A = E \cap Y$ and $B = F \cap Y$ with $E, F \in \mathcal{B}$. Then

$$A \cup B = (E \cap Y) \cup (F \cap Y) = (E \cup F) \cap Y,$$

and $E \cup F \in \mathcal{B}$, so $A \cup B \in \mathcal{B} \restriction_Y$.
Hence $\mathcal{B} \restriction_Y$ is a Boolean algebra on Y. Note that no measurability of Y was used anywhere.
Part 2: the case $Y \in \mathcal{B}$. First observe that the two right-hand descriptions agree by definition: $\mathcal{B} \cap 2^Y$ consists of those sets that both belong to $\mathcal{B}$ and are subsets of Y, which is exactly $\{E \subset Y : E \in \mathcal{B}\}$. So it suffices to prove $\mathcal{B} \restriction_Y = \mathcal{B} \cap 2^Y$.
$\subset$: let $A \in \mathcal{B} \restriction_Y$, say $A = E \cap Y$ with $E \in \mathcal{B}$. Since $Y \in \mathcal{B}$ by hypothesis, and a Boolean algebra is closed under finite intersections (as noted after Definition 1.4.1: $E \cap Y = (E^c \cup Y^c)^c$, and the right-hand side lies in $\mathcal{B}$ by closure under complement and finite union), we get $A = E \cap Y \in \mathcal{B}$. Also $A \subset Y$. Hence $A \in \mathcal{B} \cap 2^Y$.
$\supset$: let $A \in \mathcal{B} \cap 2^Y$, so $A \in \mathcal{B}$ and $A \subset Y$. Then $A = A \cap Y$ exhibits A as the restriction of the $\mathcal{B}$-measurable set A, so $A \in \mathcal{B} \restriction_Y$.
Therefore $\mathcal{B} \restriction_Y = \mathcal{B} \cap 2^Y = \{E \subset Y : E \in \mathcal{B}\}$ whenever $Y \in \mathcal{B}$.
Remark. The hypothesis $Y \in \mathcal{B}$ is genuinely needed for Part 2. For instance, take $X = \mathbb{R}$, $\mathcal{B} = \{\emptyset, X\}$ the trivial algebra, and $Y = [0, 1]$. Then $\mathcal{B} \restriction_Y = \{\emptyset, [0, 1]\}$, while $\mathcal{B} \cap 2^Y = \{\emptyset\}$, which is not even a Boolean algebra on Y.

Problem 1.4.3. Show that the non-empty atoms of an atomic algebra are determined up to relabeling. More precisely, show that if

$$X = \bigcup_{\alpha \in I} A_\alpha = \bigcup_{\alpha' \in I'} A'_{\alpha'}$$

are two partitions of X into non-empty atoms A_α, $A'_{\alpha'}$, then $\mathcal{A}((A_\alpha)_{\alpha \in I}) = \mathcal{A}((A'_{\alpha'})_{\alpha' \in I'})$ if and only if there exists a bijection $\phi : I \to I'$ such that $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$.

Solution. Recall from Example 1.4.7 that for a partition $X = \bigcup_{\alpha \in I} A_\alpha$ into disjoint sets, the atomic algebra is

$$\mathcal{A}((A_\alpha)_{\alpha \in I}) = \Big\{ \bigcup_{\alpha \in J} A_\alpha : J \subset I \Big\}.$$

Throughout, both families consist of non-empty pairwise disjoint sets covering X.
Preliminary observation A (indexing is faithful). If $\alpha \neq \beta$ in I then $A_\alpha \neq A_\beta$. Indeed $A_\alpha \cap A_\beta = \emptyset$ by disjointness, so if $A_\alpha = A_\beta$ then $A_\alpha = A_\alpha \cap A_\alpha = \emptyset$, contradicting non-emptiness. Thus $\alpha \mapsto A_\alpha$ is injective, and likewise $\alpha' \mapsto A'_{\alpha'}$ is injective.
Preliminary observation B (atoms are minimal measurable sets). Let $E \in \mathcal{A}((A_\alpha)_{\alpha \in I})$ and suppose $E \cap A_\beta \neq \emptyset$ for some $\beta \in I$. Then $A_\beta \subset E$. Indeed, write $E = \bigcup_{\alpha \in J} A_\alpha$ and pick $x \in E \cap A_\beta$. Then $x \in A_\alpha$ for some $\alpha \in J$, and $x \in A_\alpha \cap A_\beta$ forces $\alpha = \beta$ by disjointness; hence $\beta \in J$ and $A_\beta \subset E$. In particular each non-empty E in the algebra is a union of the atoms it meets.
Sufficiency ($\Leftarrow$). Suppose $\phi : I \to I'$ is a bijection with $A'_{\phi(\alpha)} = A_\alpha$ for all α. Then the two families $(A_\alpha)_{\alpha \in I}$ and $(A'_{\alpha'})_{\alpha' \in I'}$ have exactly the same members: every A_α equals $A'_{\phi(\alpha)}$, and every $A'_{\alpha'}$ equals $A_{\phi^{-1}(\alpha')}$. Consequently, given $J \subset I$ we have $\bigcup_{\alpha \in J} A_\alpha = \bigcup_{\alpha' \in \phi(J)} A'_{\alpha'}$, so $\mathcal{A}((A_\alpha)) \subset \mathcal{A}((A'_{\alpha'}))$; applying the same argument to ϕ^{-1} gives the reverse inclusion. Hence the two atomic algebras coincide.
Necessity ($\Rightarrow$). Suppose $\mathcal{A} := \mathcal{A}((A_\alpha)_{\alpha \in I}) = \mathcal{A}((A'_{\alpha'})_{\alpha' \in I'})$.
Fix $\alpha \in I$. Since $A_\alpha \in \mathcal{A}$, it also lies in $\mathcal{A}((A'_{\alpha'}))$, so $A_\alpha = \bigcup_{\alpha' \in J'} A'_{\alpha'}$ for some $J' \subset I'$. As $A_\alpha \neq \emptyset$, the index set J' is non-empty; choose $\alpha' \in J'$. Then

$$A'_{\alpha'} \subset A_\alpha.$$

Now run the same argument with the roles reversed: $A'_{\alpha'} \in \mathcal{A}$, so $A'_{\alpha'} = \bigcup_{\beta \in J} A_\beta$ for some non-empty $J \subset I$. Pick any $\beta \in J$. Then $A_\beta \subset A'_{\alpha'} \subset A_\alpha$, and $A_\beta \neq \emptyset$, so $A_\beta \cap A_\alpha \neq \emptyset$; by disjointness of the first

partition, $\beta = \alpha$. Since $\beta \in J$ was arbitrary, $J = \{\alpha\}$, whence

$$A'_{\alpha'} = A_\alpha.$$

So for each $\alpha \in I$ there is at least one $\alpha' \in I'$ with $A'_{\alpha'} = A_\alpha$, and by observation A that α' is unique. Define $\phi(\alpha)$ to be this unique index; then $\phi : I \to I'$ is a well-defined map with $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$.
ϕ is injective: if $\phi(\alpha) = \phi(\beta)$ then $A_\alpha = A'_{\phi(\alpha)} = A'_{\phi(\beta)} = A_\beta$, and observation A gives $\alpha = \beta$.
ϕ is surjective: fix $\alpha' \in I'$. Exchanging the roles of the two partitions in the argument above (which was symmetric in them), there is $\alpha \in I$ with $A_\alpha = A'_{\alpha'}$; by uniqueness in the definition of ϕ we get $\phi(\alpha) = \alpha'$.
Hence ϕ is a bijection with $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$, as required.
Conclusion: an atomic algebra determines its non-empty atoms exactly, as the minimal non-empty measurable sets it contains (observation B), and the two partitions can differ only in how those sets are labelled.

Problem 1.4.4. Show that every finite Boolean algebra is an atomic algebra. (A Boolean algebra $\mathcal{B}$ is finite if its cardinality is finite, i.e. there are only finitely many measurable sets.) Conclude that every finite Boolean algebra has a cardinality of the form 2^n for some natural number n. From this exercise and Exercise 1.4.3 we see that there is a one-to-one correspondence between finite Boolean algebras on X and finite partitions of X into non-empty sets (up to relabeling).

Solution. Let $\mathcal{B}$ be a finite Boolean algebra on a set X. Note first that $X = \emptyset^c \in \mathcal{B}$, and that $\mathcal{B}$ is closed under finite intersections (Definition 1.4.1 together with de Morgan, as remarked after that definition).
Step 1: the atom of a point. For $x \in X$ define

$$A_x := \bigcap \{E \in \mathcal{B} : x \in E\}.$$

The family being intersected is non-empty (it contains X) and finite (it is a subfamily of the finite set $\mathcal{B}$), so A_x is a finite intersection of members of $\mathcal{B}$ and therefore $A_x \in \mathcal{B}$. Clearly $x \in A_x$, so $A_x \neq \emptyset$. By construction A_x is the smallest $\mathcal{B}$-measurable set containing x: if $E \in \mathcal{B}$ and $x \in E$ then $A_x \subset E$.
Step 2: the sets A_x are pairwise equal or disjoint. We first show

$$y \in A_x \implies A_y = A_x.$$

Suppose $y \in A_x$.
($A_y \subset A_x$.) If $E \in \mathcal{B}$ and $x \in E$, then $A_x \subset E$, so $y \in E$; hence every $E \in \mathcal{B}$ containing x also contains y, so the family defining A_y contains the family defining A_x, and intersecting a larger family gives a smaller set: $A_y \subset A_x$.
($A_x \subset A_y$.) Let $E \in \mathcal{B}$ with $y \in E$; we must show $A_x \subset E$. If $x \in E$ this is immediate from Step 1. If $x \notin E$, then $x \in E^c \in \mathcal{B}$, so $A_x \subset E^c$, and since $y \in A_x$ this gives $y \in E^c$, contradicting $y \in E$. So the case $x \notin E$ cannot occur, and $A_x \subset E$ for every such E; intersecting over all of them yields $A_x \subset A_y$.
Now if $A_x \cap A_y \neq \emptyset$, pick z in the intersection; by the implication just proved, $A_z = A_x$ and $A_z = A_y$, so $A_x = A_y$. Hence any two of the sets A_x are either identical or disjoint. Since every x lies in A_x, the distinct sets among $\{A_x : x \in X\}$ form a partition of X into non-empty sets. Each is a member of the finite set $\mathcal{B}$, so there are only finitely many of them; list the distinct ones as $A_1, \ldots, A_n$ (here $n = 0$ only in the degenerate case $X = \emptyset$).
Step 3: $\mathcal{B}$ is the atomic algebra with atoms $A_1, \ldots, A_n$. Write $\mathcal{A} := \mathcal{A}((A_i)_{i=1}^n)$.
$\mathcal{A} \subset \mathcal{B}$: each $A_i \in \mathcal{B}$ by Step 1, and a member of $\mathcal{A}$ is a union of finitely many of the A_i, hence lies in $\mathcal{B}$ by closure under finite unions.
$\mathcal{B} \subset \mathcal{A}$: let $E \in \mathcal{B}$. For $x \in E$ we have $A_x \subset E$ by Step 1, and $x \in A_x$; therefore

$$E = \bigcup_{x \in E} A_x,$$

which is a union of some subcollection of $\{A_1, \ldots, A_n\}$, hence $E \in \mathcal{A}$.

So $\mathcal{B} = \mathcal{A}((A_i)_{i=1}^n)$ is an atomic algebra.
Step 4: cardinality. Consider the map $\Phi : 2^{\{1,\dots,n\}} \to \mathcal{B}$, $\Phi(J) := \bigcup_{i\in J} A_i$. It is surjective by Step 3. It is injective: if $J \neq J'$, say $i \in J \setminus J'$, then $A_i \subset \Phi(J)$ while $A_i \cap \Phi(J') = \bigcup_{j\in J'}(A_i \cap A_j) = \emptyset$ by disjointness of distinct atoms; since $A_i \neq \emptyset$ this gives $\Phi(J) \neq \Phi(J')$. Hence Φ is a bijection and

$$|\mathcal{B}| = |2^{\{1,\dots,n\}}| = 2^n.$$

Step 5: the correspondence. Steps 1 to 3 assign to each finite Boolean algebra $\mathcal{B}$ on X a finite partition $P(\mathcal{B}) := \{A_1, \dots, A_n\}$ of X into non-empty sets whose atomic algebra is $\mathcal{B}$; conversely, by Example 1.4.7 every finite partition P of X into non-empty sets generates a Boolean algebra $\mathcal{A}(P)$, which is finite (of cardinality $2^{|P|}$, by the bijection Φ of Step 4, whose injectivity used only that the members of P are non-empty and pairwise disjoint).
These two constructions are mutually inverse. In one direction, Step 3 says $\mathcal{A}(P(\mathcal{B})) = \mathcal{B}$. In the other, let $P = (A_\alpha)_{\alpha\in I}$ be a finite partition of X into non-empty sets and put $\mathcal{B} := \mathcal{A}(P)$; we claim $P(\mathcal{B}) = P$. Indeed, fix $x \in X$ and let A_β be the unique member of P containing x. Then $A_\beta \in \mathcal{B}$ and $x \in A_\beta$, so $A_x \subset A_\beta$ by the minimality in Step 1. Conversely, if $E \in \mathcal{B}$ and $x \in E$, then E is a union of members of P and meets A_β, so $A_\beta \subset E$ by observation B of Exercise 1.4.3; taking the intersection over all such E gives $A_\beta \subset A_x$. Hence $A_x = A_\beta$, and the distinct sets among $\{A_x : x \in X\}$ are exactly the members of P, i.e. $P(\mathcal{B}) = P$.
Finally, by Exercise 1.4.3 two finite partitions of X into non-empty sets generate the same algebra precisely when they agree up to a relabeling bijection, so the correspondence is well defined on relabeling classes and is injective on them. Hence finite Boolean algebras on X correspond one-to-one with finite partitions of X into non-empty sets, up to relabeling.

Problem 1.4.5. Show that the elementary, Jordan, Lebesgue, and null algebras are not atomic algebras. (Hint: argue by contradiction. If these algebras were atomic, what must the atoms be?)

Solution. Fix $d \geq 1$. The four algebras in question are $\mathcal{E}[\mathbb{R}^d]$ (Exercise 1.4.1), $\mathcal{J}[\mathbb{R}^d]$ (Example 1.4.4), $\mathcal{L}[\mathbb{R}^d]$ (Example 1.4.5) and $\mathcal{N}(\mathbb{R}^d)$ (Example 1.4.6). All four contain every singleton and none of them is the discrete algebra $2^{\mathbb{R}^d}$; the following lemma shows that these two facts are incompatible with being atomic.
Key Lemma. Let $\mathcal{B}$ be an atomic algebra on a set X which contains every singleton $\{x\}$, $x \in X$. Then $\mathcal{B} = 2^X$ is the discrete algebra.
Proof of the Lemma. Write $\mathcal{B} = \mathcal{A}((A_\alpha)_{\alpha\in I})$ for a partition $X = \bigcup_{\alpha\in I} A_\alpha$ into disjoint atoms; by the remark at the end of Example 1.4.7 we may discard the empty atoms and assume every A_α is non-empty. Fix $x \in X$ and let A_β be the (unique) atom containing x. By hypothesis $\{x\} \in \mathcal{B}$, so $\{x\} = \bigcup_{\alpha\in J} A_\alpha$ for some $J \subset I$; since $x \in A_\beta$ and the atoms are disjoint, this union must include A_β, i.e. $A_\beta \subset \{x\}$. (Equivalently: an atom meeting a measurable set is contained in it, as in Exercise 1.4.3.) As $A_\beta \ni x$ we conclude $A_\beta = \{x\}$. Thus every atom is a singleton and every singleton is an atom, so the partition is the discrete partition $X = \bigcup_{x\in X}\{x\}$, whose atomic algebra is 2^X (Example 1.4.7). $\square$
It therefore suffices, for each of the four algebras $\mathcal{B}$, to check that (a) $\{x\} \in \mathcal{B}$ for every $x \in \mathbb{R}^d$, and (b) $\mathcal{B} \neq 2^{\mathbb{R}^d}$, i.e. some subset of $\mathbb{R}^d$ is not $\mathcal{B}$-measurable. Then $\mathcal{B}$ cannot be atomic.
Case A: the elementary algebra $\mathcal{E}[\mathbb{R}^d]$.
(a) By Definition 1.1.1 intervals may be degenerate, so $\{x\} = [x_1, x_1] \times \cdots \times [x_d, x_d]$ is a box, hence an elementary set, hence lies in $\mathcal{E}[\mathbb{R}^d]$.
(b) Every box is bounded, so every elementary set (a finite union of boxes) is bounded. Let $H := \{x \in \mathbb{R}^d : x_1 > 0\}$. Both H and $H^c = \{x \in \mathbb{R}^d : x_1 \leq 0\}$ are unbounded, hence neither is elementary; so H is neither elementary nor co-elementary and $H \notin \mathcal{E}[\mathbb{R}^d]$.
Case B: the Jordan algebra $\mathcal{J}[\mathbb{R}^d]$.
(a) Singletons are elementary by Case A(a), and elementary sets are Jordan measurable (with Jordan measure equal to their elementary measure), so $\{x\} \in \mathcal{J}[\mathbb{R}^d]$.
(b) Let $E := [0,1]^d \cap \mathbb{Q}^d$. Its closure is $[0,1]^d$: it is contained in the closed set $[0,1]^d$, and conversely

every $y \in [0,1]^d$ is a limit of points of $\mathbb{Q}^d \cap [0,1]^d$, since each coordinate $y_i \in [0,1]$ is a limit of rationals in $[0,1]$. Its interior is empty: any ball about any point contains a point all of whose coordinates are irrational, and such a point is not in E. Hence the topological boundary is $\partial E = \overline{E} \setminus E^\circ = [0,1]^d$, which is an elementary set of elementary measure 1 and so has Jordan outer measure $1 \neq 0$. By Exercise 1.1.18(3), E is not Jordan measurable. (For $d = 2$ this is exactly the set of bullets in Exercise 1.1.18(4).) Moreover E^c is unbounded, and Jordan measurable sets are by definition bounded, so E^c is not Jordan measurable either. Hence $E \notin \mathcal{J}[\mathbb{R}^d]$.
Case C: the Lebesgue algebra $\mathcal{L}[\mathbb{R}^d]$.
(a) A singleton has Lebesgue outer measure zero, hence is a null set, hence is Lebesgue measurable; so $\{x\} \in \mathcal{L}[\mathbb{R}^d]$.
(b) For $d = 1$ this is Proposition 1.2.18. For general $d \geq 1$ we repeat that proof verbatim with $\mathbb{Q}$ replaced by the countable dense additive subgroup $\mathbb{Q}^d$ of $\mathbb{R}^d$. Partition $\mathbb{R}^d$ into the disjoint cosets $x + \mathbb{Q}^d$. Each coset is dense in $\mathbb{R}^d$ (a translate of a dense set), hence meets $[0,1]^d$; by the axiom of choice pick $x_C \in C \cap [0,1]^d$ for each coset C, and set $E := \{x_C : C \in \mathbb{R}^d/\mathbb{Q}^d\} \subset [0,1]^d$.
The translates $E + q$, $q \in \mathbb{Q}^d$, are pairwise disjoint: if $x_C + q = x_{C'} + q'$ with $q, q' \in \mathbb{Q}^d$, then $x_C - x_{C'} \in \mathbb{Q}^d$, so $C = C'$, so $x_C = x_{C'}$ and $q = q'$. Moreover, if $y \in [0,1]^d$ then y lies in some coset C, and $y - x_C \in \mathbb{Q}^d \cap [-1,1]^d$ since both y and x_C lie in $[0,1]^d$. Writing $Q := \mathbb{Q}^d \cap [-1,1]^d$, which is countable, we therefore have

$$[0,1]^d \subset \bigcup_{q \in Q} (E + q) \subset [-1,2]^d.$$

Suppose E were Lebesgue measurable. Each $E + q$ is then measurable with $m(E + q) = m(E)$ by translation invariance (Exercise 1.2.20), and the union above is a countable disjoint union, so by countable additivity and monotonicity (Lemma 1.2.15 and Exercise 1.2.12)

$$1 = m([0,1]^d) \leq \sum_{q \in Q} m(E) \leq m([-1,2]^d) = 3^d.$$

If $m(E) = 0$ the middle quantity is 0, contradicting the left inequality; so $m(E) > 0$. But Q is countably infinite, so the middle quantity is then a sum of infinitely many copies of a strictly positive number, hence $+\infty$, contradicting the right inequality. So E is not Lebesgue measurable, and $\mathcal{L}[\mathbb{R}^d] \neq 2^{\mathbb{R}^d}$.
Case D: the null algebra $\mathcal{N}(\mathbb{R}^d)$.
(a) Singletons are null, hence lie in $\mathcal{N}(\mathbb{R}^d)$.
(b) Let $E := [0,1]^d$. Then $m(E) = 1 \neq 0$, so E is not null; and $E^c \supset [2,3]^d$, so $m(E^c) \geq 1 \neq 0$ and E^c is not null either, i.e. E is not co-null. Hence $E \notin \mathcal{N}(\mathbb{R}^d)$.
In each of the four cases the Key Lemma would force the algebra to be the discrete algebra, contradicting (b). Therefore none of the elementary, Jordan, Lebesgue and null algebras is atomic.
Remark. The content of the argument is exactly the hint: in each of these algebras points are measurable, so the only candidate atoms are the singletons, and the corresponding atomic algebra is the discrete algebra, which is strictly finer than all four. (By contrast, the dyadic algebras $\mathcal{D}_n(\mathbb{R}^d)$ of Example 1.4.8 are atomic; consistently with the Key Lemma, they do not contain singletons, since their atoms are the dyadic cubes of side 2^{-n}.)

Problem 1.4.6. (Intersection of algebras) Let $(\mathcal{B}_\alpha)_{\alpha \in I}$ be a family of Boolean algebras on a set X, indexed by a (possibly infinite or uncountable) label set I. Show that the intersection $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha := \bigcap_{\alpha \in I} \mathcal{B}_\alpha$ of these algebras is still a Boolean algebra, and is the finest Boolean algebra that is coarser than all of the $\mathcal{B}_\alpha$. (If I is empty, we adopt the convention that $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha$ is the discrete algebra.)

Solution. Write $\mathcal{B} := \bigcap_{\alpha \in I} \mathcal{B}_\alpha = \{E \subset X : E \in \mathcal{B}_\alpha \text{ for all } \alpha \in I\}$.
The case $I = \emptyset$. Every subset of X satisfies the (vacuous) condition, so the intersection is 2^X, which is the stated convention; and 2^X is a Boolean algebra (Example 1.4.3), trivially coarser than all members of the empty family, and it is the finest Boolean algebra of all, so the assertion holds. Assume from now on that $I \neq \emptyset$ (though the verification below does not actually need this).
Step 1: $\mathcal{B}$ is a Boolean algebra. Each element of $\mathcal{B}$ is a subset of X, so $\mathcal{B}$ is a collection of subsets of

X. We check Definition 1.4.1.
(i) Empty set. For every $\alpha \in I$ we have $\emptyset \in \mathcal{B}_\alpha$, since each $\mathcal{B}_\alpha$ is a Boolean algebra. Hence $\emptyset \in \mathcal{B}$.
(ii) Complement. Let $E \in \mathcal{B}$. Fix $\alpha \in I$. Then $E \in \mathcal{B}_\alpha$, so $E^c = X \setminus E \in \mathcal{B}_\alpha$. As α was arbitrary, $E^c \in \mathcal{B}$.
(iii) Finite unions. Let $E, F \in \mathcal{B}$ and fix $\alpha \in I$. Then $E, F \in \mathcal{B}_\alpha$, so $E \cup F \in \mathcal{B}_\alpha$. As α was arbitrary, $E \cup F \in \mathcal{B}$.
Thus $\mathcal{B}$ is a Boolean algebra on X.
Step 2: $\mathcal{B}$ is coarser than every $\mathcal{B}_\alpha$. By the definition of an intersection, $\mathcal{B} \subset \mathcal{B}_\alpha$ for each $\alpha \in I$, which by the terminology of Definition 1.4.1 says precisely that $\mathcal{B}$ is coarser than $\mathcal{B}_\alpha$ (equivalently, each $\mathcal{B}_\alpha$ is a refinement of $\mathcal{B}$).
Step 3: $\mathcal{B}$ is the finest such algebra. Let $\mathcal{C}$ be any Boolean algebra on X that is coarser than all the $\mathcal{B}_\alpha$, i.e. $\mathcal{C} \subset \mathcal{B}_\alpha$ for every $\alpha \in I$. Then every $E \in \mathcal{C}$ lies in every $\mathcal{B}_\alpha$, so $E \in \mathcal{B}$; that is, $\mathcal{C} \subset \mathcal{B}$, i.e. $\mathcal{B}$ is finer than $\mathcal{C}$. So $\mathcal{B}$ is itself coarser than all the $\mathcal{B}_\alpha$ (Step 2) and finer than every other Boolean algebra with that property; hence it is the finest such algebra. Being the maximum of the family of such algebras, it is also unique with this property: if $\mathcal{B}'$ were another finest one, then $\mathcal{B} \subset \mathcal{B}'$ and $\mathcal{B}' \subset \mathcal{B}$, so $\mathcal{B}' = \mathcal{B}$.
This is exactly what is needed to make Definition 1.4.10 legitimate: for any family $\mathcal{F}$ of subsets of X, the collection of Boolean algebras containing $\mathcal{F}$ is non-empty (it contains 2^X), and its intersection $\langle \mathcal{F} \rangle_{\text{bool}}$ is by the above a Boolean algebra, indeed the coarsest one containing $\mathcal{F}$.

Problem 1.4.7. Show that the elementary algebra $\mathcal{E}(\mathbb{R}^d)$ is generated by the collection of boxes in $\mathbb{R}^d$.

Solution. Let $\mathcal{F} := \{B \subset \mathbb{R}^d : B \text{ is a box}\}$ be the collection of boxes of Definition 1.1.1, and let $\mathcal{E}[\mathbb{R}^d]$ be the elementary algebra of Exercise 1.4.1, i.e. the collection of sets that are elementary or co-elementary. By Definition 1.4.10, $\langle \mathcal{F} \rangle_{\text{bool}}$ is the intersection of all Boolean algebras on $\mathbb{R}^d$ containing $\mathcal{F}$, equivalently the coarsest Boolean algebra containing every box. We must show

$$\langle \mathcal{F} \rangle_{\text{bool}} = \mathcal{E}[\mathbb{R}^d].$$

Inclusion $\subset$. By Exercise 1.4.1, $\mathcal{E}[\mathbb{R}^d]$ is a Boolean algebra, and it contains every box B, since a box is in particular an elementary set (a union of one box). Hence $\mathcal{E}[\mathbb{R}^d]$ is one of the algebras appearing in the intersection defining $\langle \mathcal{F} \rangle_{\text{bool}}$, and therefore $\langle \mathcal{F} \rangle_{\text{bool}} \subset \mathcal{E}[\mathbb{R}^d]$.
Inclusion $\supset$. Let $\mathcal{C}$ be any Boolean algebra on $\mathbb{R}^d$ with $\mathcal{F} \subset \mathcal{C}$; we show $\mathcal{E}[\mathbb{R}^d] \subset \mathcal{C}$. Let $E \in \mathcal{E}[\mathbb{R}^d]$.
If E is elementary, then by Definition 1.1.1 we may write $E = B_1 \cup \cdots \cup B_k$ as a finite union of boxes (with $E = \emptyset \in \mathcal{C}$ in the degenerate case $k = 0$). Each $B_i \in \mathcal{F} \subset \mathcal{C}$, and $\mathcal{C}$ is closed under finite unions, so by induction on k we get $E \in \mathcal{C}$.
If instead E is co-elementary, write $E = F^c$ with F elementary. By the previous paragraph $F \in \mathcal{C}$, and $\mathcal{C}$ is closed under complements, so $E = F^c \in \mathcal{C}$.
Hence $\mathcal{E}[\mathbb{R}^d] \subset \mathcal{C}$ for every Boolean algebra $\mathcal{C}$ containing all boxes. Intersecting over all such $\mathcal{C}$ gives $\mathcal{E}[\mathbb{R}^d] \subset \langle \mathcal{F} \rangle_{\text{bool}}$.
Combining the two inclusions, $\langle \mathcal{F} \rangle_{\text{bool}} = \mathcal{E}[\mathbb{R}^d]$: the elementary algebra is the Boolean algebra generated by the boxes.
Remark. The proof of the second inclusion is precisely the statement that adjoining to the boxes their finite unions and then the complements of those already yields a Boolean algebra; in the notation of Exercise 1.4.9 this says that the recursion stabilises immediately, $\mathcal{F}_1 = \mathcal{F}_2 = \cdots = \langle \mathcal{F} \rangle_{\text{bool}}$, which is exactly the content of Cases 1 to 3 in the solution to Exercise 1.4.1.

Problem 1.4.8. Let n be a natural number. Show that if $\mathcal{F}$ is a finite collection of n sets, then $\langle \mathcal{F} \rangle_{\text{bool}}$ is a finite Boolean algebra of cardinality at most 2^{2^n} (in particular, finite sets generate finite algebras). Give an example to show that this bound is best possible. (Hint: for the latter, it may be convenient to use a discrete ambient space such as the discrete cube $X = \{0, 1\}^n$.)

Solution. Write $\mathcal{F} = \{A_1, \dots, A_n\}$, a collection of subsets of an ambient set X. For a subset $A \subset X$ we use the superscript notation

$$A^{(1)} := A, \qquad A^{(0)} := X \setminus A.$$

For each $\omega = (\omega_1, \dots, \omega_n) \in \{0,1\}^n$ define

$$A_\omega := \bigcap_{i=1}^{n} A_i^{(\omega_i)}.$$

The sets A_ω partition X. They are pairwise disjoint: if $\omega \neq \omega'$ then $\omega_i \neq \omega_i'$ for some i, and then $A_\omega \subset A_i^{(\omega_i)}$ while $A_{\omega'} \subset A_i^{(\omega_i')} = X \setminus A_i^{(\omega_i)}$, so the two sets are disjoint. They cover X: given $x \in X$, define $\omega_i := 1$ if $x \in A_i$ and $\omega_i := 0$ otherwise; then $x \in A_i^{(\omega_i)}$ for every i, i.e. $x \in A_\omega$. (Some of the A_ω may of course be empty; this is permitted in Example 1.4.7.)
The atomic algebra of this partition contains $\mathcal{F}$. Let $\mathcal{A} := \mathcal{A}\big((A_\omega)_{\omega\in\{0,1\}^n}\big)$ be the atomic algebra generated by the partition $X = \bigcup_\omega A_\omega$, as in Example 1.4.7; it is a Boolean algebra. For each i we have

$$A_i = \bigcup_{\omega\,:\,\omega_i=1} A_\omega,$$

since the right-hand side is the union of exactly those atoms contained in A_i, and every point of A_i lies in the atom indexed by its own digit string, which has i-th digit 1. Hence $A_i \in \mathcal{A}$ for all i, so $\mathcal{F} \subset \mathcal{A}$. Since $\langle\mathcal{F}\rangle_{\text{bool}}$ is by Definition 1.4.10 the coarsest Boolean algebra containing $\mathcal{F}$, we get $\langle\mathcal{F}\rangle_{\text{bool}} \subset \mathcal{A}$. Conversely $\mathcal{A} \subset \langle\mathcal{F}\rangle_{\text{bool}}$. A Boolean algebra is closed under complement and finite union, hence (by de Morgan) under finite intersection as well, as noted in Remark 1.4.1. Each $A_i^{(\omega_i)}$ is either $A_i \in \langle\mathcal{F}\rangle_{\text{bool}}$ or its complement, so each atom A_ω, being a finite intersection of such sets, lies in $\langle\mathcal{F}\rangle_{\text{bool}}$; and then every union of atoms, being a finite union (there are only 2^n atoms), lies in $\langle\mathcal{F}\rangle_{\text{bool}}$ too. Thus

$$\langle\mathcal{F}\rangle_{\text{bool}} = \mathcal{A}\big((A_\omega)_{\omega\in\{0,1\}^n}\big).$$

The cardinality bound. Every element of $\mathcal{A}$ has the form $\bigcup_{\omega\in J} A_\omega$ for some $J \subset \{0,1\}^n$, so the map $J \mapsto \bigcup_{\omega\in J} A_\omega$ from the power set of $\{0,1\}^n$ onto $\mathcal{A}$ is surjective. Since $\{0,1\}^n$ has 2^n elements, its power set has 2^{2^n} elements, and therefore

$$\#\langle\mathcal{F}\rangle_{\text{bool}} = \#\mathcal{A} \leq 2^{2^n} < \infty.$$

In particular $\langle\mathcal{F}\rangle_{\text{bool}}$ is a finite Boolean algebra.
Sharpness. Take $X := \{0,1\}^n$ (the discrete cube) and, for $i = 1, \dots, n$,

$$A_i := \{x \in X : x_i = 1\},$$

so that $\mathcal{F} = \{A_1, \dots, A_n\}$ is a collection of n sets. (For $n \geq 1$ these are genuinely distinct sets, so $\#\mathcal{F} = n$.) With the notation above, for $\omega \in \{0,1\}^n$,

$$A_\omega = \bigcap_{i=1}^{n} \{x \in X : x_i = \omega_i\} = \{\omega\},$$

a singleton. Thus all 2^n atoms are non-empty and distinct, and the map $J \mapsto \bigcup_{\omega\in J}\{\omega\} = J$ is a bijection from the power set of X onto $\mathcal{A}$. Hence $\langle\mathcal{F}\rangle_{\text{bool}} = 2^X$ is the discrete algebra on X, of cardinality exactly $2^{\#X} = 2^{2^n}$. So the bound 2^{2^n} is attained, and is therefore best possible.
(For $n = 0$ the statement reads: $\mathcal{F} = \emptyset$ generates $\langle\emptyset\rangle_{\text{bool}} = \{\emptyset, X\}$, the trivial algebra, of cardinality $2 = 2^{2^0}$, in agreement with the above, the unique atom being $A_{()} = X$.)

Problem 1.4.9. (Recursive description of a generated Boolean algebra) Let $\mathcal{F}$ be a collection of sets in a set X. Define the sets $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \dots$ recursively as follows:

- (i) $\mathcal{F}_0 := \mathcal{F}$.
- (ii) For each $n \geq 1$, we define $\mathcal{F}_n$ to be the collection of all sets that are either the union of a finite number of sets in $\mathcal{F}_{n-1}$ (including the empty union $\emptyset$), or the complement of such a union.

Show that $\langle \mathcal{F} \rangle_{\text{bool}} = \bigcup_{n=0}^{\infty} \mathcal{F}_n$.

Solution. Write $\mathcal{G} := \bigcup_{n=0}^{\infty} \mathcal{F}_n$. We prove the two inclusions.
Step 1: the sequence $(\mathcal{F}_n)$ is increasing. We claim $\mathcal{F}_{n-1} \subset \mathcal{F}_n$ for every $n \geq 1$. Indeed, if $E \in \mathcal{F}_{n-1}$, then E is the union of the finite (one-element) family $\{E\}$ of sets in $\mathcal{F}_{n-1}$, and so $E \in \mathcal{F}_n$ by clause (ii). Consequently $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$, and in particular any two elements of $\mathcal{G}$ lie in a common $\mathcal{F}_n$.
Step 2: $\mathcal{G}$ is a Boolean algebra containing $\mathcal{F}$. Clearly $\mathcal{F} = \mathcal{F}_0 \subset \mathcal{G}$. We verify the three axioms of a Boolean algebra.
Empty set: $\emptyset$ is the empty union of sets in $\mathcal{F}_0$, so $\emptyset \in \mathcal{F}_1 \subset \mathcal{G}$.
Complement: let $E \in \mathcal{G}$, say $E \in \mathcal{F}_n$. If $n = 0$, then $E \in \mathcal{F}_0$ and $X \setminus E$ is the complement of the one-fold union E of sets in $\mathcal{F}_0$, hence $X \setminus E \in \mathcal{F}_1 \subset \mathcal{G}$. If $n \geq 1$, then by definition either $E = U$ or $E = X \setminus U$, where U is a finite union of sets in $\mathcal{F}_{n-1}$. In the first case $X \setminus E = X \setminus U$ is the complement of such a union, so $X \setminus E \in \mathcal{F}_n$; in the second case $X \setminus E = U$ is itself such a union, so again $X \setminus E \in \mathcal{F}_n \subset \mathcal{G}$.
Finite union: it suffices (by induction on the number of sets, the empty union being handled above) to treat the union of two sets. Let $E, E' \in \mathcal{G}$. By Step 1 there is a single n with $E, E' \in \mathcal{F}_n$. Then $E \cup E'$ is a finite union of sets in $\mathcal{F}_n$, hence $E \cup E' \in \mathcal{F}_{n+1} \subset \mathcal{G}$.
Thus $\mathcal{G}$ is a Boolean algebra on X containing $\mathcal{F}$. Since $\langle \mathcal{F} \rangle_{\text{bool}}$ is by Definition 1.4.10 the intersection of all Boolean algebras containing $\mathcal{F}$ (equivalently, the coarsest such algebra), we conclude

$$\langle \mathcal{F} \rangle_{\text{bool}} \subset \mathcal{G}.$$

Step 3: $\mathcal{G} \subset \langle \mathcal{F} \rangle_{\text{bool}}$. We show $\mathcal{F}_n \subset \langle \mathcal{F} \rangle_{\text{bool}}$ for all n by induction on n. For $n = 0$ this is $\mathcal{F} \subset \langle \mathcal{F} \rangle_{\text{bool}}$, which holds by definition. Suppose $\mathcal{F}_{n-1} \subset \langle \mathcal{F} \rangle_{\text{bool}}$ for some $n \geq 1$, and let $E \in \mathcal{F}_n$. Then $E = U$ or $E = X \setminus U$ where $U = B_1 \cup \cdots \cup B_k$ with $k \geq 0$ and $B_1, \ldots, B_k \in \mathcal{F}_{n-1} \subset \langle \mathcal{F} \rangle_{\text{bool}}$. A Boolean algebra contains $\emptyset$ and is closed under finite unions, so $U \in \langle \mathcal{F} \rangle_{\text{bool}}$ (this covers the case $k = 0$, where $U = \emptyset$); and it is closed under complements, so $X \setminus U \in \langle \mathcal{F} \rangle_{\text{bool}}$ as well. Either way $E \in \langle \mathcal{F} \rangle_{\text{bool}}$. This closes the induction, and taking the union over n gives $\mathcal{G} \subset \langle \mathcal{F} \rangle_{\text{bool}}$.
Combining Steps 2 and 3, $\langle \mathcal{F} \rangle_{\text{bool}} = \bigcup_{n=0}^{\infty} \mathcal{F}_n$, as claimed.

Problem 1.4.10. Show that all atomic algebras are σ-algebras. In particular, the discrete algebra and trivial algebra are σ-algebras, as are the finite algebras and the dyadic algebras on Euclidean spaces.

Solution. Let $X = \bigcup_{\alpha \in I} A_\alpha$ be a partition of X into disjoint atoms A_α, and let

$$\mathcal{A} := \mathcal{A}\big((A_\alpha)_{\alpha \in I}\big) = \Big\{ \bigcup_{\alpha \in J} A_\alpha \ : \ J \subset I \Big\}$$

be the associated atomic algebra of Example 1.4.7. We verify the three axioms of Definition 1.4.12.
Empty set. Taking $J = \emptyset$ gives the empty union, so $\emptyset \in \mathcal{A}$.
Complement. Let $E = \bigcup_{\alpha \in J} A_\alpha$ with $J \subset I$. We claim

$$X \setminus E = \bigcup_{\alpha \in I \setminus J} A_\alpha.$$

Indeed, if $x \in X \setminus E$, then since the A_α cover X we have $x \in A_\beta$ for some $\beta \in I$; if β were in J we would get $x \in E$, a contradiction, so $\beta \in I \setminus J$ and x lies in the right-hand side. Conversely, if $x \in A_\beta$ with $\beta \in I \setminus J$ and we had $x \in E$, then $x \in A_\alpha$ for some $\alpha \in J$; since $\alpha \neq \beta$ this contradicts the disjointness of the atoms. Hence $X \setminus E \in \mathcal{A}$.
Countable unions. In fact $\mathcal{A}$ is closed under arbitrary unions. Let $(E_n)_{n=1}^{\infty}$ be elements of $\mathcal{A}$, say

$E_n = \bigcup_{\alpha\in J_n} A_\alpha$ with $J_n \subset I$. Put $J := \bigcup_{n=1}^{\infty} J_n \subset I$. Then

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} \bigcup_{\alpha\in J_n} A_\alpha = \bigcup_{\alpha\in J} A_\alpha \in \mathcal{A},$$

the middle equality being just a regrouping of a union of sets. (Note that no countability of I is needed: the index set J is allowed to be arbitrary.)
Thus $\mathcal{A}$ is a σ-algebra on X.
Now for the special cases named in the exercise. As observed in Example 1.4.7, the trivial algebra $\{\emptyset, X\}$ is the atomic algebra of the one-atom partition $X = X$, and the discrete algebra 2^X is the atomic algebra of the partition $X = \bigcup_{x\in X}\{x\}$ into singletons; so both are σ-algebras. By Exercise 1.4.4, every finite Boolean algebra is an atomic algebra, hence is a σ-algebra. Finally, the dyadic algebra $\mathcal{D}_n(\mathbb{R}^d)$ of Example 1.4.8 is by definition the atomic algebra whose atoms are the half-open dyadic cubes

$$\Big[\frac{i_1}{2^n}, \frac{i_1+1}{2^n}\Big) \times \cdots \times \Big[\frac{i_d}{2^n}, \frac{i_d+1}{2^n}\Big), \qquad i_1, \ldots, i_d \in \mathbb{Z},$$

(see Exercise 1.1.14). These do partition $\mathbb{R}^d$: given $x \in \mathbb{R}^d$, the unique index vector with x in the corresponding cube is $i_r = \lfloor 2^n x_r \rfloor$ for $r = 1, \ldots, d$, since $i_r/2^n \le x_r < (i_r+1)/2^n$ holds precisely for that i_r. Hence each $\mathcal{D}_n(\mathbb{R}^d)$ is an atomic algebra, and therefore a σ-algebra as well.

Problem 1.4.11. Show that the Lebesgue and null algebras are σ-algebras, but the elementary and Jordan algebras are not.

Solution. Throughout, m^* denotes Lebesgue outer measure and $m^{*,(J)}, m_{*,(J)}$ denote Jordan outer and inner measure.
The Lebesgue algebra. Recall from Example 1.4.5 that $\mathcal{L}[\mathbb{R}^d]$ is the collection of Lebesgue measurable subsets of $\mathbb{R}^d$. Lemma 1.2.13 states exactly the three σ-algebra axioms for this collection: $\emptyset$ is Lebesgue measurable by Lemma 1.2.13(iv); if E is Lebesgue measurable then so is $\mathbb{R}^d \setminus E$ by Lemma 1.2.13(v); and if $E_1, E_2, \ldots$ are Lebesgue measurable then so is $\bigcup_{n=1}^{\infty} E_n$ by Lemma 1.2.13(vi). Hence $\mathcal{L}[\mathbb{R}^d]$ is a σ-algebra.
The null algebra. Recall from Example 1.4.6 that $\mathcal{N}(\mathbb{R}^d)$ consists of those $E \subset \mathbb{R}^d$ that are either Lebesgue null (i.e. $m^*(E) = 0$) or co-null (i.e. $m^*(\mathbb{R}^d \setminus E) = 0$). The empty set is null, so $\emptyset \in \mathcal{N}(\mathbb{R}^d)$, and the class is symmetric under complementation by construction, so it remains to check countable unions. Let $E_1, E_2, \cdots \in \mathcal{N}(\mathbb{R}^d)$ and put $E := \bigcup_{n=1}^{\infty} E_n$.
Case 1: every E_n is null. By countable subadditivity of Lebesgue outer measure (Exercise 1.2.3(iii)),

$$m^*(E) \le \sum_{n=1}^{\infty} m^*(E_n) = 0,$$

so E is null and hence $E \in \mathcal{N}(\mathbb{R}^d)$.
Case 2: some E_{n_0} is co-null. Since $E \supset E_{n_0}$ we have $\mathbb{R}^d \setminus E \subset \mathbb{R}^d \setminus E_{n_0}$, and the latter is null; by monotonicity of outer measure (Exercise 1.2.3(ii)), $m^*(\mathbb{R}^d \setminus E) = 0$, so E is co-null and again $E \in \mathcal{N}(\mathbb{R}^d)$.
Thus $\mathcal{N}(\mathbb{R}^d)$ is a σ-algebra.
The elementary algebra is not a σ-algebra. Recall from Exercise 1.4.1 that $\mathcal{E}[\mathbb{R}^d]$ consists of the elementary sets and their complements. For $n = 1, 2, 3, \ldots$ let

$$E_n := \Big\{\tfrac{1}{n}\Big\} \times \{0\}^{d-1} \subset \mathbb{R}^d$$

(for $d = 1$ this is just the singleton $\{1/n\}$). Each E_n is a box in the sense of Definition 1.1.1, all of whose intervals are degenerate, hence is an elementary set and lies in $\mathcal{E}[\mathbb{R}^d]$. Let

$$S := \bigcup_{n=1}^{\infty} E_n = \Big\{\big(\tfrac{1}{n}, 0, \ldots, 0\big) : n \ge 1\Big\}.$$

We claim $S \notin \mathcal{E}[\mathbb{R}^d]$.
First, S is not elementary. Suppose $S = B_1 \cup \cdots \cup B_k$ were a finite union of boxes; discarding any empty B_j (which changes nothing), we may assume each B_j is non-empty. Fix j and write $B_j = I_1 \times \cdots \times I_d$, so that each I_r is a non-empty interval. Since B_j is non-empty and $B_j \subset S$, projecting onto the first coordinate gives $I_1 \subset \{1/n : n \geq 1\}$; but an interval of positive length contains uncountably many points, whereas $\{1/n : n \geq 1\}$ is countable, so I_1 has length 0 and, being non-empty, is a single point. Projecting onto the r-th coordinate for $r \geq 2$ gives likewise $I_r \subset \{0\}$, so $I_r = \{0\}$. Hence each B_j is a single point, and so S would have at most k elements. But S is infinite, a contradiction.
Second, S is not co-elementary. Its complement $\mathbb{R}^d \setminus S$ is unbounded, whereas every elementary set is a finite union of boxes and hence bounded.
So S is a countable union of members of $\mathcal{E}[\mathbb{R}^d]$ that does not lie in $\mathcal{E}[\mathbb{R}^d]$: the elementary algebra is a Boolean algebra (Exercise 1.4.1) but not a σ-algebra.
The Jordan algebra is not a σ-algebra. Recall from Example 1.4.4 that $\mathcal{J}[\mathbb{R}^d]$ consists of the Jordan measurable sets and their complements. Enumerate $\mathbb{Q} \cap [0,1] = \{q_1, q_2, q_3, \dots\}$ (possible since the rationals are countable), and set

$$E_n := \{q_n\} \times [0,1]^{d-1}, \qquad S := \bigcup_{n=1}^{\infty} E_n = (\mathbb{Q} \cap [0,1]) \times [0,1]^{d-1}$$

(again the second factor is absent when $d = 1$). Each E_n is a box (degenerate in the first coordinate), hence elementary, hence Jordan measurable by the corollary to Exercise 1.1.5; so $E_n \in \mathcal{J}[\mathbb{R}^d]$.
We show S is not Jordan measurable by computing its Jordan inner and outer measures.
Inner measure. Note first that S is bounded (it is contained in $[0,1]^d$), so Definition 1.1.4 applies. Let $A \subset S$ be elementary, say $A = B_1 \cup \cdots \cup B_k$ with boxes $B_j = I_1^{(j)} \times \cdots \times I_d^{(j)}$. If B_j is empty then $|B_j| = 0$; and if B_j is non-empty then $B_j \subset S$ forces $I_1^{(j)} \subset \mathbb{Q}$, and as every interval of positive length contains an irrational number this gives $|I_1^{(j)}| = 0$ and hence $|B_j| = 0$. By the finite subadditivity of elementary measure recorded in Section 1.1.1 (a consequence of finite additivity and Exercise 1.1.1), $m(A) \leq \sum_{j=1}^{k} |B_j| = 0$, and $m(A) \geq 0$ by non-negativity, so $m(A) = 0$. Taking the supremum over all such A gives $m_{*,(J)}(S) = 0$.
Outer measure. Let $B \supset S$ be elementary, say $B = B_1 \cup \cdots \cup B_k$ with boxes B_j. Its closure satisfies $\overline{B} = \bigcup_j \overline{B_j}$, and

$$\overline{B} \setminus B \subset \bigcup_{j=1}^{k} \left(\overline{B_j} \setminus B_j\right) \subset \bigcup_{j=1}^{k} \partial B_j.$$

Here the second inclusion holds because $\partial B_j = \overline{B_j} \setminus \operatorname{int}(B_j)$ while $\operatorname{int}(B_j) \subset B_j$. If B_j is empty then $\partial B_j = \emptyset$ and $m(\partial B_j) = 0$ trivially, so fix a non-empty box $B_j = I_1^{(j)} \times \cdots \times I_d^{(j)}$, where $I_r^{(j)}$ has endpoints $a_r \leq b_r$; then $\overline{B_j} = \prod_r \overline{I_r^{(j)}} = \prod_r [a_r, b_r]$ and $\operatorname{int}(B_j) = \prod_r (a_r, b_r)$, so a point of ∂B_j has $x_r \in \{a_r, b_r\}$ for at least one r. Hence ∂B_j is contained in the union of the $2d$ degenerate closed boxes obtained from $\prod_r [a_r, b_r]$ by replacing the r-th factor by $\{a_r\}$ or by $\{b_r\}$. Each such box has volume 0, so this union is an elementary set of elementary measure 0 by finite subadditivity, and hence $m(\partial B_j) = 0$. Since $\overline{B} \setminus B$ is elementary (each $\overline{B_j}$ is a closed box, so $\overline{B}$ is elementary, and differences of elementary sets are elementary by Exercise 1.1.1) and is contained in the elementary set $\bigcup_j \partial B_j$ of measure 0, monotonicity gives $m(\overline{B} \setminus B) = 0$. Now $\overline{B} = B \cup (\overline{B} \setminus B)$, so finite subadditivity gives $m(\overline{B}) \leq m(B) + 0 = m(B)$, while monotonicity gives $m(B) \leq m(\overline{B})$; hence $m(\overline{B}) = m(B)$. Finally, $S \subset B$ implies $[0,1]^d = \overline{S} \subset \overline{B}$, so by monotonicity $m(B) = m(\overline{B}) \geq m([0,1]^d) = 1$. Taking the infimum over all elementary $B \supset S$ gives $m^{*,(J)}(S) \geq 1$; conversely $[0,1]^d$ is itself an elementary set containing S, so $m^{*,(J)}(S) \leq m([0,1]^d) = 1$. Hence $m^{*,(J)}(S) = 1$.
Since $m_{*,(J)}(S) = 0 \neq 1 = m^{*,(J)}(S)$, the set S is not Jordan measurable. Nor is its complement $\mathbb{R}^d \setminus S$, which is unbounded and therefore not Jordan measurable by the convention of Definition 1.1.4. Hence $S \notin \mathcal{J}[\mathbb{R}^d]$, and the Jordan algebra, though a Boolean algebra, is not a σ-algebra.

Problem 1.4.12. Show that any restriction $\mathcal{B}|_Y$ of a σ-algebra $\mathcal{B}$ to a subspace Y of X (as defined in Exercise 1.4.2) is again a σ-algebra on the subspace Y.

Solution. Recall from Exercise 1.4.2 that

$$\mathcal{B}|_Y := \{E \cap Y : E \in \mathcal{B}\},$$

a collection of subsets of Y (here $Y \subset X$ need not be $\mathcal{B}$-measurable). We check the three axioms of Definition 1.4.12, with Y playing the role of the ambient set, so that complements are taken relative to Y.
Empty set. Since $\mathcal{B}$ is a σ-algebra we have $\emptyset \in \mathcal{B}$, and $\emptyset = \emptyset \cap Y$, so $\emptyset \in \mathcal{B}|_Y$.
Complement. Let $F \in \mathcal{B}|_Y$, say $F = E \cap Y$ with $E \in \mathcal{B}$. Then

$$Y \setminus F = Y \setminus (E \cap Y) = \{y \in Y : y \notin E\} = (X \setminus E) \cap Y.$$

Since $\mathcal{B}$ is closed under complements, $X \setminus E \in \mathcal{B}$, and therefore $Y \setminus F \in \mathcal{B}|_Y$. (Note that this computation does not require Y itself to lie in $\mathcal{B}$.)
Countable unions. Let $F_1, F_2, \cdots \in \mathcal{B}|_Y$, and choose for each n a set $E_n \in \mathcal{B}$ with $F_n = E_n \cap Y$. By the distributivity of intersection over arbitrary unions,

$$\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} (E_n \cap Y) = \Big(\bigcup_{n=1}^{\infty} E_n \Big) \cap Y.$$

Since $\mathcal{B}$ is closed under countable unions, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$, and hence $\bigcup_{n=1}^{\infty} F_n \in \mathcal{B}|_Y$.
All three axioms hold, so $\mathcal{B}|_Y$ is a σ-algebra on Y.

Problem 1.4.13. (Intersection of σ-algebras) Show that the intersection $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha := \bigcap_{\alpha \in I} \mathcal{B}_\alpha$ of an arbitrary (and possibly infinite or uncountable) number of σ-algebras $\mathcal{B}_\alpha$ is again a σ-algebra, and is the finest σ-algebra that is coarser than all of the $\mathcal{B}_\alpha$.

Solution. Let $(\mathcal{B}_\alpha)_{\alpha \in I}$ be a family of σ-algebras on a common set X, and put $\mathcal{B} := \bigcap_{\alpha \in I} \mathcal{B}_\alpha$, i.e. $E \in \mathcal{B}$ precisely when $E \in \mathcal{B}_\alpha$ for every $\alpha \in I$. (If I is empty we adopt, as in Exercise 1.4.6, the convention that the intersection is the discrete algebra 2^X, which is a σ-algebra by Exercise 1.4.10; so we may assume I non-empty below, although in fact the verifications go through verbatim.)
$\mathcal{B}$ is a σ-algebra. We check Definition 1.4.12.
Empty set. Each $\mathcal{B}_\alpha$ contains $\emptyset$, so $\emptyset \in \mathcal{B}$.
Complement. Let $E \in \mathcal{B}$. For each α we have $E \in \mathcal{B}_\alpha$, and since $\mathcal{B}_\alpha$ is closed under complements, $X \setminus E \in \mathcal{B}_\alpha$. As this holds for every α, we get $X \setminus E \in \mathcal{B}$.
Countable unions. Let $E_1, E_2, \cdots \in \mathcal{B}$. Fix $\alpha \in I$. Then $E_n \in \mathcal{B}_\alpha$ for every n, and since $\mathcal{B}_\alpha$ is closed under countable unions, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_\alpha$. As α was arbitrary, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$.
Hence $\mathcal{B}$ is a σ-algebra on X.
$\mathcal{B}$ is the finest σ-algebra coarser than all the $\mathcal{B}_\alpha$. Recall that $\mathcal{B}'$ is coarser than $\mathcal{B}''$ precisely when $\mathcal{B}' \subset \mathcal{B}''$. By construction $\mathcal{B} \subset \mathcal{B}_\alpha$ for every $\alpha \in I$, so $\mathcal{B}$ is indeed coarser than each $\mathcal{B}_\alpha$. Now suppose $\mathcal{B}'$ is any σ-algebra on X that is coarser than every $\mathcal{B}_\alpha$, i.e. $\mathcal{B}' \subset \mathcal{B}_\alpha$ for all $\alpha \in I$. Then every $E \in \mathcal{B}'$ lies in every $\mathcal{B}_\alpha$, i.e. $E \in \mathcal{B}$; thus $\mathcal{B}' \subset \mathcal{B}$, meaning $\mathcal{B}$ is finer than $\mathcal{B}'$. So $\mathcal{B}$ is the finest such σ-algebra (and it is unique with this property, since two σ-algebras each finer than the other are equal).

Problem 1.4.14. Show that the Borel σ-algebra $\mathcal{B}[\mathbb{R}^d]$ of a Euclidean set is generated by any of the following collections of sets:

- (i) The open subsets of $\mathbb{R}^d$.
- (ii) The closed subsets of $\mathbb{R}^d$.
- (iii) The compact subsets of $\mathbb{R}^d$.

- (iv) The open balls of $\mathbb{R}^d$.
- (v) The boxes in $\mathbb{R}^d$.
- (vi) The elementary sets in $\mathbb{R}^d$.

(Hint: To show that two families $\mathcal{F}, \mathcal{F}'$ of sets generate the same σ-algebra, it suffices to show that every σ-algebra that contains $\mathcal{F}$, contains $\mathcal{F}'$ also, and conversely.)

Solution. Write $\mathcal{O}, \mathcal{C}, \mathcal{K}, \mathcal{U}, \mathcal{X}, \mathcal{E}$ for the collections in (i)–(vi) respectively: open sets, closed sets, compact sets, open balls, boxes, elementary sets. By Definition 1.4.16, $\mathcal{B}[\mathbb{R}^d] = \langle \mathcal{O} \rangle$, so we must show $\langle \mathcal{F} \rangle = \langle \mathcal{O} \rangle$ for each of the six families $\mathcal{F}$.

A preliminary reduction. If $\mathcal{F}, \mathcal{F}'$ are families of subsets of $\mathbb{R}^d$ with $\mathcal{F} \subset \langle \mathcal{F}' \rangle$, then $\langle \mathcal{F}' \rangle$ is a σ-algebra containing $\mathcal{F}$, and since by Definition 1.4.14 $\langle \mathcal{F} \rangle$ is the intersection of all σ-algebras containing $\mathcal{F}$, we get $\langle \mathcal{F} \rangle \subset \langle \mathcal{F}' \rangle$. Consequently:

$$\mathcal{F} \subset \langle \mathcal{F}' \rangle \text{ and } \mathcal{F}' \subset \langle \mathcal{F} \rangle \quad \Longrightarrow \quad \langle \mathcal{F} \rangle = \langle \mathcal{F}' \rangle.$$

This is the hint's criterion, and it is what we verify in each case. Throughout, we freely use that a σ-algebra is closed under complements, countable unions, countable intersections (by de Morgan) and finite unions and intersections (pad with $\emptyset$ or the whole space).

(i) Open sets. There is nothing to prove: $\langle \mathcal{O} \rangle = \mathcal{B}[\mathbb{R}^d]$ is exactly Definition 1.4.16 applied to the metric space $\mathbb{R}^d$.

(ii) Closed sets. Every closed set C is the complement $\mathbb{R}^d \setminus U$ of an open set U, so $C \in \langle \mathcal{O} \rangle$; thus $\mathcal{C} \subset \langle \mathcal{O} \rangle$. Symmetrically, every open set is the complement of a closed set, so $\mathcal{O} \subset \langle \mathcal{C} \rangle$. Hence $\langle \mathcal{C} \rangle = \langle \mathcal{O} \rangle$.

(iii) Compact sets. By the Heine–Borel theorem, a subset of $\mathbb{R}^d$ is compact if and only if it is closed and bounded; in particular $\mathcal{K} \subset \mathcal{C} \subset \langle \mathcal{C} \rangle$. Conversely, let C be closed and let $\overline{B}(0, n) := \{x : |x| \leq n\}$ be the closed ball of radius n about the origin. Then

$$C = \bigcup_{n=1}^{\infty} \big(C \cap \overline{B}(0, n)\big),$$

since every point of C has finite norm; and each $C \cap \overline{B}(0, n)$ is closed (an intersection of two closed sets) and bounded, hence compact. So C is a countable union of compact sets, whence $C \in \langle \mathcal{K} \rangle$ and $\mathcal{C} \subset \langle \mathcal{K} \rangle$. By the criterion and (ii), $\langle \mathcal{K} \rangle = \langle \mathcal{C} \rangle = \langle \mathcal{O} \rangle$.

(iv) Open balls. Open balls are open, so $\mathcal{U} \subset \mathcal{O} \subset \langle \mathcal{O} \rangle$. Conversely, let $U \subset \mathbb{R}^d$ be open; we claim U is a countable union of open balls with rational centre and rational radius. Let

$$\mathcal{U}_{\mathbb{Q}} := \{B(q, s) : q \in \mathbb{Q}^d,\ s \in \mathbb{Q},\ s > 0\},$$

a countable family (it is indexed by the countable set $\mathbb{Q}^d \times (\mathbb{Q} \cap (0, \infty))$). Given $x \in U$, openness supplies $r > 0$ with $B(x, r) \subset U$. Choose a rational point $q \in \mathbb{Q}^d$ with $|q - x| < r/4$ (possible since $\mathbb{Q}^d$ is dense in $\mathbb{R}^d$) and a rational s with $r/4 < s < r/2$. Then $|x - q| < r/4 < s$, so $x \in B(q, s)$; and for any $y \in B(q, s)$,

$$|y - x| \leq |y - q| + |q - x| < s + \tfrac{r}{4} < \tfrac{r}{2} + \tfrac{r}{4} < r,$$

so $B(q, s) \subset B(x, r) \subset U$. Therefore U is the union of those members of $\mathcal{U}_{\mathbb{Q}}$ that are contained in U; this is an at most countable union of open balls, so $U \in \langle \mathcal{U} \rangle$. Hence $\mathcal{O} \subset \langle \mathcal{U} \rangle$ and $\langle \mathcal{U} \rangle = \langle \mathcal{O} \rangle$.

(v) Boxes. First we show every box lies in $\langle \mathcal{O} \rangle$. Let $B = I_1 \times \cdots \times I_d$ with intervals I_r as in Definition 1.1.1, and let $\pi_r : \mathbb{R}^d \to \mathbb{R}$ be the r-th coordinate map, so that

$$B = \bigcap_{r=1}^{d} \pi_r^{-1}(I_r),$$

a finite intersection. Each interval I_r with endpoints $a \leq b$ satisfies $I_r = (a, b) \cup P_r$ where $P_r \subset \{a, b\}$ consists of whichever endpoints are included (and $I_r = \emptyset$ or $\{a\}$ in the degenerate case $a = b$). Hence

$$\pi_r^{-1}(I_r) = \pi_r^{-1}\big((a, b)\big) \cup \bigcup_{c \in P_r} \pi_r^{-1}(\{c\}),$$

where $\pi_r^{-1}((a,b)) = \{x : a < x_r < b\}$ is open and each $\pi_r^{-1}(\{c\}) = \{x : x_r = c\}$ is a closed hyperplane. Since $\langle\mathcal{O}\rangle$ contains the open sets and (by (ii)) the closed sets, and is closed under finite unions and finite intersections, we get $\pi_r^{-1}(I_r) \in \langle\mathcal{O}\rangle$ for each r and hence $B \in \langle\mathcal{O}\rangle$. Thus $\mathcal{X} \subset \langle\mathcal{O}\rangle$.
Conversely, let $U \subset \mathbb{R}^d$ be open. Repeating the argument of (iv) with cubes in place of balls: let

$$\mathcal{X}_{\mathbb{Q}} := \Big\{ \textstyle\prod_{r=1}^{d}(a_r, b_r) : a_r, b_r \in \mathbb{Q},\ a_r < b_r \Big\},$$

a countable family of open boxes. Given $x \in U$, choose $r > 0$ with $B(x,r) \subset U$, and then rationals $a_j < x_j < b_j$ with $b_j - a_j < r/\sqrt{d}$ for each j; setting $Q := \prod_j (a_j, b_j)$ we have $x \in Q$, and every $y \in Q$ satisfies $|y - x| = (\sum_j |y_j - x_j|^2)^{1/2} < (d \cdot r^2/d)^{1/2} = r$, so $Q \subset B(x,r) \subset U$. Hence U is the union of those members of $\mathcal{X}_{\mathbb{Q}}$ contained in U, an at most countable union of boxes, so $U \in \langle\mathcal{X}\rangle$. Thus $\mathcal{O} \subset \langle\mathcal{X}\rangle$, and $\langle\mathcal{X}\rangle = \langle\mathcal{O}\rangle$.
(vi) Elementary sets. Every box is an elementary set (a union of one box), so $\mathcal{X} \subset \mathcal{E} \subset \langle\mathcal{E}\rangle$, whence $\langle\mathcal{X}\rangle \subset \langle\mathcal{E}\rangle$. Conversely, every elementary set is by Definition 1.1.1 a finite union of boxes, and $\langle\mathcal{X}\rangle$ is closed under finite unions, so $\mathcal{E} \subset \langle\mathcal{X}\rangle$ and $\langle\mathcal{E}\rangle \subset \langle\mathcal{X}\rangle$. Hence $\langle\mathcal{E}\rangle = \langle\mathcal{X}\rangle = \langle\mathcal{O}\rangle$ by (v).
Collecting the six cases,

$$\langle\mathcal{O}\rangle = \langle\mathcal{C}\rangle = \langle\mathcal{K}\rangle = \langle\mathcal{U}\rangle = \langle\mathcal{X}\rangle = \langle\mathcal{E}\rangle = \mathcal{B}[\mathbb{R}^d],$$

as required.

Problem 1.4.15. (Recursive description of a generated σ-algebra). (This exercise requires familiarity with the theory of ordinals, which is reviewed in Section 2.4 of *An epsilon of room*, Vol. I. Recall that we are assuming the axiom of choice throughout this text.) Let $\mathcal{F}$ be a collection of sets in a set X, and let ω_1 be the first uncountable ordinal. Define the sets $\mathcal{F}_\alpha$ for every countable ordinal $\alpha \in \omega_1$ via transfinite induction as follows:

- (i) $\mathcal{F}_0 := \mathcal{F}$.
- (ii) For each countable successor ordinal $\alpha = \beta + 1$, we define $\mathcal{F}_\alpha$ to be the collection of all sets that are either the union of an at most countable number of sets in $\mathcal{F}_\beta$ (including the empty union $\emptyset$), or the complement of such a union.
- (iii) For each countable limit ordinal $\alpha = \sup_{\beta<\alpha} \beta$, we define $\mathcal{F}_\alpha := \bigcup_{\beta<\alpha} \mathcal{F}_\beta$.

Show that $\langle\mathcal{F}\rangle = \bigcup_{\alpha\in\omega_1} \mathcal{F}_\alpha$.

Solution. Write $G := \bigcup_{\alpha\in\omega_1} \mathcal{F}_\alpha$. We prove the two inclusions separately.
Step 1: the hierarchy is increasing. We first check that $\mathcal{F}_\beta \subset \mathcal{F}_{\beta+1}$ for every $\beta < \omega_1$. Indeed, if $E \in \mathcal{F}_\beta$, then E is the union of the at most countable family $\{E\}$ of sets in $\mathcal{F}_\beta$, hence $E \in \mathcal{F}_{\beta+1}$ by clause (ii). Now we show $\mathcal{F}_\beta \subset \mathcal{F}_\alpha$ whenever $\beta \le \alpha < \omega_1$, by transfinite induction on α (with β fixed). The case $\alpha = \beta$ is trivial. If $\alpha = \gamma + 1 > \beta$, then $\beta \le \gamma$, so by the induction hypothesis $\mathcal{F}_\beta \subset \mathcal{F}_\gamma$, and by the previous paragraph $\mathcal{F}_\gamma \subset \mathcal{F}_{\gamma+1} = \mathcal{F}_\alpha$. If $\alpha > \beta$ is a limit ordinal, then $\beta < \alpha$ and clause (iii) gives $\mathcal{F}_\beta \subset \bigcup_{\gamma<\alpha} \mathcal{F}_\gamma = \mathcal{F}_\alpha$ directly. This completes the induction.
Step 2: $G \subset \langle\mathcal{F}\rangle$. We show by transfinite induction that $\mathcal{F}_\alpha \subset \langle\mathcal{F}\rangle$ for every $\alpha \in \omega_1$. For $\alpha = 0$ this is the containment $\mathcal{F} \subset \langle\mathcal{F}\rangle$ from Definition 1.4.14. If $\alpha = \beta + 1$ and $\mathcal{F}_\beta \subset \langle\mathcal{F}\rangle$, then any at most countable union of sets in $\mathcal{F}_\beta$ is an at most countable union of sets in the σ-algebra $\langle\mathcal{F}\rangle$, hence lies in $\langle\mathcal{F}\rangle$ (padding a finite or empty union out to a countable one with copies of $\emptyset$, which lies in $\langle\mathcal{F}\rangle$ by axiom (i) of Definition 1.4.12); and the complement of such a set again lies in $\langle\mathcal{F}\rangle$ by the complement axiom. Thus $\mathcal{F}_\alpha \subset \langle\mathcal{F}\rangle$. If α is a limit ordinal and $\mathcal{F}_\beta \subset \langle\mathcal{F}\rangle$ for all $\beta < \alpha$, then $\mathcal{F}_\alpha = \bigcup_{\beta<\alpha} \mathcal{F}_\beta \subset \langle\mathcal{F}\rangle$. Taking the union over all $\alpha \in \omega_1$ gives $G \subset \langle\mathcal{F}\rangle$.
Step 3: G is a σ-algebra containing $\mathcal{F}$. Clearly $\mathcal{F} = \mathcal{F}_0 \subset G$. The empty set is the empty union of sets of $\mathcal{F}_0$, so $\emptyset \in \mathcal{F}_1 \subset G$.
Complements: let $E \in G$, say $E \in \mathcal{F}_\alpha$ with $\alpha < \omega_1$. Then E is the union of the at most countable family $\{E\} \subset \mathcal{F}_\alpha$, so E^c is the complement of an at most countable union of sets in $\mathcal{F}_\alpha$, whence $E^c \in \mathcal{F}_{\alpha+1}$. Since ω_1 is a limit ordinal, $\alpha + 1 < \omega_1$, so $E^c \in G$.
Countable unions: let $E_1, E_2, \ldots \in G$, and for each n let α_n be the *least* ordinal $< \omega_1$ with $E_n \in \mathcal{F}_{\alpha_n}$

(such an ordinal exists since $E_n \in G$, and it is well defined without any appeal to choice, since the ordinals are well-ordered). Put $\alpha := \sup_n \alpha_n$. Identifying each ordinal with the set of its predecessors, $\alpha = \bigcup_n \alpha_n$, and each α_n is an at most countable set; a countable union of at most countable sets is at most countable (this does use the axiom of countable choice, which we are assuming), so α is a countable ordinal, i.e. $\alpha < \omega_1$. By Step 1, $E_n \in \mathcal{F}_\alpha$ for every n, so $\bigcup_{n=1}^{\infty} E_n$ is an at most countable union of sets in $\mathcal{F}_\alpha$ and therefore lies in $\mathcal{F}_{\alpha+1} \subset G$ (again $\alpha + 1 < \omega_1$).
Thus G is a σ-algebra containing $\mathcal{F}$, and since $\langle \mathcal{F} \rangle$ is the coarsest such σ-algebra, $\langle \mathcal{F} \rangle \subset G$. Combining with Step 2 we conclude $\langle \mathcal{F} \rangle = G = \bigcup_{\alpha \in \omega_1} \mathcal{F}_\alpha$.
The role of ω_1 is exactly the boundedness used above: the reason the process stabilises at the first uncountable ordinal is that any countable sequence of countable ordinals is bounded by a countable ordinal, so no countable sequence of sets can escape the union G.

Problem 1.4.16. (This exercise requires familiarity with the theory of cardinals.) Let $\mathcal{F}$ be an infinite family of subsets of X of cardinality κ (thus κ is an infinite cardinal). Show that $\langle \mathcal{F} \rangle$ has cardinality at most $\kappa^{\aleph_0}$. (Hint: use Exercise 1.4.15.) In particular, show that the Borel σ-algebra $\mathcal{B}[\mathbb{R}^d]$ has cardinality at most $\mathfrak{c} := 2^{\aleph_0}$.
Conclude that there exist Jordan measurable (and hence Lebesgue measurable) subsets of $\mathbb{R}^d$ which are not Borel measurable. (Hint: How many subsets of the Cantor set are there?) Use this to place the Borel σ-algebra on the diagram that you drew for Exercise 1.4.8.

Solution. The cardinality bound. Write $\lambda := \kappa^{\aleph_0}$. Two standard facts about λ will be used repeatedly:

$$\lambda^{\aleph_0} = (\kappa^{\aleph_0})^{\aleph_0} = \kappa^{\aleph_0 \cdot \aleph_0} = \kappa^{\aleph_0} = \lambda, \qquad \lambda \geq 2^{\aleph_0} \geq \aleph_1,$$

the second because $\kappa \geq 2$ and because $\aleph_1$ is the least cardinal above $\aleph_0$ while $2^{\aleph_0} > \aleph_0$ by Cantor's theorem.
Let $\mathcal{F}_\alpha$, $\alpha \in \omega_1$, be the transfinite hierarchy of Exercise 1.4.15. We claim $|\mathcal{F}_\alpha| \leq \lambda$ for every countable ordinal α, by transfinite induction.
For $\alpha = 0$ we have $|\mathcal{F}_0| = \kappa \leq \kappa^{\aleph_0} = \lambda$.
Suppose $\alpha = \beta + 1$ and $|\mathcal{F}_\beta| \leq \lambda$. Every at most countable subfamily of $\mathcal{F}_\beta$ is either empty or the range of some function $\omega \to \mathcal{F}_\beta$, so the number of at most countable subfamilies is at most $1 + |\mathcal{F}_\beta|^{\aleph_0} \leq 1 + \lambda^{\aleph_0} = \lambda$. Each such subfamily gives one union, and $\mathcal{F}_{\beta+1}$ consists of these unions together with their complements, so $|\mathcal{F}_{\beta+1}| \leq 2\lambda = \lambda$.
Suppose α is a limit ordinal and $|\mathcal{F}_\beta| \leq \lambda$ for all $\beta < \alpha$. Since $\alpha < \omega_1$, the index set $\{\beta : \beta < \alpha\}$ is at most countable, so $|\mathcal{F}_\alpha| \leq \aleph_0 \cdot \lambda = \lambda$.
Finally, by Exercise 1.4.15,

$$|\langle \mathcal{F} \rangle| = \Big| \bigcup_{\alpha \in \omega_1} \mathcal{F}_\alpha \Big| \leq \aleph_1 \cdot \lambda = \lambda = \kappa^{\aleph_0},$$

using $\aleph_1 \leq \lambda$.
The Borel algebra. By Exercise 1.4.14(v), $\mathcal{B}[\mathbb{R}^d]$ is generated by the family $\mathcal{F}$ of boxes in $\mathbb{R}^d$. A box is determined by its $2d$ endpoints together with a choice of open/closed at each end, so $|\mathcal{F}| \leq \mathfrak{c}^{2d} \cdot 4^d = \mathfrak{c}$, and clearly $|\mathcal{F}| \geq \mathfrak{c}$; thus $\kappa = \mathfrak{c}$. Hence

$$|\mathcal{B}[\mathbb{R}^d]| \leq \mathfrak{c}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0} = \mathfrak{c}.$$

(Conversely $\mathcal{B}[\mathbb{R}^d]$ contains all $\mathfrak{c}$ singletons, so its cardinality is exactly $\mathfrak{c}$.)
Jordan measurable sets that are not Borel. Let $C \subset \mathbb{R}$ be the middle-thirds Cantor set and put $K := C \times \{0\}^{d-1} \subset \mathbb{R}^d$ (so $K = C$ when $d = 1$). For each n, C is covered by 2^n closed intervals $I_{n,1}, \ldots, I_{n,2^n}$ of length 3^{-n}, so for every $\varepsilon > 0$ the set K is covered by the elementary set $\bigcup_i I_{n,i} \times [-\varepsilon, \varepsilon]^{d-1}$, of elementary measure $(2/3)^n (2\varepsilon)^{d-1}$. Letting $n \to \infty$ (with $\varepsilon = 1$ say, when $d \geq 2$; when $d = 1$ the factor is absent) we see that the Jordan outer measure of K is 0. Consequently every subset $A \subset K$ has Jordan outer measure 0 by monotonicity, and since the Jordan inner measure is squeezed between 0 and the outer measure, A is Jordan measurable with Jordan measure 0. In particular every subset of K is Lebesgue measurable.

Now $|K| = |C| = \mathfrak{c}$ (the Cantor set is in bijection with $\{0,1\}^{\mathbb{N}}$ via ternary expansions using only the digits 0 and 2), so K has $2^{\mathfrak{c}}$ subsets, all of them Jordan measurable. By Cantor's theorem $2^{\mathfrak{c}} > \mathfrak{c} = |\mathcal{B}[\mathbb{R}^d]|$, so the collection of subsets of K cannot be contained in $\mathcal{B}[\mathbb{R}^d]$. Hence there exists $A \subset K$ which is Jordan measurable (and therefore Lebesgue measurable) but not Borel measurable. In fact "most" subsets of K are of this kind, since only at most $\mathfrak{c}$ of the $2^{\mathfrak{c}}$ of them are Borel.

Placement on the diagram. (The diagram in question is the one requested in Example 1.4.8, comparing the dyadic algebras with the elementary, Jordan, Lebesgue, null, discrete and trivial algebras.) Every box is Borel and $\mathcal{B}[\mathbb{R}^d]$ is a σ-algebra, so every elementary set is Borel, and a co-elementary set is Borel as well; hence the elementary algebra is coarser than $\mathcal{B}[\mathbb{R}^d]$. Each dyadic algebra $\mathcal{D}_n(\mathbb{R}^d)$ is also coarser than $\mathcal{B}[\mathbb{R}^d]$, but this does *not* follow from the previous sentence, since a dyadic algebra contains infinite unions of dyadic cubes and is therefore not contained in the elementary algebra; instead one argues directly: there are only countably many dyadic cubes at scale 2^{-n}, so every element of $\mathcal{D}_n(\mathbb{R}^d)$ is an at most countable union of boxes, hence Borel. Both containments are strict: $\mathbb{Q}^d \cap [0,1]^d$ is a countable set, hence a countable union of singletons and so Borel, yet it is neither elementary nor co-elementary, and it is not a union of dyadic cubes at any fixed scale. Conversely every open set is Lebesgue measurable, so $\mathcal{B}[\mathbb{R}^d]$ is coarser than the Lebesgue σ-algebra, and the previous paragraph shows this containment is strict as well. Thus in the diagram the Borel algebra sits strictly above the elementary and dyadic algebras, strictly below the Lebesgue algebra, and strictly below the discrete algebra (strictly, since $|\mathcal{B}[\mathbb{R}^d]| = \mathfrak{c} < 2^{\mathfrak{c}} = |2^{\mathbb{R}^d}|$), and strictly above the trivial algebra.

The Borel algebra is however not comparable with the Jordan algebra or with the null algebra. The set A above is Jordan measurable but not Borel, while $\mathbb{Q}^d \cap [0,1]^d$ is Borel (it is countable, hence a countable union of singletons) but does not lie in the Jordan algebra: it has Jordan inner measure 0 and Jordan outer measure 1, so it is not Jordan measurable, and its complement is not Jordan measurable either (a Jordan measurable set is by definition bounded, and $\mathbb{R}^d \setminus (\mathbb{Q}^d \cap [0,1]^d)$ is unbounded), so it is not co-Jordan measurable. Likewise A is a null set (so it lies in the null algebra) but is not Borel, while $[0,1]^d$ is Borel and is neither null nor co-null.

Problem 1.4.17. Let E, F be Borel measurable subsets of $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}$ respectively. Show that $E \times F$ is a Borel measurable subset of $\mathbb{R}^{d_1+d_2}$. (Hint: first establish this in the case when F is a box, by using Remark 1.4.15. To obtain the general case, apply Remark 1.4.15 yet again.)

Solution. Throughout we use two elementary closure facts about the σ-algebra $\mathcal{B}[\mathbb{R}^{d_1+d_2}]$: it is closed under finite intersections and hence under differences, since $A \cap B = (A^c \cup B^c)^c$ and $A \setminus B = A \cap B^c$ (finite unions being countable unions padded with $\emptyset$).

Step 0: two auxiliary Borel sets. If $U \subset \mathbb{R}^{d_1}$ is open then $U \times \mathbb{R}^{d_2}$ is open in $\mathbb{R}^{d_1+d_2}$, hence Borel. If $B \subset \mathbb{R}^{d_2}$ is a box, then

$$\mathbb{R}^{d_1} \times B = \bigcup_{n=1}^{\infty} \big([-n,n]^{d_1} \times B\big),$$

a countable union of boxes in $\mathbb{R}^{d_1+d_2}$, hence Borel (boxes are Borel by Exercise 1.4.14(v), or directly since a box is an intersection of an open box with finitely many closed half-spaces; in any case each $[-n,n]^{d_1} \times B$ is a box in $\mathbb{R}^{d_1+d_2}$).

Step 1: E Borel, $F = B$ a box. Fix a box $B \subset \mathbb{R}^{d_2}$ and let $P(E)$ be the property that $E \times B$ is a Borel subset of $\mathbb{R}^{d_1+d_2}$, for $E \subset \mathbb{R}^{d_1}$. We verify the four hypotheses of Remark 1.4.15, taking $\mathcal{F}$ to be the collection of open subsets of $\mathbb{R}^{d_1}$ (so that $\langle \mathcal{F} \rangle = \mathcal{B}[\mathbb{R}^{d_1}]$ by Definition 1.4.16).

- $P(\emptyset)$ holds, since $\emptyset \times B = \emptyset$.
- If U is open, then $U \times B = (U \times \mathbb{R}^{d_2}) \cap (\mathbb{R}^{d_1} \times B)$ is an intersection of two Borel sets by Step 0, hence Borel; so $P(U)$ holds.
- If $P(E)$ holds, then $(\mathbb{R}^{d_1} \setminus E) \times B = (\mathbb{R}^{d_1} \times B) \setminus (E \times B)$ is a difference of Borel sets, hence Borel; so $P(\mathbb{R}^{d_1} \setminus E)$ holds.
- If $P(E_n)$ holds for all n, then $\big(\bigcup_n E_n\big) \times B = \bigcup_n (E_n \times B)$ is a countable union of Borel sets, hence Borel.

By Remark 1.4.15, $P(E)$ holds for every $E \in \mathcal{B}[\mathbb{R}^{d_1}]$: that is, $E \times B$ is Borel for every Borel E and

every box B.
Step 2: E Borel, F Borel. Fix a Borel set $E \subset \mathbb{R}^{d_1}$ and let $Q(F)$ be the property that $E \times F$ is Borel in $\mathbb{R}^{d_1+d_2}$, for $F \subset \mathbb{R}^{d_2}$. This time apply Remark 1.4.15 with $\mathcal{F}$ the collection of boxes in $\mathbb{R}^{d_2}$, which generates $\mathcal{B}[\mathbb{R}^{d_2}]$ by Exercise 1.4.14(v).

- $Q(\emptyset)$ holds, since $E \times \emptyset = \emptyset$.
- $Q(B)$ holds for every box B, by Step 1.
- Before checking complements, note that $E \times \mathbb{R}^{d_2} = \bigcup_{n=1}^{\infty}(E \times [-n,n]^{d_2})$ is Borel, each term being Borel by Step 1. Hence if $Q(F)$ holds, then $E \times (\mathbb{R}^{d_2} \setminus F) = (E \times \mathbb{R}^{d_2}) \setminus (E \times F)$ is a difference of Borel sets, so $Q(\mathbb{R}^{d_2} \setminus F)$ holds.
- If $Q(F_n)$ holds for all n, then $E \times \bigcup_n F_n = \bigcup_n (E \times F_n)$ is Borel.

By Remark 1.4.15, $Q(F)$ holds for all $F \in \langle \mathcal{F} \rangle = \mathcal{B}[\mathbb{R}^{d_2}]$. Since E was an arbitrary Borel subset of $\mathbb{R}^{d_1}$, this proves that $E \times F$ is Borel in $\mathbb{R}^{d_1+d_2}$ whenever E and F are Borel.

Problem 1.4.18. Let E be a Borel measurable subset of $\mathbb{R}^{d_1+d_2}$.

- (i) Show that for any $x_1 \in \mathbb{R}^{d_1}$, the slice $\{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\}$ is a Borel measurable subset of $\mathbb{R}^{d_2}$. Similarly, show that for every $x_2 \in \mathbb{R}^{d_2}$, the slice $\{x_1 \in \mathbb{R}^{d_1} : (x_1, x_2) \in E\}$ is a Borel measurable subset of $\mathbb{R}^{d_1}$.
- (ii) Give a counterexample to show that this claim is not true if "Borel" is replaced with "Lebesgue" throughout. (Hint: the Cartesian product of any set with a point is a null set, even if the first set was not measurable.)

Solution. (i) Fix $x_1 \in \mathbb{R}^{d_1}$ and for $E \subset \mathbb{R}^{d_1+d_2}$ write

$$E_{x_1} := \{x_2 \in \mathbb{R}^{d_2} : (x_1, x_2) \in E\},$$

which is exactly $\iota^{-1}(E)$ for the map $\iota : \mathbb{R}^{d_2} \to \mathbb{R}^{d_1+d_2}$, $\iota(x_2) := (x_1, x_2)$. Since taking preimages commutes with all set operations, the slicing operation satisfies

$$\emptyset_{x_1} = \emptyset, \qquad (E^c)_{x_1} = (E_{x_1})^c, \qquad \Big(\bigcup_{n=1}^{\infty} E_n\Big)_{x_1} = \bigcup_{n=1}^{\infty} (E_n)_{x_1}.$$

Let $P(E)$ be the property that E_{x_1} is a Borel subset of $\mathbb{R}^{d_2}$, and apply Remark 1.4.15 with $\mathcal{F}$ the collection of open subsets of $\mathbb{R}^{d_1+d_2}$, so that $\langle \mathcal{F} \rangle = \mathcal{B}[\mathbb{R}^{d_1+d_2}]$.

- $P(\emptyset)$ holds, as $\emptyset_{x_1} = \emptyset$ is Borel.
- If U is open, then $U_{x_1} = \iota^{-1}(U)$ is open, because ι is continuous (it is an isometry of $\mathbb{R}^{d_2}$ onto the affine slice $\{x_1\} \times \mathbb{R}^{d_2}$); concretely, if $(x_1, x_2) \in U$ then some ball $B((x_1, x_2), r)$ lies in U, whence $B(x_2, r) \subset U_{x_1}$. So $P(U)$ holds.
- If $P(E)$ holds, then $(E^c)_{x_1} = (E_{x_1})^c$ is Borel, so $P(E^c)$ holds.
- If $P(E_n)$ holds for all n, then $\big(\bigcup_n E_n\big)_{x_1} = \bigcup_n (E_n)_{x_1}$ is a countable union of Borel sets, hence Borel.

By Remark 1.4.15, $P(E)$ holds for every Borel E, which is the first claim. The second claim is identical after exchanging the roles of the two factors (fix x_2 and use the continuous map $x_1 \mapsto (x_1, x_2)$).
(ii) The analogous statement for Lebesgue measurability fails. By Proposition 1.2.18 there is a non-Lebesgue-measurable subset of $\mathbb{R}$, and the same argument run with the subgroup $\mathbb{Q}^{d_1}$ of $\mathbb{R}^{d_1}$ in place of $\mathbb{Q}$ produces a non-Lebesgue-measurable set $A \subset [0,1]^{d_1}$ for any $d_1 \geq 1$. In detail: using the axiom of choice, pick one representative $x_S \in S \cap [0,1]^{d_1}$ from each coset $S \in \mathbb{R}^{d_1}/\mathbb{Q}^{d_1}$ (each coset is dense, so meets $[0,1]^{d_1}$), and let A be the set of these representatives. The countably many translates $A+q$ with $q \in \mathbb{Q}^{d_1} \cap [-1,1]^{d_1}$ are pairwise disjoint (as A meets each coset once), cover $[0,1]^{d_1}$, and are contained in $[-1,2]^{d_1}$. Were A measurable, countable additivity and translation invariance (Exercise 1.2.20) would give $1 \leq \sum_q m(A) \leq 3^{d_1}$, which is absurd, since that sum is 0 if $m(A) = 0$ and $+\infty$ if $m(A) > 0$.
Now set

$$E := A \times \{0\}^{d_2} \subset \mathbb{R}^{d_1+d_2}.$$

Then E is contained in the hyperplane-like set $H := \mathbb{R}^{d_1} \times \{0\}^{d_2}$, and H is a Lebesgue null set: indeed $H = \bigcup_{n=1}^{\infty} \big([-n,n]^{d_1} \times \{0\}^{d_2}\big)$, and for each n and each $\varepsilon > 0$ the set $[-n,n]^{d_1} \times \{0\}^{d_2}$ is contained in

the box $[-n,n]^{d_1} \times [-\varepsilon,\varepsilon]^{d_2}$ of volume $(2n)^{d_1}(2\varepsilon)^{d_2}$, so it has Lebesgue outer measure 0; a countable union of null sets is null. Hence E has Lebesgue outer measure 0, so E is Lebesgue measurable (every null set is Lebesgue measurable) with $m(E) = 0$.
On the other hand, the slice of E at $x_2 = 0 \in \mathbb{R}^{d_2}$ is

$$\{x_1 \in \mathbb{R}^{d_1} : (x_1, 0) \in E\} = A,$$

which is not Lebesgue measurable. The other slicing direction fails just as badly: taking $B \subset [0,1]^{d_2}$ non-Lebesgue-measurable by the same construction and setting $E' := \{0\}^{d_1} \times B$, the set E' is again contained in a null hyperplane $\{0\}^{d_1} \times \mathbb{R}^{d_2}$, hence null and Lebesgue measurable, while its slice $\{x_2 \in \mathbb{R}^{d_2} : (0, x_2) \in E'\} = B$ is not. So a Lebesgue measurable subset of $\mathbb{R}^{d_1+d_2}$ can have non-Lebesgue-measurable slices in either variable, and part (i) genuinely uses the Borel structure. (The point is that the Lebesgue σ-algebra is obtained from the Borel one by throwing in all null sets, and a null set in the product can have wild slices.)

Problem 1.4.19. Show that the Lebesgue σ-algebra on $\mathbb{R}^d$ is generated by the union of the Borel σ-algebra and the null σ-algebra.

Solution. Write $\mathcal{L}[\mathbb{R}^d]$ for the Lebesgue σ-algebra, $\mathcal{B}[\mathbb{R}^d]$ for the Borel σ-algebra, and $\mathcal{N}(\mathbb{R}^d)$ for the null algebra, i.e. the collection of subsets of $\mathbb{R}^d$ that are either Lebesgue null or Lebesgue co-null (Example 1.4.6). Both $\mathcal{L}[\mathbb{R}^d]$ and $\mathcal{N}(\mathbb{R}^d)$ are σ-algebras by Exercise 1.4.11. We must show

$$\langle \mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d) \rangle = \mathcal{L}[\mathbb{R}^d].$$

The inclusion $\subset$. Every open set is Lebesgue measurable, so $\mathcal{L}[\mathbb{R}^d]$ is a σ-algebra containing the open sets, and hence $\mathcal{B}[\mathbb{R}^d] \subset \mathcal{L}[\mathbb{R}^d]$ by the minimality in Definition 1.4.14. Also every null set is Lebesgue measurable (a set of Lebesgue outer measure zero is Lebesgue measurable, since it is contained in an open set of arbitrarily small measure), and complements of Lebesgue measurable sets are Lebesgue measurable, so $\mathcal{N}(\mathbb{R}^d) \subset \mathcal{L}[\mathbb{R}^d]$. Thus $\mathcal{L}[\mathbb{R}^d]$ is a σ-algebra containing $\mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d)$, and since $\langle \mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d) \rangle$ is the coarsest such σ-algebra, the inclusion $\subset$ follows.
The inclusion $\supset$. Let $E \subset \mathbb{R}^d$ be Lebesgue measurable. By Exercise 1.2.19 (the equivalence (i) $\Leftrightarrow$ (iii)), we may write

$$E = F \cup N,$$

where F is an F_σ set and N is a null set. An F_σ set is a countable union of closed sets, and closed sets are complements of open sets, so $F \in \mathcal{B}[\mathbb{R}^d]$. The null set N lies in $\mathcal{N}(\mathbb{R}^d)$ by definition. Hence both F and N lie in $\mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d) \subset \langle \mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d) \rangle$, and since a σ-algebra is closed under finite unions (pad the union $F \cup N$ out to a countable one with copies of $\emptyset$), we conclude $E \in \langle \mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d) \rangle$.
Combining the two inclusions gives $\mathcal{L}[\mathbb{R}^d] = \langle \mathcal{B}[\mathbb{R}^d] \cup \mathcal{N}(\mathbb{R}^d) \rangle$, as required. (Equivalently: the Lebesgue σ-algebra is the completion of the Borel σ-algebra with respect to Lebesgue measure; one could also use the alternative description in Exercise 1.2.19(ii), writing $E = G \setminus N$ with G a G_δ set and N null, and noting that $G \setminus N = G \cap N^c$ with N^c co-null, hence in $\mathcal{N}(\mathbb{R}^d)$.)

Problem 1.4.20. Let $\mu : \mathcal{B} \to [0, +\infty]$ be a finitely additive measure on a Boolean σ-algebra $\mathcal{B}$. Establish the following facts:

- (i) (Monotonicity) If E, F are $\mathcal{B}$-measurable and $E \subset F$, then $\mu(E) \leq \mu(F)$.
- (ii) (Finite additivity) If k is a natural number, and $E_1, \ldots, E_k$ are $\mathcal{B}$-measurable and disjoint, then $\mu(E_1 \cup \ldots \cup E_k) = \mu(E_1) + \ldots + \mu(E_k)$.
- (iii) (Finite subadditivity) If k is a natural number, and $E_1, \ldots, E_k$ are $\mathcal{B}$-measurable, then $\mu(E_1 \cup \ldots \cup E_k) \leq \mu(E_1) + \ldots + \mu(E_k)$.
- (iv) (Inclusion-exclusion for two sets) If E, F are $\mathcal{B}$-measurable, then $\mu(E \cup F) + \mu(E \cap F) = \mu(E) + \mu(F)$.

(Caution: remember that the cancellation law $a + c = b + c \implies a = b$ does not hold in $[0, +\infty]$ if c

is infinite, and so the use of cancellation (or subtraction) should be avoided if possible.)

Solution. Throughout we use that a Boolean algebra is closed under intersections and set differences (Remark 1.4.1): $E \cap F = (E^c \cup F^c)^c$ and $F \setminus E = F \cap E^c$. No subtraction or cancellation in $[0, +\infty]$ will be used anywhere below.
(i) Monotonicity. Let $E \subset F$ be $\mathcal{B}$-measurable. Then $F \setminus E$ is $\mathcal{B}$-measurable, and $F = E \cup (F \setminus E)$ is a disjoint union. By finite additivity (axiom (ii) of Definition 1.4.19),

$$\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E),$$

since $\mu(F \setminus E) \in [0, +\infty]$ is non-negative and addition in $[0, +\infty]$ is monotone.
(ii) Finite additivity. We induct on k. For $k = 0$ the union is empty and the sum is empty, and $\mu(\emptyset) = 0$ by axiom (i) of Definition 1.4.19. For $k = 1$ the claim is trivial. Suppose the claim holds for some $k \geq 1$, and let $E_1, \ldots, E_{k+1}$ be disjoint $\mathcal{B}$-measurable sets. The set $G := E_1 \cup \ldots \cup E_k$ is $\mathcal{B}$-measurable, and $G \cap E_{k+1} = \bigcup_{j \leq k} (E_j \cap E_{k+1}) = \emptyset$ by disjointness. Hence by the two-set additivity axiom and the induction hypothesis,

$$\mu(E_1 \cup \ldots \cup E_{k+1}) = \mu(G) + \mu(E_{k+1}) = \big(\mu(E_1) + \ldots + \mu(E_k)\big) + \mu(E_{k+1}),$$

completing the induction.
(iii) Finite subadditivity. Let $E_1, \ldots, E_k$ be $\mathcal{B}$-measurable, and disjointify them in the standard way: put

$$F_1 := E_1, \qquad F_j := E_j \setminus (E_1 \cup \ldots \cup E_{j-1}) \quad (2 \leq j \leq k).$$

Each F_j is $\mathcal{B}$-measurable, and $F_j \subset E_j$. The F_j are pairwise disjoint: if $i < j$ then $F_i \subset E_i$ while $F_j \cap E_i = \emptyset$ by construction. Moreover $\bigcup_{j \leq k} F_j = \bigcup_{j \leq k} E_j$: the inclusion $\subset$ is clear from $F_j \subset E_j$, and conversely if $x \in \bigcup_{j \leq k} E_j$ then letting j be the least index with $x \in E_j$ we get $x \in F_j$. Therefore, by (ii) and then (i),

$$\mu(E_1 \cup \ldots \cup E_k) = \mu(F_1 \cup \ldots \cup F_k) = \sum_{j=1}^{k} \mu(F_j) \leq \sum_{j=1}^{k} \mu(E_j),$$

where the last step uses $\mu(F_j) \leq \mu(E_j)$ from monotonicity, together with the fact that addition in $[0, +\infty]$ preserves the order in each variable.
(iv) Inclusion-exclusion. Let E, F be $\mathcal{B}$-measurable. The sets E and $F \setminus E$ are disjoint and $\mathcal{B}$-measurable with union $E \cup F$, so

$$\mu(E \cup F) = \mu(E) + \mu(F \setminus E).$$

The sets $E \cap F$ and $F \setminus E$ are disjoint and $\mathcal{B}$-measurable with union F, so

$$\mu(F) = \mu(E \cap F) + \mu(F \setminus E).$$

Adding $\mu(E \cap F)$ to the first identity and using commutativity and associativity of addition in $[0, +\infty]$,

$$\mu(E \cup F) + \mu(E \cap F) = \mu(E) + \big(\mu(F \setminus E) + \mu(E \cap F)\big) = \mu(E) + \mu(F).$$

This is the desired identity; note that we only ever added quantities, so the argument is valid even when some of the measures involved are infinite.

Problem 1.4.21. Let $\mathcal{B}$ be a finite Boolean algebra, generated by a finite family $A_1, \ldots, A_k$ of non-empty atoms. Show that for every finitely additive measure μ on $\mathcal{B}$ there exists $c_1, \ldots, c_k \in [0, +\infty]$ such that

$$\mu(E) = \sum_{1 \leq j \leq k:\, A_j \subset E} c_j.$$

Equivalently, if x_j is a point in A_j for each $1 \leq j \leq k$, then

$$\mu = \sum_{j=1}^{k} c_j \delta_{x_j}.$$

Furthermore, show that the $c_1, \ldots, c_k$ are uniquely determined by μ.

Solution. The hypothesis that $\mathcal{B}$ is generated by the non-empty atoms $A_1, \ldots, A_k$ means precisely that $\mathcal{B} = \mathcal{A}((A_j)_{1 \leq j \leq k})$ is the atomic algebra of Example 1.4.7 attached to a partition $X = A_1 \cup \ldots \cup A_k$ of the ambient space into disjoint non-empty sets; explicitly, $\mathcal{B}$ consists exactly of the sets $\bigcup_{j \in J} A_j$ with $J \subset \{1, \ldots, k\}$. (By Exercise 1.4.4 every finite Boolean algebra is of this form, and by Exercise 1.4.3 its non-empty atoms are determined up to relabelling, so no generality is lost.)
Step 1: every $\mathcal{B}$-measurable set is the union of the atoms it contains. Let $E \in \mathcal{B}$, say $E = \bigcup_{j \in J} A_j$. I claim that

$$J' := \{j : A_j \subset E\} \quad \text{satisfies} \quad E = \bigcup_{j \in J'} A_j.$$

Clearly $J \subset J'$, so $E \subset \bigcup_{j \in J'} A_j$. Conversely, if $j \in J'$ then $A_j \subset E$ by definition, so $\bigcup_{j \in J'} A_j \subset E$. Hence $E = \bigcup_{j \in J'} A_j$, a disjoint union of $\mathcal{B}$-measurable sets. (In fact $J = J'$: if $j \notin J$, then A_j is disjoint from every A_i with $i \in J$, so $A_j \cap E = \emptyset$; since $A_j \neq \emptyset$, this forces $A_j \not\subset E$.)
Step 2: the formula. Let μ be a finitely additive measure on $\mathcal{B}$ and set

$$c_j := \mu(A_j) \in [0, +\infty] \qquad (1 \leq j \leq k).$$

Let $E \in \mathcal{B}$. By Step 1, E is the disjoint union of the finitely many atoms A_j with $A_j \subset E$, so by finite additivity (Exercise 1.4.20(ii)),

$$\mu(E) = \sum_{1 \leq j \leq k:\, A_j \subset E} \mu(A_j) = \sum_{1 \leq j \leq k:\, A_j \subset E} c_j,$$

which is the asserted formula (with the convention that an empty sum equals 0, consistent with $\mu(\emptyset) = 0$).
Step 3: the reformulation in terms of Dirac measures. Fix $x_j \in A_j$ for each j (possible since each atom is non-empty). I claim that for $E \in \mathcal{B}$,

$$A_j \subset E \iff x_j \in E.$$

If $A_j \subset E$ then $x_j \in A_j \subset E$. Conversely suppose $x_j \in E$. By Step 1, $E = \bigcup_{i:\, A_i \subset E} A_i$, so $x_j \in A_i$ for some i with $A_i \subset E$; since the atoms are disjoint and $x_j \in A_j$, we must have $i = j$, so $A_j \subset E$.
Consequently, recalling $\delta_{x_j}(E) = 1_E(x_j)$ from Example 1.4.22, and using the convention $(+\infty) \cdot 0 = 0$ for scalar multiples of measures (Example 1.4.24),

$$\sum_{j=1}^{k} c_j \delta_{x_j}(E) = \sum_{j=1}^{k} c_j 1_E(x_j) = \sum_{1 \leq j \leq k:\, x_j \in E} c_j = \sum_{1 \leq j \leq k:\, A_j \subset E} c_j = \mu(E),$$

i.e. $\mu = \sum_{j=1}^{k} c_j \delta_{x_j}$. (Conversely, any such combination is a finitely additive measure, by Examples 1.4.22 and 1.4.24, so this describes all finitely additive measures on $\mathcal{B}$.)
Step 4: uniqueness. Suppose $c_1, \ldots, c_k \in [0, +\infty]$ are such that $\mu(E) = \sum_{j:\, A_j \subset E} c_j$ for all $E \in \mathcal{B}$. Fix $1 \leq i \leq k$ and take $E := A_i \in \mathcal{B}$. If $j \neq i$ and $A_j \subset A_i$, then $A_j = A_j \cap A_i = \emptyset$ by disjointness of the atoms, contradicting the assumption that the atoms are non-empty; so the only index j with $A_j \subset A_i$ is $j = i$. Hence the formula gives

$$\mu(A_i) = c_i.$$

Thus the coefficients are forced to be $c_i = \mu(A_i)$, which is exactly the choice made in Step 2; in particular they are uniquely determined by μ. (Note this is where the hypothesis that the atoms are non-empty is essential: an empty atom would have c_j completely undetermined, since δ_{x_j} would not even be defined.)

Problem 1.4.22. (Countable combinations of measures) Let $(X, \mathcal{B})$ be a measurable space.

- (i) If μ is a countably additive measure on $\mathcal{B}$, and $c \in [0, +\infty]$, then $c\mu$ is also countably additive.
- (ii) If $\mu_1, \mu_2, \ldots$ are a sequence of countably additive measures on $\mathcal{B}$, then the sum $\sum_{n=1}^{\infty} \mu_n : E \mapsto \sum_{n=1}^{\infty} \mu_n(E)$ is also a countably additive measure.

Solution. Throughout we use the arithmetic of $[0, +\infty]$, with the convention $0 \cdot \infty = \infty \cdot 0 = 0$, and we recall that a sum $\sum_{n=1}^{\infty} a_n$ of elements $a_n \in [0, +\infty]$ is by definition $\lim_{N\to\infty} \sum_{n=1}^{N} a_n = \sup_N \sum_{n=1}^{N} a_n$, the limit existing in $[0, +\infty]$ because the partial sums are non-decreasing.

Two elementary facts about sums in $[0, +\infty]$.

Fact 1. For any $c \in [0, +\infty]$ and any non-decreasing sequence $s_1 \le s_2 \le \cdots$ in $[0, +\infty]$ we have $c \sup_N s_N = \sup_N (c s_N)$. Indeed, if $c = 0$ both sides are 0 by the convention. If $0 < c < \infty$, this is the usual homogeneity of suprema in $[0, +\infty]$ (multiplication by c is an order-isomorphism of $[0, +\infty]$). If $c = +\infty$, then the left side is $+\infty$ when $\sup_N s_N > 0$ and 0 when all $s_N = 0$; the right side is likewise $+\infty$ when some $s_N > 0$ and 0 otherwise; and $\sup_N s_N > 0$ holds precisely when some $s_N > 0$. Consequently $c \sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} c a_n$ for every $c \in [0, +\infty]$ and every sequence $a_n \in [0, +\infty]$, since finite distributivity $c \sum_{n=1}^{N} a_n = \sum_{n=1}^{N} c a_n$ holds in $[0, +\infty]$ (check the cases where some term is 0 or ∞ directly).

Fact 2 (Tonelli for series). For any doubly indexed family $a_{n,m} \in [0, +\infty]$, $n, m \ge 1$, one has

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m} = S, \qquad S := \sup_{F} \sum_{(n,m)\in F} a_{n,m},$$

the supremum being over all finite subsets F of $\mathbb{N} \times \mathbb{N}$. By symmetry it suffices to prove the first sum equals S. Every finite F is contained in a square $\{1, \ldots, N\}^2$, and $\sum_{(n,m)\in F} a_{n,m} \le \sum_{n=1}^{N} \sum_{m=1}^{N} a_{n,m} \le \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m}$ by monotonicity of sums of non-negative terms; hence $S \le \sum_n \sum_m a_{n,m}$. Conversely, fix N. For each $n \le N$ the partial sums $\sum_{m=1}^{M} a_{n,m}$ increase to $\sum_{m=1}^{\infty} a_{n,m}$ as $M \to \infty$, and a sum of finitely many non-decreasing sequences in $[0, +\infty]$ converges to the sum of their limits (immediate by induction from the two-term case, which is valid in $[0, +\infty]$ including infinite limits). Hence

$$\sum_{n=1}^{N} \sum_{m=1}^{\infty} a_{n,m} = \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} \le S,$$

the inequality holding because each $\sum_{n\le N} \sum_{m\le M} a_{n,m}$ is a sum over the finite set $\{1, \ldots, N\} \times \{1, \ldots, M\}$. Letting $N \to \infty$ gives $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} \le S$, as required.

(i). Let $\nu := c\mu$, i.e. $\nu(E) := c \cdot \mu(E)$. Then $\nu(\emptyset) = c \cdot \mu(\emptyset) = c \cdot 0 = 0$, so the empty set axiom of Definition 1.4.27 holds. If $E_1, E_2, \ldots \in \mathcal{B}$ are disjoint, then by countable additivity of μ and Fact 1,

$$\nu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = c\,\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = c \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} c\,\mu(E_n) = \sum_{n=1}^{\infty} \nu(E_n).$$

Thus $c\mu$ is countably additive.

(ii). Let $\nu(E) := \sum_{n=1}^{\infty} \mu_n(E)$, which is a well-defined element of $[0, +\infty]$ for each $E \in \mathcal{B}$. Since $\mu_n(\emptyset) = 0$ for every n, we get $\nu(\emptyset) = 0$. Now let $E_1, E_2, \ldots \in \mathcal{B}$ be disjoint and put $E := \bigcup_{m=1}^{\infty} E_m$. Applying countable additivity of each μ_n and then Fact 2 to the array $a_{n,m} := \mu_n(E_m)$,

$$\nu(E) = \sum_{n=1}^{\infty} \mu_n(E) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu_n(E_m) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(E_m) = \sum_{m=1}^{\infty} \nu(E_m).$$

Hence $\sum_{n=1}^{\infty} \mu_n$ is a countably additive measure on $\mathcal{B}$.

Problem 1.4.23. Let $(X, \mathcal{B}, \mu)$ be a measure space.

- (i) (Countable subadditivity) If $E_1, E_2, \ldots$ are $\mathcal{B}$-measurable, then $\mu(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu(E_n)$.
- (ii) (Upwards monotone convergence) If $E_1 \subset E_2 \subset \ldots$ are $\mathcal{B}$-measurable, then
$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \mu(E_n) = \sup_n \mu(E_n).$$
- (iii) (Downwards monotone convergence) If $E_1 \supset E_2 \supset \ldots$ are $\mathcal{B}$-measurable, and $\mu(E_n) < \infty$ for at least one n, then
$$\mu\Big(\bigcap_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \mu(E_n) = \inf_n \mu(E_n).$$

Show that the downward monotone convergence claim can fail if the hypothesis that $\mu(E_n) < \infty$ for at least one n is dropped. (Hint: mimic the solution to Exercise 1.2.11.)

Solution. We freely use that a countably additive measure is finitely additive (pad a finite disjoint union out to a countable one with copies of $\emptyset$), and hence, by Exercise 1.4.20, monotone and finitely subadditive.

(i). Disjointify: set
$$F_n := E_n \setminus (E_1 \cup \cdots \cup E_{n-1}) \qquad (n \geq 1),$$
with the convention that $F_1 = E_1$. Each F_n lies in $\mathcal{B}$, since $\mathcal{B}$ is closed under finite unions and complements (hence under set difference). The F_n are pairwise disjoint: if $m < n$ then F_n is disjoint from $E_m \supset F_m$. Moreover $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n$: the inclusion $\subset$ is clear as $F_n \subset E_n$; conversely if $x \in \bigcup_n E_n$, let n be the least index with $x \in E_n$, and then $x \in F_n$. Hence, by countable additivity and then monotonicity ($F_n \subset E_n$),
$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} \mu(F_n) \leq \sum_{n=1}^{\infty} \mu(E_n).$$

(ii). Now the E_n are increasing, so with $E_0 := \emptyset$ the sets $F_n := E_n \setminus E_{n-1}$ are $\mathcal{B}$-measurable, pairwise disjoint, satisfy $\bigcup_{n=1}^{N} F_n = E_N$ for each N, and $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n$ (same argument as in (i)). By finite additivity, $\mu(E_N) = \sum_{n=1}^{N} \mu(F_n)$, and by countable additivity,
$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} \mu(F_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(F_n) = \lim_{N\to\infty} \mu(E_N).$$
Finally, by monotonicity the sequence $\mu(E_n)$ is non-decreasing, so its limit equals its supremum.

(iii). Suppose $\mu(E_{n_0}) < \infty$ for some n_0. Since $E_1 \supset E_2 \supset \cdots$, the sequence $\mu(E_n)$ is non-increasing (monotonicity), so both $\lim_n \mu(E_n)$ and $\inf_n \mu(E_n)$ are unchanged if we discard the first $n_0 - 1$ terms; the intersection $\bigcap_n E_n$ is also unchanged. So we may assume $n_0 = 1$, i.e. $\mu(E_1) < \infty$; then $\mu(E_n) \leq \mu(E_1) < \infty$ for all n, and also $\mu(\bigcap_n E_n) \leq \mu(E_1) < \infty$.

The sets $E_1 \setminus E_n$ are $\mathcal{B}$-measurable and increase in n, with
$$\bigcup_{n=1}^{\infty} (E_1 \setminus E_n) = E_1 \setminus \bigcap_{n=1}^{\infty} E_n$$
by De Morgan. By part (ii),
$$\mu\Big(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \mu(E_1 \setminus E_n).$$
Since $E_n \subset E_1$, finite additivity applied to the disjoint decomposition $E_1 = E_n \cup (E_1 \setminus E_n)$ gives $\mu(E_1) = \mu(E_n) + \mu(E_1 \setminus E_n)$; as all quantities here are finite we may subtract to obtain $\mu(E_1 \setminus E_n) = \mu(E_1) - \mu(E_n)$. The same computation with E_n replaced by the (finite-measure, $\mathcal{B}$-measurable) set $\bigcap_n E_n \subset E_1$ gives $\mu(E_1 \setminus \bigcap_n E_n) = \mu(E_1) - \mu(\bigcap_n E_n)$. Substituting both into the displayed identity,
$$\mu(E_1) - \mu\Big(\bigcap_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \big(\mu(E_1) - \mu(E_n)\big) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n),$$

and cancelling the finite quantity $\mu(E_1)$ yields $\mu(\bigcap_n E_n) = \lim_n \mu(E_n)$, which equals $\inf_n \mu(E_n)$ because the sequence is non-increasing.

Failure without the finiteness hypothesis. Take $X = \mathbb{R}$ with the Lebesgue σ-algebra and Lebesgue measure m, and let $E_n := [n, +\infty)$. These are Lebesgue measurable (indeed Borel), decreasing, and $m(E_n) = +\infty$ for every n, so no E_n has finite measure. Yet $\bigcap_{n=1}^{\infty} E_n = \emptyset$, so

$$m\Big(\bigcap_{n=1}^{\infty} E_n\Big) = 0 \neq +\infty = \lim_{n\to\infty} m(E_n).$$

(Exactly as in Exercise 1.2.11; a purely abstract variant is counting measure $\#$ on $X = \mathbb{N}$ with the discrete σ-algebra and $E_n := \{n, n+1, \ldots\}$, where again every $\#(E_n) = +\infty$ but $\bigcap_n E_n = \emptyset$.)

Problem 1.4.24. (Dominated convergence for sets) Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $E_1, E_2, \ldots$ be a sequence of $\mathcal{B}$-measurable sets that converge to another set E, in the sense that 1_{E_n} converges pointwise to 1_E.

- (i) Show that E is also $\mathcal{B}$-measurable.
- (ii) If there exists a $\mathcal{B}$-measurable set F of finite measure (i.e. $\mu(F) < \infty$) that contains all of the E_n, show that $\lim_{n\to\infty} \mu(E_n) = \mu(E)$. (Hint: Apply downward monotonicity to the sets $\bigcup_{n>N}(E_n \Delta E)$.)
- (iii) Show that the previous part of this exercise can fail if the hypothesis that all the E_n are contained in a set of finite measure is omitted.

Solution. (i). We claim that

$$E = \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} E_n.$$

Indeed, the indicator functions take only the values 0 and 1, so a sequence $1_{E_n}(x)$ converges to $1_E(x)$ if and only if it is eventually equal to $1_E(x)$ (once the terms are within distance $1/2$ of the limit they must equal it). Hence, for $x \in E$: $1_{E_n}(x) = 1$ for all n beyond some N, i.e. $x \in \bigcap_{n\geq N} E_n$ for some N. Conversely, if $x \in \bigcap_{n\geq N} E_n$ for some N, then $1_{E_n}(x) = 1$ for all $n \geq N$, so the limit $1_E(x)$ equals 1 and $x \in E$. This proves the claim; and the right-hand side is $\mathcal{B}$-measurable, since a σ-algebra is closed under countable unions and (via complements and De Morgan) under countable intersections. Thus $E \in \mathcal{B}$.

(ii). Suppose $E_n \subset F$ for all n with $F \in \mathcal{B}$, $\mu(F) < \infty$. By the identity in (i), $E \subset \bigcup_N \bigcap_{n\geq N} E_n \subset F$ as well, so E and every $E_n \Delta E$ are $\mathcal{B}$-measurable subsets of F and therefore have finite measure by monotonicity.

Following the hint, set

$$G_N := \bigcup_{n>N} (E_n \Delta E) \in \mathcal{B}, \qquad N \geq 1.$$

Then $G_1 \supset G_2 \supset \cdots$, and $G_N \subset F$, so $\mu(G_1) \leq \mu(F) < \infty$. Moreover $\bigcap_{N=1}^{\infty} G_N = \emptyset$: if x belonged to every G_N, then for each N there would be some $n > N$ with $x \in E_n \Delta E$, i.e. $1_{E_n}(x) \neq 1_E(x)$; thus $1_{E_n}(x) \neq 1_E(x)$ for infinitely many n, contradicting the pointwise convergence $1_{E_n}(x) \to 1_E(x)$ (which, as noted in (i), forces eventual equality). Applying downwards monotone convergence, Exercise 1.4.23(iii), which is legitimate because $\mu(G_1) < \infty$, we conclude

$$\lim_{N\to\infty} \mu(G_N) = \mu\Big(\bigcap_{N=1}^{\infty} G_N\Big) = \mu(\emptyset) = 0.$$

Now observe that for any two $\mathcal{B}$-measurable subsets A, B of F we have $|\mu(A) - \mu(B)| \leq \mu(A \Delta B)$: indeed $A \subset B \cup (A \setminus B)$, so finite subadditivity (Exercise 1.4.20(iii)) gives $\mu(A) \leq \mu(B) + \mu(A \setminus B)$, and since all three quantities are finite we may rearrange to get $\mu(A) - \mu(B) \leq \mu(A \setminus B) \leq \mu(A \Delta B)$; swapping the roles of A and B gives the other inequality.

Applying this with $A = E_n$, $B = E$, and using monotonicity together with $E_n \Delta E \subset G_N$ for $n > N$, we obtain

$$|\mu(E_n) - \mu(E)| \leq \mu(E_n \Delta E) \leq \mu(G_N) \qquad \text{for all } n > N.$$

Given $\varepsilon > 0$, choose N with $\mu(G_N) < \varepsilon$; then $|\mu(E_n) - \mu(E)| < \varepsilon$ for all $n > N$. Hence $\lim_{n\to\infty} \mu(E_n) = \mu(E)$.

(iii). Take $X = \mathbb{R}$ with Lebesgue measure m on the Lebesgue σ-algebra, and let $E_n := [n, n+1]$. For each fixed $x \in \mathbb{R}$ we have $x \notin E_n$ as soon as $n > x$, so $1_{E_n}(x) \to 0$ for every x; that is, $1_{E_n} \to 1_E$ pointwise with $E = \emptyset$. But $m(E_n) = 1$ for every n, so

$$\lim_{n\to\infty} m(E_n) = 1 \neq 0 = m(E).$$

Of course the E_n here are not contained in any single set of finite measure: any measurable F containing all of them contains $[1, +\infty)$, whence $m(F) = +\infty$.

Problem 1.4.25. Let X be an at most countable set with the discrete σ-algebra. Show that every measure μ on this measurable space can be uniquely represented in the form

$$\mu = \sum_{x\in X} c_x \delta_x$$

for some $c_x \in [0, +\infty]$, thus

$$\mu(E) = \sum_{x\in E} c_x$$

for all $E \subset X$. (This claim fails in the uncountable case, although showing this is slightly tricky.)

Solution. Here δ_x is the Dirac measure of Example 1.4.22, $\delta_x(E) = 1_E(x)$, which is countably additive by Example 1.4.29; the discrete σ-algebra is the full power set 2^X, so every subset of X is measurable.
Unordered sums. For an at most countable index set I and numbers $a_i \in [0, +\infty]$, define

$$\sum_{i\in I} a_i := \sup\Big\{ \sum_{i\in F} a_i : F \subset I \text{ finite} \Big\} \in [0, +\infty],$$

with the empty sum equal to 0. If I is infinite and $i_1, i_2, \ldots$ is any enumeration of I, then this agrees with the ordinary series $\sum_{n=1}^{\infty} a_{i_n}$: each partial sum $\sum_{n\leq N} a_{i_n}$ is a finite-subset sum, so the series (i.e. the supremum of the partial sums) is at most the supremum above; conversely every finite $F \subset I$ is contained in an initial segment $\{i_1, \ldots, i_N\}$, so every finite-subset sum is at most a partial sum. In particular the value of $\sum_{i\in I} a_i$ does not depend on the enumeration chosen, and the expressions in the statement of the exercise are unambiguous. Note also that with this convention $\sum_{x\in X} c_x \delta_x$, interpreted as the map $E \mapsto \sum_{x\in X} c_x \delta_x(E) = \sum_{x\in X} c_x 1_E(x)$, is exactly the map $E \mapsto \sum_{x\in E} c_x$.
Existence. Let μ be a measure on 2^X and set

$$c_x := \mu(\{x\}) \in [0, +\infty] \qquad (x \in X).$$

Let $E \subset X$ be arbitrary. Since X is at most countable, so is E, and E is the disjoint union of its singletons. If E is infinite, enumerate $E = \{x_1, x_2, \ldots\}$ with the x_n distinct; the sets $\{x_n\}$ are disjoint and measurable with union E, so countable additivity gives

$$\mu(E) = \sum_{n=1}^{\infty} \mu(\{x_n\}) = \sum_{n=1}^{\infty} c_{x_n} = \sum_{x\in E} c_x,$$

the last equality by the discussion of unordered sums. If E is finite, the same conclusion follows from finite additivity (Exercise 1.4.20(ii)), or equivalently by padding the finite disjoint union out to a countable one with copies of $\emptyset$ and using $\mu(\emptyset) = 0$; and for $E = \emptyset$ both sides are 0. Hence $\mu(E) = \sum_{x\in E} c_x$ for all $E \subset X$, i.e. $\mu = \sum_{x\in X} c_x \delta_x$.

(Conversely, for any choice of constants $c_x \in [0,+\infty]$ the formula $E \mapsto \sum_{x\in E} c_x$ does define a measure on 2^X: it is the countable sum $\sum_{x\in X} c_x\delta_x$ of the countably additive measures $c_x\delta_x$, each of which is countably additive by Exercise 1.4.22(i) applied to the Dirac measure δ_x (countably additive by Example 1.4.29), and the sum of countably many countably additive measures is countably additive by Exercise 1.4.22(ii). (If X is finite, enumerate the family as a sequence padded out with copies of the zero measure of Example 1.4.23, which is countably additive, so that Exercise 1.4.22(ii) still applies.) So the representation describes exactly the measures on $(X, 2^X)$.)
Uniqueness. Suppose $\mu = \sum_{x\in X} c_x\delta_x = \sum_{x\in X} c'_x\delta_x$ for two families $(c_x)_{x\in X}$, $(c'_x)_{x\in X}$ in $[0,+\infty]$. Evaluating both representations at the singleton $E = \{y\}$ gives

$$c_y = \sum_{x\in\{y\}} c_x = \mu(\{y\}) = \sum_{x\in\{y\}} c'_x = c'_y$$

for every $y \in X$. Hence the coefficients are uniquely determined by μ (necessarily $c_x = \mu(\{x\})$), completing the proof.

Problem 1.4.26. (Completion) Let $(X, \mathcal{B}, \mu)$ be a measure space. Show that there exists a unique refinement $(X, \overline{\mathcal{B}}, \overline{\mu})$, known as the *completion* of $(X, \mathcal{B}, \mu)$, which is the coarsest refinement of $(X, \mathcal{B}, \mu)$ that is complete. Furthermore, show that $\overline{\mathcal{B}}$ consists precisely of those sets that differ from a $\mathcal{B}$-measurable set by a $\mathcal{B}$-subnull set.

Solution. Recall the conventions: a refinement of $(X, \mathcal{B}, \mu)$ is a measure space $(X, \mathcal{B}', \mu')$ with $\mathcal{B} \subset \mathcal{B}'$ and $\mu'|_{\mathcal{B}} = \mu$; a null set is a $\mathcal{B}$-measurable set of measure zero, a subnull set is a subset of a null set, and the space is complete if every subnull set is null (Definition 1.4.31). We will use repeatedly the group structure of symmetric difference: Δ is commutative and associative on subsets of X, with identity $\emptyset$ and $A\Delta A = \emptyset$; this is immediate from $1_{A\Delta B} = 1_A + 1_B \bmod 2$. In particular $E\Delta A = Z$ if and only if $E = A\Delta Z$, so "E differs from A by Z" is an unambiguous statement.
The candidate. Define

$$\mathcal{B}^* := \{A\Delta Z : A \in \mathcal{B},\ Z \text{ a } \mathcal{B}\text{-subnull set}\},$$

that is, the collection of sets that differ from a $\mathcal{B}$-measurable set by a $\mathcal{B}$-subnull set.
Step 1: $\mathcal{B}^*$ is a σ-algebra containing $\mathcal{B}$. Taking $Z = \emptyset$ (subnull, as $\emptyset$ is null) shows $\mathcal{B} \subset \mathcal{B}^*$; in particular $\emptyset \in \mathcal{B}^*$. If $E = A\Delta Z \in \mathcal{B}^*$ then, since $1_{A^c\Delta Z} = 1 + 1_A + 1_Z = 1 + 1_{A\Delta Z} \bmod 2$, we have $X \setminus E = (X \setminus A)\Delta Z \in \mathcal{B}^*$. For countable unions, let $E_n = A_n\Delta Z_n$ with $A_n \in \mathcal{B}$ and $Z_n \subset N_n$ for null sets N_n. Put $E := \bigcup_n E_n$, $A := \bigcup_n A_n \in \mathcal{B}$ and $N := \bigcup_n N_n \in \mathcal{B}$; by countable subadditivity (Exercise 1.4.23(i)), $\mu(N) \le \sum_n \mu(N_n) = 0$, so N is null. We claim $E\Delta A \subset \bigcup_n (E_n\Delta A_n) = \bigcup_n Z_n \subset N$. Indeed, if $x \in E \setminus A$ then $x \in E_n$ for some n while $x \notin A_n$ for all n, so $x \in E_n\Delta A_n$; symmetrically if $x \in A \setminus E$. Hence $Z := E\Delta A$ is $\mathcal{B}$-subnull and $E = A\Delta Z \in \mathcal{B}^*$.
Step 2: a well-defined measure on $\mathcal{B}^*$. For $E \in \mathcal{B}^*$ set $\mu^*(E) := \mu(A)$ whenever $E = A\Delta Z$ with $A \in \mathcal{B}$ and Z subnull. This is well defined: if $A\Delta Z = A'\Delta Z'$ with $Z \subset N$, $Z' \subset N'$ null, then applying $\Delta A\Delta Z'$ to both sides and using the group laws gives $A\Delta A' = Z\Delta Z' \subset N \cup N'$, a null set (its measure is at most $\mu(N) + \mu(N') = 0$ by finite subadditivity). Since $A \subset A' \cup (A\Delta A')$, monotonicity and finite subadditivity give $\mu(A) \le \mu(A') + 0 = \mu(A')$, and by symmetry $\mu(A') \le \mu(A)$; so $\mu(A) = \mu(A')$. Taking $Z = \emptyset$ shows $\mu^*|_{\mathcal{B}} = \mu$, so in particular $\mu^*(\emptyset) = 0$.
To see μ^* is countably additive, let $E_1, E_2, \ldots \in \mathcal{B}^*$ be disjoint, and write $E_n = A_n\Delta Z_n$ with $Z_n \subset N_n$ null. Let $N := \bigcup_n N_n$, a null set as in Step 1, and set $A'_n := A_n \setminus N \in \mathcal{B}$. Since $E_n\Delta A_n = Z_n \subset N$, the sets E_n and A_n agree outside N, hence

$$A'_n = A_n \setminus N = E_n \setminus N.$$

As the E_n are disjoint, so are the A'_n. Also $E_n\Delta A'_n \subset N$, so $\mu^*(E_n) = \mu(A'_n)$ by the well-definedness just proved. Finally, putting $E := \bigcup_n E_n$ and $A' := \bigcup_n A'_n$, we have $E\Delta A' \subset N$ (both sets agree outside N), so by countable additivity of μ,

$$\mu^*(E) = \mu(A') = \sum_{n=1}^{\infty} \mu(A'_n) = \sum_{n=1}^{\infty} \mu^*(E_n).$$

Thus $(X, \mathcal{B}^*, \mu^*)$ is a measure space, and it is a refinement of $(X, \mathcal{B}, \mu)$.
Step 3: $(X, \mathcal{B}^*, \mu^*)$ *is complete.* Let $E \in \mathcal{B}^*$ with $\mu^*(E) = 0$, and let $S \subset E$. Write $E = A \Delta Z$ with $Z \subset N$ null; then $\mu(A) = \mu^*(E) = 0$ and $E \subset A \cup Z \subset A \cup N$, where $\mu(A \cup N) \leq \mu(A) + \mu(N) = 0$. So $S \subset A \cup N$ is $\mathcal{B}$-subnull, whence $S = \emptyset \Delta S \in \mathcal{B}^*$ and $\mu^*(S) = \mu(\emptyset) = 0$. Every $\mathcal{B}^*$-subnull set is therefore $\mathcal{B}^*$-null.
Step 4: $(X, \mathcal{B}^*, \mu^*)$ *is coarser than every complete refinement.* Let $(X, \mathcal{B}', \mu')$ be any complete refinement of $(X, \mathcal{B}, \mu)$. Let $E = A \Delta Z \in \mathcal{B}^*$, with $A \in \mathcal{B}$ and $Z \subset N$ for some $\mathcal{B}$-null N. Then $N \in \mathcal{B} \subset \mathcal{B}'$ and $\mu'(N) = \mu(N) = 0$, so N is $\mathcal{B}'$-null and Z is $\mathcal{B}'$-subnull; by completeness of $\mathcal{B}'$ we get $Z \in \mathcal{B}'$ with $\mu'(Z) = 0$. Hence $E = A \Delta Z \in \mathcal{B}'$, so $\mathcal{B}^* \subset \mathcal{B}'$. Moreover, since $E \subset A \cup Z$ and $A \subset E \cup Z$, finite subadditivity for μ' gives $\mu'(E) \leq \mu'(A) + 0$ and $\mu'(A) \leq \mu'(E) + 0$, so

$$\mu'(E) = \mu'(A) = \mu(A) = \mu^*(E).$$

Thus μ' restricted to $\mathcal{B}^*$ equals μ^*, i.e. $(X, \mathcal{B}', \mu')$ is a refinement of $(X, \mathcal{B}^*, \mu^*)$.
Step 5: existence and uniqueness of the completion. Steps 2-4 show that $(X, \mathcal{B}^*, \mu^*)$ is a complete refinement of $(X, \mathcal{B}, \mu)$ which every complete refinement refines; that is, it is a coarsest complete refinement. If $(X, \mathcal{B}_1, \mu_1)$ is another coarsest complete refinement, then since $(X, \mathcal{B}^*, \mu^*)$ is a complete refinement we have $\mathcal{B}_1 \subset \mathcal{B}^*$ with $\mu^*|_{\mathcal{B}_1} = \mu_1$, and by Step 4 applied to $(X, \mathcal{B}_1, \mu_1)$ we have $\mathcal{B}^* \subset \mathcal{B}_1$ with $\mu_1|_{\mathcal{B}^*} = \mu^*$. Hence $\mathcal{B}_1 = \mathcal{B}^*$ and $\mu_1 = \mu^*$. So the completion $(X, \overline{\mathcal{B}}, \overline{\mu})$ exists and is unique, and by construction

$$\overline{\mathcal{B}} = \mathcal{B}^* = \{A \Delta Z : A \in \mathcal{B},\ Z\ \mathcal{B}\text{-subnull}\},$$

which is exactly the assertion that $\overline{\mathcal{B}}$ consists precisely of those sets differing from a $\mathcal{B}$-measurable set by a $\mathcal{B}$-subnull set, with $\overline{\mu}(A \Delta Z) = \mu(A)$.

Problem 1.4.27. Show that the Lebesgue measure space $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d], m)$ is the completion of the Borel measure space $(\mathbb{R}^d, \mathcal{B}[\mathbb{R}^d], m)$.

Solution. Write $\mathcal{L} := \mathcal{L}[\mathbb{R}^d]$ and $\mathcal{B} := \mathcal{B}[\mathbb{R}^d]$. By Exercise 1.4.26 the completion of $(\mathbb{R}^d, \mathcal{B}, m)$ is the unique measure space $(\mathbb{R}^d, \overline{\mathcal{B}}, \overline{m})$ with

$$\overline{\mathcal{B}} = \{A \Delta Z : A \in \mathcal{B},\ Z \subset N \text{ for some } N \in \mathcal{B} \text{ with } m(N) = 0\}, \qquad \overline{m}(A \Delta Z) = m(A).$$

So it suffices to show that $\mathcal{L} = \overline{\mathcal{B}}$ and that m on $\mathcal{L}$ is the measure $\overline{m}$.
First note that $(\mathbb{R}^d, \mathcal{L}, m)$ is indeed a refinement of $(\mathbb{R}^d, \mathcal{B}, m)$: every open set is Lebesgue measurable by Lemma 1.2.13(i) and $\mathcal{L}$ is a σ-algebra by Lemma 1.2.13(iv),(v),(vi), so the σ-algebra $\mathcal{B}$ generated by the open sets is contained in $\mathcal{L}$; and the measure is literally the same function m restricted.
A preliminary remark: null sets sit inside Borel null sets. If $Z \subset \mathbb{R}^d$ has Lebesgue outer measure $m^*(Z) = 0$, then there is a Borel set $N \supset Z$ with $m(N) = 0$. Indeed, for each $k \geq 1$, by definition of Lebesgue outer measure we may cover Z by countably many boxes $B_1, B_2, \ldots$ with $\sum_j |B_j| \leq 1/(2k)$; enlarging each B_j slightly to an open box $B_j' \supset B_j$ with $|B_j'| \leq |B_j| + 2^{-j}/(2k)$, the set $V_k := \bigcup_j B_j'$ is open, contains Z, and by countable subadditivity of outer measure $m^*(V_k) \leq \sum_j |B_j'| \leq 1/k$. Then $N := \bigcap_{k \geq 1} V_k$ is a G_δ set, hence Borel, contains Z, and $m(N) \leq m^*(V_k) \leq 1/k$ for every k, so $m(N) = 0$.
$\overline{\mathcal{B}} \subset \mathcal{L}$. Let $E = A \Delta Z$ with $A \in \mathcal{B}$ and $Z \subset N$ for a Borel set N of measure zero. Then $m^*(Z) \leq m^*(N) = 0$, so Z is Lebesgue measurable by Lemma 1.2.13(iii). Since $A \in \mathcal{B} \subset \mathcal{L}$ and $\mathcal{L}$ is a σ-algebra (hence closed under the Boolean operations making up $A \Delta Z = (A \setminus Z) \cup (Z \setminus A)$), we get $E \in \mathcal{L}$.
$\mathcal{L} \subset \overline{\mathcal{B}}$. Let $E \in \mathcal{L}$. By the criterion for measurability in Exercise 1.2.7 (outer approximation by open sets), for each $n \geq 1$ there is an open set $U_n \supset E$ with $m^*(U_n \setminus E) \leq 1/n$. Put $G := \bigcap_{n \geq 1} U_n$. Then G is a G_δ, hence Borel, and $E \subset G$. Moreover $G \setminus E \subset U_n \setminus E$ for every n, so by monotonicity of outer measure $m^*(G \setminus E) \leq 1/n$ for all n, i.e. $m^*(G \setminus E) = 0$. By the preliminary remark there is a Borel set $N \supset G \setminus E$ with $m(N) = 0$. Since $G \setminus E \subset G$ and $E \subset G$, we have

$$G \Delta (G \setminus E) = G \cap E = E,$$

exhibiting E as a Borel set G symmetric-differenced with the $\mathcal{B}$-subnull set $G \setminus E$. Hence $E \in \overline{\mathcal{B}}$.

The measures agree. We have shown $\mathcal{L} = \overline{\mathcal{B}}$. If $E \in \mathcal{L}$ is written as $E = A \Delta Z$ with A Borel and $Z \subset N$ Borel null, then $E \subset A \cup Z$ and $A \subset E \cup Z$, and $m(Z) = 0$ (Lemma 1.2.13(iii) again, plus monotonicity), so by finite subadditivity of Lebesgue measure $m(E) \leq m(A)$ and $m(A) \leq m(E)$; thus $m(E) = m(A) = \overline{m}(E)$.
Therefore $(\mathbb{R}^d, \mathcal{L}, m)$ coincides with the completion $(\mathbb{R}^d, \overline{\mathcal{B}}, \overline{m})$ of the Borel measure space, as claimed. (Consistently with Definition 1.4.31 and the discussion following it: the Lebesgue space is complete, while the Borel space is not, so the completion genuinely enlarges $\mathcal{B}$.)

Problem 1.4.28. (Approximation by an algebra) Let $\mathcal{A}$ be a Boolean algebra on X, and let μ be a measure on $\langle \mathcal{A} \rangle$.

- (i) If $\mu(X) < \infty$, show that for every $E \in \langle \mathcal{A} \rangle$ and $\varepsilon > 0$ there exists $F \in \mathcal{A}$ such that $\mu(E \Delta F) < \varepsilon$.
- (ii) More generally, if $X = \bigcup_{n=1}^{\infty} A_n$ for some $A_1, A_2, \ldots \in \mathcal{A}$ with $\mu(A_n) < \infty$ for all n, $E \in \langle \mathcal{A} \rangle$ has finite measure, and $\varepsilon > 0$, show that there exists $F \in \mathcal{A}$ such that $\mu(E \Delta F) < \varepsilon$.

Solution. Here $\langle \mathcal{A} \rangle$ denotes the σ-algebra generated by $\mathcal{A}$, i.e. the intersection of all σ-algebras on X containing $\mathcal{A}$; it is characterised by the property that any σ-algebra containing $\mathcal{A}$ contains $\langle \mathcal{A} \rangle$. We use throughout the "triangle inequality" for symmetric differences,

$$A \Delta C \subset (A \Delta B) \cup (B \Delta C),$$

valid for all sets A, B, C (if x lies in exactly one of A, C then it cannot agree with B in membership on both sides), and the identity $(X \setminus A) \Delta (X \setminus B) = A \Delta B$.
(i). Assume $\mu(X) < \infty$ and let

$$\mathcal{C} := \{E \in \langle \mathcal{A} \rangle : \text{for every } \varepsilon > 0 \text{ there is } F \in \mathcal{A} \text{ with } \mu(E \Delta F) < \varepsilon\}.$$

We show $\mathcal{C}$ is a σ-algebra on X containing $\mathcal{A}$; since $\mathcal{C} \subset \langle \mathcal{A} \rangle$ by definition, this forces $\mathcal{C} = \langle \mathcal{A} \rangle$ and proves the claim.
$\mathcal{A} \subset \mathcal{C}$*:* for $E \in \mathcal{A}$ take $F := E$, so that $\mu(E \Delta F) = \mu(\emptyset) = 0 < \varepsilon$. In particular $\emptyset, X \in \mathcal{C}$.
Closure under complements: if $E \in \mathcal{C}$ and $\varepsilon > 0$, pick $F \in \mathcal{A}$ with $\mu(E \Delta F) < \varepsilon$. Then $X \setminus F \in \mathcal{A}$ (Boolean algebras are closed under complement) and $\mu((X \setminus E) \Delta (X \setminus F)) = \mu(E \Delta F) < \varepsilon$. Also $X \setminus E \in \langle \mathcal{A} \rangle$. Hence $X \setminus E \in \mathcal{C}$.
Closure under countable unions: let $E_1, E_2, \ldots \in \mathcal{C}$ and $E := \bigcup_{n=1}^{\infty} E_n \in \langle \mathcal{A} \rangle$. Fix $\varepsilon > 0$. Set $E^{(N)} := \bigcup_{n=1}^{N} E_n$; these increase to E, so by upwards monotone convergence (Exercise 1.4.23(ii)) $\mu(E^{(N)}) \to \mu(E)$. Since $E^{(N)} \subset E$ and all measures involved are at most $\mu(X) < \infty$, finite additivity gives $\mu(E \setminus E^{(N)}) = \mu(E) - \mu(E^{(N)}) \to 0$; note $E \Delta E^{(N)} = E \setminus E^{(N)}$. Choose N with $\mu(E \Delta E^{(N)}) < \varepsilon/2$. Next, for each $1 \leq n \leq N$ choose $F_n \in \mathcal{A}$ with $\mu(E_n \Delta F_n) < \varepsilon 2^{-n-1}$, and set $F := \bigcup_{n=1}^{N} F_n \in \mathcal{A}$ (Boolean algebras are closed under finite unions). Since

$$\Big(\bigcup_{n \leq N} E_n \Big) \Delta \Big(\bigcup_{n \leq N} F_n \Big) \subset \bigcup_{n \leq N} (E_n \Delta F_n)$$

(the same verification as in Step 1 of Exercise 1.4.26), finite subadditivity gives $\mu(E^{(N)} \Delta F) \leq \sum_{n \leq N} \varepsilon 2^{-n-1} < \varepsilon/2$. By the triangle inequality and finite subadditivity,

$$\mu(E \Delta F) \leq \mu(E \Delta E^{(N)}) + \mu(E^{(N)} \Delta F) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Hence $E \in \mathcal{C}$, and $\mathcal{C}$ is a σ-algebra, completing part (i).
(ii). Now suppose $X = \bigcup_{n=1}^{\infty} A_n$ with $A_n \in \mathcal{A}$, $\mu(A_n) < \infty$, and let $E \in \langle \mathcal{A} \rangle$ with $\mu(E) < \infty$ and $\varepsilon > 0$. Replacing A_n by $A_1 \cup \cdots \cup A_n$ (still in $\mathcal{A}$, and of finite measure by finite subadditivity) we may assume $A_1 \subset A_2 \subset \cdots$, still with union X.
Reduction to a finite piece. The sets $E \cap A_n$ lie in $\langle \mathcal{A} \rangle$ and increase to E, so by Exercise 1.4.23(ii), $\mu(E \cap A_n) \to \mu(E)$; as $\mu(E) < \infty$ we may subtract to get $\mu(E \setminus A_n) = \mu(E) - \mu(E \cap A_n) \to 0$. Fix N with $\mu(E \setminus A_N) < \varepsilon/2$, and write $Y := A_N$, so $\mu(Y) < \infty$.

Restricting to Y. Let $\mathcal{A}|_Y := \{A \cap Y : A \in \mathcal{A}\}$, a Boolean algebra on the set Y (it is closed under finite unions, and the complement of $A \cap Y$ within Y is $(X \setminus A) \cap Y$). Likewise $\langle\mathcal{A}\rangle|_Y := \{B \cap Y : B \in \langle\mathcal{A}\rangle\} = \{B \in \langle\mathcal{A}\rangle : B \subset Y\}$ (the second description uses $Y \in \mathcal{A} \subset \langle\mathcal{A}\rangle$), which is a σ-algebra on Y. I claim that the σ-algebra on Y generated by $\mathcal{A}|_Y$ is exactly $\langle\mathcal{A}\rangle|_Y$. Indeed, $\langle\mathcal{A}\rangle|_Y$ is a σ-algebra on Y containing $\mathcal{A}|_Y$, so it contains $\langle\mathcal{A}|_Y\rangle$. Conversely, the collection $\mathcal{D} := \{B \subset X : B \cap Y \in \langle\mathcal{A}|_Y\rangle\}$ is a σ-algebra on X (intersection with Y commutes with countable unions, and $(X \setminus B) \cap Y = Y \setminus (B \cap Y)$) which contains $\mathcal{A}$; hence $\langle\mathcal{A}\rangle \subset \mathcal{D}$, i.e. $B \cap Y \in \langle\mathcal{A}|_Y\rangle$ for every $B \in \langle\mathcal{A}\rangle$, giving $\langle\mathcal{A}\rangle|_Y \subset \langle\mathcal{A}|_Y\rangle$.
The restriction $\mu|_Y$ of μ to $\langle\mathcal{A}\rangle|_Y = \langle\mathcal{A}|_Y\rangle$ is a countably additive measure on this σ-algebra (Example 1.4.30), and its total mass is $\mu(Y) < \infty$. Since $E \cap Y \in \langle\mathcal{A}|_Y\rangle$, part (i) applied to the measure space $(Y, \langle\mathcal{A}|_Y\rangle, \mu|_Y)$ and the Boolean algebra $\mathcal{A}|_Y$ furnishes some $F_0 \in \mathcal{A}$ with

$$\mu\big((E \cap Y)\Delta(F_0 \cap Y)\big) < \varepsilon/2.$$

Conclusion. Put $F := F_0 \cap Y \in \mathcal{A}$ (Boolean algebras are closed under intersection, by De Morgan). Then $F \subset Y$, and

$$E\Delta F \subset (E \setminus Y) \cup \big((E \cap Y)\Delta F\big),$$

since a point of $E \setminus F$ either lies outside Y, and then lies in $E \setminus Y$, or lies in Y, and then lies in $(E \cap Y) \setminus F$; while a point of $F \setminus E$ lies in Y and hence in $F \setminus (E \cap Y)$. By finite subadditivity,

$$\mu(E\Delta F) \leq \mu(E \setminus A_N) + \mu\big((E \cap Y)\Delta F\big) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

as required.

Problem 1.4.29. Let $(X, \mathcal{B})$ be a measurable space.

- (i) Show that a function $f : X \to [0, +\infty]$ is measurable if and only if the level sets $\{x \in X : f(x) > \lambda\}$ are $\mathcal{B}$-measurable.
- (ii) Show that an indicator function 1_E of a set $E \subset X$ is measurable if and only if E itself is $\mathcal{B}$-measurable.
- (iii) Show that a function $f : X \to [0, +\infty]$ or $f : X \to \mathbb{C}$ is measurable if and only if $f^{-1}(E)$ is $\mathcal{B}$-measurable for every Borel-measurable subset E of $[0, +\infty]$ or $\mathbb{C}$.
- (iv) Show that a function $f : X \to \mathbb{C}$ is measurable if and only if its real and imaginary parts are measurable.
- (v) Show that a function $f : X \to \mathbb{R}$ is measurable if and only if the magnitudes $f_+ := \max(f, 0)$, $f_- := \max(-f, 0)$ of its positive and negative parts are measurable.
- (vi) If $f_n : X \to [0, +\infty]$ are a sequence of measurable functions that converge pointwise to a limit $f : X \to [0, +\infty]$, then show that f is also measurable. Obtain the same claim if $[0, +\infty]$ is replaced by $\mathbb{C}$.
- (vii) If $f : X \to [0, +\infty]$ is measurable and $\varphi : [0, +\infty] \to [0, +\infty]$ is continuous, show that $\varphi \circ f$ is measurable. Obtain the same claim if $[0, +\infty]$ is replaced by $\mathbb{C}$.
- (viii) Show that the sum or product of two measurable functions in $[0, +\infty]$ or $\mathbb{C}$ is still measurable.

Solution. Throughout we use Definition 1.4.32: f is measurable iff $f^{-1}(U) \in \mathcal{B}$ for every open U in the target space. Two elementary remarks will be used repeatedly. First, preimages commute with all Boolean operations: $f^{-1}(Y \setminus E) = X \setminus f^{-1}(E)$ and $f^{-1}(\bigcup_n E_n) = \bigcup_n f^{-1}(E_n)$. Second, the topology of $[0, +\infty]$ has as a base the sets $[0, b)$, (a, b), $(a, +\infty]$ with $0 \leq a < b < \infty$ rational, and the topology of $\mathbb{C}$ has as a base the open discs with rational centre and radius; both bases are countable, so every open set in either space is a countable union of basic sets.
(i) Suppose f is measurable. For $\lambda \in [0, +\infty)$ the set $(\lambda, +\infty]$ is open in $[0, +\infty]$, so $\{f > \lambda\} = f^{-1}((\lambda, +\infty])$ lies in $\mathcal{B}$. (For $\lambda < 0$ the level set is all of X, which is in $\mathcal{B}$.)
Conversely suppose every $\{f > \lambda\}$ is in $\mathcal{B}$. For $b \in (0, \infty)$ we have

$$\{f \geq b\} = \bigcap_{n=1}^{\infty} \{f > b - \tfrac{1}{n}\},$$

a countable intersection of $\mathcal{B}$-measurable sets, hence in $\mathcal{B}$; therefore $f^{-1}([0,b)) = \{f < b\} = X \setminus \{f \geq b\}$ is in $\mathcal{B}$. Also $f^{-1}((a,+\infty]) = \{f > a\} \in \mathcal{B}$, and $f^{-1}((a,b)) = \{f > a\} \cap \{f < b\} \in \mathcal{B}$. Thus the preimage of every basic open set lies in $\mathcal{B}$, and since every open $U \subset [0,+\infty]$ is a countable union of basic sets, $f^{-1}(U) \in \mathcal{B}$. Hence f is measurable.
(ii) If $E \in \mathcal{B}$ and U is open, then $1_E^{-1}(U)$ is one of $\emptyset$, E, $X \setminus E$, X according to which of $0, 1$ lie in U; all four are in $\mathcal{B}$, so 1_E is measurable. Conversely, if 1_E is measurable then $E = \{1_E > 1/2\}$ is in $\mathcal{B}$ by (i) (or directly, $E = 1_E^{-1}((1/2,+\infty])$).
(iii) Let Y denote $[0,+\infty]$ or $\mathbb{C}$, and set

$$\mathcal{C} := \{E \subset Y : f^{-1}(E) \in \mathcal{B}\}.$$

Since $f^{-1}(\emptyset) = \emptyset$, and preimages commute with complements and countable unions as noted above, $\mathcal{C}$ is a σ-algebra on Y. If f is measurable then $\mathcal{C}$ contains every open subset of Y, hence contains the σ-algebra generated by the open sets, which is by definition the Borel σ-algebra $\mathcal{B}[Y]$. Thus $f^{-1}(E) \in \mathcal{B}$ for every Borel E. The converse is immediate, since open sets are Borel.
(iv) Write $f = u + iv$ with $u = \operatorname{Re} f$, $v = \operatorname{Im} f$, regarded as complex-valued functions taking real values. If f is measurable and $U \subset \mathbb{C}$ is open, then $u^{-1}(U) = u^{-1}(U \cap \mathbb{R})$, and $U \cap \mathbb{R}$ is open in $\mathbb{R}$, hence a countable union of intervals (a,b). Since $u^{-1}((a,b)) = f^{-1}(\{z \in \mathbb{C} : a < \operatorname{Re} z < b\})$ and the latter set is open in $\mathbb{C}$, each such preimage lies in $\mathcal{B}$, so $u^{-1}(U) \in \mathcal{B}$. The same works for v.
Conversely, suppose u, v are measurable. Write $S_{a,b} := \{z \in \mathbb{C} : a < \operatorname{Re} z < b\}$, an open strip. For an open rectangle $R = \{z : a < \operatorname{Re} z < b,\ c < \operatorname{Im} z < d\}$ we have, since u and v are real-valued, $f^{-1}(R) = \{x : a < u(x) < b\} \cap \{x : c < v(x) < d\} = u^{-1}(S_{a,b}) \cap v^{-1}(S_{c,d}) \in \mathcal{B}$. Every open $U \subset \mathbb{C}$ is a countable union of such rectangles with rational corners, so $f^{-1}(U) \in \mathcal{B}$.
(v) Here, as in (iv), a real-valued function is regarded as a complex-valued one (Definition 1.4.32 assigns a meaning to "measurable" only for targets $[0,+\infty]$ and $\mathbb{C}$), while f_+ and f_- are regarded as unsigned functions with values in $[0,+\infty]$. Suppose $f : X \to \mathbb{R}$ is measurable. For $\lambda \geq 0$ we have $\{f_+ > \lambda\} = \{f > \lambda\} = f^{-1}(\{z : \operatorname{Re} z > \lambda\}) \in \mathcal{B}$, and for $\lambda < 0$, $\{f_+ > \lambda\} = X$. By (i), f_+ is measurable; the same argument applied to $-f$ (which is measurable, since $(-f)^{-1}(U) = f^{-1}(-U)$ and $-U$ is open) shows f_- is measurable.
Conversely suppose f_+, f_- are measurable. For $\lambda \geq 0$ we have $\{f > \lambda\} = \{f_+ > \lambda\} \in \mathcal{B}$. For $\lambda < 0$ we have $\{f > \lambda\} = \{f_- < -\lambda\}$, which is in $\mathcal{B}$ because $\{f_- < c\} = X \setminus \bigcap_n \{f_- > c - \frac{1}{n}\}$ for $c > 0$. Hence $\{f > \lambda\} \in \mathcal{B}$ for all real λ. Now let $U \subset \mathbb{C}$ be open; as in (iv), $f^{-1}(U) = f^{-1}(U \cap \mathbb{R})$ and $U \cap \mathbb{R}$ is a countable union of intervals (a,b), while $f^{-1}((a,b)) = \{f > a\} \setminus \bigcap_n \{f > b - \frac{1}{n}\}$ lies in $\mathcal{B}$. So f is measurable.
(vi) Unsigned case. We claim that for every $\lambda \in [0,\infty)$,

$$\{f > \lambda\} = \bigcup_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \geq N} \{f_n > \lambda + \tfrac{1}{m}\}.$$

If $f(x) > \lambda$, pick m with $\lambda + \frac{2}{m} < f(x)$ (possible whether $f(x)$ is finite or $+\infty$); since $f_n(x) \to f(x)$, there is N with $f_n(x) > \lambda + \frac{1}{m}$ for all $n \geq N$, so x lies in the right-hand side. Conversely, if $f_n(x) > \lambda + \frac{1}{m}$ for all $n \geq N$, then $f(x) = \lim_n f_n(x) \geq \lambda + \frac{1}{m} > \lambda$. The right-hand side is a countable combination of sets which lie in $\mathcal{B}$ by (i), so $\{f > \lambda\} \in \mathcal{B}$, and f is measurable by (i).
Complex case. Write $f_n = u_n + iv_n$ and $f = u + iv$; by (iv) each u_n, v_n is measurable, and $u_n \to u$, $v_n \to v$ pointwise. The maps $t \mapsto t_+$ and $t \mapsto t_-$ are continuous on $\mathbb{R}$, so $(u_n)_\pm \to u_\pm$ pointwise; by (v) each $(u_n)_\pm$ is measurable, so by the unsigned case $u_\pm$ are measurable, and by (v) again u is measurable. Likewise v is measurable, and by (iv) so is f.
(vii) If U is open in the target, then $\varphi^{-1}(U)$ is open by continuity of φ, so $(\varphi \circ f)^{-1}(U) = f^{-1}(\varphi^{-1}(U)) \in \mathcal{B}$ by measurability of f. The argument uses nothing about the particular spaces, so it applies verbatim with $[0,+\infty]$ replaced by $\mathbb{C}$ (and indeed to any continuous φ between the two spaces).
(viii) Unsigned case. Let $f, g : X \to [0,+\infty]$ be measurable. For $\lambda \in [0,\infty)$ we claim

$$\{f + g > \lambda\} = \bigcup_{q \in \mathbb{Q}} \left(\{f > q\} \cap \{g > \lambda - q\}\right),$$

where $\{f > q\} := X$ when $q < 0$, and similarly for g. Indeed, if $f(x) > q$ and $g(x) > \lambda - q$ then $f(x) + g(x) > \lambda$ by monotonicity of addition. Conversely, if $f(x) + g(x) > \lambda$ then $f(x) > \lambda - g(x)$ as extended reals (reading $\lambda - g(x)$ as $-\infty$ when $g(x) = +\infty$), so by density of the rationals there is $q \in \mathbb{Q}$ with $\lambda - g(x) < q < f(x)$, and then $g(x) > \lambda - q$. Each set on the right lies in $\mathcal{B}$ by (i), and the union is countable, so $f + g$ is measurable by (i).
For the product (with the convention $0 \cdot \infty = 0$) we claim, for $\lambda \in [0, \infty)$,

$$\{fg > \lambda\} = \bigcup_{q \in \mathbb{Q},\, q > 0} (\{f > q\} \cap \{g > \lambda/q\}).$$

If $f(x) > q > 0$ and $g(x) > \lambda/q \geq 0$, then $f(x)g(x) > \lambda$: when $\lambda = 0$ this says the product of two strictly positive extended reals is strictly positive, and when $\lambda > 0$ it follows from strict monotonicity of multiplication of positive quantities. Conversely, suppose $f(x)g(x) > \lambda \geq 0$; then $f(x)g(x) > 0$, so (with the stated convention) both $f(x) > 0$ and $g(x) > 0$. Now $\lambda/g(x) < f(x)$, reading $\lambda/g(x) = 0$ when $g(x) = +\infty$, so we may pick a rational q with $\max(0, \lambda/g(x)) < q < f(x)$; then $g(x) > \lambda/q$. Again the right-hand side lies in $\mathcal{B}$, so fg is measurable by (i).
Complex case. Let $f, g : X \to \mathbb{C}$ be measurable and consider $F := (f, g) : X \to \mathbb{C}^2$. For open discs $D, D' \subset \mathbb{C}$ we have $F^{-1}(D \times D') = f^{-1}(D) \cap g^{-1}(D') \in \mathcal{B}$, and every open subset of $\mathbb{C}^2$ is a countable union of products of discs with rational data; hence $F^{-1}(U) \in \mathcal{B}$ for every open $U \subset \mathbb{C}^2$. Addition $\sigma(z, w) = z + w$ and multiplication $\pi(z, w) = zw$ are continuous $\mathbb{C}^2 \to \mathbb{C}$, so as in (vii), $(f + g)^{-1}(U) = F^{-1}(\sigma^{-1}(U)) \in \mathcal{B}$ and $(fg)^{-1}(U) = F^{-1}(\pi^{-1}(U)) \in \mathcal{B}$ for every open $U \subset \mathbb{C}$. Thus $f + g$ and fg are measurable.

Problem 1.4.30. Let $(X, \mathcal{B})$ be a measurable space that is atomic, thus $\mathcal{B} = \mathcal{A}((A_\alpha)_{\alpha \in I})$ for some partition $X = \bigcup_{\alpha \in I} A_\alpha$ of X into disjoint non-empty atoms. Show that a function $f : X \to [0, +\infty]$ or $f : X \to \mathbb{C}$ is measurable if and only if it is constant on each atom, or equivalently if one has a representation of the form

$$f = \sum_{\alpha \in I} c_\alpha 1_{A_\alpha}$$

for some constants c_α in $[0, +\infty]$ or in $\mathbb{C}$ as appropriate. Furthermore, the c_α are uniquely determined by f.

Solution. Let Y denote the target space, either $[0, +\infty]$ or $\mathbb{C}$. Recall from Example 1.4.7 that $\mathcal{B} = \mathcal{A}((A_\alpha)_{\alpha \in I})$ consists exactly of the sets of the form $\bigcup_{\alpha \in J} A_\alpha$ with $J \subset I$.
We first record the key structural fact: if $E \in \mathcal{B}$ and E meets an atom A_α, then $A_\alpha \subset E$. Indeed, write $E = \bigcup_{\beta \in J} A_\beta$ and pick $x \in E \cap A_\alpha$. Then $x \in A_\beta$ for some $\beta \in J$; since the A_β are disjoint and $x \in A_\alpha$, we must have $A_\beta = A_\alpha$, whence $A_\alpha \subset E$. So each $\mathcal{B}$-measurable set either contains a given atom or is disjoint from it.
Constant on atoms $\Rightarrow$ measurable. Suppose f is constant on each atom, say $f \equiv c_\alpha$ on A_α. Let $E \subset Y$ be **any** subset. Then

$$f^{-1}(E) = \bigcup \{A_\alpha : \alpha \in I,\ c_\alpha \in E\},$$

because every $x \in X$ lies in exactly one atom A_α (the A_α partition X) and $f(x) = c_\alpha$ there. This is a union of atoms, hence lies in $\mathcal{B}$. In particular $f^{-1}(U) \in \mathcal{B}$ for every open $U \subset Y$, so f is measurable.
Measurable $\Rightarrow$ constant on atoms. Suppose f is measurable, and suppose for contradiction that f is not constant on some atom A_α: there are $x, y \in A_\alpha$ with $f(x) \neq f(y)$. The space Y is Hausdorff, so we may choose an open $U \subset Y$ containing $f(x)$ but not $f(y)$. (Explicitly: if $Y = \mathbb{C}$, take an open disc about $f(x)$ of radius $|f(x) - f(y)|$; if $Y = [0, +\infty]$ and, say, $f(x) < f(y)$, take $U = [0, r)$ for any finite r with $f(x) < r < f(y)$, while if $f(y) < f(x)$ take $U = (r, +\infty]$ with $f(y) < r < f(x)$.) Then $E := f^{-1}(U)$ lies in $\mathcal{B}$ and contains $x \in A_\alpha$; by the structural fact above, $A_\alpha \subset E$, so $y \in E$ and $f(y) \in U$, a contradiction. Hence f is constant on every atom.
Equivalence with the representation. If f is constant on each atom, define c_α to be the common value of f on A_α; this is well defined precisely because each atom is non-empty. For $x \in X$ there is exactly one $\alpha_0 \in I$ with $x \in A_{\alpha_0}$, so in the sum $\sum_{\alpha \in I} c_\alpha 1_{A_\alpha}(x)$ every term with $\alpha \neq \alpha_0$ vanishes (using the

convention $0 \cdot \infty = 0$ in the unsigned case) and the sum equals $c_{\alpha_0} = f(x)$; in particular the possibly uncountable sum is unambiguously defined, having at most one non-zero term at each point. Thus $f = \sum_{\alpha \in I} c_\alpha 1_{A_\alpha}$. Conversely, if f has such a representation then by the same computation $f \equiv c_\alpha$ on A_α, so f is constant on each atom. Combining with the two implications above, all three conditions (measurable; constant on each atom; admitting such a representation) are equivalent.
Uniqueness. Suppose $\sum_{\alpha \in I} c_\alpha 1_{A_\alpha} = \sum_{\alpha \in I} c'_\alpha 1_{A_\alpha}$ as functions on X. Fix $\alpha \in I$ and choose $x \in A_\alpha$, which is possible since the atoms are non-empty. Evaluating both sides at x gives $c_\alpha = c'_\alpha$. Hence the coefficients are uniquely determined by f (namely, c_α is the value of f at any point of A_α). Note that this is where the non-emptiness hypothesis is essential: an empty atom would carry an entirely arbitrary coefficient.

Problem 1.4.31. (Egorov's theorem) Let $(X, \mathcal{B}, \mu)$ be a finite measure space (so $\mu(X) < \infty$), and let $f_n : X \to \mathbb{C}$ be a sequence of measurable functions that converge pointwise almost everywhere to a limit $f : X \to \mathbb{C}$, and let $\varepsilon > 0$. Show that there exists a measurable set E of measure at most ε such that f_n converges uniformly to f outside of E. Give an example to show that the claim can fail when the measure μ is not finite.

Solution. Setting up. By hypothesis the set $Z := \{x \in X : f_n(x) \not\to f(x)\}$ is a sub-null set, so (Definition 1.4.31) there is a null set $N \in \mathcal{B}$ with $\mu(N) = 0$ and $Z \subset N$.
We first arrange that everything in sight is measurable. Put

$$\tilde{f}_n := f_n 1_{X \setminus N}, \qquad \tilde{f} := f 1_{X \setminus N}.$$

Since $N \in \mathcal{B}$, the indicator $1_{X \setminus N}$ is measurable by Exercise 1.4.29(ii), and hence each $\tilde{f}_n$ is measurable by Exercise 1.4.29(viii). Moreover $\tilde{f}_n \to \tilde{f}$ pointwise on all of X: on N both sides are 0, and off N we have $x \notin Z$, so $f_n(x) \to f(x)$. Therefore $\tilde{f}$ is measurable by Exercise 1.4.29(vi). Consequently $h_n := \tilde{f}_n - \tilde{f}$ is measurable: the constant function -1 is measurable (the preimage of any open set is $\emptyset$ or X), so $(-1)\tilde{f}$ is measurable by Exercise 1.4.29(viii), and then so is $h_n = \tilde{f}_n + (-1)\tilde{f}$ by Exercise 1.4.29(viii) again. Write

$$g_n := |h_n| = |\tilde{f}_n - \tilde{f}| : X \to [0, \infty).$$

We shall only ever need the sets $\{g_n > 1/m\}$, and these lie in $\mathcal{B}$ directly from Definition 1.4.32: the set $\{z \in \mathbb{C} : |z| > 1/m\}$ is open in $\mathbb{C}$, and

$$\{x \in X : g_n(x) > \tfrac{1}{m}\} = h_n^{-1}(\{z \in \mathbb{C} : |z| > \tfrac{1}{m}\}) \in \mathcal{B}.$$

The main construction. For integers $M, m \geq 1$ set

$$E_{M,m} := \bigcup_{n \geq M} \{x \in X : g_n(x) > \tfrac{1}{m}\},$$

which lies in $\mathcal{B}$ as a countable union of the measurable sets just described.
Fix m. The sets $E_{M,m}$ are decreasing in M: $E_{1,m} \supset E_{2,m} \supset \cdots$. Their intersection is empty: if $x \in \bigcap_M E_{M,m}$ then $g_n(x) > 1/m$ for infinitely many n, contradicting $g_n(x) \to 0$, which holds for every $x \in X$ since $\tilde{f}_n \to \tilde{f}$ pointwise everywhere. Since $\mu(X) < \infty$, in particular $\mu(E_{1,m}) < \infty$, so downward monotone convergence (Exercise 1.4.23(iii)) gives

$$\lim_{M \to \infty} \mu(E_{M,m}) = \mu\Big(\bigcap_{M=1}^{\infty} E_{M,m} \Big) = \mu(\emptyset) = 0.$$

This is precisely where finiteness of μ is used. Hence we may choose M_m with

$$\mu(E_{M_m,m}) \leq \varepsilon 2^{-m}.$$

Now define

$$E := N \cup \bigcup_{m=1}^{\infty} E_{M_m,m} \in \mathcal{B}.$$

By countable subadditivity (Exercise 1.4.23(i)),

$$\mu(E) \leq \mu(N) + \sum_{m=1}^{\infty} \mu(E_{M_m,m}) \leq 0 + \sum_{m=1}^{\infty} \varepsilon 2^{-m} = \varepsilon.$$

Uniform convergence off E. Let $\delta > 0$ and choose m with $1/m < \delta$. If $x \in X \setminus E$ then $x \notin E_{M_m,m}$, so $g_n(x) \leq 1/m$ for every $n \geq M_m$. Also $x \notin N$, so $\tilde{f}_n(x) = f_n(x)$ and $\tilde{f}(x) = f(x)$, i.e. $g_n(x) = |f_n(x) - f(x)|$. Therefore

$$\sup_{x \in X \setminus E} |f_n(x) - f(x)| \leq \tfrac{1}{m} < \delta \qquad \text{for all } n \geq M_m.$$

As $\delta > 0$ was arbitrary, $f_n \to f$ uniformly on $X \setminus E$, and $\mu(E) \leq \varepsilon$ as required.
Failure for infinite measure. Take $X = \mathbb{R}$ with Lebesgue measure, here written Leb to avoid a clash with the index m above, on the Lebesgue σ-algebra, and let $f_n := 1_{[n,n+1]}$. Each f_n is measurable, and $f_n \to 0$ pointwise everywhere (for fixed x, $f_n(x) = 0$ once $n > x$), so the hypotheses hold with limit $f := 0$. Let $\varepsilon := 1/2$ and suppose E were a measurable set with $\mathrm{Leb}(E) \leq 1/2$ such that $f_n \to 0$ uniformly on $\mathbb{R} \setminus E$. Applying the definition of uniform convergence with tolerance $1/2$, there is M such that $|f_n(x)| \leq 1/2$ for all $n \geq M$ and all $x \notin E$. Since $f_n \equiv 1$ on $[n, n+1]$, this forces $[n, n+1] \subset E$ for every $n \geq M$, whence by monotonicity

$$\mathrm{Leb}(E) \geq \mathrm{Leb}\Big(\bigcup_{n \geq M} [n, n+1] \Big) = \mathrm{Leb}([M, \infty)) = \infty,$$

contradicting $\mathrm{Leb}(E) \leq 1/2$. So no such E exists, and Egorov's theorem genuinely fails without the finiteness hypothesis. (The obstruction is exactly the failure of downward monotone convergence used above: with $m = 2$, the set $E_{M,2}$ contains $[M, \infty)$ and so has infinite measure for every M, even though the sets $E_{M,2}$ decrease to the empty set.)

Problem 1.4.32. Propose a definition for the simple integral for absolutely convergent complex-valued functions on a measurable space with a finite σ-algebra.

Solution. The proposed definition. Let $(X, \mathcal{B}, \mu)$ be a measure space with $\mathcal{B}$ finite. By Exercise 1.4.4, X is partitioned into finitely many non-empty atoms $A_1, \ldots, A_n$ of $\mathcal{B}$. Let $f : X \to \mathbb{C}$ be measurable; by Exercise 1.4.30 it has a unique representation

$$f = \sum_{i=1}^{n} c_i 1_{A_i}, \qquad c_1, \ldots, c_n \in \mathbb{C}.$$

Call f **absolutely convergent** (or absolutely integrable) if the unsigned simple integral of $|f|$ is finite, i.e.

$$\mathrm{Simp} \int_X |f| \, d\mu = \sum_{i=1}^{n} |c_i| \mu(A_i) < \infty,$$

which, since the sum is finite and all terms are unsigned, happens if and only if $c_i = 0$ for every i with $\mu(A_i) = \infty$. For such f we define

$$\mathrm{Simp} \int_X f \, d\mu := \sum_{i=1}^{n} c_i \mu(A_i),$$

with the convention $0 \cdot \infty = 0$. The absolute convergence hypothesis guarantees that every term is a well-defined complex number (a term $c_i \mu(A_i)$ with $\mu(A_i) = \infty$ has $c_i = 0$ and so equals 0), so the sum is a well-defined element of $\mathbb{C}$.
Well-definedness. The definition does not depend on the enumeration of the atoms: by Exercise 1.4.3 the non-empty atoms of a finite Boolean algebra are determined up to relabelling, and both $\sum_i |c_i| \mu(A_i)$

and $\sum_i c_i \mu(A_i)$ are invariant under permuting the index set. The coefficients c_i themselves are uniquely determined by f, again by Exercise 1.4.30.

Consistency with the unsigned integral. If f happens to take values in $[0, +\infty)$ and is absolutely convergent, then all $c_i \in [0, \infty)$ and the formula above is literally the formula of Definition 1.4.34. So the two notions agree wherever both are defined, and $\mathrm{Simp} \int_X f \, d\mu$ is then finite.

Consistency with the real/imaginary decomposition. This is the definition one would also arrive at by mimicking Section 1.3: for real-valued absolutely convergent f one sets

$$\mathrm{Simp} \int_X f \, d\mu := \mathrm{Simp} \int_X f_+ \, d\mu - \mathrm{Simp} \int_X f_- \, d\mu,$$

and for complex-valued $f = u + iv$,

$$\mathrm{Simp} \int_X f \, d\mu := \mathrm{Simp} \int_X u \, d\mu + i \, \mathrm{Simp} \int_X v \, d\mu.$$

These agree with the formula proposed above. Indeed, write $c_i = a_i + ib_i$; then by uniqueness in Exercise 1.4.30, $u = \sum_i a_i 1_{A_i}$ and $v = \sum_i b_i 1_{A_i}$, and $u_\pm = \sum_i (a_i)_\pm 1_{A_i}$. Absolute convergence gives $|a_i|, |b_i| \leq |c_i|$, hence $a_i = b_i = 0$ on every atom of infinite measure, so all four unsigned integrals $\mathrm{Simp} \int u_\pm$, $\mathrm{Simp} \int v_\pm$ are finite and

$$\mathrm{Simp} \int_X u \, d\mu = \sum_i (a_i)_+ \mu(A_i) - \sum_i (a_i)_- \mu(A_i) = \sum_i a_i \mu(A_i),$$

the cancellation being legitimate because every term is finite. Similarly for v, and adding gives $\sum_i (a_i + ib_i)\mu(A_i) = \sum_i c_i \mu(A_i)$.

Basic properties. With this definition the following are immediate from the formula, using the uniqueness of the atomic representation (Exercise 1.4.30) and the fact that if $f = \sum_i c_i 1_{A_i}$ and $g = \sum_i d_i 1_{A_i}$ then $f + g = \sum_i (c_i + d_i) 1_{A_i}$ and $cf = \sum_i (cc_i) 1_{A_i}$:

- (Complex linearity) If f, g are absolutely convergent and $c \in \mathbb{C}$, then so are $f + g$ and cf (since $|c_i + d_i| \leq |c_i| + |d_i|$ and $|cc_i| = |c||c_i|$), and
$$\mathrm{Simp} \int_X (f + g) \, d\mu = \mathrm{Simp} \int_X f \, d\mu + \mathrm{Simp} \int_X g \, d\mu, \qquad \mathrm{Simp} \int_X cf \, d\mu = c \, \mathrm{Simp} \int_X f \, d\mu.$$
- (Compatibility with measure) For $\mathcal{B}$-measurable E of finite measure, $\mathrm{Simp} \int_X 1_E \, d\mu = \mu(E)$, since E is a union of atoms.
- (Triangle inequality) $\left|\mathrm{Simp} \int_X f \, d\mu\right| \leq \sum_{i=1}^n |c_i| \mu(A_i) = \mathrm{Simp} \int_X |f| \, d\mu$, by the triangle inequality in $\mathbb{C}$.

Finally, exactly as in Exercise 1.4.33 for the unsigned case, this integral is unchanged on passing to a finite refinement $(X, \mathcal{B}', \mu')$: each atom A_i of $\mathcal{B}$ splits into finitely many atoms $A'_{i,1}, \ldots, A'_{i,k_i}$ of $\mathcal{B}'$ on each of which f still takes the value c_i, and finite additivity of μ' gives $\sum_j \mu'(A'_{i,j}) = \mu'(A_i) = \mu(A_i)$; multiplying by c_i (all terms finite, by absolute convergence) and summing over i gives the claim. Hence the definition extends, exactly as in Definition 1.4.35, to absolutely convergent complex-valued simple functions on an arbitrary measure space, by integrating with respect to the restriction of μ to any finite sub-σ-algebra with respect to which f is measurable.

Problem 1.4.33. (Simple integral unaffected by refinements) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $(X, \mathcal{B}', \mu')$ be a refinement of $(X, \mathcal{B}, \mu)$, which means that $\mathcal{B}'$ contains $\mathcal{B}$ and $\mu' : \mathcal{B}' \to [0, +\infty]$ agrees with $\mu : \mathcal{B} \to [0, +\infty]$ on $\mathcal{B}$. Suppose that both $\mathcal{B}, \mathcal{B}'$ are finite, and let $f : X \to [0, +\infty]$ be $\mathcal{B}$-measurable. Show that

$$\mathrm{Simp} \int_X f \, d\mu = \mathrm{Simp} \int_X f \, d\mu'.$$

Solution. By Exercise 1.4.4 the finite Boolean algebras $\mathcal{B}$ and $\mathcal{B}'$ are atomic: let $A_1, \ldots, A_n$ be the non-empty atoms of $\mathcal{B}$ and $A'_1, \ldots, A'_m$ the non-empty atoms of $\mathcal{B}'$, so that X is partitioned by the A_i

and also by the A'_j.
Step 1: each $\mathcal{B}'$-atom sits inside exactly one $\mathcal{B}$-atom. Fix j and pick $x \in A'_j$ (possible, the atom being non-empty). There is a unique i with $x \in A_i$. Now $A_i \in \mathcal{B} \subset \mathcal{B}'$, so A_i is a union of $\mathcal{B}'$-atoms, and it meets A'_j (at x); by the argument recorded in the solution to Exercise 1.4.30 (a member of an atomic algebra that meets an atom must contain it), $A'_j \subset A_i$. Uniqueness of i is clear since the A_i are disjoint and $A'_j \neq \emptyset$.
Consequently, setting $J_i := \{j \in \{1,\dots,m\} : A'_j \subset A_i\}$, the sets $J_1,\dots,J_n$ partition $\{1,\dots,m\}$, and since the A'_j cover X we get the disjoint decomposition

$$A_i = \bigcup_{j \in J_i} A'_j \qquad (i = 1,\dots,n).$$

Step 2: the two atomic representations of f. Since f is $\mathcal{B}$-measurable, Exercise 1.4.30 gives a unique representation $f = \sum_{i=1}^n c_i 1_{A_i}$ with $c_i \in [0,+\infty]$, namely c_i is the constant value of f on A_i. Because $\mathcal{B} \subset \mathcal{B}'$, every preimage $f^{-1}(U) \in \mathcal{B}$ also lies in $\mathcal{B}'$, so f is $\mathcal{B}'$-measurable as well, and its $\mathcal{B}'$-atomic representation is $f = \sum_{j=1}^m c'_j 1_{A'_j}$ where c'_j is the constant value of f on A'_j. By Step 1, if $j \in J_i$ then $A'_j \subset A_i$ and hence $c'_j = c_i$.
Step 3: computation. A countably additive measure is finitely additive (pad a finite union out to a countable one with empty sets), so applying finite additivity of μ' to the disjoint decomposition of A_i from Step 1 gives $\sum_{j \in J_i} \mu'(A'_j) = \mu'(A_i)$. Also $\mu'(A_i) = \mu(A_i)$, since $A_i \in \mathcal{B}$ and μ' agrees with μ on $\mathcal{B}$. Therefore, using Definition 1.4.34 twice,

$$\begin{aligned}\operatorname{Simp}\int_X f\,d\mu' &= \sum_{j=1}^m c'_j \mu'(A'_j) = \sum_{i=1}^n \sum_{j \in J_i} c_i \mu'(A'_j) \\ &= \sum_{i=1}^n c_i \Big(\sum_{j \in J_i} \mu'(A'_j) \Big) = \sum_{i=1}^n c_i \mu'(A_i) = \sum_{i=1}^n c_i \mu(A_i) = \operatorname{Simp}\int_X f\,d\mu.\end{aligned}$$

The pulling of c_i out of the inner sum is legitimate in the unsigned arithmetic of $[0,+\infty]$, where $c(a+b) = ca + cb$ holds for all $c,a,b \in [0,+\infty]$ under the convention $0 \cdot \infty = 0$: if $c = 0$ both sides are 0; if $c = \infty$ and $a = b = 0$ both sides are 0; and if $c > 0$ with at least one of a, b positive, or c finite and positive, the identity is the usual distributive law (both sides being $+\infty$ exactly when $c > 0$ and $a + b = \infty$).
Finally, note (as Tao remarks after Definition 1.4.34) that the answer does not depend on the labelling of the atoms, by Exercise 1.4.3, so both simple integrals are unambiguously defined and we have shown them to be equal.

Problem 1.4.34. (Basic properties of the simple integral) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to [0,+\infty]$ be simple functions.

- (i) (Monotonicity) If $f \leq g$ pointwise, then $\operatorname{Simp}\int_X f\,d\mu \leq \operatorname{Simp}\int_X g\,d\mu$.
- (ii) (Compatibility with measure) For every $\mathcal{B}$-measurable set E, we have $\operatorname{Simp}\int_X 1_E\,d\mu = \mu(E)$.
- (iii) (Homogeneity) For every $c \in [0,+\infty]$, one has $\operatorname{Simp}\int_X cf\,d\mu = c \times \operatorname{Simp}\int_X f\,d\mu$.
- (iv) (Finite additivity) $\operatorname{Simp}\int_X (f+g)\,d\mu = \operatorname{Simp}\int_X f\,d\mu + \operatorname{Simp}\int_X g\,d\mu$.
- (v) (Insensitivity to refinement) If $(X, \mathcal{B}', \mu')$ is a refinement of $(X, \mathcal{B}, \mu)$ (as defined in Exercise 1.4.33), then $\operatorname{Simp}\int_X f\,d\mu = \operatorname{Simp}\int_X f\,d\mu'$.
- (vi) (Almost everywhere equivalence) If $f(x) = g(x)$ for μ-almost every $x \in X$, then $\operatorname{Simp}\int_X f\,d\mu = \operatorname{Simp}\int_X g\,d\mu$.
- (vii) (Finiteness) $\operatorname{Simp}\int_X f\,d\mu < \infty$ if and only if f is finite almost everywhere, and is supported on a set of finite measure.
- (viii) (Vanishing) $\operatorname{Simp}\int_X f\,d\mu = 0$ if and only if f is zero almost everywhere.

Solution. A common finite algebra. We begin with a lemma that reduces everything to a single finite sub-σ-algebra.

Lemma. Let $f_1, \dots, f_r : X \to [0, +\infty]$ be simple functions on $(X, \mathcal{B})$. Then there is a finite sub-σ-algebra $\mathcal{B}'' \subset \mathcal{B}$ with respect to which every f_s is measurable, and for each s, if $A_1, \dots, A_n$ are the non-empty atoms of $\mathcal{B}''$ and $f_s = \sum_{i=1}^{n} c_i^{(s)} 1_{A_i}$ is the representation from Exercise 1.4.30, then

$$\operatorname{Simp} \int_X f_s \, d\mu = \sum_{i=1}^{n} c_i^{(s)} \mu(A_i).$$

Proof. As in Definition 1.4.35, each f_s takes finitely many values $a_1^{(s)}, \dots, a_{k_s}^{(s)}$ and is measurable with respect to the finite σ-algebra $\mathcal{B}_s := \langle f_s^{-1}(\{a_1^{(s)}\}), \dots, f_s^{-1}(\{a_{k_s}^{(s)}\}) \rangle \subset \mathcal{B}$ (these preimages lie in $\mathcal{B}$ by Exercise 1.4.29(iii), singletons being Borel). Let $\mathcal{B}'' := \langle \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r \rangle \subset \mathcal{B}$. Each $\mathcal{B}_s$ is finite by Exercise 1.4.8, so $\mathcal{B}''$ is generated by a finite family of sets and is therefore a finite Boolean algebra, again by Exercise 1.4.8; a finite Boolean algebra is automatically a σ-algebra, since a countable union of its members takes only finitely many distinct values and so reduces to a finite union. Each f_s is $\mathcal{B}''$-measurable since $\mathcal{B}_s \subset \mathcal{B}''$. Moreover $(X, \mathcal{B}'', \mu \restriction_{\mathcal{B}''})$ is a refinement of $(X, \mathcal{B}_s, \mu \restriction_{\mathcal{B}_s})$ in the sense of Exercise 1.4.33, and both algebras are finite, so by that exercise the simple integrals of f_s computed in the two agree. By Definition 1.4.35 the left-hand side above may thus be computed in $\mathcal{B}''$, and Definition 1.4.34 gives the stated formula (the atoms A_i lie in $\mathcal{B}'' \subset \mathcal{B}$, so $\mu(A_i)$ makes sense). $\square$
Throughout we use the arithmetic of $[0, +\infty]$ with the convention $0 \cdot \infty = 0$; as noted in Exercise 1.4.33 this arithmetic is commutative, associative and distributive, and multiplication is monotone.
(i) Monotonicity. Apply the Lemma to f, g, obtaining atoms $A_1, \dots, A_n$ and $f = \sum_i c_i 1_{A_i}$, $g = \sum_i d_i 1_{A_i}$. Each A_i is non-empty; picking $x \in A_i$ and using $f(x) \le g(x)$ gives $c_i \le d_i$. Hence $c_i \mu(A_i) \le d_i \mu(A_i)$ for every i, and summing,

$$\operatorname{Simp} \int_X f \, d\mu = \sum_i c_i \mu(A_i) \le \sum_i d_i \mu(A_i) = \operatorname{Simp} \int_X g \, d\mu.$$

(ii) Compatibility with measure. Let $E \in \mathcal{B}$. By Exercise 1.4.29(ii) the function 1_E is measurable, and it takes at most two values, so it is simple. If $E \neq \emptyset$ and $E \neq X$, then $\mathcal{B}' := \{\emptyset, E, X \setminus E, X\}$ is a finite σ-algebra contained in $\mathcal{B}$ with non-empty atoms $E, X \setminus E$, and 1_E is $\mathcal{B}'$-measurable with representation $1_E = 1 \cdot 1_E + 0 \cdot 1_{X \setminus E}$. So by Definitions 1.4.35 and 1.4.34,

$$\operatorname{Simp} \int_X 1_E \, d\mu = 1 \cdot \mu(E) + 0 \cdot \mu(X \setminus E) = \mu(E),$$

the second term being 0 even if $\mu(X \setminus E) = \infty$. In the degenerate cases we may use the trivial algebra $\{\emptyset, X\}$, whose single atom is X: if $E = \emptyset$ then $1_E = 0 \cdot 1_X$ and the integral is $0 = \mu(\emptyset)$; if $E = X$ then $1_E = 1 \cdot 1_X$ and the integral is $\mu(X)$. (If $X = \emptyset$ everything is trivially 0.)
(iii) Homogeneity. Let $c \in [0, +\infty]$. Apply the Lemma to f: $f = \sum_i c_i 1_{A_i}$ on the atoms of some finite $\mathcal{B}'' \subset \mathcal{B}$. Then $cf = \sum_i (cc_i) 1_{A_i}$ is constant on each atom, hence $\mathcal{B}''$-measurable by Exercise 1.4.30, and takes finitely many values, so it is simple; by Exercise 1.4.33 and Definition 1.4.35 its integral may be computed in $\mathcal{B}''$. Using associativity and distributivity of unsigned arithmetic,

$$\operatorname{Simp} \int_X cf \, d\mu = \sum_i (cc_i) \mu(A_i) = \sum_i c \left(c_i \mu(A_i) \right) = c \sum_i c_i \mu(A_i) = c \times \operatorname{Simp} \int_X f \, d\mu.$$

(iv) Finite additivity. Apply the Lemma to f, g simultaneously: $f = \sum_i c_i 1_{A_i}$, $g = \sum_i d_i 1_{A_i}$ on the atoms of a common finite $\mathcal{B}'' \subset \mathcal{B}$. Then $f + g = \sum_i (c_i + d_i) 1_{A_i}$ is constant on each atom and takes finitely many values, so it is a simple function, $\mathcal{B}''$-measurable, and

$$\operatorname{Simp} \int_X (f + g) \, d\mu = \sum_i (c_i + d_i) \mu(A_i) = \sum_i c_i \mu(A_i) + \sum_i d_i \mu(A_i) = \operatorname{Simp} \int_X f \, d\mu + \operatorname{Simp} \int_X g \, d\mu,$$

using distributivity in $[0, +\infty]$ and rearrangement of a finite sum of unsigned quantities.
(v) Insensitivity to refinement. Let $\mathcal{B}_f \subset \mathcal{B}$ be a finite σ-algebra with respect to which f is measurable, as in Definition 1.4.35. Since $\mathcal{B} \subset \mathcal{B}'$, we also have $\mathcal{B}_f \subset \mathcal{B}'$, so f is a simple function on $(X, \mathcal{B}')$ too and

its μ'-integral may likewise be computed in $\mathcal{B}_f$. But μ' agrees with μ on $\mathcal{B}$, hence on the subcollection $\mathcal{B}_f$, so $\mu' \restriction_{\mathcal{B}_f} = \mu \restriction_{\mathcal{B}_f}$ and the two computations are literally the same:

$$\text{Simp} \int_X f \, d\mu = \text{Simp} \int_X f \, d(\mu \restriction_{\mathcal{B}_f}) = \text{Simp} \int_X f \, d(\mu' \restriction_{\mathcal{B}_f}) = \text{Simp} \int_X f \, d\mu'.$$

(That each of these is independent of the choice of $\mathcal{B}_f$ is exactly the content of Exercise 1.4.33.)
(vi) Almost everywhere equivalence. By definition of "almost everywhere", the set $\{x : f(x) \neq g(x)\}$ is a sub-null set, so there is $N \in \mathcal{B}$ with $\mu(N) = 0$ and $f = g$ on $X \setminus N$. Apply the Lemma to f, g to get a finite $\mathcal{B}'' \subset \mathcal{B}$ with non-empty atoms $A_1, \ldots, A_n$ on which both are constant. Now refine: the non-empty sets among $A_i \cap N$ and $A_i \setminus N$ form a finite partition of X into sets belonging to $\mathcal{B}$, and the atomic algebra $\mathcal{B}'''$ they generate is a finite σ-algebra (Exercise 1.4.10) contained in $\mathcal{B}$ and refining $\mathcal{B}''$. Both f and g are constant on each of these smaller atoms (being constant on the A_i), hence $\mathcal{B}'''$-measurable by Exercise 1.4.30, and by Exercise 1.4.33 their integrals may be computed in $\mathcal{B}'''$.
Write the atoms of $\mathcal{B}'''$ as $B_1, \ldots, B_k$, so that $f = \sum_j c'_j 1_{B_j}$ and $g = \sum_j d'_j 1_{B_j}$; by construction each B_j is either contained in N or disjoint from N. If $B_j \subset X \setminus N$, then $f = g$ on B_j and hence $c'_j = d'_j$ (the atom being non-empty). If $B_j \subset N$, then $\mu(B_j) \leq \mu(N) = 0$ by monotonicity, so $c'_j \mu(B_j) = 0 = d'_j \mu(B_j)$ even when the coefficients are $+\infty$. Summing term by term,

$$\text{Simp} \int_X f \, d\mu = \sum_j c'_j \mu(B_j) = \sum_j d'_j \mu(B_j) = \text{Simp} \int_X g \, d\mu.$$

(vii) Finiteness. Apply the Lemma to f alone: $f = \sum_{i=1}^n c_i 1_{A_i}$ on the non-empty atoms of a finite $\mathcal{B}'' \subset \mathcal{B}$, and $\text{Simp} \int_X f \, d\mu = \sum_i c_i \mu(A_i)$. A finite sum of unsigned quantities is finite if and only if each term is, so

$$\text{Simp} \int_X f \, d\mu < \infty \iff c_i \mu(A_i) < \infty \text{ for all } i.$$

Note also that $\{f = \infty\} = \bigcup_{i : c_i = \infty} A_i$ and $\{f \neq 0\} = \bigcup_{i : c_i \neq 0} A_i$ are disjoint unions of atoms, hence lie in $\mathcal{B}$, with $\mu(\{f = \infty\}) = \sum_{i : c_i = \infty} \mu(A_i)$ and $\mu(\{f \neq 0\}) = \sum_{i : c_i \neq 0} \mu(A_i)$ by finite additivity.
Suppose the integral is finite. If $c_i = \infty$ then $\infty \cdot \mu(A_i) < \infty$ forces $\mu(A_i) = 0$; summing, $\mu(\{f = \infty\}) = 0$, i.e. f is finite almost everywhere. If $c_i \neq 0$ then $c_i > 0$ and $c_i \mu(A_i) < \infty$ forces $\mu(A_i) < \infty$; summing over the finitely many such i, $\mu(\{f \neq 0\}) < \infty$, so f is supported on a set of finite measure.
Conversely, suppose f is finite almost everywhere and vanishes outside a set $F \in \mathcal{B}$ with $\mu(F) < \infty$; then $\{f \neq 0\} \subset F$ has finite measure, and $\{f = \infty\}$, being a $\mathcal{B}$-measurable subset of a null set, has measure 0. Fix i. If $c_i = 0$ the term $c_i \mu(A_i)$ is 0. If $0 < c_i < \infty$, then $A_i \subset \{f \neq 0\}$ so $\mu(A_i) \leq \mu(\{f \neq 0\}) < \infty$ and the term is finite. If $c_i = \infty$, then $A_i \subset \{f = \infty\}$ so $\mu(A_i) = 0$ and the term is $\infty \cdot 0 = 0$. Hence every term is finite and so is the integral.
(viii) Vanishing. With the same notation, $\sum_i c_i \mu(A_i) = 0$ if and only if $c_i \mu(A_i) = 0$ for every i, i.e. if and only if for each i either $c_i = 0$ or $\mu(A_i) = 0$. If this holds, then $\mu(\{f \neq 0\}) = \sum_{i : c_i \neq 0} \mu(A_i) = 0$, so f vanishes almost everywhere. Conversely, if $f = 0$ almost everywhere, then $\{f \neq 0\}$ is a $\mathcal{B}$-measurable subset of a null set and hence $\mu(\{f \neq 0\}) = 0$; for each i with $c_i \neq 0$ we have $A_i \subset \{f \neq 0\}$, so $\mu(A_i) = 0$ by monotonicity, and every term $c_i \mu(A_i)$ vanishes. Thus $\text{Simp} \int_X f \, d\mu = 0$.

Problem 1.4.35. (Inclusion-exclusion principle) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $A_1, \ldots, A_n$ be $\mathcal{B}$-measurable sets of finite measure. Show that

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{J \subset \{1,\ldots,n\} : J \neq \emptyset} (-1)^{|J|-1} \mu\left(\bigcap_{i \in J} A_i\right).$$

(Hint: Compute $\text{Simp} \int_X (1 - \prod_{i=1}^n (1 - 1_{A_i})) \, d\mu$ in two different ways.)

Solution. Write $A := \bigcup_{i=1}^n A_i$ and, for non-empty $J \subset \{1,\dots,n\}$, $A_J := \bigcap_{i\in J} A_i$. All these sets lie in $\mathcal{B}$, and all have finite measure: for $J \neq \emptyset$ pick $i_0 \in J$, then $A_J \subset A_{i_0}$ gives $\mu(A_J) \leq \mu(A_{i_0}) < \infty$ by monotonicity, while finite subadditivity gives $\mu(A) \leq \sum_{i=1}^n \mu(A_i) < \infty$. This finiteness will let us subtract at the end.
Step 1: the pointwise identity. We claim that as functions on X,

$$1 - \prod_{i=1}^n \left(1 - 1_{A_i}\right) = 1_A = \sum_{J\neq\emptyset} (-1)^{|J|-1} 1_{A_J}.$$

For the first equality, fix $x \in X$: the product $\prod_i (1 - 1_{A_i}(x))$ is a product of factors each equal to 0 or 1, and it equals 1 precisely when $x \notin A_i$ for every i, i.e. precisely when $x \notin A$; otherwise it is 0. So the product equals $1_{X\setminus A}(x)$, and $1 - 1_{X\setminus A} = 1_A$.
For the second equality, expand the product multilinearly:

$$\prod_{i=1}^n \left(1 - 1_{A_i}\right) = \sum_{J\subset\{1,\dots,n\}} (-1)^{|J|} \prod_{i\in J} 1_{A_i} = \sum_{J\subset\{1,\dots,n\}} (-1)^{|J|} 1_{A_J},$$

using $\prod_{i\in J} 1_{A_i} = 1_{\bigcap_{i\in J} A_i}$ and the convention that the empty intersection is X, so that the $J = \emptyset$ term is $1_X = 1$. Subtracting the $J = \emptyset$ term and negating gives

$$1 - \prod_{i=1}^n (1 - 1_{A_i}) = -\sum_{J\neq\emptyset} (-1)^{|J|} 1_{A_J} = \sum_{J\neq\emptyset} (-1)^{|J|-1} 1_{A_J},$$

as claimed. (Equivalently, one can verify the identity directly at a point x: with $S := \{i : x \in A_i\}$ and $s := |S|$, the right-hand side equals $\sum_{\emptyset\neq J\subset S} (-1)^{|J|-1} = 1 - \sum_{J\subset S} (-1)^{|J|} = 1 - (1-1)^s$, which is 1 if $s \geq 1$ and 0 if $s = 0$, matching $1_A(x)$.)
Step 2: separating the signs. To integrate the identity we avoid signed integrals by moving the negative terms to the other side. Define the unsigned functions

$$P := \sum_{J\neq\emptyset,\ |J| \text{ odd}} 1_{A_J}, \qquad N := \sum_{J\neq\emptyset,\ |J| \text{ even}} 1_{A_J}.$$

Each 1_{A_J} is measurable by Exercise 1.4.29(ii), so P and N are measurable by Exercise 1.4.29(viii), and each takes only finitely many values (they are sums of at most 2^n indicator functions, so take values in $\{0, 1, \dots, 2^n\}$). Hence P and N are simple functions in the sense of Definition 1.4.35. Step 1 rearranges to the pointwise identity of unsigned functions

$$1_A + N = P.$$

Step 3: integrating. Apply the simple integral to both sides. By finite additivity (Exercise 1.4.34(iv), applied repeatedly) and compatibility with measure (Exercise 1.4.34(ii)),

$$\text{Simp}\int_X (1_A + N)\, d\mu = \mu(A) + \sum_{J\neq\emptyset,\ |J| \text{ even}} \mu(A_J), \qquad \text{Simp}\int_X P\, d\mu = \sum_{J\neq\emptyset,\ |J| \text{ odd}} \mu(A_J).$$

Since $1_A + N = P$ as functions, these two quantities are equal:

$$\mu(A) + \sum_{J\neq\emptyset,\ |J| \text{ even}} \mu(A_J) = \sum_{J\neq\emptyset,\ |J| \text{ odd}} \mu(A_J).$$

Every term appearing here is finite, as observed at the outset, and there are finitely many of them; so we may subtract the even-cardinality sum from both sides to obtain

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{J\neq\emptyset,\ |J| \text{ odd}} \mu(A_J) - \sum_{J\neq\emptyset,\ |J| \text{ even}} \mu(A_J) = \sum_{J\subset\{1,\dots,n\}: J\neq\emptyset} (-1)^{|J|-1} \mu\left(\bigcap_{i\in J} A_i\right),$$

which is the inclusion-exclusion principle.
Remark (the hint, phrased with a signed integral). If one has the absolutely convergent simple integral of Exercise 1.4.32 available, the two computations of Simp $\int_X (1 - \prod_i (1 - 1_{A_i}))\, d\mu$ are exactly Steps 1 and 3 above: by the first form of the pointwise identity the integrand is 1_A, whose integral is $\mu(A)$; by the second form it is $\sum_{J \neq \emptyset} (-1)^{|J|-1} 1_{A_J}$, whose integral is $\sum_{J \neq \emptyset} (-1)^{|J|-1} \mu(A_J)$ by linearity. All the functions involved are absolutely convergent, since they are bounded by 2^n and supported in A, a set of finite measure. Equating the two gives the identity. Step 2 above is just the standard device for deducing this from the unsigned theory alone.

Problem 1.4.36. (Easy properties of the unsigned integral) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to [0, +\infty]$ be measurable.

- (i) (Almost everywhere equivalence) If $f = g$ μ-almost everywhere, then $\int_X f\, d\mu = \int_X g\, d\mu$.
- (ii) (Monotonicity) If $f \leq g$ μ-almost everywhere, then $\int_X f\, d\mu \leq \int_X g\, d\mu$.
- (iii) (Homogeneity) We have $\int_X cf\, d\mu = c \int_X f\, d\mu$ for every $c \in [0, +\infty]$.
- (iv) (Superadditivity) We have $\int_X (f + g)\, d\mu \geq \int_X f\, d\mu + \int_X g\, d\mu$.
- (v) (Compatibility with the simple integral) If f is simple, then $\int_X f\, d\mu = \text{Simp} \int_X f\, d\mu$.
- (vi) (Markov's inequality) For any $0 < \lambda < \infty$, one has
$$\mu(\{x \in X : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_X f\, d\mu.$$

In particular, if $\int_X f\, d\mu < \infty$, then the sets $\{x \in X : f(x) \geq \lambda\}$ have finite measure for each $\lambda > 0$.

- (vii) (Finiteness) If $\int_X f\, d\mu < \infty$, then $f(x)$ is finite for μ-almost every x.
- (viii) (Vanishing) If $\int_X f\, d\mu = 0$, then $f(x)$ is zero for μ-almost every x.
- (ix) (Vertical truncation) We have $\lim_{n \to \infty} \int_X \min(f, n)\, d\mu = \int_X f\, d\mu$.
- (x) (Horizontal truncation) If $E_1 \subset E_2 \subset \ldots$ is an increasing sequence of $\mathcal{B}$-measurable sets, then
$$\lim_{n \to \infty} \int_X f 1_{E_n}\, d\mu = \int_X f 1_{\bigcup_{n=1}^\infty E_n}\, d\mu.$$
- (xi) (Restriction) If Y is a measurable subset of X, then $\int_X f 1_Y\, d\mu = \int_Y f \restriction_Y\, d\mu \restriction_Y$, where $f \restriction_Y : Y \to [0, +\infty]$ is the restriction of $f : X \to [0, +\infty]$ to Y, and the restriction $\mu \restriction_Y$ was defined in Example 1.4.25. We will often abbreviate $\int_Y f \restriction_Y d\mu \restriction_Y$ (by slight abuse of notation) as $\int_Y f\, d\mu$.

Solution. Throughout, "simple" means as in Definition 1.4.35: a measurable function $X \to [0, +\infty]$ taking only finitely many values. Recall Definition 1.4.37:

$$\int_X f\, d\mu = \sup_{0 \leq h \leq f;\ h \text{ simple}} \text{Simp} \int_X h\, d\mu,$$

the supremum being over all simple h with $0 \leq h \leq f$ pointwise. We record two closure facts used repeatedly: if h, h' are simple and E is $\mathcal{B}$-measurable and $c \in [0, +\infty]$, then $h + h'$, ch and $h1_E$ are again simple, since each is measurable and takes only finitely many values.
We prove the parts in an order that avoids circularity: (ii), (i), (v), (vi), (vii), (viii), (iii), (iv).
(ii) Monotonicity. Suppose $f \leq g$ outside a sub-null set; that is, there is a $\mathcal{B}$-measurable set N with $\mu(N) = 0$ such that $f(x) \leq g(x)$ for all $x \in X \setminus N$. Let h be simple with $0 \leq h \leq f$, and put $h' := h1_{X \setminus N}$, which is again simple. For $x \notin N$ we have $h'(x) = h(x) \leq f(x) \leq g(x)$, and for $x \in N$ we have $h'(x) = 0 \leq g(x)$; so $h' \leq g$ pointwise, and therefore h' is one of the competitors in the supremum defining $\int_X g\, d\mu$, giving

$$\text{Simp} \int_X h'\, d\mu \leq \int_X g\, d\mu.$$

On the other hand $h = h'$ outside N, hence $h = h'$ μ-almost everywhere, so by Exercise 1.4.34(vi) Simp $\int_X h\, d\mu =$ Simp $\int_X h'\, d\mu$. Combining, Simp $\int_X h\, d\mu \leq \int_X g\, d\mu$ for every simple $0 \leq h \leq f$, and taking the supremum over such h gives $\int_X f\, d\mu \leq \int_X g\, d\mu$.
(i) Almost everywhere equivalence. If $f = g$ μ-almost everywhere then both $f \leq g$ and $g \leq f$ hold μ-almost everywhere, so (ii) applied twice gives $\int_X f\, d\mu \leq \int_X g\, d\mu \leq \int_X f\, d\mu$.

(v) Compatibility with the simple integral. Let f be simple. Since f is itself a competitor in the supremum (as $0 \le f \le f$), we get $\int_X f\,d\mu \ge \operatorname{Simp}\int_X f\,d\mu$. Conversely, if h is simple with $0 \le h \le f$ pointwise, then monotonicity of the simple integral (Exercise 1.4.34(i)) gives $\operatorname{Simp}\int_X h\,d\mu \le \operatorname{Simp}\int_X f\,d\mu$; taking the supremum over h yields $\int_X f\,d\mu \le \operatorname{Simp}\int_X f\,d\mu$. Hence the two agree.
(vi) Markov's inequality. Fix $0 < \lambda < \infty$ and set $E := \{x \in X : f(x) \ge \lambda\}$. This set is $\mathcal{B}$-measurable: by Exercise 1.4.29(i) each level set $\{f > \lambda - 1/m\}$ ($m \ge 1$) is $\mathcal{B}$-measurable, and $E = \bigcap_{m=1}^{\infty}\{x \in X : f(x) > \lambda - 1/m\}$, a countable intersection of $\mathcal{B}$-measurable sets. The function $h := \lambda 1_E$ is simple and satisfies $0 \le h \le f$ pointwise: on E we have $h = \lambda \le f$, and off E we have $h = 0 \le f$. By Exercise 1.4.34(iii) and (ii),

$$\operatorname{Simp}\int_X \lambda 1_E\,d\mu = \lambda \operatorname{Simp}\int_X 1_E\,d\mu = \lambda\mu(E),$$

and by Definition 1.4.37 this is at most $\int_X f\,d\mu$. Thus $\lambda\mu(E) \le \int_X f\,d\mu$, and dividing by the finite positive number λ gives $\mu(E) \le \frac{1}{\lambda}\int_X f\,d\mu$. In particular if $\int_X f\,d\mu < \infty$ then $\mu(\{f \ge \lambda\}) < \infty$ for every $\lambda > 0$.
(vii) Finiteness. Suppose $\int_X f\,d\mu < \infty$. The set $F := \{x \in X : f(x) = +\infty\}$ is measurable, and $F \subset \{f \ge \lambda\}$ for every $\lambda \in (0, \infty)$, so by (vi) and monotonicity of μ (Exercise 1.4.20(i)),

$$\mu(F) \le \frac{1}{\lambda}\int_X f\,d\mu \qquad \text{for all } \lambda > 0.$$

Letting $\lambda \to \infty$ and using $\int_X f\,d\mu < \infty$ forces $\mu(F) = 0$. So f is finite μ-almost everywhere.
(viii) Vanishing. Suppose $\int_X f\,d\mu = 0$. For each natural number $n \ge 1$, (vi) with $\lambda = 1/n$ gives $\mu(\{f \ge 1/n\}) \le n\int_X f\,d\mu = 0$. Since

$$\{x \in X : f(x) > 0\} = \bigcup_{n=1}^{\infty}\{x \in X : f(x) \ge 1/n\},$$

countable subadditivity (Exercise 1.4.23(i)) gives $\mu(\{f > 0\}) \le \sum_{n=1}^{\infty} 0 = 0$. Hence $f = 0$ μ-almost everywhere.
(iii) Homogeneity. We use the convention $0 \cdot \infty = \infty \cdot 0 = 0$.
Case $0 < c < \infty$. The map $h \mapsto ch$ is a bijection from the set of simple functions h with $0 \le h \le f$ onto the set of simple functions k with $0 \le k \le cf$, with inverse $k \mapsto c^{-1}k$ (multiplication by a finite positive constant preserves simplicity, and $h \le f$ if and only if $ch \le cf$, both being pointwise inequalities in $[0, +\infty]$). By Exercise 1.4.34(iii), $\operatorname{Simp}\int_X ch\,d\mu = c\operatorname{Simp}\int_X h\,d\mu$. Therefore

$$\int_X cf\,d\mu = \sup_{0\le h\le f;\ h \text{ simple}} c\operatorname{Simp}\int_X h\,d\mu = c\sup_{0\le h\le f;\ h \text{ simple}} \operatorname{Simp}\int_X h\,d\mu = c\int_X f\,d\mu,$$

where the middle equality is because multiplication by a fixed $c \in (0, \infty)$ is an order-isomorphism of $[0, +\infty]$ and hence commutes with suprema.
Case $c = 0$. Then cf is identically 0 (using $0 \cdot \infty = 0$), and the only simple h with $0 \le h \le 0$ is $h = 0$, whose simple integral is 0; so $\int_X cf\,d\mu = 0$. Also $c\int_X f\,d\mu = 0 \cdot \int_X f\,d\mu = 0$ by the same convention. The two agree.
Case $c = +\infty$. Here $cf(x) = +\infty$ when $f(x) > 0$ and $cf(x) = 0$ when $f(x) = 0$. Write $E := \{x : f(x) > 0\}$, a measurable set. If $\mu(E) = 0$ then $f = 0$ μ-almost everywhere, hence also $cf = 0$ μ-almost everywhere, and by (i) both $\int_X f\,d\mu$ and $\int_X cf\,d\mu$ equal $\int_X 0\,d\mu = 0$; so both sides of the claimed identity are 0 (using $\infty \cdot 0 = 0$). If instead $\mu(E) > 0$, then $\int_X f\,d\mu > 0$ (indeed $\int_X f\,d\mu = 0$ would force $\mu(E) = 0$ by (viii)), so the right-hand side $c\int_X f\,d\mu = +\infty$. For the left-hand side, for every natural number n the simple function $n1_E$ satisfies $n1_E \le cf$ pointwise, so

$$\int_X cf\,d\mu \ge \operatorname{Simp}\int_X n1_E\,d\mu = n\mu(E) \xrightarrow[n\to\infty]{} \infty,$$

whence $\int_X cf\,d\mu = +\infty$ as well.

(iv) Superadditivity. Let u, v be simple with $0 \le u \le f$ and $0 \le v \le g$. Then $u+v$ is simple and $0 \le u+v \le f+g$ pointwise, so by Definition 1.4.37 and finite additivity of the simple integral (Exercise 1.4.34(iv)),

$$\int_X (f+g)\,d\mu \ge \mathrm{Simp}\int_X (u+v)\,d\mu = \mathrm{Simp}\int_X u\,d\mu + \mathrm{Simp}\int_X v\,d\mu.$$

Since u and v range independently, we may take the supremum over u first and then over v (addition on $[0,+\infty]$ is continuous and increasing in each variable, so suprema pass through), obtaining

$$\int_X (f+g)\,d\mu \ge \sup_u \mathrm{Simp}\int_X u\,d\mu + \sup_v \mathrm{Simp}\int_X v\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu.$$

This proves superadditivity.

Before the last three parts we record a convenient formula for the simple integral. Let $h : X \to [0,+\infty]$ be simple and let $A_1, \dots, A_k \in \mathcal{B}$ be an arbitrary finite partition of X on each piece of which h is constant, say $h \equiv a_j$ on A_j. Then

$$\mathrm{Simp}\int_X h\,d\mu = \sum_{j=1}^{k} a_j\,\mu(A_j).$$

Indeed, let $\mathcal{B}_0 := \langle A_1, \dots, A_k \rangle$, a finite sub-$\sigma$-algebra of $\mathcal{B}$ whose atoms are exactly the nonempty A_j; the function h is $\mathcal{B}_0$-measurable, so by Definition 1.4.34 and Exercise 1.4.30 its simple integral against $\mu \restriction_{\mathcal{B}_0}$ is $\sum_j a_j \mu(A_j)$ (empty A_j contribute $a_j\mu(\emptyset) = 0$). By Definition 1.4.35, together with the observation recorded there (a consequence of Exercise 1.4.33) that any finite sub-σ-algebra with respect to which h is measurable computes the same value, this is $\mathrm{Simp}\int_X h\,d\mu$.

(ix) Vertical truncation. Since $\min(f,n) \le \min(f,n+1) \le f$ pointwise, part (ii) shows that the quantities $\int_X \min(f,n)\,d\mu$ are non-decreasing in n and bounded above by $\int_X f\,d\mu$; so the limit $L := \lim_{n\to\infty} \int_X \min(f,n)\,d\mu$ exists in $[0,+\infty]$ and $L \le \int_X f\,d\mu$.

For the converse, let h be simple with $0 \le h \le f$, with distinct values $a_1, \dots, a_k$ on the level sets $A_1, \dots, A_k$ (a finite measurable partition of X). For each n the function $h_n := \min(h,n)$ is simple, is constant with value $\min(a_j, n)$ on A_j, and satisfies $h_n \le \min(f,n)$ pointwise. Hence

$$\sum_{j=1}^{k} \min(a_j,n)\,\mu(A_j) = \mathrm{Simp}\int_X h_n\,d\mu \le \int_X \min(f,n)\,d\mu \le L.$$

Now let $n \to \infty$. For each fixed j we have $\min(a_j,n)\mu(A_j) \to a_j\mu(A_j)$ in $[0,+\infty]$: this is clear if $a_j < \infty$ and $\mu(A_j) < \infty$; if $a_j = 0$ both sides are 0; if $a_j > 0$ and $\mu(A_j) = \infty$ both sides are $+\infty$ for n large; and if $a_j = \infty$ then $\min(a_j,n)\mu(A_j) = n\mu(A_j)$, which tends to $+\infty = a_j\mu(A_j)$ when $\mu(A_j) > 0$ and is identically $0 = \infty \cdot 0$ when $\mu(A_j) = 0$. Since the sum is finite, we may pass to the limit termwise and obtain $\mathrm{Simp}\int_X h\,d\mu = \sum_j a_j\mu(A_j) \le L$. Taking the supremum over h gives $\int_X f\,d\mu \le L$, so the two are equal.

(x) Horizontal truncation. Write $E := \bigcup_{n=1}^{\infty} E_n$. Since $f1_{E_n} \le f1_{E_{n+1}} \le f1_E$ pointwise, part (ii) shows the integrals $\int_X f1_{E_n}\,d\mu$ are non-decreasing and bounded by $\int_X f1_E\,d\mu$, so $L := \lim_{n\to\infty} \int_X f1_{E_n}\,d\mu$ exists and $L \le \int_X f1_E\,d\mu$.

Conversely, let h be simple with $0 \le h \le f1_E$; note this forces $h = h1_E$, since h vanishes wherever $f1_E$ does. Let $a_1, \dots, a_k$ be the distinct values of h with level sets $A_1, \dots, A_k$. For each n, the function $h1_{E_n}$ is simple and is constant on each member of the finite measurable partition $\{A_1 \cap E_n, \dots, A_k \cap E_n, X \setminus E_n\}$, with values $a_1, \dots, a_k, 0$ respectively; hence

$$\mathrm{Simp}\int_X h1_{E_n}\,d\mu = \sum_{j=1}^{k} a_j\,\mu(A_j \cap E_n).$$

Also $h1_{E_n} \le f1_{E_n}$, so this quantity is at most $\int_X f1_{E_n}\,d\mu \le L$. Now $A_j \cap E_n$ increases to $A_j \cap E$ as $n \to \infty$, so $\mu(A_j \cap E_n) \to \mu(A_j \cap E)$ by upwards monotone convergence (Exercise 1.4.23(ii)); and then $a_j\mu(A_j \cap E_n) \to a_j\mu(A_j \cap E)$ in $[0,+\infty]$ (the only case needing comment is $a_j = \infty$: if $\mu(A_j \cap E) = 0$

then all $\mu(A_j \cap E_n) = 0$ and both sides vanish, while if $\mu(A_j \cap E) > 0$ then $\mu(A_j \cap E_n) > 0$ for large n and both sides are $+\infty$). Passing to the limit in the finite sum,

$$\sum_{j=1}^{k} a_j \, \mu(A_j \cap E) \leq L.$$

But the left-hand side is exactly $\mathrm{Simp} \int_X h1_E \, d\mu = \mathrm{Simp} \int_X h \, d\mu$, since $h = h1_E$. Taking the supremum over h gives $\int_X f1_E \, d\mu \leq L$, hence equality.
(xi) Restriction. Recall $\mathcal{B} \restriction_Y = \{E \in \mathcal{B} : E \subset Y\}$ and $\mu \restriction_Y$ is the restriction of μ to this σ-algebra (Example 1.4.25). We first note that extension by zero matches simple integrals: if $w : Y \to [0, +\infty]$ is simple on $(Y, \mathcal{B} \restriction_Y)$ and $\tilde{w} : X \to [0, +\infty]$ denotes its extension by 0 to X, then $\tilde{w}$ is simple on $(X, \mathcal{B})$ and

$$\mathrm{Simp} \int_X \tilde{w} \, d\mu = \mathrm{Simp} \int_Y w \, d\mu \restriction_Y .$$

Indeed, let $b_1, \ldots, b_r$ be the distinct values of w, with level sets $C_1, \ldots, C_r \subset Y$; these lie in $\mathcal{B} \restriction_Y \subset \mathcal{B}$, so $\tilde{w}$ is $\mathcal{B}$-measurable and takes finitely many values on X, being constant on each member of the finite $\mathcal{B}$-measurable partition $\{C_1, \ldots, C_r, X \setminus Y\}$ of X, with value 0 on $X \setminus Y$. Hence, by the partition formula recorded above applied on X and on Y respectively,

$$\mathrm{Simp} \int_X \tilde{w} \, d\mu = \sum_{i=1}^{r} b_i \mu(C_i) + 0 \cdot \mu(X \setminus Y) = \sum_{i=1}^{r} b_i \, \mu \restriction_Y (C_i) = \mathrm{Simp} \int_Y w \, d\mu \restriction_Y .$$

Conversely, if h is simple on X with $h = h1_Y$, then $h \restriction_Y$ is $\mathcal{B} \restriction_Y$-measurable (for Borel U, $(h \restriction_Y)^{-1}(U) = h^{-1}(U) \cap Y \in \mathcal{B} \restriction_Y$) with finitely many values, hence simple on Y, and h is its extension by zero, so the same identity applies.
Now: if h is simple on X with $0 \leq h \leq f1_Y$, then $h = h1_Y$ and $h \restriction_Y$ is simple on Y with $0 \leq h \restriction_Y \leq f \restriction_Y$, so

$$\mathrm{Simp} \int_X h \, d\mu = \mathrm{Simp} \int_Y h \restriction_Y \, d\mu \restriction_Y \leq \int_Y f \restriction_Y \, d\mu \restriction_Y;$$

taking suprema gives $\int_X f1_Y \, d\mu \leq \int_Y f \restriction_Y \, d\mu \restriction_Y$. In the other direction, if w is simple on Y with $0 \leq w \leq f \restriction_Y$, its extension $\tilde{w}$ by zero is simple on X with $0 \leq \tilde{w} \leq f1_Y$, so

$$\mathrm{Simp} \int_Y w \, d\mu \restriction_Y = \mathrm{Simp} \int_X \tilde{w} \, d\mu \leq \int_X f1_Y \, d\mu;$$

taking suprema gives the reverse inequality. Hence $\int_X f1_Y \, d\mu = \int_Y f \restriction_Y \, d\mu \restriction_Y$, completing the exercise.

Problem 1.4.37. (Linearity in μ) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f : X \to [0, +\infty]$ be measurable.

- (i) Show that $\int_X f \, d(c\mu) = c \times \int_X f \, d\mu$ for every $c \in [0, +\infty]$.
- (ii) If $\mu_1, \mu_2, \ldots$ are a sequence of measures on $\mathcal{B}$, show that

$$\int_X f \, d\Big(\sum_{n=1}^{\infty} \mu_n\Big) = \sum_{n=1}^{\infty} \int_X f \, d\mu_n.$$

Solution. By Exercise 1.4.22(i) the set function $c\mu$ is again a countably additive measure on $\mathcal{B}$, and by Exercise 1.4.22(ii) so is $\lambda := \sum_{n=1}^{\infty} \mu_n$ (the series being one of unsigned terms, so convergent in $[0, +\infty]$); so both sides of both identities are defined. Note that all the measures in play live on the **same** σ-algebra $\mathcal{B}$, so the class of simple functions is the same for all of them, and in particular the family

$$\mathcal{S}_f := \{h : X \to [0, +\infty] \ : \ h \text{ simple}, \ 0 \leq h \leq f \text{ pointwise}\}$$

over which the supremum in Definition 1.4.37 is taken does not depend on which measure we integrate against. $\mathcal{S}_f$ is nonempty (it contains $h = 0$).

A formula for the simple integral. Let h be simple, taking the distinct values $a_1, \dots, a_k \in [0, +\infty]$. By Definition 1.4.35, $\operatorname{Simp} \int_X h \, d\nu$ is computed in the finite σ-algebra $\mathcal{B}_h$ generated by the sets $A_j := h^{-1}(\{a_j\})$. These sets are nonempty, disjoint and cover X, so they are exactly the atoms of $\mathcal{B}_h$, and $h = \sum_{j=1}^{k} a_j 1_{A_j}$ is the unique representation of Definition 1.4.34. Hence for any measure ν on $\mathcal{B}$,

$$\operatorname{Simp} \int_X h \, d\nu = \sum_{j=1}^{k} a_j \, \nu(A_j).$$

The crucial point is that the sets $A_1, \dots, A_k$ depend only on h, not on ν.
(i). Fix $c \in [0, +\infty]$ and let $h \in \mathcal{S}_f$ have the representation above. Then

$$\operatorname{Simp} \int_X h \, d(c\mu) = \sum_{j=1}^{k} a_j \, (c\mu)(A_j) = \sum_{j=1}^{k} c \, a_j \mu(A_j) = c \sum_{j=1}^{k} a_j \mu(A_j) = c \operatorname{Simp} \int_X h \, d\mu,$$

using that multiplication distributes over finite sums in $[0, +\infty]$ (with the convention $0 \cdot \infty = \infty \cdot 0 = 0$; this is checked directly in the degenerate cases $c = 0$ and $c = \infty$). Therefore

$$\int_X f \, d(c\mu) = \sup_{h \in \mathcal{S}_f} c \operatorname{Simp} \int_X h \, d\mu = c \sup_{h \in \mathcal{S}_f} \operatorname{Simp} \int_X h \, d\mu = c \int_X f \, d\mu,$$

where the middle step uses the elementary fact that $\sup_{a \in A} ca = c \sup A$ for any nonempty $A \subset [0, +\infty]$ and $c \in [0, +\infty]$. (For $0 < c < \infty$ this is because $a \mapsto ca$ is an increasing bijection of $[0, +\infty]$; for $c = 0$ both sides are 0; for $c = +\infty$, both sides are 0 if $\sup A = 0$ – since then every $a \in A$ is 0 – and both sides are $+\infty$ otherwise.)
(ii). Recall $\lambda = \sum_{n=1}^{\infty} \mu_n$, so that the assertion to be proved is $\int_X f \, d\lambda = \sum_{n=1}^{\infty} \int_X f \, d\mu_n$. Let h be any simple function, with representation $h = \sum_{j=1}^{k} a_j 1_{A_j}$ as above. Then

$$\begin{aligned} \operatorname{Simp} \int_X h \, d\lambda &= \sum_{j=1}^{k} a_j \lambda(A_j) = \sum_{j=1}^{k} a_j \sum_{n=1}^{\infty} \mu_n(A_j) = \sum_{j=1}^{k} \sum_{n=1}^{\infty} a_j \mu_n(A_j) \\ &= \sum_{n=1}^{\infty} \sum_{j=1}^{k} a_j \mu_n(A_j) = \sum_{n=1}^{\infty} \operatorname{Simp} \int_X h \, d\mu_n, \end{aligned}$$

where we used distributivity of multiplication over sums in $[0, +\infty]$ and then interchanged a finite sum with an infinite sum of unsigned quantities (legitimate in $[0, +\infty]$: for unsigned terms the double series may be summed in either order, as each partial sum of one order is dominated by the total of the other).
Upper bound. For any $h \in \mathcal{S}_f$, the display above and Definition 1.4.37 applied to each μ_n give

$$\operatorname{Simp} \int_X h \, d\lambda = \sum_{n=1}^{\infty} \operatorname{Simp} \int_X h \, d\mu_n \leq \sum_{n=1}^{\infty} \int_X f \, d\mu_n.$$

Taking the supremum over $h \in \mathcal{S}_f$ yields $\int_X f \, d\lambda \leq \sum_{n=1}^{\infty} \int_X f \, d\mu_n$.
Lower bound. Fix a natural number N and let $h_1, \dots, h_N \in \mathcal{S}_f$ be arbitrary. Put $h := \max(h_1, \dots, h_N)$. This is measurable: for every $\alpha \in [0, +\infty)$ one has $\{x : h(x) > \alpha\} = \bigcup_{n=1}^{N} \{x : h_n(x) > \alpha\}$, a finite union of $\mathcal{B}$-measurable sets, so Exercise 1.4.29(i) applies. It takes only finitely many values, each value of h being a value of some h_n; so h is simple. Also $0 \leq h \leq f$ pointwise, since each $h_n \leq f$, so $h \in \mathcal{S}_f$. Since $h \geq h_n$ pointwise, monotonicity of the simple integral (Exercise 1.4.34(i)) gives $\operatorname{Simp} \int_X h \, d\mu_n \geq \operatorname{Simp} \int_X h_n \, d\mu_n$ for each n. Hence

$$\int_X f \, d\lambda \;\geq\; \operatorname{Simp} \int_X h \, d\lambda \;=\; \sum_{n=1}^{\infty} \operatorname{Simp} \int_X h \, d\mu_n \;\geq\; \sum_{n=1}^{N} \operatorname{Simp} \int_X h_n \, d\mu_n.$$

The left-hand side does not depend on the h_n, and the h_n range independently over $\mathcal{S}_f$; since for nonempty subsets $A_1, \dots, A_N$ of $[0, +\infty]$ one has $\sup\{a_1 + \cdots + a_N : a_n \in A_n\} = \sup A_1 + \cdots + \sup A_N$, taking suprema gives

$$\int_X f \, d\lambda \;\geq\; \sum_{n=1}^{N} \int_X f \, d\mu_n.$$

Letting $N \to \infty$ gives $\int_X f \, d\lambda \geq \sum_{n=1}^{\infty} \int_X f \, d\mu_n$. Combined with the upper bound, this proves (ii).

Problem 1.4.38. (Change of variables formula) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $\phi : X \to Y$ be a measurable morphism (as defined in Remark 1.4.33) from $(X, \mathcal{B})$ to another measurable space $(Y, \mathcal{C})$. Define the pushforward $\phi_* \mu : \mathcal{C} \to [0, +\infty]$ of μ by ϕ by the formula $\phi_* \mu(E) := \mu(\phi^{-1}(E))$.

- (i) Show that $\phi_* \mu$ is a measure on $\mathcal{C}$, so that $(Y, \mathcal{C}, \phi_* \mu)$ is a measure space.
- (ii) If $f : Y \to [0, +\infty]$ is measurable, show that $\int_Y f \, d\phi_* \mu = \int_X (f \circ \phi) \, d\mu$.

(Hint: the quickest proof here is via the monotone convergence theorem (Theorem 1.4.44) below, but it is also possible to prove the exercise without this theorem.)

Solution. Write $\nu := \phi_* \mu$ throughout.
(i). First, ν is well defined: since ϕ is a measurable morphism, $\phi^{-1}(E) \in \mathcal{B}$ for every $E \in \mathcal{C}$, so $\mu(\phi^{-1}(E))$ makes sense and lies in $[0, +\infty]$.
Since $\phi^{-1}(\emptyset) = \emptyset$, we get $\nu(\emptyset) = \mu(\emptyset) = 0$.
For countable additivity, let $E_1, E_2, \cdots \in \mathcal{C}$ be disjoint. Preimages preserve disjointness and unions: if $n \neq m$ then $\phi^{-1}(E_n) \cap \phi^{-1}(E_m) = \phi^{-1}(E_n \cap E_m) = \emptyset$, and $\phi^{-1}\big(\bigcup_n E_n\big) = \bigcup_n \phi^{-1}(E_n)$. Hence, by countable additivity of μ,

$$\nu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} \phi^{-1}(E_n)\Big) = \sum_{n=1}^{\infty} \mu(\phi^{-1}(E_n)) = \sum_{n=1}^{\infty} \nu(E_n).$$

So ν is a countably additive measure on $\mathcal{C}$ and $(Y, \mathcal{C}, \nu)$ is a measure space.
(ii). We give a proof that does not use the monotone convergence theorem. Note first that $f \circ \phi$ is measurable: f is a measurable morphism from $(Y, \mathcal{C})$ to $[0, +\infty]$ with its Borel σ-algebra, ϕ is a measurable morphism from $(X, \mathcal{B})$ to $(Y, \mathcal{C})$, and compositions of measurable morphisms are measurable morphisms (Remark 1.4.33). So both sides of the asserted identity are defined.
We use throughout the formula for the simple integral: if h is simple with distinct values $a_1, \dots, a_k$ and $A_j := h^{-1}(\{a_j\})$, then the A_j are measurable, disjoint and cover the space, they are the atoms of the finite σ-algebra generated by them (with respect to which h is measurable), and so by Definitions 1.4.34 and 1.4.35,

$$\mathrm{Simp} \int h \, d\lambda = \sum_{j=1}^{k} a_j \lambda(A_j)$$

for any measure λ on the ambient σ-algebra. (By Exercise 1.4.33 it does not matter which finite sub-σ-algebra is used; a partition into measurable sets on each of which h is constant may always be used, empty pieces and pieces where h vanishes contributing nothing.)
Step 1: the identity for simple h on Y. Let $h : Y \to [0, +\infty]$ be simple, with distinct values $a_1, \dots, a_k$ and $E_j := h^{-1}(\{a_j\})$. Then $h \circ \phi$ takes values among $a_1, \dots, a_k$ and $(h \circ \phi)^{-1}(\{a_j\}) = \phi^{-1}(E_j)$; the sets $\phi^{-1}(E_j)$ are measurable, disjoint and cover X, and $h \circ \phi$ is constant on each. So $h \circ \phi$ is simple and

$$\mathrm{Simp} \int_X (h \circ \phi) \, d\mu = \sum_{j=1}^{k} a_j \, \mu(\phi^{-1}(E_j)) = \sum_{j=1}^{k} a_j \, \nu(E_j) = \mathrm{Simp} \int_Y h \, d\nu.$$

Step 2: $\int_X (f \circ \phi) \, d\mu \geq \int_Y f \, d\nu$. If h is simple on Y with $0 \leq h \leq f$, then $h \circ \phi$ is simple on X with $0 \leq h \circ \phi \leq f \circ \phi$, so by Definition 1.4.37 and Step 1,

$$\int_X (f \circ \phi) \, d\mu \;\geq\; \mathrm{Simp} \int_X (h \circ \phi) \, d\mu \;=\; \mathrm{Simp} \int_Y h \, d\nu.$$

Taking the supremum over all such h gives the claim.
Step 3: $\int_X (f\circ\phi)\,d\mu \le \int_Y f\,d\nu$. Let $g : X \to [0,+\infty]$ be simple with $0 \le g \le f\circ\phi$. It suffices to produce a simple h on Y with $0 \le h \le f$ and $\operatorname{Simp}\int_X g\,d\mu \le \operatorname{Simp}\int_Y h\,d\nu$, for then $\operatorname{Simp}\int_X g\,d\mu \le \int_Y f\,d\nu$, and taking the supremum over g finishes the proof.
List the distinct values of g in decreasing order as $a_1 > a_2 > \cdots > a_k \ge 0$ and set $A_j := g^{-1}(\{a_j\})$, so that the A_j partition X. Put

$$B_j := \{y \in Y : f(y) \ge a_j\} \in \mathcal{C} \qquad (j = 1,\dots,k),$$

so that $B_1 \subset B_2 \subset \cdots \subset B_k$ because the a_j are decreasing. Define

$$h := a_1 1_{B_1} + \sum_{j=2}^{k} a_j 1_{B_j \setminus B_{j-1}}.$$

This is simple. Moreover $h \le f$ pointwise: on B_1 we have $h = a_1 \le f$; on $B_j \setminus B_{j-1}$ $(j \ge 2)$ we have $h = a_j \le f$; and off B_k we have $h = 0 \le f$.
Write $D_1 := B_1$ and $D_j := B_j \setminus B_{j-1}$ for $j \ge 2$; these are disjoint with $B_j = D_1 \cup \cdots \cup D_j$, so $\nu(B_j) = \sum_{i\le j}\nu(D_i)$ by finite additivity. Setting $a_{k+1} := 0$, the Abel summation identity

$$\sum_{j=1}^{k} a_j t_j = \sum_{j=1}^{k} (a_j - a_{j+1})\big(t_1 + \cdots + t_j\big) \qquad (t_1,\dots,t_k \in [0,+\infty])$$

holds in $[0,+\infty]$ – every coefficient $a_j - a_{j+1}$ is nonnegative, and expanding the right side and collecting the coefficient of t_i gives the telescoping sum $\sum_{j\ge i}(a_j - a_{j+1}) = a_i$ (also when $a_i = +\infty$, which can only happen for $i = 1$). Applying it with $t_j := \nu(D_j)$ and with $t_j := \mu(A_j)$ respectively gives

$$\operatorname{Simp}\int_Y h\,d\nu = \sum_{j=1}^{k} a_j \nu(D_j) = \sum_{j=1}^{k} (a_j - a_{j+1})\,\nu(B_j),$$

$$\operatorname{Simp}\int_X g\,d\mu = \sum_{j=1}^{k} a_j \mu(A_j) = \sum_{j=1}^{k} (a_j - a_{j+1})\,\mu(A_1 \cup \cdots \cup A_j).$$

Finally we compare the two right-hand sides termwise. If $x \in A_i$ with $i \le j$, then $g(x) = a_i \ge a_j$ and $g \le f\circ\phi$, so $f(\phi(x)) \ge a_j$, i.e. $\phi(x) \in B_j$. Hence $A_1 \cup \cdots \cup A_j \subset \phi^{-1}(B_j)$, and by monotonicity of μ (Exercise 1.4.20(i)),

$$\mu(A_1 \cup \cdots \cup A_j) \le \mu(\phi^{-1}(B_j)) = \nu(B_j).$$

Multiplying by $a_j - a_{j+1} \ge 0$ and summing gives $\operatorname{Simp}\int_X g\,d\mu \le \operatorname{Simp}\int_Y h\,d\nu$, as required.
Combining Steps 2 and 3, $\int_Y f\,d\phi_*\mu = \int_X (f\circ\phi)\,d\mu$.
(For comparison: with the monotone convergence theorem available, Step 3 is not needed. Take the standard simple approximations on Y, namely $h_n(y) := \min\big(2^{-n}\lfloor 2^n f(y)\rfloor,\, n\big)$ with the convention $h_n(y) := n$ when $f(y) = +\infty$; these are simple, non-decreasing in n, and increase pointwise to f. Then $h_n \circ \phi$ are simple on X and increase pointwise to $f\circ\phi$, so Theorem 1.4.44 applied on $(Y,\mathcal{C},\nu)$ and on $(X,\mathcal{B},\mu)$, together with Step 1, gives $\int_Y f\,d\nu = \lim_n \operatorname{Simp}\int_Y h_n\,d\nu = \lim_n \operatorname{Simp}\int_X (h_n\circ\phi)\,d\mu = \int_X (f\circ\phi)\,d\mu$.)

Problem 1.4.39. Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be an invertible linear transformation, and let m be Lebesgue measure on $\mathbb{R}^d$. Show that $T_*m = \frac{1}{|\det T|}m$, where the pushforward T_*m of m was defined in Exercise 1.4.38.

Solution. We work on the Lebesgue measure space $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d], m)$.
Since T is invertible, $\det T \ne 0$, and the inverse map $S := T^{-1} : \mathbb{R}^d \to \mathbb{R}^d$ is again an invertible linear

transformation, with

$$\det S = \frac{1}{\det T},$$

because det is multiplicative and $(\det T)(\det S) = \det(TS) = \det(\mathrm{id}) = 1$.
T is a measurable morphism. Let $E \subset \mathbb{R}^d$ be Lebesgue measurable. Because T is a bijection with inverse S, the preimage of E under T coincides with the image of E under S:

$$T^{-1}(E) = \{x \in \mathbb{R}^d : Tx \in E\} = \{Sy : y \in E\} = S(E),$$

the middle equality using that $x = S(Tx)$ and $T(Sy) = y$. By Exercise 1.2.21 applied to the linear transformation S, the set $S(E)$ is Lebesgue measurable. Hence $T^{-1}(E) \in \mathcal{L}[\mathbb{R}^d]$ for every Lebesgue measurable E, so T is a measurable morphism from $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d])$ to itself, and the pushforward T_*m of Exercise 1.4.38 is defined; by Exercise 1.4.38(i) it is a measure on $\mathcal{L}[\mathbb{R}^d]$.
Computation of the pushforward. For every Lebesgue measurable E, the same Exercise 1.2.21 gives the value of the measure of $S(E)$, so that

$$T_*m(E) = m(T^{-1}(E)) = m(S(E)) = |\det S|\, m(E) = \frac{1}{|\det T|}\, m(E).$$

Since $E \in \mathcal{L}[\mathbb{R}^d]$ was arbitrary, the two measures T_*m and $\frac{1}{|\det T|}m$ agree on every measurable set, i.e.

$$T_*m = \frac{1}{|\det T|}\, m$$

as measures on $\mathcal{L}[\mathbb{R}^d]$.
Remark. The same argument works verbatim on the Borel σ-algebra $\mathcal{B}[\mathbb{R}^d]$: an invertible linear map is a homeomorphism, so $T^{-1}(E) = S(E)$ is Borel whenever E is Borel, and the identity $m(S(E)) = |\det S| m(E)$ from Exercise 1.2.21 applies in particular to Borel sets.
Combining with Exercise 1.4.38(ii) and Exercise 1.4.37(i), one recovers the familiar change of variables formula for unsigned measurable $f : \mathbb{R}^d \to [0, +\infty]$:

$$\int_{\mathbb{R}^d} f(Tx)\, dm(x) = \int_{\mathbb{R}^d} f\, d(T_*m) = \frac{1}{|\det T|} \int_{\mathbb{R}^d} f(y)\, dm(y).$$

Problem 1.4.40. (Sums as integrals) Let X be an arbitrary set (with the discrete σ-algebra), let $\#$ be counting measure (see Exercise 1.4.26), and let $f : X \to [0, +\infty]$ be an arbitrary unsigned function. Show that f is measurable with

$$\int_X f\, d\# = \sum_{x \in X} f(x).$$

Solution. Here the discrete σ-algebra is $\mathcal{B} = 2^X$, the collection of **all** subsets of X, and $\#(E)$ is the cardinality of E if E is finite and $+\infty$ otherwise; $\#$ is countably additive (Example 1.4.29). We use the standard definition of an unsigned sum over an arbitrary index set,

$$\sum_{x \in X} f(x) := \sup \Big\{ \sum_{x \in F} f(x) \ : \ F \subset X \text{ finite} \Big\} \in [0, +\infty],$$

which agrees with the usual value of a series of unsigned terms when X is countable.
Measurability. Every subset of X lies in $\mathcal{B} = 2^X$. In particular $f^{-1}(U) \in \mathcal{B}$ for every set $U \subset [0, +\infty]$, so a fortiori for every Borel U. Hence f is measurable, with no hypotheses on f at all. Note also that $\sum_{x \in X} f(x)$ is monotone in f: if $0 \le h \le f$ pointwise then $\sum_{x \in F} h(x) \le \sum_{x \in F} f(x)$ for every finite F, hence $\sum_{x \in X} h(x) \le \sum_{x \in X} f(x)$.

Step 1: the identity for simple functions. We claim that for every simple $h : X \to [0, +\infty]$,

$$\mathrm{Simp} \int_X h \, d\# = \sum_{x \in X} h(x).$$

Let $a_1, \ldots, a_k$ be the distinct values of h and $A_j := h^{-1}(\{a_j\})$, so the A_j are disjoint and cover X and are the atoms of the finite σ-algebra they generate; by Definitions 1.4.34 and 1.4.35,

$$\mathrm{Simp} \int_X h \, d\# = \sum_{j=1}^{k} a_j \, \#(A_j).$$

Discard the indices j with $a_j = 0$, which contribute 0 to both sides (note $0 \cdot \#(A_j) = 0$ even when $\#(A_j) = \infty$, by the convention $0 \cdot \infty = 0$), and let $J := \{j : a_j > 0\}$, so $\{x : h(x) > 0\} = \bigcup_{j \in J} A_j$.
Suppose first that $a_j < \infty$ and A_j is finite for every $j \in J$. Then $F_0 := \bigcup_{j \in J} A_j$ is a finite set and

$$\sum_{j=1}^{k} a_j \#(A_j) = \sum_{j \in J} a_j |A_j| = \sum_{x \in F_0} h(x).$$

On the other hand, for any finite $F \subset X$ we have $\sum_{x \in F} h(x) = \sum_{x \in F \cap F_0} h(x) \le \sum_{x \in F_0} h(x)$, since h vanishes off F_0; so the supremum defining $\sum_{x \in X} h(x)$ is attained at $F = F_0$ and equals $\sum_{x \in F_0} h(x)$. The two sides agree.
Otherwise there is some $j \in J$ with $a_j = \infty$ or with A_j infinite; in either case the left-hand side $\sum_j a_j \#(A_j)$ is $+\infty$, since $a_j \#(A_j) = \infty$ (in the first case $A_j \neq \emptyset$ so $\#(A_j) \ge 1$, and $\infty \cdot \#(A_j) = \infty$; in the second $\#(A_j) = \infty$ and $a_j > 0$). We check the right-hand side is $+\infty$ too. If $a_j = \infty$, pick any $x_0 \in A_j$; then $\sum_{x \in X} h(x) \ge h(x_0) = \infty$. If instead $0 < a_j < \infty$ and A_j is infinite, then for every natural number N we may choose a finite $F \subset A_j$ with $|F| = N$, giving $\sum_{x \in X} h(x) \ge \sum_{x \in F} h(x) = N a_j$; letting $N \to \infty$ gives $\sum_{x \in X} h(x) = \infty$. In all cases the claim holds.
Step 2: $\int_X f \, d\# \le \sum_{x \in X} f(x)$. By Definition 1.4.37, Step 1, and monotonicity of the unsigned sum, every simple h with $0 \le h \le f$ satisfies

$$\mathrm{Simp} \int_X h \, d\# = \sum_{x \in X} h(x) \le \sum_{x \in X} f(x),$$

and taking the supremum over such h gives the inequality.
Step 3: $\int_X f \, d\# \ge \sum_{x \in X} f(x)$. Let $F \subset X$ be finite. The function $h_F := f 1_F$ is measurable (everything is) and takes at most $|F| + 1$ values, so it is simple, and $0 \le h_F \le f$. By Step 1,

$$\int_X f \, d\# \;\ge\; \mathrm{Simp} \int_X h_F \, d\# \;=\; \sum_{x \in X} h_F(x) \;=\; \sum_{x \in F} f(x).$$

Taking the supremum over all finite $F \subset X$ gives $\int_X f \, d\# \ge \sum_{x \in X} f(x)$.
Steps 2 and 3 together give $\int_X f \, d\# = \sum_{x \in X} f(x)$.

Problem 1.4.41. Let $(X, \mathcal{B}, \mu)$ be a measure space.

- (i) Show that $L^1(X, \mathcal{B}, \mu)$ is a complex vector space.
- (ii) Show that the integration map $f \mapsto \int_X f \, d\mu$ is a complex-linear map from $L^1(X, \mathcal{B}, \mu)$ to $\mathbb{C}$.
- (iii) Establish the triangle inequality $\|f + g\|_{L^1(\mu)} \le \|f\|_{L^1(\mu)} + \|g\|_{L^1(\mu)}$ and the homogeneity property $\|cf\|_{L^1(\mu)} = |c| \|f\|_{L^1(\mu)}$ for all $f, g \in L^1(X, \mathcal{B}, \mu)$ and $c \in \mathbb{C}$.
- (iv) Show that if $f, g \in L^1(X, \mathcal{B}, \mu)$ are such that $f(x) = g(x)$ for μ-almost every $x \in X$, then $\int_X f \, d\mu = \int_X g \, d\mu$.
- (v) If $f \in L^1(X, \mathcal{B}, \mu)$, and $(X, \mathcal{B}', \mu')$ is a refinement of $(X, \mathcal{B}, \mu)$, then $f \in L^1(X, \mathcal{B}', \mu')$, and $\int_X f \, d\mu = \int_X f \, d\mu'$. (Hint: it is easy to get one inequality. To get the other inequality, first work

in the case when f is both bounded and has finite measure support (i.e. is both vertically and horizontally truncated).)

- (vi) Show that if $f \in L^1(X, \mathcal{B}, \mu)$, then $\|f\|_{L^1(\mu)} = 0$ if and only if f is zero μ-almost everywhere.
- (vii) If $Y \subset X$ is $\mathcal{B}$-measurable and $f \in L^1(X, \mathcal{B}, \mu)$, then $f \restriction_Y \in L^1(Y, \mathcal{B} \restriction_Y, \mu \restriction_Y)$ and $\int_Y f \restriction_Y d\mu \restriction_Y = \int_X f 1_Y \, d\mu$. As before, by abuse of notation we write $\int_Y f \, d\mu$ for $\int_Y f \restriction_Y d\mu \restriction_Y$.

Solution. Throughout we use Definition 1.4.39, Exercise 1.4.36, and Theorem 1.4.38 (additivity of the unsigned integral: $\int_X (F+G)\,d\mu = \int_X F\,d\mu + \int_X G\,d\mu$ for unsigned measurable F, G), which by induction extends to any finite number of unsigned summands. Measurability of sums, products, absolute values, real and imaginary parts and positive/negative parts is provided by Exercise 1.4.29(iv), (v), (vii), (viii).
(i). The set of all functions $X \to \mathbb{C}$ is a complex vector space under pointwise operations, so it suffices to check that $L^1(X, \mathcal{B}, \mu)$ is a nonempty subset closed under addition and scalar multiplication. The zero function is measurable with $\int_X 0\,d\mu = 0$, so $0 \in L^1$.
Let $f, g \in L^1$ and $c \in \mathbb{C}$. Then $f+g$ and cf are measurable by Exercise 1.4.29(viii), and $|f+g|, |cf|$ are measurable by Exercise 1.4.29(vii). Pointwise $|f+g| \le |f| + |g|$, so by monotonicity (Exercise 1.4.36(ii)) and Theorem 1.4.38,

$$\int_X |f+g|\,d\mu \le \int_X (|f|+|g|)\,d\mu = \int_X |f|\,d\mu + \int_X |g|\,d\mu < \infty,$$

so $f+g \in L^1$. Also $|cf| = |c|\,|f|$ pointwise, so by homogeneity (Exercise 1.4.36(iii)), $\int_X |cf|\,d\mu = |c| \int_X |f|\,d\mu < \infty$ and $cf \in L^1$. Hence $L^1(X, \mathcal{B}, \mu)$ is a complex vector space.
(ii). We build up linearity in stages.
Additivity, real case. Let $f, g : X \to \mathbb{R}$ be absolutely integrable and set $h := f + g$, which is absolutely integrable by (i). All of f, g, h are finite-valued, so from $h_+ - h_- = h = f + g = (f_+ - f_-) + (g_+ - g_-)$ we get the pointwise identity of unsigned finite-valued functions

$$h_+ + f_- + g_- = h_- + f_+ + g_+.$$

Integrating and applying Theorem 1.4.38 to both sides,

$$\int_X h_+\,d\mu + \int_X f_-\,d\mu + \int_X g_-\,d\mu = \int_X h_-\,d\mu + \int_X f_+\,d\mu + \int_X g_+\,d\mu.$$

Every one of these six unsigned integrals is finite: $f_\pm \le |f|$, $g_\pm \le |g|$ and $h_\pm \le |h|$ pointwise, and $|f|, |g|, |h|$ have finite integral, so monotonicity applies. We may therefore rearrange the (finite, real) terms to get

$$\int_X h_+\,d\mu - \int_X h_-\,d\mu = \Big(\int_X f_+\,d\mu - \int_X f_-\,d\mu\Big) + \Big(\int_X g_+\,d\mu - \int_X g_-\,d\mu\Big),$$

which by Definition 1.4.39 says $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.
Homogeneity, real scalars and real functions. Let $f : X \to \mathbb{R}$ be absolutely integrable. If $c \ge 0$ then $(cf)_+ = cf_+$ and $(cf)_- = cf_-$ pointwise, so by Exercise 1.4.36(iii),

$$\int_X cf\,d\mu = c\int_X f_+\,d\mu - c\int_X f_-\,d\mu = c\int_X f\,d\mu.$$

For $c = -1$ we have $(-f)_+ = f_-$ and $(-f)_- = f_+$, so $\int_X (-f)\,d\mu = \int_X f_-\,d\mu - \int_X f_+\,d\mu = -\int_X f\,d\mu$. A general real c is $\pm|c|$, so combining the two cases gives $\int_X cf\,d\mu = c\int_X f\,d\mu$ for all $c \in \mathbb{R}$.
Additivity, complex case. Let $f, g \in L^1$ be complex-valued. Their real and imaginary parts are measurable (Exercise 1.4.29(iv)) and absolutely integrable, since $|\operatorname{Re} f| \le |f|$ and $|\operatorname{Im} f| \le |f|$ pointwise. Since $\operatorname{Re}(f+g) = \operatorname{Re} f + \operatorname{Re} g$ and likewise for imaginary parts, Definition 1.4.39 and the real case give

$$\begin{aligned}\int_X (f+g)\,d\mu &= \int_X \operatorname{Re}(f+g)\,d\mu + i\int_X \operatorname{Im}(f+g)\,d\mu \\ &= \Big(\int_X \operatorname{Re} f\,d\mu + \int_X \operatorname{Re} g\,d\mu\Big) + i\Big(\int_X \operatorname{Im} f\,d\mu + \int_X \operatorname{Im} g\,d\mu\Big) \\ &= \int_X f\,d\mu + \int_X g\,d\mu.\end{aligned}$$

Homogeneity, complex scalars. For real c, $\mathrm{Re}(cf) = c\,\mathrm{Re}\,f$ and $\mathrm{Im}(cf) = c\,\mathrm{Im}\,f$, so the real homogeneity above gives $\int_X cf\,d\mu = c\int_X f\,d\mu$. For $c = i$ we have $\mathrm{Re}(if) = -\mathrm{Im}\,f$ and $\mathrm{Im}(if) = \mathrm{Re}\,f$, so

$$\int_X if\,d\mu = -\int_X \mathrm{Im}\,f\,d\mu + i\int_X \mathrm{Re}\,f\,d\mu = i\Big(\int_X \mathrm{Re}\,f\,d\mu + i\int_X \mathrm{Im}\,f\,d\mu\Big) = i\int_X f\,d\mu.$$

Finally, for $c = a + ib$ with $a, b \in \mathbb{R}$, writing $cf = af + b(if)$ and using complex additivity together with the two previous cases,

$$\int_X cf\,d\mu = a\int_X f\,d\mu + b\,i\int_X f\,d\mu = c\int_X f\,d\mu.$$

Thus $f \mapsto \int_X f\,d\mu$ is complex-linear on $L^1(X, \mathcal{B}, \mu)$; it takes values in $\mathbb{C}$ because each of the four unsigned integrals $\int_X (\mathrm{Re}\,f)_\pm\,d\mu$, $\int_X (\mathrm{Im}\,f)_\pm\,d\mu$ is finite.
(iii). Both assertions were established in the course of proving (i). The chain

$$\int_X |f+g|\,d\mu \le \int_X (|f| + |g|)\,d\mu = \int_X |f|\,d\mu + \int_X |g|\,d\mu$$

is exactly $\|f+g\|_{L^1(\mu)} \le \|f\|_{L^1(\mu)} + \|g\|_{L^1(\mu)}$, and the identity $\int_X |cf|\,d\mu = |c|\int_X |f|\,d\mu$ is exactly $\|cf\|_{L^1(\mu)} = |c|\,\|f\|_{L^1(\mu)}$. Note the latter is an equality, not merely an inequality, because Exercise 1.4.36(iii) is an equality.
(iv). Suppose $f = g$ μ-almost everywhere. Then $\mathrm{Re}\,f = \mathrm{Re}\,g$ and $\mathrm{Im}\,f = \mathrm{Im}\,g$ μ-almost everywhere, and hence $(\mathrm{Re}\,f)_\pm = (\mathrm{Re}\,g)_\pm$ and $(\mathrm{Im}\,f)_\pm = (\mathrm{Im}\,g)_\pm$ μ-almost everywhere (all outside the same null set). By almost everywhere equivalence for the unsigned integral (Exercise 1.4.36(i)) each of the four corresponding unsigned integrals agrees, and hence so do $\int_X f\,d\mu$ and $\int_X g\,d\mu$ by Definition 1.4.39.
(v). Let $(X, \mathcal{B}', \mu')$ be a refinement of $(X, \mathcal{B}, \mu)$: thus $\mathcal{B} \subset \mathcal{B}'$ and μ' agrees with μ on $\mathcal{B}$. Any $\mathcal{B}$-measurable function is automatically $\mathcal{B}'$-measurable, since its preimages lie in $\mathcal{B} \subset \mathcal{B}'$.
Claim. If $F : X \to [0, +\infty]$ is $\mathcal{B}$-measurable with $\int_X F\,d\mu < \infty$, then $\int_X F\,d\mu' = \int_X F\,d\mu$.
Easy inequality. Every $\mathcal{B}$-simple h with $0 \le h \le F$ is also $\mathcal{B}'$-simple, and by insensitivity of the simple integral to refinements (Exercise 1.4.34(v)) $\mathrm{Simp}\int_X h\,d\mu = \mathrm{Simp}\int_X h\,d\mu'$, which is at most $\int_X F\,d\mu'$. Taking the supremum over such h gives $\int_X F\,d\mu \le \int_X F\,d\mu'$.
The truncated case. Suppose in addition that $0 \le F \le M$ for some finite $M > 0$ and that F vanishes outside some $E \in \mathcal{B}$ with $\mu(E) < \infty$. Fix $n \ge 1$ and set, for $0 \le j \le n-1$,

$$A_j := \Big\{x \in E : \tfrac{jM}{n} \le F(x) < \tfrac{(j+1)M}{n}\Big\} \quad (j < n-1), \qquad A_{n-1} := \Big\{x \in E : \tfrac{(n-1)M}{n} \le F(x) \le M\Big\}.$$

These are $\mathcal{B}$-measurable and partition E. Put $u_n := \sum_{j=0}^{n-1} \frac{jM}{n} 1_{A_j}$ and $v_n := u_n + \frac{M}{n} 1_E$. Both are $\mathcal{B}$-simple, $u_n \le F$ everywhere, and $F \le v_n$ everywhere (on E by construction, and off E both sides are 0). Now let g be any $\mathcal{B}'$-simple function with $0 \le g \le F$. Then $g \le v_n$, so by monotonicity and linearity of the simple integral (Exercise 1.4.34(i),(ii),(iii),(iv)) and insensitivity to refinement (Exercise 1.4.34(v)),

$$\mathrm{Simp}\int_X g\,d\mu' \le \mathrm{Simp}\int_X v_n\,d\mu' = \mathrm{Simp}\int_X v_n\,d\mu = \mathrm{Simp}\int_X u_n\,d\mu + \frac{M}{n}\mu(E) \le \int_X F\,d\mu + \frac{M}{n}\mu(E).$$

Since $\mu(E) < \infty$ and $M < \infty$, letting $n \to \infty$ gives $\mathrm{Simp}\int_X g\,d\mu' \le \int_X F\,d\mu$; taking the supremum over g gives $\int_X F\,d\mu' \le \int_X F\,d\mu$, and with the easy inequality, equality.
The general case of the Claim. Let F be $\mathcal{B}$-measurable with $\int_X F\,d\mu < \infty$. By Markov's inequality (Exercise 1.4.36(vi)) the $\mathcal{B}$-measurable sets $E_k := \{x \in X : F(x) > 1/k\}$ satisfy $\mu(E_k) < \infty$; they increase with $\bigcup_k E_k = \{F > 0\}$. For each n, k the function $\min(F, n)1_{E_k}$ is $\mathcal{B}$-measurable, bounded by n, and vanishes off E_k, so by the truncated case

$$\int_X \min(F,n)1_{E_k}\,d\mu' = \int_X \min(F,n)1_{E_k}\,d\mu =: c_{n,k}.$$

Fix n and let $k \to \infty$. By horizontal truncation (Exercise 1.4.36(x)), applied in the measure space $(X, \mathcal{B}', \mu')$ and in $(X, \mathcal{B}, \mu)$ respectively,

$$\lim_{k\to\infty} c_{n,k} = \int_X \min(F, n) 1_{\{F>0\}}\, d\mu' = \int_X \min(F, n)\, d\mu',$$

and equally $\lim_{k\to\infty} c_{n,k} = \int_X \min(F, n)\, d\mu$ (we used $\min(F, n)1_{\{F>0\}} = \min(F, n)$, both sides vanishing where $F = 0$). Hence $\int_X \min(F, n)\, d\mu' = \int_X \min(F, n)\, d\mu$ for every n. Letting $n \to \infty$ and using vertical truncation (Exercise 1.4.36(ix)) on each side gives $\int_X F\, d\mu' = \int_X F\, d\mu$, proving the Claim.
Conclusion of (v). Let $f \in L^1(X, \mathcal{B}, \mu)$. Then f is $\mathcal{B}'$-measurable, and applying the Claim to $F = |f|$ gives $\int_X |f|\, d\mu' = \int_X |f|\, d\mu < \infty$, so $f \in L^1(X, \mathcal{B}', \mu')$ with the same L^1 norm. Applying the Claim to each of the four $\mathcal{B}$-measurable unsigned functions $(\operatorname{Re} f)_\pm$, $(\operatorname{Im} f)_\pm$ (each dominated by $|f|$, hence of finite μ-integral) shows that all four unsigned integrals are unchanged when μ is replaced by μ'. By Definition 1.4.39, $\int_X f\, d\mu' = \int_X f\, d\mu$.
(vi). If $\|f\|_{L^1(\mu)} = \int_X |f|\, d\mu = 0$, then by vanishing for the unsigned integral (Exercise 1.4.36(viii)) applied to $|f|$, we get $|f| = 0$ μ-almost everywhere, i.e. $f = 0$ μ-almost everywhere. Conversely if $f = 0$ μ-almost everywhere then $|f| = 0$ μ-almost everywhere, so by Exercise 1.4.36(i), $\int_X |f|\, d\mu = \int_X 0\, d\mu = 0$.
(vii). Since $Y \in \mathcal{B}$, the restricted space $(Y, \mathcal{B}\restriction_Y, \mu\restriction_Y)$ of Example 1.4.25 is a measure space, where $\mathcal{B}\restriction_Y = \{E \in \mathcal{B} : E \subset Y\}$. The restriction $f\restriction_Y$ is measurable, because for any Borel $U \subset \mathbb{C}$ we have $(f\restriction_Y)^{-1}(U) = f^{-1}(U) \cap Y \in \mathcal{B}\restriction_Y$ (using Exercise 1.4.29(iii)). Applying the restriction property of the unsigned integral (Exercise 1.4.36(xi)) to $|f|$ and using monotonicity,

$$\int_Y |f\restriction_Y|\, d\mu\restriction_Y = \int_X |f| 1_Y\, d\mu \le \int_X |f|\, d\mu < \infty,$$

so $f\restriction_Y \in L^1(Y, \mathcal{B}\restriction_Y, \mu\restriction_Y)$; likewise $f1_Y \in L^1(X, \mathcal{B}, \mu)$. Now observe that taking real parts, imaginary parts and positive/negative parts all commute with restriction to Y and with multiplication by 1_Y: for instance $\operatorname{Re}(f1_Y) = (\operatorname{Re} f)1_Y$, and $((\operatorname{Re} f)1_Y)_+ = (\operatorname{Re} f)_+ 1_Y$ since $\max(a1_Y, 0) = \max(a, 0)1_Y$; and $(\operatorname{Re}(f\restriction_Y))_+ = ((\operatorname{Re} f)_+)\restriction_Y$. Applying Exercise 1.4.36(xi) to each of the four unsigned functions $(\operatorname{Re} f)_\pm$ and $(\operatorname{Im} f)_\pm$ gives

$$\int_X (\operatorname{Re} f)_\pm 1_Y\, d\mu = \int_Y ((\operatorname{Re} f)_\pm)\restriction_Y\, d\mu\restriction_Y, \qquad \int_X (\operatorname{Im} f)_\pm 1_Y\, d\mu = \int_Y ((\operatorname{Im} f)_\pm)\restriction_Y\, d\mu\restriction_Y,$$

all eight quantities being finite. Assembling these four identities according to Definition 1.4.39 yields $\int_X f1_Y\, d\mu = \int_Y f\restriction_Y\, d\mu\restriction_Y$.

Problem 1.4.42. (Uniform convergence on a finite measure space) Suppose that $(X, \mathcal{B}, \mu)$ is a finite measure space (so $\mu(X) < \infty$), and $f_n : X \to [0, +\infty]$ (resp. $f_n : X \to \mathbb{C}$) are a sequence of unsigned measurable functions (resp. absolutely integrable functions) that converge uniformly to a limit f. Show that $\int_X f_n\, d\mu$ converges to $\int_X f\, d\mu$.

Solution. The whole point is that a constant function is integrable on a finite measure space: for any $\varepsilon \in (0, \infty)$ the function $\varepsilon 1_X$ is simple, so by Exercise 1.4.36(v) and Exercise 1.4.34(ii),(iii),

$$\int_X \varepsilon\, d\mu = \operatorname{Simp} \int_X \varepsilon 1_X\, d\mu = \varepsilon\mu(X) < \infty.$$

Note also that uniform convergence implies pointwise convergence, so in both cases the limit f is measurable by Exercise 1.4.29(vi).
The unsigned case. Here uniform convergence of $[0, +\infty]$-valued functions is to be read in the natural way: for every $\varepsilon > 0$ there is N such that for all $n \ge N$ and all $x \in X$,

$$f(x) - \varepsilon \;\le\; f_n(x) \;\le\; f(x) + \varepsilon,$$

with the conventions $\infty + \varepsilon = \infty$ and $\infty - \varepsilon = \infty$ (equivalently, $|f_n(x) - f(x)| \leq \varepsilon$ whenever both values are finite, and $f_n(x) = \infty$ exactly when $f(x) = \infty$, for $n \geq N$).
Fix $\varepsilon > 0$ and let N be as above, and let $n \geq N$. From $f_n \leq f + \varepsilon$ pointwise, monotonicity (Exercise 1.4.36(ii)) and additivity (Theorem 1.4.38) give

$$\int_X f_n \, d\mu \;\leq\; \int_X (f + \varepsilon) \, d\mu \;=\; \int_X f \, d\mu + \varepsilon\mu(X),$$

and symmetrically, from $f \leq f_n + \varepsilon$ pointwise,

$$\int_X f \, d\mu \;\leq\; \int_X f_n \, d\mu + \varepsilon\mu(X).$$

Suppose first that $\int_X f \, d\mu = \infty$. The second display, together with $\varepsilon\mu(X) < \infty$, forces $\int_X f_n \, d\mu = \infty$ for all $n \geq N$ (a finite quantity plus a finite quantity cannot be infinite). Hence $\int_X f_n \, d\mu = \infty = \int_X f \, d\mu$ for all large n, and in particular the integrals converge to $\int_X f \, d\mu$.
Suppose instead that $\int_X f \, d\mu < \infty$. Then the first display shows $\int_X f_n \, d\mu < \infty$ for $n \geq N$, and the two displays combine (all quantities now being finite real numbers) to give

$$\Big| \int_X f_n \, d\mu - \int_X f \, d\mu \Big| \;\leq\; \varepsilon\mu(X) \qquad (n \geq N).$$

As $\mu(X) < \infty$ and $\varepsilon > 0$ was arbitrary, $\int_X f_n \, d\mu \to \int_X f \, d\mu$.
The absolutely integrable case. Now let $f_n \in L^1(X, \mathcal{B}, \mu)$ converge uniformly to $f : X \to \mathbb{C}$, so that for every $\varepsilon > 0$ there is N with $|f_n(x) - f(x)| \leq \varepsilon$ for all $n \geq N$ and all $x \in X$.
First, $f \in L^1(X, \mathcal{B}, \mu)$. Indeed f is measurable, and taking $\varepsilon = 1$ and a corresponding N, the pointwise bound $|f| \leq |f_N| + 1$ together with monotonicity and Theorem 1.4.38 gives

$$\int_X |f| \, d\mu \;\leq\; \int_X |f_N| \, d\mu + \mu(X) \;<\; \infty.$$

Now fix $\varepsilon > 0$ and N as above, and let $n \geq N$. Put $g := f_n - f$, which lies in L^1 by Exercise 1.4.41(i) and satisfies $|g| \leq \varepsilon$ pointwise. By complex linearity of the integral (Exercise 1.4.41(ii)),

$$\int_X f_n \, d\mu - \int_X f \, d\mu = \int_X g \, d\mu.$$

We estimate $\int_X g \, d\mu$ directly from Definition 1.4.39. Since $|\operatorname{Re} g| \leq |g| \leq \varepsilon$ pointwise, both $(\operatorname{Re} g)_+$ and $(\operatorname{Re} g)_-$ are bounded by ε, so by monotonicity each of $\int_X (\operatorname{Re} g)_\pm \, d\mu$ lies in $[0, \varepsilon\mu(X)]$, and therefore

$$\Big| \int_X \operatorname{Re} g \, d\mu \Big| = \Big| \int_X (\operatorname{Re} g)_+ \, d\mu - \int_X (\operatorname{Re} g)_- \, d\mu \Big| \leq \varepsilon\mu(X).$$

The same argument applied to $\operatorname{Im} g$ gives $|\int_X \operatorname{Im} g \, d\mu| \leq \varepsilon\mu(X)$. Hence, by the triangle inequality in $\mathbb{C}$,

$$\Big| \int_X f_n \, d\mu - \int_X f \, d\mu \Big| = \Big| \int_X \operatorname{Re} g \, d\mu + i \int_X \operatorname{Im} g \, d\mu \Big| \leq 2\varepsilon\mu(X) \qquad (n \geq N).$$

Since $\mu(X) < \infty$ and $\varepsilon > 0$ was arbitrary, $\int_X f_n \, d\mu \to \int_X f \, d\mu$.
Remark. Finiteness of μ is essential: in Example 1.4.41 the functions $f_n = \frac{1}{n} 1_{[0,n]}$ converge uniformly to 0 on the real line with Lebesgue measure, yet $\int_{\mathbb{R}} f_n \, dm = 1$ for every n. What fails is precisely the bound $\int_X \varepsilon \, d\mu = \varepsilon\mu(X) < \infty$ used above.

Problem 1.4.43. Give an example to show that this corollary can fail if the f_n are assumed to be absolutely integrable rather than unsigned measurable, even if the sum $\sum_{n=1}^{\infty} f_n(x)$ is absolutely convergent for each x. (Hint: think about the three escapes to infinity.)

Solution. Here "this corollary" is Corollary 1.4.46 (Tonelli's theorem for sums and integrals), which asserts $\int_X \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu$ for unsigned measurable f_n.
Take $X = \mathbb{R}$ with Lebesgue measure m (restricted to the Lebesgue σ-algebra), and for $n = 1, 2, 3, \ldots$ set

$$f_n := 1_{[n,\, n+1)} - 1_{[n+1,\, n+2)}.$$

Each f_n is a simple function, hence measurable, and it is absolutely integrable since

$$\int_{\mathbb{R}} |f_n| \, dm = m([n, n+2)) = 2 < \infty.$$

Fix $x \in \mathbb{R}$. The intervals $[n, n+1)$, $n \in \mathbb{Z}$, are disjoint, so x lies in at most one of them; consequently $f_n(x) \neq 0$ for at most two values of n (namely the n with $x \in [n, n+1)$ and the n with $x \in [n+1, n+2)$). In particular $\sum_{n=1}^{\infty} |f_n(x)| \leq 2 < \infty$, so the series $\sum_{n=1}^{\infty} f_n(x)$ is absolutely convergent for every x, as required.
The partial sums telescope: for every $N \geq 1$,

$$\sum_{n=1}^{N} f_n = 1_{[1,2)} - 1_{[N+1,\, N+2)}.$$

For fixed x the subtracted term vanishes as soon as $N + 1 > x$, so letting $N \to \infty$ gives the pointwise identity

$$\sum_{n=1}^{\infty} f_n = 1_{[1,2)} \qquad \text{everywhere on } \mathbb{R}.$$

Hence

$$\int_{\mathbb{R}} \sum_{n=1}^{\infty} f_n \, dm = m([1, 2)) = 1.$$

On the other hand each individual integral vanishes: by linearity of the integral for absolutely integrable functions,

$$\int_{\mathbb{R}} f_n \, dm = m([n, n+1)) - m([n+1, n+2)) = 1 - 1 = 0,$$

so $\sum_{n=1}^{\infty} \int_{\mathbb{R}} f_n \, dm = 0$.
Thus $1 = \int_{\mathbb{R}} \sum_n f_n \, dm \neq \sum_n \int_{\mathbb{R}} f_n \, dm = 0$, and Corollary 1.4.46 fails for absolutely integrable f_n.
The mechanism is exactly the escape to horizontal infinity of Example 1.4.40: the mass $+1$ deposited on $[1, 2)$ by the first summand is paid for by a compensating mass -1 which is pushed off to $+\infty$ and is never accounted for in the pointwise limit.
The construction is an instance of a general telescoping recipe, which shows that any of the three escapes to infinity will do. Suppose $a_1, a_2, \ldots : X \to [0, +\infty)$ are unsigned measurable with $\int_X a_n \, d\mu = 1$ for every n and $a_n \to 0$ pointwise. Put $f_n := a_n - a_{n+1}$. Each f_n is measurable and absolutely integrable, since $\int_X |f_n| \, d\mu \leq \int_X a_n \, d\mu + \int_X a_{n+1} \, d\mu = 2$, and by linearity $\int_X f_n \, d\mu = 1 - 1 = 0$, so $\sum_{n=1}^{\infty} \int_X f_n \, d\mu = 0$. The partial sums telescope to $\sum_{n=1}^{N} f_n = a_1 - a_{N+1}$, which converges pointwise to a_1; hence $\int_X \sum_{n=1}^{\infty} f_n \, d\mu = \int_X a_1 \, d\mu = 1 \neq 0$.
Our example above is the case $a_n := 1_{[n,n+1)}$ (Example 1.4.40). Taking instead $a_n := \frac{1}{n} 1_{[0,n)}$ on $\mathbb{R}$ realises the escape to width infinity (Example 1.4.41), and $a_n := n 1_{(0,1/n]}$ on $[0, 1]$ realises the escape to vertical infinity (Example 1.4.42); in both cases $\int a_n \, dm = 1$ and $a_n \to 0$ pointwise. Absolute convergence of $\sum_n f_n(x)$ holds in all three cases: for the horizontal and vertical examples, $f_n(x) = 0$ for all sufficiently large n (once $n > x$, respectively once $1/n < x$), while for the width example $f_n(x) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$ for all $n > x$, which is summable.

Problem 1.4.44. (Borel-Cantelli lemma) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $E_1, E_2, E_3, \ldots$ be a sequence of $\mathcal{B}$-measurable sets such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Show that almost every $x \in X$ is

contained in at most finitely many of the E_n (i.e. $\{n \in \mathbb{N} : x \in E_n\}$ is finite for almost every $x \in X$). (Hint: Apply Tonelli's theorem to the indicator functions 1_{E_n}.)

Solution. Each indicator function $1_{E_n} : X \to [0, +\infty]$ is unsigned and $\mathcal{B}$-measurable, since E_n is $\mathcal{B}$-measurable. Define

$$F := \sum_{n=1}^{\infty} 1_{E_n},$$

where the series is a sum of unsigned quantities and hence converges in $[0, +\infty]$ at every point; equivalently $F = \lim_{N\to\infty} \sum_{n=1}^{N} 1_{E_n}$ is the increasing limit of measurable functions (each partial sum being measurable by Exercise 1.4.29(ii), (viii)), so $F : X \to [0, +\infty]$ is unsigned measurable by Exercise 1.4.29(vi).

The key observation is that F simply counts memberships: for each $x \in X$,

$$F(x) = \sum_{n=1}^{\infty} 1_{E_n}(x) = \#\{n \in \mathbb{N} : x \in E_n\},$$

with the convention that the right-hand side is $+\infty$ when the set is infinite. Indeed, the n-th summand is 1 precisely when $x \in E_n$ and 0 otherwise, so the partial sums of the series count the elements of $\{n \leq N : x \in E_n\}$; letting $N \to \infty$ gives the claim, the series diverging to $+\infty$ exactly when x lies in infinitely many E_n.

Now apply Corollary 1.4.46 (Tonelli's theorem for sums and integrals) to the unsigned measurable functions 1_{E_n}:

$$\int_X F \, d\mu = \int_X \sum_{n=1}^{\infty} 1_{E_n} \, d\mu = \sum_{n=1}^{\infty} \int_X 1_{E_n} \, d\mu = \sum_{n=1}^{\infty} \mu(E_n) < \infty,$$

where we used that $\int_X 1_{E_n} \, d\mu = \mu(E_n)$ (the integral of an indicator function is the measure of the set) and the hypothesis that $\sum_n \mu(E_n) < \infty$.

Thus F is an unsigned measurable function with finite integral. By the finiteness property of the unsigned integral (Exercise 1.4.36(vii)), $F(x)$ is finite for μ-almost every $x \in X$; that is, the set

$$N := \{x \in X : F(x) = +\infty\}$$

is a μ-null set. (Concretely, this is Markov's inequality, Exercise 1.4.36(vi): for every $\lambda > 0$ we have $\mu(\{F \geq \lambda\}) \leq \frac{1}{\lambda} \int_X F \, d\mu$, and since $N \subset \{F \geq \lambda\}$ for every λ, monotonicity gives $\mu(N) \leq \frac{1}{\lambda} \int_X F \, d\mu \to 0$ as $\lambda \to \infty$, so $\mu(N) = 0$.)

By the counting identity above, for every $x \in X \setminus N$ the set $\{n \in \mathbb{N} : x \in E_n\}$ has finite cardinality $F(x)$. Hence almost every $x \in X$ lies in at most finitely many of the E_n, as claimed.

Problem 1.4.45.

- (i) Give an alternate proof of the Borel-Cantelli lemma (Exercise 1.4.44) that does not go through any of the convergence theorems, but instead exploits the more basic properties of measure from Exercise 1.4.23.
- (ii) Give a counterexample that shows that the Borel-Cantelli lemma can fail if the condition $\sum_{n=1}^{\infty} \mu(E_n) < \infty$ is relaxed to $\lim_{n\to\infty} \mu(E_n) = 0$.

Solution. (i) Let $(X, \mathcal{B}, \mu)$ be a measure space and $E_1, E_2, \ldots \in \mathcal{B}$ with $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. The set of points lying in infinitely many E_n is

$$E := \limsup_{n\to\infty} E_n = \bigcap_{N=1}^{\infty} F_N, \qquad F_N := \bigcup_{n=N}^{\infty} E_n.$$

Indeed, $x \in \bigcap_N F_N$ if and only if for every N there is some $n \geq N$ with $x \in E_n$, which is exactly the statement that $\{n : x \in E_n\}$ is unbounded, i.e. infinite. Each F_N is a countable union of $\mathcal{B}$-measurable

sets and hence $\mathcal{B}$-measurable, and so is E (a countable intersection of measurable sets, using that $\mathcal{B}$ is a σ-algebra).
We must show $\mu(E) = 0$. Fix $N \geq 1$. By countable subadditivity (Exercise 1.4.23(i)) applied to the sets $E_N, E_{N+1}, \ldots$,

$$\mu(F_N) = \mu\Big(\bigcup_{n=N}^{\infty} E_n\Big) \leq \sum_{n=N}^{\infty} \mu(E_n) =: t_N.$$

Since $E \subset F_N$, monotonicity of measure gives $\mu(E) \leq \mu(F_N) \leq t_N$. (Monotonicity is immediate from finite additivity: if $A \subset B$ are measurable then $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$, as all quantities are unsigned.)
Now t_N is the tail of the convergent series $\sum_{n=1}^{\infty} \mu(E_n)$, so $t_N \to 0$ as $N \to \infty$. Since $\mu(E) \leq t_N$ for every N, we conclude $\mu(E) = 0$. Thus almost every $x \in X$ fails to lie in E, i.e. lies in only finitely many E_n.
No convergence theorem was used: only the fact that $\mathcal{B}$ is a σ-algebra, countable subadditivity from Exercise 1.4.23(i), and monotonicity.
(One may alternatively finish with downwards monotone convergence, Exercise 1.4.23(iii): the F_N are decreasing with $\mu(F_1) \leq \sum_{n\geq 1} \mu(E_n) < \infty$, so $\mu(E) = \lim_{N\to\infty} \mu(F_N) \leq \lim_{N\to\infty} t_N = 0$. The finiteness hypothesis needed for (iii) is supplied by the convergence of the series.)
(ii) The hypothesis cannot be weakened to $\mu(E_n) \to 0$. Take $X = [0,1]$ with Lebesgue measure m, and use the "typewriter" sequence of dyadic intervals. Every integer $n \geq 1$ has a unique representation $n = 2^k + j$ with $k \geq 0$ an integer and $0 \leq j < 2^k$; define

$$E_n := \Big[\frac{j}{2^k},\ \frac{j+1}{2^k}\Big].$$

Thus $E_1 = [0,1]$; E_2, E_3 are the two dyadic halves; $E_4, \ldots, E_7$ are the four dyadic quarters; and so on, each generation k sweeping across $[0,1]$ from left to right.
Each E_n is a closed interval, hence Lebesgue measurable, and $m(E_n) = 2^{-k}$. Since $n = 2^k + j < 2^{k+1}$, we have $2^{-k} < 2/n$, so

$$m(E_n) < \frac{2}{n} \longrightarrow 0 \qquad \text{as } n \to \infty,$$

and the relaxed hypothesis $\lim_{n\to\infty} \mu(E_n) = 0$ holds.
However the conclusion of the Borel-Cantelli lemma fails as badly as possible, since every single point of $[0,1]$ lies in infinitely many E_n. Indeed, fix $x \in [0,1]$ and $k \geq 0$. The 2^k intervals $[j2^{-k}, (j+1)2^{-k}]$ for $0 \leq j < 2^k$ cover $[0,1]$, so there is some j with $x \in E_{2^k+j}$, and $2^k \leq 2^k + j < 2^{k+1}$. Hence for each k there is an index $n \in [2^k, 2^{k+1})$ with $x \in E_n$; these indices are distinct for distinct k, so $\{n : x \in E_n\}$ is infinite. Thus the set of x lying in infinitely many E_n is all of $[0,1]$, which has measure 1, not 0.
Consistently with (i), the series $\sum_n m(E_n)$ diverges here: generation k contributes 2^k terms each of size 2^{-k}, so $\sum_{n=1}^{\infty} m(E_n) = \sum_{k=0}^{\infty} 2^k \cdot 2^{-k} = \sum_{k=0}^{\infty} 1 = \infty$.

Problem 1.4.46. Under the hypotheses of the dominated convergence theorem (Theorem 1.4.49), establish also that $\|f_n - f\|_{L^1} \to 0$ as $n \to \infty$.

Solution. Recall the hypotheses of Theorem 1.4.49: $(X, \mathcal{B}, \mu)$ is a measure space, $f_1, f_2, \ldots : X \to \mathbb{C}$ are measurable and converge pointwise μ-almost everywhere to a measurable $f : X \to \mathbb{C}$, and there is an unsigned absolutely integrable $G : X \to [0, +\infty]$ with $|f_n| \leq G$ pointwise μ-almost everywhere for each n.
Step 1 (reduce to everywhere statements). For each n let A_n be the null set off which $|f_n| \leq G$, and let A_0 be the null set off which $f_n \to f$ pointwise. Then $A := A_0 \cup \bigcup_{n\geq 1} A_n$ is a countable union of null sets, hence null by countable subadditivity (Exercise 1.4.23(i)). Replacing f_n and f by $f_n 1_{X\setminus A}$ and $f 1_{X\setminus A}$ changes neither $\|f_n - f\|_{L^1}$ nor any of the hypotheses (functions agreeing μ-almost everywhere have the same integrals), so we may assume that $f_n \to f$ pointwise everywhere and $|f_n| \leq G$ pointwise everywhere.
Step 2 (the limit is dominated too). For every $x \in X$ we have $|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq G(x)$, by

continuity of the absolute value and the pointwise bound. Hence $|f| \leq G$ pointwise, and in particular f is absolutely integrable, since $\int_X |f|\,d\mu \leq \int_X G\,d\mu < \infty$ by monotonicity of the unsigned integral.
Step 3 (apply the dominated convergence theorem to the differences). Set

$$h_n := |f_n - f|.$$

Each h_n is measurable (as $f_n - f$ is measurable and $|\cdot|$ is continuous) and takes values in $[0, \infty)$, since f_n and f are complex-valued and hence finite. We check the two hypotheses of Theorem 1.4.49 for the sequence (h_n):

- Pointwise convergence: for every $x \in X$, $f_n(x) \to f(x)$ in $\mathbb{C}$, so $h_n(x) = |f_n(x) - f(x)| \to 0$. Thus $h_n \to 0$ pointwise everywhere, and the zero function is measurable.
- Domination: by the triangle inequality and Steps 1 and 2,
$$h_n = |f_n - f| \leq |f_n| + |f| \leq G + G = 2G$$

pointwise on X. The function $2G$ is unsigned, measurable, and absolutely integrable, since $\int_X 2G\,d\mu = 2\int_X G\,d\mu < \infty$ by homogeneity of the unsigned integral.
Theorem 1.4.49 applied to the sequence $h_1, h_2, \ldots$ with limit 0 and dominating function $2G$ therefore gives

$$\lim_{n\to\infty} \int_X h_n\,d\mu = \int_X 0\,d\mu = 0,$$

which is precisely the assertion

$$\|f_n - f\|_{L^1(\mu)} = \int_X |f_n - f|\,d\mu \longrightarrow 0 \qquad (n \to \infty).$$

This is a strengthening of Theorem 1.4.49 rather than a separate fact: by the triangle inequality for the integral,

$$\Big|\int_X f_n\,d\mu - \int_X f\,d\mu\Big| = \Big|\int_X (f_n - f)\,d\mu\Big| \leq \int_X |f_n - f|\,d\mu = \|f_n - f\|_{L^1(\mu)},$$

so L^1 convergence immediately recovers the convergence of the integrals. (Convergence in L^1 is genuinely stronger: for $f_n := 1_{[n,n+1]} - 1_{[n+1,n+2]}$ on $\mathbb{R}$ the integrals all vanish while the L^1 distance to the pointwise limit 0 is 2 for every n; of course no dominating function exists there.)

Problem 1.4.47. (Almost dominated convergence) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots : X \to \mathbb{C}$ be a sequence of measurable functions that converge pointwise μ-almost everywhere to a measurable limit $f : X \to \mathbb{C}$. Suppose that there are unsigned absolutely integrable functions $G, g_1, g_2, \ldots : X \to [0, +\infty]$ such that the $|f_n|$ are pointwise μ-almost everywhere bounded by $G + g_n$, and that $\int_X g_n\,d\mu \to 0$ as $n \to \infty$. Show that

$$\lim_{n\to\infty} \int_X f_n\,d\mu = \int_X f\,d\mu.$$

Solution. As in Exercise 1.4.46 we first discard a null set. Let A_0 be the null set off which $f_n \to f$ pointwise, and for each n let A_n be the null set off which $|f_n| \leq G + g_n$. Then $A := \bigcup_{n\geq 0} A_n$ is null, by countable subadditivity (Exercise 1.4.23(i)). Multiplying every function in sight by $1_{X\setminus A}$ changes no integral and no hypothesis, so we may and do assume that

$$f_n \to f \text{ pointwise on } X, \qquad |f_n| \leq G + g_n \text{ pointwise on } X \text{ for every } n.$$

Note first that each f_n is absolutely integrable: $\int_X |f_n|\,d\mu \leq \int_X (G + g_n)\,d\mu = \int_X G\,d\mu + \int_X g_n\,d\mu < \infty$, using monotonicity and additivity of the unsigned integral. So all the integrals in the statement are well defined.
Step 1 (the limit is dominated by G alone). Since $\int_X g_n\,d\mu \to 0$, we may choose a subsequence $n_1 < n_2 < \cdots$ with $\int_X g_{n_k}\,d\mu \leq 2^{-k}$ for all k. Then $\sum_{k=1}^{\infty} \int_X g_{n_k}\,d\mu < \infty$, so by Corollary 1.4.46

(Tonelli's theorem for sums and integrals),

$$\int_X \sum_{k=1}^{\infty} g_{n_k}\, d\mu = \sum_{k=1}^{\infty} \int_X g_{n_k}\, d\mu < \infty.$$

By the finiteness property of the unsigned integral (Exercise 1.4.36(vii)), $\sum_k g_{n_k}(x) < \infty$ for μ-almost every x, and for such x the terms of a convergent series tend to zero, so $g_{n_k}(x) \to 0$ as $k \to \infty$. Letting $k \to \infty$ in the bound $|f_{n_k}(x)| \leq G(x) + g_{n_k}(x)$ and using $f_{n_k}(x) \to f(x)$, we obtain

$$|f(x)| \leq G(x) \qquad \text{for } \mu\text{-almost every } x \in X.$$

Discarding this further null set as before, we may assume $|f| \leq G$ pointwise on X; in particular f is absolutely integrable.

Step 2 (splitting the error at the level $2G$). Put $h_n := |f_n - f|$, a measurable function with values in $[0, \infty)$ (finite, since f_n, f are complex-valued). By the triangle inequality and Step 1,

$$h_n \leq |f_n| + |f| \leq (G + g_n) + G = 2G + g_n \qquad \text{pointwise on } X.$$

Now decompose h_n at the truncation level $2G$: for any real $a \geq 0$ and any $b \in [0, +\infty]$ we have the identity $a = \min(a, b) + (a - b)^+$, where $(a - b)^+ := \max(a - b, 0)$ is interpreted as 0 when $b = +\infty$. Applying this pointwise with $a = h_n(x)$, $b = 2G(x)$,

$$h_n = \min(h_n, 2G) + (h_n - 2G)^+.$$

Moreover $h_n \leq 2G + g_n$ forces $(h_n - 2G)^+ \leq g_n$ pointwise: where $G(x) = +\infty$ the left side is 0, and where $G(x) < \infty$ we may subtract $2G(x)$ from the bound to get $h_n(x) - 2G(x) \leq g_n(x)$, whence $(h_n(x) - 2G(x))^+ \leq g_n(x)$ since $g_n \geq 0$. Both terms of the decomposition are unsigned and measurable, so by additivity (Theorem 1.4.38) and monotonicity (Exercise 1.4.36(ii)) of the unsigned integral,

$$\int_X h_n\, d\mu = \int_X \min(h_n, 2G)\, d\mu + \int_X (h_n - 2G)^+\, d\mu \leq \int_X \min(h_n, 2G)\, d\mu + \int_X g_n\, d\mu.$$

Step 3 (dominated convergence for the truncated part). The functions $H_n := \min(h_n, 2G)$ are measurable, real-valued and finite (being bounded above by $h_n < \infty$), and they satisfy $0 \leq H_n \leq 2G$ pointwise, where $2G$ is unsigned and absolutely integrable ($\int_X 2G\, d\mu = 2\int_X G\, d\mu < \infty$). Since $f_n \to f$ pointwise, $h_n \to 0$ pointwise, and hence $0 \leq H_n \leq h_n \to 0$ pointwise. By the dominated convergence theorem (Theorem 1.4.49) applied to (H_n) with dominating function $2G$,

$$\int_X \min(h_n, 2G)\, d\mu \longrightarrow \int_X 0\, d\mu = 0.$$

Step 4 (conclusion). Combining Steps 2 and 3 with the hypothesis $\int_X g_n\, d\mu \to 0$,

$$\|f_n - f\|_{L^1(\mu)} = \int_X h_n\, d\mu \leq \int_X \min(h_n, 2G)\, d\mu + \int_X g_n\, d\mu \longrightarrow 0.$$

Finally, by linearity of the integral on absolutely integrable functions and the triangle inequality for integrals,

$$\Big|\int_X f_n\, d\mu - \int_X f\, d\mu\Big| = \Big|\int_X (f_n - f)\, d\mu\Big| \leq \int_X |f_n - f|\, d\mu \longrightarrow 0,$$

so $\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu$, as required. (As in Exercise 1.4.46, we have in fact proved the stronger conclusion $\|f_n - f\|_{L^1(\mu)} \to 0$.)

Problem 1.4.48. (Defect version of Fatou's lemma) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots : X \to [0, +\infty]$ be a sequence of unsigned absolutely integrable functions that converges

pointwise to an absolutely integrable limit f. Show that

$$\int_X f_n\,d\mu - \int_X f\,d\mu - \|f-f_n\|_{L^1(\mu)} \to 0$$

as $n \to \infty$. (Hint: Apply the dominated convergence theorem (Theorem 1.4.49) to $\min(f_n, f)$.) Informally, this tells us that the gap between the left and right hand sides of Fatou's lemma can be measured by the quantity $\|f-f_n\|_{L^1(\mu)}$.

Solution. Write $h_n := \min(f_n, f)$, which is unsigned and measurable (the pointwise minimum of two measurable functions is measurable).

Step 1 (everything is finite almost everywhere). Since each f_n and f is unsigned absolutely integrable, i.e. $\int_X f_n\,d\mu < \infty$ and $\int_X f\,d\mu < \infty$, the finiteness property of the unsigned integral (Exercise 1.4.36(vii)) says that each of $f_1, f_2, \dots, f$ is finite outside a null set. The union A of these countably many null sets is null by countable subadditivity (Exercise 1.4.23(i)). Replacing every function by its product with $1_{X\setminus A}$ alters no integral and preserves the hypotheses (pointwise convergence still holds, now everywhere), so we may assume that f_n and f take values in $[0,\infty)$ everywhere on X.

Step 2 (the algebraic identity). For real numbers $a, b \geq 0$ one has

$$a - \min(a,b) = (a-b)^+, \qquad b - \min(a,b) = (b-a)^+, \qquad |a-b| = (a-b)^+ + (b-a)^+,$$

each verified by checking the two cases $a \leq b$ and $a > b$. Adding the first two identities and comparing with the third gives $|a-b| = (a-\min(a,b)) + (b-\min(a,b))$. Applying this pointwise with $a = f_n(x)$, $b = f(x)$ yields

$$|f_n - f| = (f_n - h_n) + (f - h_n) \qquad \text{pointwise on } X,$$

all three terms being unsigned (as $h_n \leq f_n$ and $h_n \leq f$) and measurable. Integrating and using additivity of the unsigned integral (Theorem 1.4.38),

$$\|f-f_n\|_{L^1(\mu)} = \int_X (f_n - h_n)\,d\mu + \int_X (f - h_n)\,d\mu.$$

Since $0 \leq h_n \leq f$ we have $\int_X h_n\,d\mu \leq \int_X f\,d\mu < \infty$, so $\int_X h_n\,d\mu$ is finite; combined with $\int_X f_n\,d\mu < \infty$ and additivity ($\int_X (f_n - h_n)\,d\mu + \int_X h_n\,d\mu = \int_X f_n\,d\mu$, and likewise for f) we may rearrange to get

$$\|f-f_n\|_{L^1(\mu)} = \Big(\int_X f_n\,d\mu - \int_X h_n\,d\mu\Big) + \Big(\int_X f\,d\mu - \int_X h_n\,d\mu\Big) = \int_X f_n\,d\mu + \int_X f\,d\mu - 2\int_X h_n\,d\mu.$$

Step 3 (dominated convergence for h_n). The functions $h_n = \min(f_n, f)$ are measurable and real-valued (finite, by Step 1), they satisfy the pointwise bound $0 \leq h_n \leq f$ with f unsigned absolutely integrable, and they converge pointwise to $\min(f,f) = f$, since $f_n \to f$ pointwise and min is continuous. Hence the dominated convergence theorem (Theorem 1.4.49) applies to the sequence (h_n) with dominating function f and gives

$$\lim_{n\to\infty} \int_X h_n\,d\mu = \int_X f\,d\mu.$$

Step 4 (conclusion). Substituting the expression from Step 2 into the quantity of interest, and noting that all the integrals involved are finite so the cancellations are legitimate,

$$\begin{aligned}\int_X f_n\,d\mu - \int_X f\,d\mu - \|f-f_n\|_{L^1(\mu)} &= \int_X f_n\,d\mu - \int_X f\,d\mu - \int_X f_n\,d\mu - \int_X f\,d\mu + 2\int_X h_n\,d\mu \\ &= 2\Big(\int_X h_n\,d\mu - \int_X f\,d\mu\Big).\end{aligned}$$

By Step 3 the right-hand side tends to 0 as $n \to \infty$, which is exactly the assertion to be proved.

Two remarks. First, since $h_n \leq f$ gives $\int_X h_n\,d\mu \leq \int_X f\,d\mu$, the quantity above is always ≤ 0; that is, $\int_X f_n\,d\mu \leq \int_X f\,d\mu + \|f-f_n\|_{L^1(\mu)}$ for every n, with the deficit tending to zero. Second, this indeed refines Fatou's lemma (Corollary 1.4.47) in the present setting: the identity of Step 4 says

$$\int_X f_n\,d\mu = \int_X f\,d\mu + \|f-f_n\|_{L^1(\mu)} + o(1),$$

so taking lim inf recovers $\liminf_{n\to\infty} \int_X f_n \, d\mu \geq \int_X f \, d\mu = \int_X \lim_{n\to\infty} f_n \, d\mu$, and the gap between the two sides of Fatou's inequality is precisely $\liminf_{n\to\infty} \|f - f_n\|_{L^1(\mu)}$.

Problem 1.4.49. Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $g : X \to [0, +\infty]$ be measurable. Show that the function $\mu_g : \mathcal{B} \to [0, +\infty]$ defined by the formula

$$\mu_g(E) := \int_X 1_E g \, d\mu = \int_E g \, d\mu$$

is a measure. (Such measures are studied in greater detail in Section 1.2 of *An epsilon of room, Vol. I.*)

Solution. Recall that a measure on the measurable space $(X, \mathcal{B})$ is a map $\nu : \mathcal{B} \to [0, +\infty]$ which is well defined (takes unsigned values on every $\mathcal{B}$-measurable set), satisfies $\nu(\emptyset) = 0$, and is countably additive: $\nu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \nu(E_n)$ whenever $E_1, E_2, \ldots \in \mathcal{B}$ are pairwise disjoint. We verify these three points for μ_g.

Well-definedness. Let $E \in \mathcal{B}$. Then 1_E is $\mathcal{B}$-measurable by Exercise 1.4.29(ii), hence so is the product $1_E g$ by Exercise 1.4.29(viii), and this product is unsigned (a product of two unsigned functions, with the convention $0 \cdot \infty = 0$ in force for the unsigned calculus). Concretely,

$$(1_E g)(x) = \begin{cases} g(x), & x \in E, \\ 0, & x \notin E. \end{cases}$$

Therefore the unsigned integral $\int_X 1_E g \, d\mu$ is defined and lies in $[0, +\infty]$; by the restriction property of the unsigned integral (Exercise 1.4.36(xi)) it equals $\int_E g \, d\mu$, so the two formulas in the statement do define the same quantity. Hence $\mu_g : \mathcal{B} \to [0, +\infty]$ is a well-defined map.

Empty set. Taking $E = \emptyset$ gives $1_\emptyset g = 0$ identically, and the unsigned integral of the zero function is 0 (it is the supremum of the integrals of simple functions dominated by 0, all of which are 0). Hence $\mu_g(\emptyset) = 0$.

Countable additivity. Let $E_1, E_2, \ldots \in \mathcal{B}$ be pairwise disjoint and put $E := \bigcup_{n=1}^{\infty} E_n$, which lies in $\mathcal{B}$ since $\mathcal{B}$ is a σ-algebra. We claim the pointwise identity

$$1_E g = \sum_{n=1}^{\infty} 1_{E_n} g \qquad \text{on } X,$$

where the right-hand side is a sum of unsigned terms, hence well defined in $[0, +\infty]$. To see this, fix $x \in X$. If $x \notin E$ then $x \notin E_n$ for every n, so both sides equal 0. If $x \in E$ then, by disjointness, there is exactly one index m with $x \in E_m$; the m-th summand equals $g(x)$ and all the others vanish, so the right-hand side equals $g(x) = 1_E(x) g(x)$, which is the left-hand side. (This is valid even when $g(x) = +\infty$, since a sum in $[0, +\infty]$ with a single nonzero term $+\infty$ equals $+\infty$.)

Each function $1_{E_n} g$ is unsigned and measurable, so Corollary 1.4.46 (Tonelli's theorem for sums and integrals) applies and yields

$$\mu_g(E) = \int_X 1_E g \, d\mu = \int_X \sum_{n=1}^{\infty} 1_{E_n} g \, d\mu = \sum_{n=1}^{\infty} \int_X 1_{E_n} g \, d\mu = \sum_{n=1}^{\infty} \mu_g(E_n).$$

Thus μ_g is unsigned, vanishes on the empty set, and is countably additive, i.e. μ_g is a measure on $(X, \mathcal{B})$.

Two consistency checks worth recording. Taking $g \equiv 1$ recovers $\mu_g = \mu$, since $\int_X 1_E \, d\mu = \mu(E)$. And if E is μ-null then $1_E g$ vanishes μ-almost everywhere, so $\mu_g(E) = 0$; that is, μ_g is absolutely continuous with respect to μ, which is the point of departure for the theory of such measures referred to in the statement.

Problem 1.4.50. (Characterisation of the unsigned integral). Let $(X, \mathcal{B})$ be a measurable space. Let $I : f \mapsto I(f)$ be a map from the space $\mathcal{U}(X, \mathcal{B})$ of unsigned measurable functions $f : X \to [0, +\infty]$ to $[0, +\infty]$ that obeys the following axioms:

- (i) (Homogeneity) For every $f \in \mathcal{U}(X, \mathcal{B})$ and $c \in [0, +\infty]$, one has $I(cf) = cI(f)$.
- (ii) (Finite additivity) For every $f, g \in \mathcal{U}(X, \mathcal{B})$, one has $I(f + g) = I(f) + I(g)$.
- (iii) (Monotone convergence) If $0 \leq f_1 \leq f_2 \leq \ldots$ are a non-decreasing sequence of unsigned measurable functions, then $I(\lim_{n\to\infty} f_n) = \lim_{n\to\infty} I(f_n)$.

Then there exists a unique measure μ on $(X, \mathcal{B})$ such that $I(f) = \int_X f \, d\mu$ for all $f \in \mathcal{U}(X, \mathcal{B})$. Furthermore, μ is given by the formula $\mu(E) := I(1_E)$ for all $\mathcal{B}$-measurable sets E.

Solution. Throughout we use the convention $0 \cdot \infty = \infty \cdot 0 = 0$, which is the convention in force for the unsigned integral.

Uniqueness, and the formula for μ. Suppose μ is any measure on $(X, \mathcal{B})$ with $I(f) = \int_X f \, d\mu$ for all $f \in \mathcal{U}(X, \mathcal{B})$. For a $\mathcal{B}$-measurable set E the indicator 1_E is unsigned measurable (Exercise 1.4.29(ii)) and simple, so by Exercise 1.4.36(v) and Exercise 1.4.34(ii),

$$I(1_E) \;=\; \int_X 1_E \, d\mu \;=\; \mathrm{Simp}\int_X 1_E \, d\mu \;=\; \mu(E).$$

Thus any such μ must be given by $\mu(E) = I(1_E)$, and in particular there is at most one such measure. It remains to show that this formula does define a measure, and that it represents I.

The formula defines a measure. Define $\mu : \mathcal{B} \to [0, +\infty]$ by $\mu(E) := I(1_E)$.

First, $\mu(\emptyset) = I(0)$, where 0 denotes the zero function. Applying homogeneity (i) to $f = 0$ and $c = 0$ gives $I(0) = I(0 \cdot 0) = 0 \cdot I(0) = 0$, using the convention $0 \cdot \infty = 0$ (so the argument is valid even if a priori $I(0) = +\infty$). Hence $\mu(\emptyset) = 0$.

Next, let $E_1, E_2, \cdots \in \mathcal{B}$ be pairwise disjoint, and put $E := \bigcup_{n=1}^{\infty} E_n$. For each N, disjointness gives the pointwise identity

$$1_{E_1 \cup \cdots \cup E_N} = \sum_{n=1}^{N} 1_{E_n},$$

so by finite additivity (ii) and induction on N,

$$\mu\Big(\bigcup_{n=1}^{N} E_n\Big) = I\Big(\sum_{n=1}^{N} 1_{E_n}\Big) = \sum_{n=1}^{N} I(1_{E_n}) = \sum_{n=1}^{N} \mu(E_n).$$

The functions $1_{\bigcup_{n\leq N} E_n}$ are unsigned measurable (Exercise 1.4.29(ii)) and non-decreasing in N, and they converge pointwise to 1_E: indeed $1_{\bigcup_{n\leq N} E_n}(x) = 1$ for some (hence all larger) N precisely when $x \in E$. Hence by monotone convergence (iii),

$$\mu(E) = I(1_E) = \lim_{N\to\infty} I\big(1_{\bigcup_{n\leq N} E_n}\big) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(E_n) = \sum_{n=1}^{\infty} \mu(E_n).$$

So μ is countably additive and vanishes on $\emptyset$, i.e. $(X, \mathcal{B}, \mu)$ is a measure space.

Step 1: I agrees with the integral on simple functions. Let $f : X \to [0, +\infty]$ be an unsigned simple function, so by Definition 1.4.35 it is measurable and takes finitely many values, which we list without repetition as $a_1, \ldots, a_k \in [0, +\infty]$. Writing $A_i := f^{-1}(\{a_i\})$, the A_i are non-empty, pairwise disjoint $\mathcal{B}$-measurable sets partitioning X (measurable by Exercise 1.4.29(iii), each singleton $\{a_i\}$ being a Borel subset of $[0, +\infty]$), and

$$f = \sum_{i=1}^{k} a_i 1_{A_i}$$

pointwise. Each summand $a_i 1_{A_i}$ is again simple, so finite additivity of the simple integral (Exercise 1.4.34(iv), iterated $k - 1$ times), followed by homogeneity (Exercise 1.4.34(iii)) and compatibility with

the measure (Exercise 1.4.34(ii)), gives

$$\mathrm{Simp}\int_X f\,d\mu = \sum_{i=1}^{k} a_i\,\mathrm{Simp}\int_X 1_{A_i}\,d\mu = \sum_{i=1}^{k} a_i\mu(A_i).$$

On the other hand, by finite additivity (ii) (iterated $k-1$ times) and homogeneity (i),

$$I(f) = \sum_{i=1}^{k} I(a_i 1_{A_i}) = \sum_{i=1}^{k} a_i I(1_{A_i}) = \sum_{i=1}^{k} a_i\mu(A_i).$$

Therefore $I(f) = \mathrm{Simp}\int_X f\,d\mu = \int_X f\,d\mu$, the last equality by Exercise 1.4.36(v).
Step 2: approximation of a general f from below by simple functions. Let $f \in \mathcal{U}(X,\mathcal{B})$ be arbitrary. For each $n \geq 1$ define

$$f_n := \sum_{k=1}^{n2^n} 2^{-n} 1_{E_{n,k}}, \qquad E_{n,k} := \{x \in X : f(x) \geq k2^{-n}\}.$$

Each $E_{n,k}$ is $\mathcal{B}$-measurable (Exercise 1.4.29(i) applied to the level sets, noting $\{f \geq \lambda\} = \bigcap_m \{f > \lambda - 1/m\}$), so f_n is a finite nonnegative combination of indicators of measurable sets; it is measurable and takes only the finitely many values $0, 2^{-n}, 2\cdot 2^{-n}, \ldots, n$, hence is simple. Unwinding the sum, f_n has the closed form

$$f_n(x) = \min\big(2^{-n}\lfloor 2^n f(x)\rfloor,\ n\big),$$

with the convention that this equals n when $f(x) = +\infty$.
The sequence is non-decreasing: since $\lfloor 2^{n+1}t\rfloor \geq 2\lfloor 2^n t\rfloor$ for all $t \in [0,+\infty]$, we have $2^{-(n+1)}\lfloor 2^{n+1}t\rfloor \geq 2^{-n}\lfloor 2^n t\rfloor$, and $n+1 \geq n$; as min is non-decreasing in each argument, $f_{n+1} \geq f_n$ pointwise.
The sequence converges pointwise to f: if $f(x) = +\infty$ then $f_n(x) = n \to +\infty = f(x)$; if $f(x) < \infty$ then for all $n > f(x)$ we have $f_n(x) = 2^{-n}\lfloor 2^n f(x)\rfloor \in [f(x) - 2^{-n}, f(x)]$, so $f_n(x) \to f(x)$. Also $0 \leq f_n \leq f$ throughout.
Step 3: conclusion. Applying the monotone convergence axiom (iii) to the non-decreasing sequence $0 \leq f_1 \leq f_2 \leq \ldots$ with pointwise limit f gives

$$I(f) = \lim_{n\to\infty} I(f_n),$$

while the monotone convergence theorem (Theorem 1.4.44) applied to the same sequence gives

$$\int_X f\,d\mu = \lim_{n\to\infty}\int_X f_n\,d\mu.$$

By Step 1, $I(f_n) = \int_X f_n\,d\mu$ for every n, so the two limits coincide, and $I(f) = \int_X f\,d\mu$.
Thus $\mu(E) := I(1_E)$ defines a measure representing I, and by the uniqueness argument above it is the only such measure. □
Remark. All three axioms are needed. Homogeneity and finite additivity alone determine I on simple functions from the set function $E \mapsto I(1_E)$, but only guarantee that this set function is finitely additive; the monotone convergence axiom is what upgrades it to countable additivity and simultaneously what transfers the identity from simple functions to all of $\mathcal{U}(X,\mathcal{B})$.

Problem 1.4.51. Let $(X,\mathcal{B},\mu)$ be a finite measure space (i.e. $\mu(X) < \infty$), and let $f : X \to \mathbb{R}$ be a bounded function. Suppose that μ is complete (see Definition 1.4.31). Suppose that the upper integral

$$\overline{\int_X} f\,d\mu := \inf_{g \geq f;\, g \text{ simple}} \int_X g\,d\mu$$

and lower integral

$$\underline{\int_X} f\,d\mu := \sup_{h\le f;\, h \text{ simple}} \int_X h\,d\mu$$

agree. Show that f is measurable. (This is a converse to Exercise 1.3.11.)

Solution. Throughout, a simple function means a real-valued simple function, i.e. a measurable $g : X \to \mathbb{R}$ taking only finitely many values; this is the right reading here, since f is allowed to take negative values and there would otherwise be no simple functions below f at all. Such a g is bounded, hence absolutely integrable (as $\mu(X) < \infty$), so $\int_X g\,d\mu$ is a well-defined real number in the sense of Definition 1.4.39, and it agrees with the simple integral of Definition 1.4.35 when g is unsigned (Exercise 1.4.36(v)). Both classes of competitor are non-empty: if $|f| \le M$ then the constants M and $-M$ are simple functions lying above and below f respectively.

We record once and for all the monotonicity we shall use: if $u \le v$ are real-valued and absolutely integrable, then $v - u$ is unsigned and absolutely integrable, so $\int_X (v-u)\,d\mu \ge 0$, and by linearity of the integral (Exercise 1.4.41(ii)) we get $\int_X v\,d\mu = \int_X u\,d\mu + \int_X (v-u)\,d\mu \ge \int_X u\,d\mu$.

Reduction to the unsigned case. Fix $M \in (0,\infty)$ with $|f| \le M$ and set $F := f + M$, so $0 \le F \le C$ where $C := 2M$. The map $g \mapsto g + M$ is a bijection from the simple functions $\ge f$ onto the simple functions $\ge F$, and likewise from the simple functions $\le f$ onto those $\le F$; by linearity of the absolutely convergent integral (Exercise 1.4.41(ii)) and finiteness of μ,

$$\int_X (g+M)\,d\mu = \int_X g\,d\mu + M\mu(X),$$

a finite shift. Hence the upper and lower integrals of F are those of f shifted by the same finite constant $M\mu(X)$, and so they agree as well.

It suffices to prove that F is measurable. Indeed, suppose $F : X \to [0,C]$ is measurable, so that all the sets $\{F > c\}$ are $\mathcal{B}$-measurable (Exercise 1.4.29(i)); then so are all the sets $\{F \ge c\} = \bigcap_{m\ge 1}\{F > c - \frac{1}{m}\}$ and $\{F < c\} = X \setminus \{F \ge c\}$. The positive and negative parts of $f = F - M$ then satisfy, for $\lambda \ge 0$,

$$\{\max(f,0) > \lambda\} = \{F > \lambda + M\}, \qquad \{\max(-f,0) > \lambda\} = \{F < M - \lambda\},$$

while for $\lambda < 0$ both level sets are all of X; so $\max(f,0)$ and $\max(-f,0)$ are measurable by Exercise 1.4.29(i), and hence f is measurable by Exercise 1.4.29(v). Write

$$S := \overline{\int_X} F\,d\mu = \underline{\int_X} F\,d\mu,$$

which lies in $[0, C\mu(X)]$ and in particular is finite.

Normalising the competitors. If g is simple with $g \ge F$ then automatically $g \ge 0$, and $\min(g,C)$ is again simple, still satisfies $\min(g,C) \ge F$ (as $F \le C$), and satisfies $\min(g,C) \le g$, so by the monotonicity recorded above its integral is no larger. Similarly if h is simple with $h \le F$ then $\max(h,0)$ is simple, still satisfies $\max(h,0) \le F$ (as $F \ge 0$), and has integral no smaller. (Both are simple: each takes only finitely many values, $\min(g,C)$ is measurable by the minimum rule recorded in the next paragraph applied to the unsigned functions g and the constant C, and $\max(h,0) = h_+$ is measurable by Exercise 1.4.29(v).) So the infimum defining $\overline{\int_X} F\,d\mu$ and the supremum defining $\underline{\int_X} F\,d\mu$ are unchanged if we restrict to simple functions taking values in $[0,C]$.

Accordingly, for each $n \ge 1$ choose simple functions $g_n, h_n : X \to [0,C]$ with

$$h_n \le F \le g_n, \qquad \int_X g_n\,d\mu \le S + \tfrac{1}{n}, \qquad \int_X h_n\,d\mu \ge S - \tfrac{1}{n}.$$

Making the sequences monotone. Put

$$G_n := \min(g_1,\dots,g_n), \qquad H_n := \max(h_1,\dots,h_n).$$

A pointwise minimum or maximum of finitely many unsigned measurable functions is measurable: $\{\min(u,v) > \lambda\} = \{u > \lambda\} \cap \{v > \lambda\}$ and $\{\max(u,v) > \lambda\} = \{u > \lambda\} \cup \{v > \lambda\}$, so Exercise 1.4.29(i)

applies. (All the g_i, h_i take values in $[0, C]$, so they are unsigned.) A minimum or maximum of finitely many simple functions also takes only finitely many values; hence G_n and H_n are simple, with values in $[0, C]$. Moreover G_n is non-increasing in n, H_n is non-decreasing in n, and

$$H_n \leq F \leq G_n$$

pointwise. Since $G_n \leq g_n$ and $H_n \geq h_n$, monotonicity of the integral gives

$$\int_X G_n \, d\mu \leq S + \tfrac{1}{n}, \qquad \int_X H_n \, d\mu \geq S - \tfrac{1}{n}.$$

The function $G_n - H_n$ is unsigned and simple, and $H_n + (G_n - H_n) = G_n$, so by finite additivity of the simple integral (Exercise 1.4.34(iv)) together with the finiteness of $\int_X H_n \, d\mu \leq C\mu(X) < \infty$,

$$\int_X (G_n - H_n) \, d\mu = \int_X G_n \, d\mu - \int_X H_n \, d\mu \leq \tfrac{2}{n}.$$

Passing to the limit. Since G_n is non-increasing and H_n is non-decreasing, both bounded in $[0, C]$, the pointwise limits

$$G := \lim_{n\to\infty} G_n = \inf_n G_n, \qquad H := \lim_{n\to\infty} H_n = \sup_n H_n$$

exist, take values in $[0, C]$, and are measurable as pointwise limits of unsigned measurable functions (Exercise 1.4.29(vi)). Passing to the limit in $H_n \leq F \leq G_n$ gives

$$H \leq F \leq G \quad \text{pointwise.}$$

The difference $G - H$ is measurable: G and H are bounded real-valued measurable functions, hence measurable as $\mathbb{C}$-valued functions (Exercise 1.4.29(iv)), so $G + (-1)H$ is measurable by Exercise 1.4.29(viii); being unsigned it is measurable as a $[0, +\infty]$-valued function. For every n we have $0 \leq G - H \leq G_n - H_n$ pointwise (since $G \leq G_n$ and $H \geq H_n$), so by monotonicity of the unsigned integral (Exercise 1.4.36(ii)),

$$\int_X (G - H) \, d\mu \leq \int_X (G_n - H_n) \, d\mu \leq \tfrac{2}{n}$$

for all n, whence $\int_X (G - H) \, d\mu = 0$. Set

$$N := \{x \in X : G(x) > H(x)\} = \bigcup_{m=1}^{\infty} \{x \in X : (G - H)(x) \geq \tfrac{1}{m}\},$$

a $\mathcal{B}$-measurable set; by Markov's inequality (Exercise 1.4.36(vi)) each set in the union has measure at most $m \int_X (G - H) \, d\mu = 0$, so $\mu(N) = 0$ by countable subadditivity. (Alternatively, invoke the vanishing property, Exercise 1.4.36(viii).) Thus $G(x) = H(x)$ for all $x \notin N$, and since $H \leq F \leq G$ we conclude

$$F(x) = G(x) \qquad \text{for all } x \in X \setminus N.$$

Using completeness. Fix $\lambda \in \mathbb{R}$ and set $P := \{x \in X : G(x) > \lambda\}$, which is $\mathcal{B}$-measurable since G is, and $Q := \{x \in X : F(x) > \lambda\}$. If $x \notin N$ then $F(x) = G(x)$, so $x \in P$ if and only if $x \in Q$; hence

$$P \setminus Q \subseteq N \qquad \text{and} \qquad Q \setminus P \subseteq N.$$

Both $P \setminus Q$ and $Q \setminus P$ are therefore sub-null sets, and by completeness of μ (Definition 1.4.31) each is a null set, in particular $\mathcal{B}$-measurable. Now

$$Q = (P \cap Q) \cup (Q \setminus P) = \big(P \setminus (P \setminus Q)\big) \cup (Q \setminus P),$$

and each of P, $P \setminus Q$, $Q \setminus P$ is $\mathcal{B}$-measurable, so Q is $\mathcal{B}$-measurable.
Since $\lambda \in \mathbb{R}$ was arbitrary and F takes values in $[0, C] \subseteq [0, +\infty]$, all level sets $\{F > \lambda\}$ are measurable (for $\lambda < 0$ the level set is all of X), so F is measurable by Exercise 1.4.29(i). By the reduction above, $f = F - M$ is measurable. $\square$

Remark. Completeness cannot be dropped. Take $X = [0,1]$ with the Borel σ-algebra and Lebesgue measure, and let $K \subseteq [0,1]$ be the Cantor set, which is Borel and has measure zero. The set K has cardinality $2^{\aleph_0}$, hence $2^{2^{\aleph_0}}$ subsets, while there are only $2^{\aleph_0}$ Borel subsets of $[0,1]$; so some $A \subseteq K$ is not Borel. Then $f := 1_A$ satisfies $0 \le f \le 1_K$ and is bounded. Taking the simple competitor $g := 1_K \ge f$ gives $\overline{\int_X} f\, d\mu \le \int_X 1_K\, d\mu = 0$, while $h := 0 \le f$ gives $\underline{\int_X} f\, d\mu \ge 0$; since the lower integral never exceeds the upper one (by the monotonicity recorded at the start), both equal 0 and so they agree. Yet f is not Borel measurable, since $f^{-1}(\{1\}) = A$. The only hypothesis that fails here is completeness. Finiteness of μ, together with boundedness of f, was used to guarantee that all the integrals appearing above are finite, so that the differences $\int_X G_n\, d\mu - \int_X H_n\, d\mu$ make sense.

2.5 §1.5 — Modes of Convergence

Problem 1.5.1. (Linearity of convergence) Let $(X, \mathcal{B}, \mu)$ be a measure space, let $f_n, g_n : X \to \mathbb{C}$ be sequences of measurable functions, and let $f, g : X \to \mathbb{C}$ be measurable functions.
(i) Show that f_n converges to f along one of the above seven modes of convergence if and only if $|f_n - f|$ converges to 0 along the same mode.
(ii) If f_n converges to f along one of the above seven modes of convergence, and g_n converges to g along the same mode, show that $f_n + g_n$ converges to $f + g$ along the same mode, and that cf_n converges to cf along the same mode for any $c \in \mathbb{C}$.
(iii) (Squeeze test) If f_n converges to 0 along one of the above seven modes, and $|g_n| \le f_n$ pointwise for each n, show that g_n converges to 0 along the same mode.

Solution. The seven modes are: pointwise, uniform, pointwise almost everywhere, in L^∞ norm (uniformly almost everywhere), almost uniform, in L^1 norm, and in measure.
Part (i). Write $h_n := |f_n - f|$, an unsigned measurable function. The point is simply that

$$|h_n(x) - 0| = \big|\, |f_n(x) - f(x)| - 0 \,\big| = |f_n(x) - f(x)| \qquad \text{for every } x \in X,$$

while each of the seven definitions is a statement purely about the family of numbers $|f_n(x) - f(x)|$, $x \in X$, $n \in \mathbb{N}$. Indeed:

- $f_n \to f$ pointwise means: for all x and all $\varepsilon > 0$ there is N with $|f_n(x) - f(x)| \le \varepsilon$ for $n \ge N$;
- $f_n \to f$ uniformly means: for all $\varepsilon > 0$ there is N with $|f_n(x) - f(x)| \le \varepsilon$ for all $n \ge N$ and all x;
- pointwise a.e. convergence is the first condition restricted to x outside a null set;
- L^∞ convergence means: for all $\varepsilon > 0$ there is N such that for each $n \ge N$ one has $|f_n(x) - f(x)| \le \varepsilon$ for a.e. x;
- almost uniform convergence means: for all $\varepsilon > 0$ there is $E \in \mathcal{B}$ with $\mu(E) \le \varepsilon$ such that the uniform condition holds for x ranging over $X \setminus E$;
- L^1 convergence means $\int_X |f_n - f|\, d\mu \to 0$;
- convergence in measure means $\mu(\{x : |f_n(x) - f(x)| \ge \varepsilon\}) \to 0$ for each $\varepsilon > 0$.

Substituting the pair $(h_n, 0)$ for the pair (f_n, f) leaves every one of these quantities unchanged, by the displayed identity. Hence $f_n \to f$ in a given mode if and only if $h_n = |f_n - f| \to 0$ in that same mode.
Part (ii). We use the pointwise triangle inequality

$$|(f_n + g_n) - (f + g)|(x) \le |f_n(x) - f(x)| + |g_n(x) - g(x)|$$

and the identity $|cf_n(x) - cf(x)| = |c|\, |f_n(x) - f(x)|$.
Sums, mode by mode.
Pointwise. Fix x and $\varepsilon > 0$. Choose N_1 with $|f_n(x) - f(x)| \le \varepsilon/2$ for $n \ge N_1$ and N_2 with $|g_n(x) - g(x)| \le \varepsilon/2$ for $n \ge N_2$. For $n \ge \max(N_1, N_2)$ the triangle inequality gives $|(f_n + g_n) - (f + g)|(x) \le \varepsilon$.
Uniform. The same argument, with N_1, N_2 now chosen independently of x.
Pointwise a.e. Let A be a null set off which $f_n \to f$ pointwise and B a null set off which $g_n \to g$ pointwise. Then $A \cup B$ is null (monotonicity and finite subadditivity of μ, Exercise 1.4.23(i)) and off $A \cup B$ the pointwise argument applies.
L^∞. Fix $\varepsilon > 0$ and choose N_1, N_2 as in the definition. Fix $n \ge \max(N_1, N_2)$. There are null sets A_n, B_n with $|f_n - f| \le \varepsilon/2$ off A_n and $|g_n - g| \le \varepsilon/2$ off B_n; off the null set $A_n \cup B_n$ we get $|(f_n + g_n) - (f + g)| \le \varepsilon$. As $n \ge \max(N_1, N_2)$ was arbitrary, this is L^∞ convergence.

Almost uniform. Given $\varepsilon > 0$, choose E with $\mu(E) \leq \varepsilon/2$ and $f_n \to f$ uniformly on $X \setminus E$, and F with $\mu(F) \leq \varepsilon/2$ and $g_n \to g$ uniformly on $X \setminus F$. Then $\mu(E \cup F) \leq \varepsilon$ by subadditivity, and on $X \setminus (E \cup F)$ both sequences converge uniformly, hence so does the sum by the uniform case.
L^1. By monotonicity and additivity of the unsigned integral (Exercise 1.4.36(ii) and Theorem 1.4.38),

$$\int_X |(f_n + g_n) - (f + g)|\, d\mu \leq \int_X |f_n - f|\, d\mu + \int_X |g_n - g|\, d\mu \to 0.$$

In measure. For $\varepsilon > 0$ the triangle inequality gives the inclusion

$$\{|(f_n + g_n) - (f + g)| \geq \varepsilon\} \subseteq \{|f_n - f| \geq \varepsilon/2\} \cup \{|g_n - g| \geq \varepsilon/2\},$$

so by finite subadditivity the measure of the left-hand set is at most the sum of two quantities each tending to 0.
Scalar multiples. If $c = 0$ then $cf_n = cf = 0$ and every mode is trivially satisfied. If $c \neq 0$, then $|cf_n - cf| = |c|\,|f_n - f|$ pointwise. For the first five modes one simply replaces the tolerance ε by $\varepsilon/|c|$ in the definition; for L^1 one uses homogeneity of the unsigned integral (Exercise 1.4.36(iii)) to get $\int |cf_n - cf|\, d\mu = |c| \int |f_n - f|\, d\mu \to 0$; and for convergence in measure one has the exact identity of sets $\{|cf_n - cf| \geq \varepsilon\} = \{|f_n - f| \geq \varepsilon/|c|\}$, whose measure tends to 0.
Part (iii). The hypothesis $|g_n| \leq f_n$ pointwise forces $f_n \geq 0$, so $|f_n - 0| = f_n$, and we have the pointwise bound

$$|g_n(x) - 0| \leq |f_n(x) - 0| \qquad \text{for every } x \in X.$$

So it suffices to observe that each of the seven modes is monotone: if $0 \leq h_n \leq k_n$ pointwise and $k_n \to 0$ in a given mode, then $h_n \to 0$ in that mode. Indeed:

- pointwise and uniform: if $k_n(x) \leq \varepsilon$ (for the relevant x and $n \geq N$) then $h_n(x) \leq \varepsilon$ as well;
- pointwise a.e.: apply the pointwise statement off the exceptional null set;
- L^∞: if $k_n \leq \varepsilon$ outside a null set then $h_n \leq \varepsilon$ outside that same null set;
- almost uniform: if $\mu(E) \leq \varepsilon$ and $k_n \to 0$ uniformly on $X \setminus E$, then by the uniform case $h_n \to 0$ uniformly on $X \setminus E$ with the same exceptional set E;
- L^1: by monotonicity of the unsigned integral (Exercise 1.4.36(ii)), $\int_X h_n\, d\mu \leq \int_X k_n\, d\mu \to 0$;
- in measure: $\{h_n \geq \varepsilon\} \subseteq \{k_n \geq \varepsilon\}$, so by monotonicity of μ the measure of the former tends to 0.

Applying this with $h_n := |g_n|$ and $k_n := f_n$, and then invoking part (i) to convert "$|g_n| \to 0$ in the mode" back into "$g_n \to 0$ in the mode", we conclude that g_n converges to 0 along the same mode.

Problem 1.5.2. (Easy implications) Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_n : X \to \mathbb{C}$ and $f : X \to \mathbb{C}$ be measurable functions.
(i) If f_n converges to f uniformly, then f_n converges to f pointwise.
(ii) If f_n converges to f uniformly, then f_n converges to f in L^∞ norm. Conversely, if f_n converges to f in L^∞ norm, then f_n converges to f uniformly outside of a null set (i.e. there exists a null set E such that the restriction $f_n|_{X \setminus E}$ of f_n to the complement of E converges to the restriction $f|_{X \setminus E}$ of f).
(iii) If f_n converges to f in L^∞ norm, then f_n converges to f almost uniformly.
(iv) If f_n converges to f almost uniformly, then f_n converges to f pointwise almost everywhere.
(v) If f_n converges to f pointwise, then f_n converges to f pointwise almost everywhere.
(vi) If f_n converges to f in L^1 norm, then f_n converges to f in measure.
(vii) If f_n converges to f almost uniformly, then f_n converges to f in measure.

Solution. Throughout, following Definition 1.4.31, a *null set* is a $\mathcal{B}$-measurable set of measure zero and a *sub-null set* is a subset of one; a property holds almost everywhere if it holds outside a sub-null set. We use freely that a countable union of null sets is null (countable subadditivity, Exercise 1.4.23(i)), and the following remark, which lets us avoid any completeness assumption on $(X, \mathcal{B}, \mu)$: if $S \in \mathcal{B}$ is contained in a sub-null set, then S is itself null, since S is then contained in some null set F and monotonicity gives $\mu(S) \leq \mu(F) = 0$. In particular, since f_n and f are measurable, whenever a bound such as $|f_n(x) - f(x)| \leq \varepsilon$ holds for almost every x, the *measurable* set $\{x \in X : |f_n(x) - f(x)| > \varepsilon\}$ is

a genuine null set.
Part (i). Let $\varepsilon > 0$. By uniform convergence there is N such that $|f_n(x) - f(x)| \leq \varepsilon$ for all $n \geq N$ and all $x \in X$. In particular, for each fixed x this same N witnesses the pointwise requirement at x. Hence $f_n(x) \to f(x)$ for every x.
Part (ii), forward direction. Let $\varepsilon > 0$ and take N as above. For each $n \geq N$ we have $|f_n(x) - f(x)| \leq \varepsilon$ for *every* x, and hence in particular for almost every x (the exceptional set being empty). This is exactly L^∞ convergence.
Part (ii), converse. Suppose $f_n \to f$ in L^∞ norm. For each positive integer k, apply the definition with $\varepsilon = 1/k$ to get N_k such that for every $n \geq N_k$ the bound $|f_n(x) - f(x)| \leq 1/k$ holds for almost every x. Define

$$E_{n,k} := \{x \in X : |f_n(x) - f(x)| > 1/k\} \quad (n \geq N_k), \qquad E_{n,k} := \emptyset \quad (n < N_k),$$

which are measurable because $|f_n - f|$ is a measurable function. For $n \geq N_k$ the set $E_{n,k}$ is contained in a sub-null set, hence is null by the remark above; so every $E_{n,k}$ is null. Set

$$E := \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} E_{n,k},$$

a countable union of null sets, hence a member of $\mathcal{B}$ of measure zero, i.e. a null set. Now fix $x \in X \setminus E$ and let $\eta > 0$; choose k with $1/k \leq \eta$. For every $n \geq N_k$ we have $x \notin E_{n,k}$ and therefore $|f_n(x) - f(x)| \leq 1/k \leq \eta$. Since N_k does not depend on x, this says precisely that the restrictions $f_n|_{X \setminus E}$ converge uniformly to $f|_{X \setminus E}$.
Part (iii). Suppose $f_n \to f$ in L^∞ norm, and let $\varepsilon > 0$ be given. By the converse half of part (ii) there is a null set E such that $f_n \to f$ uniformly on $X \setminus E$. Crucially, the set produced there lies in $\mathcal{B}$, as the definition of almost uniform convergence demands of the exceptional set. Since $\mu(E) = 0 \leq \varepsilon$, this E serves as the exceptional set for the given ε (and the same E works for every ε). Hence $f_n \to f$ almost uniformly.
Part (iv). Suppose $f_n \to f$ almost uniformly. For each positive integer k, apply the definition with $\varepsilon = 1/k$ to obtain $E_k \in \mathcal{B}$ with $\mu(E_k) \leq 1/k$ such that $f_n \to f$ uniformly, hence pointwise (part (i)), on $X \setminus E_k$. Put $E := \bigcap_{k=1}^{\infty} E_k$. Then $E \in \mathcal{B}$ and, by monotonicity, $\mu(E) \leq \mu(E_k) \leq 1/k$ for every k, so $\mu(E) = 0$. If $x \in X \setminus E$ then $x \notin E_k$ for some k, and therefore $f_n(x) \to f(x)$. Thus $f_n \to f$ pointwise outside the null set E, i.e. pointwise almost everywhere.
Part (v). Immediate: pointwise convergence is pointwise almost everywhere convergence with empty exceptional set (and $\mu(\emptyset) = 0$).
Part (vi). Fix $\varepsilon > 0$ and set $A_n := \{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}$, a measurable set. Markov's inequality (Exercise 1.4.36(vi)) applied to the unsigned measurable function $|f_n - f|$ with $\lambda = \varepsilon$ gives

$$\mu(A_n) \leq \frac{1}{\varepsilon} \int_X |f_n - f| \, d\mu = \frac{1}{\varepsilon} \|f_n - f\|_{L^1(\mu)}.$$

Since $\|f_n - f\|_{L^1(\mu)} \to 0$ by hypothesis and ε is fixed, $\mu(A_n) \to 0$. As $\varepsilon > 0$ was arbitrary, $f_n \to f$ in measure.
Part (vii). Fix $\varepsilon > 0$; we must show $\mu(\{|f_n - f| \geq \varepsilon\}) \to 0$. Let $\delta > 0$ be arbitrary. By almost uniform convergence there is $E \in \mathcal{B}$ with $\mu(E) \leq \delta$ such that $f_n \to f$ uniformly on $X \setminus E$; hence there is N with $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq N$ and all $x \in X \setminus E$. Consequently

$$\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\} \subseteq E \qquad (n \geq N),$$

so by monotonicity $\mu(\{|f_n - f| \geq \varepsilon\}) \leq \mu(E) \leq \delta$ for all $n \geq N$. Since $\delta > 0$ was arbitrary, $\limsup_{n \to \infty} \mu(\{|f_n - f| \geq \varepsilon\}) = 0$, which is the desired convergence in measure.

Problem 1.5.3. (Convergence for step functions) Let the notation and assumptions be as above; thus $f_n = A_n 1_{E_n}$ where the heights $A_n > 0$ are positive reals, the sets $E_n \in \mathcal{B}$ have positive width

$\mu(E_n) > 0$, the A_n either converge to zero or are bounded away from zero, and $E_N^* := \bigcup_{n \geq N} E_n$ denotes the Nth tail support. Establish the following claims:
(i) f_n converges uniformly to zero if and only if $A_n \to 0$ as $n \to \infty$.
(ii) f_n converges in L^∞ norm to zero if and only if $A_n \to 0$ as $n \to \infty$.
(iii) f_n converges almost uniformly to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\mu(E_N^*) \to 0$ as $N \to \infty$.
(iv) f_n converges pointwise to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^\infty E_N^* = \emptyset$.
(v) f_n converges pointwise almost everywhere to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^\infty E_N^*$ is a null set.
(vi) f_n converges in measure to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\mu(E_n) \to 0$ as $n \to \infty$.
(vii) f_n converges in L^1 norm if and only if $A_n \mu(E_n) \to 0$ as $n \to \infty$.

Solution. Two remarks before we begin. First, since $\mu(E_n) > 0$ the set E_n is non-empty, and indeed has positive measure; so $\sup_{x \in X} f_n(x) = A_n$, and moreover the bound "$f_n \leq \varepsilon$ almost everywhere" fails whenever $\varepsilon < A_n$. Second, by the standing dichotomy, if $A_n \not\to 0$ then there is $c > 0$ with $A_n \geq c$ for all n; we use this in every "only if" direction below. Note also that the tail supports E_N^* are decreasing in N, and that $x \in \bigcap_N E_N^*$ if and only if x lies in E_n for infinitely many n.
Part (i). Uniform convergence to zero says: for each $\varepsilon > 0$ there is N with $\sup_x f_n(x) \leq \varepsilon$ for all $n \geq N$. As $\sup_x f_n(x) = A_n$, this is exactly the statement that for each $\varepsilon > 0$ we have $A_n \leq \varepsilon$ eventually, i.e. $A_n \to 0$.
Part (ii). L^∞ convergence to zero says: for each $\varepsilon > 0$ there is N such that for all $n \geq N$ one has $f_n(x) \leq \varepsilon$ for almost every x. Since f_n equals A_n on the positive-measure set E_n, the condition "$f_n \leq \varepsilon$ a.e." holds if and only if $A_n \leq \varepsilon$. So again the criterion is that $A_n \leq \varepsilon$ eventually for every $\varepsilon > 0$, i.e. $A_n \to 0$.
Part (iii). Suppose first that $A_n \to 0$. Then $f_n \to 0$ uniformly by (i), and uniform convergence implies almost uniform convergence (take the exceptional set $E = \emptyset$). Suppose instead that $\mu(E_N^*) \to 0$. Given $\varepsilon > 0$, choose N with $\mu(E_N^*) \leq \varepsilon$ and put $E := E_N^*$. For $n \geq N$ we have $E_n \subseteq E_N^* = E$, so f_n vanishes identically on $X \setminus E$; in particular $f_n \to 0$ uniformly on $X \setminus E$. Hence $f_n \to 0$ almost uniformly.
Conversely, suppose $f_n \to 0$ almost uniformly and $A_n \not\to 0$, so $A_n \geq c > 0$ for all n. Let $\varepsilon > 0$ and choose $E \in \mathcal{B}$ with $\mu(E) \leq \varepsilon$ and $f_n \to 0$ uniformly on $X \setminus E$. Then there is N with $f_n(x) < c$ for all $n \geq N$ and $x \in X \setminus E$. But $f_n = A_n \geq c$ on E_n, so $E_n \cap (X \setminus E) = \emptyset$, i.e. $E_n \subseteq E$, for every $n \geq N$. Therefore $E_N^* \subseteq E$ and $\mu(E_N^*) \leq \varepsilon$; since the E_M^* decrease, $\mu(E_M^*) \leq \varepsilon$ for all $M \geq N$. As $\varepsilon > 0$ was arbitrary, $\mu(E_N^*) \to 0$.
Part (iv). If $A_n \to 0$ then $f_n \to 0$ uniformly by (i), hence pointwise by Exercise 1.5.2(i). If instead $\bigcap_N E_N^* = \emptyset$, fix $x \in X$; then $x \notin E_N^*$ for some N, so $x \notin E_n$ and $f_n(x) = 0$ for every $n \geq N$. Thus $f_n(x) \to 0$ for every x.
Conversely, suppose $f_n \to 0$ pointwise and $A_n \geq c > 0$ for all n. If some x belonged to $\bigcap_N E_N^*$, then $x \in E_n$ for infinitely many n, so $f_n(x) = A_n \geq c$ for infinitely many n, contradicting $f_n(x) \to 0$. Hence $\bigcap_N E_N^* = \emptyset$.
Part (v). If $A_n \to 0$ then $f_n \to 0$ pointwise (indeed uniformly), hence pointwise almost everywhere. If $Z := \bigcap_N E_N^*$ is a null set, then for each $x \in X \setminus Z$ the argument of part (iv) gives $f_n(x) = 0$ for all sufficiently large n, so $f_n \to 0$ pointwise outside the null set Z.
Conversely, suppose $f_n \to 0$ pointwise almost everywhere and $A_n \geq c > 0$ for all n. As in part (iv), every $x \in Z := \bigcap_N E_N^*$ satisfies $f_n(x) \geq c$ for infinitely many n, so $f_n(x)$ does not converge to 0. By hypothesis the set of x at which convergence fails is contained in some null set F, so $Z \subseteq F$. Now Z is measurable, being a countable intersection of the measurable sets E_N^* (each of which is a countable union of the $E_n \in \mathcal{B}$), so monotonicity gives $\mu(Z) \leq \mu(F) = 0$. Thus Z is a null set. (Note that this step does not require μ to be complete: it is the measurability of Z that upgrades the containment in a null set to $\mu(Z) = 0$.)
Part (vi). Observe that for $\varepsilon > 0$,

$$\{x \in X : f_n(x) \geq \varepsilon\} = \begin{cases} E_n & \text{if } A_n \geq \varepsilon, \\ \emptyset & \text{if } A_n < \varepsilon. \end{cases}$$

If $A_n \to 0$, then for each fixed $\varepsilon > 0$ we have $A_n < \varepsilon$ for all large n, so the above set is empty for all large n and its measure tends to 0. If $\mu(E_n) \to 0$, then the above set has measure at most $\mu(E_n) \to 0$ in either case. Either way $f_n \to 0$ in measure.
Conversely, suppose $f_n \to 0$ in measure and $A_n \geq c > 0$ for all n. Taking $\varepsilon := c$ in the display, $A_n \geq c = \varepsilon$ for every n, so $\{f_n \geq c\} = E_n$; convergence in measure then says exactly that $\mu(E_n) \to 0$.
Part (vii). Each $f_n = A_n 1_{E_n}$ is a simple unsigned function, so by compatibility of the unsigned integral with the simple integral (Exercise 1.4.36(v)),

$$\|f_n - 0\|_{L^1(\mu)} = \int_X A_n 1_{E_n}\, d\mu = A_n \mu(E_n)$$

(with the convention $A_n \cdot (+\infty) = +\infty$, legitimate since $A_n > 0$). Hence $f_n \to 0$ in L^1 norm if and only if $A_n \mu(E_n) \to 0$. Note that this criterion needs no appeal to the dichotomy on the A_n.

Problem 1.5.4. Let X have finite measure, and let $f_n : X \to \mathbb{C}$ and $f : X \to \mathbb{C}$ be measurable functions. Show that if f_n converges to f in L^∞ norm, then f_n also converges to f in L^1 norm.

Solution. If $\mu(X) = 0$ then every measurable function has vanishing integral and there is nothing to prove, so assume $0 < \mu(X) < \infty$.
Let $\eta > 0$ be arbitrary and set $\varepsilon := \eta/\mu(X) > 0$. By L^∞ convergence there is N such that for every $n \geq N$ we have

$$|f_n(x) - f(x)| \leq \varepsilon \qquad \text{for } \mu\text{-almost every } x \in X.$$

Fix such an n. Then $|f_n - f| \leq \varepsilon 1_X$ almost everywhere, so by monotonicity of the unsigned integral up to null sets (Exercise 1.4.36(ii)) together with compatibility with the simple integral (Exercise 1.4.36(v)),

$$\|f_n - f\|_{L^1(\mu)} = \int_X |f_n - f|\, d\mu \leq \int_X \varepsilon 1_X\, d\mu = \varepsilon \mu(X) = \eta.$$

Thus $\|f_n - f\|_{L^1(\mu)} \leq \eta$ for all $n \geq N$. Since $\eta > 0$ was arbitrary, $\|f_n - f\|_{L^1(\mu)} \to 0$, i.e. $f_n \to f$ in L^1 norm.
The hypothesis $\mu(X) < \infty$ cannot be dropped: on $X = \mathbb{R}$ with Lebesgue measure the functions $f_n := \frac{1}{n} 1_{[0,n]}$ of Example 1.5.3 converge to zero uniformly, hence in L^∞ norm, yet $\|f_n\|_{L^1} = 1$ for every n (escape to width infinity).

Problem 1.5.5. (Fast L^1 convergence) Suppose that $f_n, f : X \to \mathbb{C}$ are measurable functions such that $\sum_{n=1}^\infty \|f_n - f\|_{L^1(\mu)} < \infty$; thus, not only do the quantities $\|f_n - f\|_{L^1(\mu)}$ go to zero (which would mean L^1 convergence), but they converge in an absolutely summable fashion.
(i) Show that f_n converges pointwise almost everywhere to f.
(ii) Show that f_n converges almost uniformly to f.
[Hint: If you have trouble getting started, try working first in the special case in which $f_n = A_n 1_{E_n}$ are step functions and $f = 0$ and use Exercise 1.5.3 in order to gain some intuition. The second part of the exercise implies the first, but the first is a little easier to prove and may thus serve as a useful warmup. The $\varepsilon/2^n$ trick may come in handy for the second part.]

Solution. Write $\delta_n := \|f_n - f\|_{L^1(\mu)}$, so that $\sum_{n=1}^\infty \delta_n < \infty$. For each $N \geq 1$ define the unsigned function

$$G_N := \sum_{n=N}^\infty |f_n - f| : X \to [0, +\infty],$$

which is measurable, being the increasing pointwise limit of the measurable partial sums $\sum_{n=N}^M |f_n - f|$. By Tonelli's theorem for sums and integrals (Corollary 1.4.46),

$$\int_X G_N\, d\mu = \sum_{n=N}^\infty \int_X |f_n - f|\, d\mu = \sum_{n=N}^\infty \delta_n =: T_N.$$

Since $\sum_n \delta_n$ converges, the tails satisfy $T_N \to 0$ as $N \to \infty$; in particular $T_1 < \infty$.
Part (i). Take $N = 1$: the unsigned measurable function $G_1 = \sum_{n=1}^{\infty} |f_n - f|$ has $\int_X G_1 \, d\mu = T_1 < \infty$. By the finiteness property of the unsigned integral (Exercise 1.4.36(vii)), $G_1(x) < \infty$ for almost every x. Fix such an x. Then the series of non-negative reals $\sum_{n=1}^{\infty} |f_n(x) - f(x)|$ converges, and therefore its terms tend to zero: $|f_n(x) - f(x)| \to 0$, i.e. $f_n(x) \to f(x)$. Hence $f_n \to f$ pointwise almost everywhere.
Part (ii). Let $\varepsilon > 0$ be given. Since $T_N \to 0$, we may choose, for each positive integer k, an index N_k with

$$T_{N_k} \leq \frac{\varepsilon}{k\, 2^k}.$$

Put $B_k := \{x \in X : G_{N_k}(x) \geq 1/k\}$. By Markov's inequality (Exercise 1.4.36(vi)) applied to G_{N_k} with $\lambda = 1/k$,

$$\mu(B_k) \leq k \int_X G_{N_k} \, d\mu = k\, T_{N_k} \leq \frac{\varepsilon}{2^k}.$$

Set $E := \bigcup_{k=1}^{\infty} B_k \in \mathcal{B}$. By countable subadditivity (Exercise 1.4.23(i)),

$$\mu(E) \leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$

We claim $f_n \to f$ uniformly on $X \setminus E$. Indeed, let $\eta > 0$ and choose k with $1/k < \eta$. For any $x \in X \setminus E$ we have $x \notin B_k$, so $G_{N_k}(x) < 1/k$; and since every single term of the series defining G_{N_k} is dominated by the whole sum, for every $n \geq N_k$,

$$|f_n(x) - f(x)| \leq G_{N_k}(x) < \frac{1}{k} < \eta.$$

The threshold N_k depends on η but not on $x \in X \setminus E$, so this is uniform convergence on $X \setminus E$. As $\varepsilon > 0$ was arbitrary, $f_n \to f$ almost uniformly.
(As a check, part (ii) recovers part (i) via Exercise 1.5.2(iv).) It is instructive to see the criterion in the step function case $f_n = A_n 1_{E_n}$, $f = 0$: the hypothesis reads $\sum_n A_n \mu(E_n) < \infty$, whence $\mu(E_N^*) \leq \sum_{n \geq N} \mu(E_n) \leq \frac{1}{c} \sum_{n \geq N} A_n \mu(E_n) \to 0$ when the heights are bounded below by $c > 0$, which is precisely the tail-support criterion of Exercise 1.5.3(iii) for almost uniform convergence.

Problem 1.5.6. Suppose that $f_n : X \to \mathbb{C}$ are a sequence of measurable functions that converge in measure to a limit f. Then there exists a subsequence f_{n_j} that converges almost uniformly (and hence, pointwise almost everywhere) to f.
[Hint: Choose the n_j so that the sets $\{x \in X : |f_{n_j}(x) - f(x)| > 1/j\}$ have a suitably small measure.]

Solution. Construction of the subsequence. For each positive integer j, convergence in measure applied with $\varepsilon = 1/j$ gives

$$\mu\big(\{x \in X : |f_n(x) - f(x)| \geq 1/j\}\big) \longrightarrow 0 \qquad \text{as } n \to \infty.$$

Hence we may choose indices $n_1 < n_2 < n_3 < \cdots$ recursively: having chosen $n_1, \ldots, n_{j-1}$, pick $n_j > n_{j-1}$ large enough that

$$\mu(A_j) \leq 2^{-j}, \qquad \text{where } A_j := \{x \in X : |f_{n_j}(x) - f(x)| \geq 1/j\}.$$

Each A_j is measurable, since $|f_{n_j} - f|$ is a measurable function.
Almost uniform convergence. Let $\varepsilon > 0$. Choose J so large that $2^{-J+1} \leq \varepsilon$, and set

$$E := \bigcup_{j \geq J} A_j \in \mathcal{B}.$$

By countable subadditivity (Exercise 1.4.23(i)),

$$\mu(E) \leq \sum_{j \geq J} 2^{-j} = 2^{-J+1} \leq \varepsilon.$$

Now let $x \in X \setminus E$. For every $j \geq J$ we have $x \notin A_j$, hence

$$|f_{n_j}(x) - f(x)| < \frac{1}{j}.$$

Given $\eta > 0$, choose $J' \geq J$ with $1/J' < \eta$; then $|f_{n_j}(x) - f(x)| < \eta$ for all $j \geq J'$ and all $x \in X \setminus E$, the threshold J' being independent of x. Thus $f_{n_j} \to f$ uniformly on $X \setminus E$, and since $\mu(E) \leq \varepsilon$ with $\varepsilon > 0$ arbitrary, $f_{n_j} \to f$ almost uniformly.
Finally, almost uniform convergence implies pointwise almost everywhere convergence by Exercise 1.5.2(iv), so $f_{n_j} \to f$ pointwise almost everywhere as well.
Note that this strengthens Corollary 1.5.10, since L^1 convergence implies convergence in measure by Exercise 1.5.2(vi). For the typewriter sequence (Example 1.5.5) the construction picks out one indicator function from each sweep across $[0, 1]$, of rapidly shrinking length: the resulting subsequence has tail supports of measure tending to zero, so by Exercise 1.5.3(iii) it converges almost uniformly to zero, even though the full sequence does not.

Problem 1.5.7. Let $(X, \mathcal{B}, \mu)$ be a measure space, let $f_n : X \to \mathbb{C}$ be a sequence of measurable functions converging pointwise almost everywhere as $n \to \infty$ to a measurable limit $f : X \to \mathbb{C}$, and for each n, let $f_{n,m} : X \to \mathbb{C}$ be a sequence of measurable functions converging pointwise almost everywhere as $m \to \infty$ (keeping n fixed) to f_n.
(i) If $\mu(X)$ is finite, show that there exists a sequence $m_1, m_2, \ldots$ such that f_{n,m_n} converges pointwise almost everywhere to f.
(ii) Show the same claim is true if, instead of assuming that $\mu(X)$ is finite, we merely assume that X is σ-finite, i.e. it is the countable union of sets of finite measure.
[The claim can fail if X is not σ-finite. A counterexample is if $X = \mathbb{N}^{\mathbb{N}}$ with counting measure, f_n and f are identically zero for all $n \in \mathbb{N}$, and $f_{n,m}$ is the indicator function of the space of all sequences $(a_i)_{i \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ with $a_n \geq m$.]

Solution. We first isolate the two ingredients, then run a Borel–Cantelli argument.
A Borel–Cantelli lemma. If $A_1, A_2, \cdots \in \mathcal{B}$ satisfy $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then the set $A := \bigcap_{N=1}^{\infty} \bigcup_{n \geq N} A_n$ of points lying in infinitely many A_n is null. Indeed A is measurable, and for each N, monotonicity and countable subadditivity (Exercise 1.4.23(i)) give $\mu(A) \leq \mu(\bigcup_{n \geq N} A_n) \leq \sum_{n \geq N} \mu(A_n)$, which tends to 0 as $N \to \infty$ since the series converges. Hence $\mu(A) = 0$.
Part (i). Assume $\mu(X) < \infty$. Fix n. By hypothesis $f_{n,m} \to f_n$ pointwise almost everywhere as $m \to \infty$; since X has finite measure, Egorov's theorem (Theorem 1.5.9) upgrades this to almost uniform convergence, and almost uniform convergence implies convergence in measure by Exercise 1.5.2(vii). Applying the latter with $\varepsilon := 1/n$,

$$\mu\big(\{x \in X : |f_{n,m}(x) - f_n(x)| \geq 1/n\}\big) \longrightarrow 0 \qquad \text{as } m \to \infty.$$

We may therefore choose m_n so that the measurable set

$$A_n := \{x \in X : |f_{n,m_n}(x) - f_n(x)| \geq 1/n\}$$

satisfies $\mu(A_n) \leq 2^{-n}$. Do this for every n; this defines the desired sequence $m_1, m_2, \ldots$.
Since $\sum_n \mu(A_n) \leq \sum_n 2^{-n} = 1 < \infty$, the Borel–Cantelli lemma above shows that $A := \bigcap_N \bigcup_{n \geq N} A_n$ is null. Let Z be a null set outside of which $f_n \to f$ pointwise, and put $Z' := A \cup Z$, again a null set. Let $x \in X \setminus Z'$. Since $x \notin A$, there is N with $x \notin A_n$ for all $n \geq N$, i.e. $|f_{n,m_n}(x) - f_n(x)| < 1/n$ for all $n \geq N$; in particular $f_{n,m_n}(x) - f_n(x) \to 0$. Since $x \notin Z$ we also have $f_n(x) \to f(x)$. By the triangle inequality,

$$|f_{n,m_n}(x) - f(x)| \leq |f_{n,m_n}(x) - f_n(x)| + |f_n(x) - f(x)| \longrightarrow 0.$$

Hence $f_{n,m_n} \to f$ pointwise outside the null set Z', which is the claim.
Part (ii). Now assume $X = \bigcup_{k=1}^{\infty} Y_k$ with each $\mu(Y_k) < \infty$. Replacing Y_k by $X_k := Y_1 \cup \cdots \cup Y_k$, we may assume the sets are increasing, $X_1 \subseteq X_2 \subseteq \cdots$, still of finite measure (finite subadditivity), and still exhausting X.

Fix n. Since $X_n \in \mathcal{B}$, the triple $(X_n, \{S \in \mathcal{B} : S \subseteq X_n\}, \mu|)$ is again a measure space (Example 1.4.30), and it has finite measure. On it, the functions $f_{n,m}|_{X_n}$ converge pointwise almost everywhere to $f_n|_{X_n}$ as $m \to \infty$: by hypothesis there is a null set $F \subseteq X$ off which $f_{n,m} \to f_n$ pointwise, and $F \cap X_n$ is then a null set of the restricted space off which the restricted functions converge. As in part (i), Egorov's theorem (Theorem 1.5.9) and Exercise 1.5.2(vii) give convergence in measure on X_n, so we may choose m_n with

$$\mu(A_n) \leq 2^{-n}, \qquad \text{where } A_n := \{x \in X_n : |f_{n,m_n}(x) - f_n(x)| \geq 1/n\}.$$

Note carefully that the nth set is cut down to X_n: this is what makes its measure controllable even though $\mu(X)$ may be infinite.
Again $\sum_n \mu(A_n) < \infty$, so by Borel–Cantelli the set $A := \bigcap_N \bigcup_{n \geq N} A_n$ is null; let Z be a null set off which $f_n \to f$ pointwise, and set $Z' := A \cup Z$.
Let $x \in X \setminus Z'$. Since the X_k exhaust X, there is k with $x \in X_k$, and since the X_k increase we have $x \in X_n$ for all $n \geq k$. Since $x \notin A$, there is N with $x \notin A_n$ for all $n \geq N$. For $n \geq \max(N, k)$ we then have $x \in X_n$ and $x \notin A_n$, which by the definition of A_n forces $|f_{n,m_n}(x) - f_n(x)| < 1/n$. Hence $f_{n,m_n}(x) - f_n(x) \to 0$, and as $x \notin Z$ we have $f_n(x) \to f(x)$; the triangle inequality gives $f_{n,m_n}(x) \to f(x)$ exactly as before. So $f_{n,m_n} \to f$ pointwise outside the null set Z'.
Remark on the counterexample. In $X = \mathbb{N}^{\mathbb{N}}$ with counting measure (which is not σ-finite, as X is uncountable and every finite-measure set is finite), the functions $f_{n,m} = 1_{\{(a_i):a_n \geq m\}}$ do converge pointwise to $0 = f_n$ as $m \to \infty$ for each fixed n, since any fixed sequence (a_i) has $a_n < m$ for large m. But for any choice of $m_1, m_2, \ldots$ the sequence $a := (m_i)_{i \in \mathbb{N}}$ lies in every one of the sets $\{(a_i) : a_n \geq m_n\}$, so $f_{n,m_n}(a) = 1$ for all n and $f_{n,m_n}(a) \not\to 0 = f(a)$. Under counting measure the only null set is the empty set, so the set of points of non-convergence, containing a, is not sub-null. Thus f_{n,m_n} fails to converge to $f = 0$ pointwise almost everywhere, for every choice of $m_1, m_2, \ldots$. The finite-measure hypothesis entered above precisely through Egorov's theorem, which is what fails here.

Problem 1.5.8. Let $f_n : X \to \mathbb{C}$ be a sequence of measurable functions, and let $f : X \to \mathbb{C}$ be another measurable function. Show that the following are equivalent:

- (i) f_n converges in measure to f.
- (ii) Every subsequence f_{n_j} of the f_n has a further subsequence $f_{n_{j_i}}$ that converges almost uniformly to f.

Solution. Throughout, $(X, \mathcal{B}, \mu)$ is the ambient measure space.
Proof that (i) implies (ii). Suppose f_n converges in measure to f, and let f_{n_j} be an arbitrary subsequence. Fix $\varepsilon > 0$. The real numbers $a_n := \mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\})$ converge to 0 by hypothesis, and $(a_{n_j})_j$ is a subsequence of $(a_n)_n$, hence also converges to 0. As $\varepsilon > 0$ was arbitrary, the subsequence f_{n_j} itself converges in measure to f. By Exercise 1.5.6, a sequence converging in measure has a subsequence converging almost uniformly to the same limit; applying this to f_{n_j} produces a further subsequence $f_{n_{j_i}}$ converging almost uniformly to f, as required.
Proof that (ii) implies (i). We argue by contradiction. Suppose (ii) holds but f_n does not converge in measure to f. Then there is some $\varepsilon > 0$ for which the quantities

$$a_n := \mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\})$$

do not converge to 0. Failure of convergence to 0 for a sequence in $[0, +\infty]$ means precisely that there exists $\delta > 0$ and indices $n_1 < n_2 < n_3 < \cdots$ with

$$a_{n_j} \geq \delta \qquad \text{for all } j.$$

(Indeed, if no such δ and subsequence existed, then for every $\delta > 0$ the set of indices n with $a_n \geq \delta$ would be finite, so that $a_n < \delta$ for all sufficiently large n; as $\delta > 0$ is arbitrary this is precisely $a_n \to 0$.)
Apply hypothesis (ii) to the subsequence f_{n_j}: there is a further subsequence $f_{n_{j_i}}$ converging almost uniformly to f. By Exercise 1.5.2(vii), almost uniform convergence implies convergence in measure, so

$$a_{n_{j_i}} = \mu(\{x \in X : |f_{n_{j_i}}(x) - f(x)| \geq \varepsilon\}) \to 0 \quad \text{as } i \to \infty.$$

But every term of this sequence is at least $\delta > 0$, a contradiction. Hence f_n converges in measure to f, completing the proof.
(Note that the implication (ii) implies (i) does not even need the full strength of almost uniform convergence: convergence in measure of the sub-subsequences suffices, by the same argument.)

Problem 1.5.9. Suppose that $f_n : X \to \mathbb{C}$ are a dominated sequence of measurable functions, and let $f : X \to \mathbb{C}$ be another measurable function. Show that f_n converges in L^1 norm to f if and only if f_n converges in measure to f. (Hint: one way to establish the "if" direction is first show that every subsequence of the f_n has a further subsequence that converges in L^1 to f, using Exercise 1.5.6 and the dominated convergence theorem (Theorem 1.4.49). Alternatively, use monotone convergence to find a set E of finite measure such that $\int_{X\setminus E} g\,d\mu$, and hence $\int_{X\setminus E} f_n\,d\mu$ and $\int_{X\setminus E} f\,d\mu$, are small.)

Solution. By hypothesis there is an absolutely integrable $g : X \to \mathbb{C}$ with $|f_n(x)| \le g(x)$ for every n and almost every x. Since $|f_n(x)| \ge 0$, the bound forces $g(x)$ to be a nonnegative real number for almost every x; discarding a null set (which changes no integral and no mode of convergence) we may and do assume $g : X \to [0, +\infty)$ with $\int_X g\,d\mu < \infty$ and $|f_n| \le g$ everywhere.
The "only if" direction is immediate and needs no domination: if $f_n \to f$ in L^1 norm then $f_n \to f$ in measure, by Exercise 1.5.2(vi).
For the "if" direction, suppose f_n converges in measure to f. We first record that f is also dominated by g. By Exercise 1.5.6 there is a subsequence f_{n_k} converging pointwise almost everywhere to f; since $|f_{n_k}(x)| \le g(x)$ for all k and almost every x, letting $k \to \infty$ gives $|f(x)| \le g(x)$ for almost every x. In particular f is absolutely integrable, all the differences $f_n - f$ are absolutely integrable, and

$$|f_n - f| \le |f_n| + |f| \le 2g \qquad \text{almost everywhere, for every } n.$$

Set $a_n := \|f_n - f\|_{L^1(\mu)} = \int_X |f_n - f|\,d\mu \in [0, 2\|g\|_{L^1(\mu)}]$. We must show $a_n \to 0$, and we do so by the sub-subsequence principle: it suffices to show that every subsequence of (a_n) has a further subsequence tending to 0. Indeed, if $a_n \not\to 0$ then there are $\varepsilon > 0$ and indices $n_1 < n_2 < \cdots$ with $a_{n_j} \ge \varepsilon$ for all j, and no further subsequence of (a_{n_j}) can tend to 0.
So let f_{n_j} be an arbitrary subsequence. As observed in the solution to Exercise 1.5.8, a subsequence of a sequence converging in measure to f again converges in measure to f. Hence by Exercise 1.5.6 there is a further subsequence $f_{n_{j_i}}$ converging pointwise almost everywhere to f. Consider the functions

$$h_i := |f_{n_{j_i}} - f|.$$

They are measurable, nonnegative, converge pointwise almost everywhere to 0, and satisfy $h_i \le 2g$ almost everywhere with $2g$ absolutely integrable. By the dominated convergence theorem (Theorem 1.4.49),

$$a_{n_{j_i}} = \int_X h_i\,d\mu \to \int_X 0\,d\mu = 0 \qquad \text{as } i \to \infty.$$

Thus every subsequence of (a_n) has a further subsequence tending to 0, and therefore $a_n \to 0$; that is, f_n converges to f in L^1 norm.

Problem 1.5.10.
- (i) Show that if f is an absolutely integrable function, then the constant sequence $f_n = f$ is uniformly integrable. (Hint: use the monotone convergence theorem.)
- (ii) Show that every dominated sequence of measurable functions is uniformly integrable.
- (iii) Give an example of a sequence that is uniformly integrable but not dominated.

Solution. Recall Definition 1.5.11: a sequence $f_n : X \to \mathbb{C}$ of absolutely integrable functions is uniformly integrable if (a) $\sup_n \int_X |f_n|\,d\mu < +\infty$; (b) $\sup_n \int_{|f_n|\ge M} |f_n|\,d\mu \to 0$ as $M \to +\infty$; and (c) $\sup_n \int_{|f_n|\le\delta} |f_n|\,d\mu \to 0$ as $\delta \to 0$.
Part (i). Let f be absolutely integrable, so $\int_X |f|\,d\mu < \infty$, and put $f_n := f$ for all n. Condition (a) is

immediate: the supremum equals $\|f\|_{L^1(\mu)} < \infty$.
For (b), work along integers $M = m = 1, 2, 3, \ldots$ and set $u_m := |f| 1_{\{|f|<m\}}$. These are nonnegative measurable functions, monotone increasing in m (the sets $\{|f| < m\}$ increase), and they converge pointwise to $|f|$: indeed f takes values in $\mathbb{C}$, so $|f(x)| < \infty$ for every x, and hence $|f(x)| < m$ for all sufficiently large m. By the monotone convergence theorem (Theorem 1.4.44),

$$\int_X |f| 1_{\{|f|<m\}} \, d\mu \to \int_X |f| \, d\mu \qquad (m \to \infty).$$

Since $\int_X |f| \, d\mu$ is finite, we may subtract, obtaining

$$\int_{|f|\geq m} |f| \, d\mu = \int_X |f| \, d\mu - \int_X |f| 1_{\{|f|<m\}} \, d\mu \to 0.$$

Finally, $M \mapsto \int_{|f|\geq M} |f| \, d\mu$ is nonincreasing, so convergence to 0 along the integers upgrades to convergence as $M \to +\infty$ through the reals. This is (b).
For (c), argue symmetrically with $v_m := |f| 1_{\{|f|>1/m\}}$. These are nonnegative, monotone increasing in m, and converge pointwise to $|f|$: at a point where $f(x) = 0$ every term equals $0 = |f(x)|$, and at a point where $f(x) \neq 0$ the terms equal $|f(x)|$ once $1/m < |f(x)|$. Monotone convergence (Theorem 1.4.44) and subtraction of the finite quantity $\int_X |f| \, d\mu$ give

$$\int_{|f|\leq 1/m} |f| \, d\mu \to 0 \qquad (m \to \infty),$$

and since $\delta \mapsto \int_{|f|\leq\delta} |f| \, d\mu$ is nondecreasing, this gives $\int_{|f|\leq\delta} |f| \, d\mu \to 0$ as $\delta \to 0$. This is (c). Hence the constant sequence is uniformly integrable.
Part (ii). Suppose $|f_n| \leq g$ for all n and almost every x, with g absolutely integrable. As in Exercise 1.5.9 we may modify g on a null set and assume $g : X \to [0, +\infty)$ with $\int_X g \, d\mu < \infty$ and $|f_n| \leq g$ everywhere. Each f_n is then absolutely integrable, since $\int_X |f_n| \, d\mu \leq \int_X g \, d\mu < \infty$; this also gives (a), with $\sup_n \|f_n\|_{L^1(\mu)} \leq \|g\|_{L^1(\mu)} < \infty$.
For (b): for $M > 0$ the inclusion $\{|f_n| \geq M\} \subseteq \{g \geq M\}$ holds, because $|f_n| \leq g$ pointwise. Hence, using $|f_n| \leq g$ again,

$$\int_{|f_n|\geq M} |f_n| \, d\mu \leq \int_{|f_n|\geq M} g \, d\mu \leq \int_{g\geq M} g \, d\mu,$$

and the right-hand side is independent of n and tends to 0 as $M \to +\infty$ by part (i) applied to the absolutely integrable function g. Taking suprema over n gives (b).
For (c): on the set $\{|f_n| \leq \delta\}$ we have both $|f_n| \leq \delta$ and $|f_n| \leq g$, so $|f_n| \leq \min(\delta, g)$ there, whence

$$\int_{|f_n|\leq\delta} |f_n| \, d\mu \leq \int_X \min(\delta, g) \, d\mu,$$

again a bound independent of n. Along $\delta = 1/m$, the functions $\min(1/m, g)$ are measurable, converge pointwise to 0 (as g is finite valued), and are dominated by the absolutely integrable g; so by the dominated convergence theorem (Theorem 1.4.49), $\int_X \min(1/m, g) \, d\mu \to 0$. Since $\delta \mapsto \int_X \min(\delta, g) \, d\mu$ is nondecreasing, we get $\int_X \min(\delta, g) \, d\mu \to 0$ as $\delta \to 0$, which yields (c). Hence a dominated sequence is uniformly integrable.
Part (iii). Take $X = \mathbb{R}$ with Lebesgue measure and let $f_n := 1_{[n,n+1]}$ (the escape to horizontal infinity of Example 1.5.2). Each f_n is absolutely integrable with $\int_{\mathbb{R}} |f_n| \, dm = 1$, so (a) holds with supremum 1. For (b): if $M > 1$ then $\{|f_n| \geq M\} = \emptyset$ for every n, since f_n only takes the values 0 and 1; hence $\sup_n \int_{|f_n|\geq M} |f_n| \, dm = 0$ for all $M > 1$, which certainly tends to 0. For (c): if $0 < \delta < 1$ then $\{|f_n| \leq \delta\} = \{f_n = 0\}$, on which the integrand $|f_n|$ vanishes, so $\sup_n \int_{|f_n|\leq\delta} |f_n| \, dm = 0$ for all $\delta < 1$. Thus (f_n) is uniformly integrable.
It is not dominated. Suppose g were absolutely integrable with $|f_n| \leq g$ almost everywhere for every n. For each n this gives $g \geq 1$ almost everywhere on $[n, n+1]$, i.e. the set $N_n := \{x \in [n, n+1] : g(x) < 1\}$

is null. Then $N := \bigcup_{n\geq 1} N_n$ is null, being a countable union of null sets, and $g \geq 1$ on $[1, +\infty) \setminus N$. Consequently

$$\int_{\mathbb{R}} g\, dm \geq \int_{[1,+\infty)} 1\, dm = +\infty,$$

contradicting the absolute integrability of g. (Here we used that $g \geq |f_1| \geq 0$ almost everywhere, so that $|g| = g$ almost everywhere and $\int_{\mathbb{R}} |g|\, dm = \int_{\mathbb{R}} g\, dm = +\infty$.) So the sequence is uniformly integrable but not dominated.

Problem 1.5.11. Suppose that X has finite measure, and let $f_n : X \to \mathbb{C}$ be a sequence of measurable functions. Show that f_n is uniformly integrable if and only if $\sup_n \int_{|f_n|\geq M} |f_n|\, d\mu \to 0$ as $M \to +\infty$.

Solution. Write $\mu(X) < \infty$ for the finite total measure. We must compare the single condition

$$(*) \qquad \sup_n \int_{|f_n|\geq M} |f_n|\, d\mu \to 0 \quad \text{as } M \to +\infty$$

with the three conditions of Definition 1.5.11: (a) $\sup_n \int_X |f_n|\, d\mu < +\infty$; (b) the condition $(*)$ itself; (c) $\sup_n \int_{|f_n|\leq\delta} |f_n|\, d\mu \to 0$ as $\delta \to 0$ (together with the standing requirement that each f_n be absolutely integrable).
The "only if" direction is trivial: condition $(*)$ is literally clause (b) of Definition 1.5.11, so a uniformly integrable sequence satisfies it. (This direction does not use $\mu(X) < \infty$.)
For the "if" direction, assume $(*)$. We verify absolute integrability and clauses (a) and (c).
Absolute integrability and clause (a). By $(*)$ there is a finite $M_0 > 0$ with

$$\sup_n \int_{|f_n|\geq M_0} |f_n|\, d\mu \leq 1.$$

For each n split X into $\{|f_n| \geq M_0\}$ and $\{|f_n| < M_0\}$. On the second set the integrand is bounded by M_0, and that set has measure at most $\mu(X) < \infty$, so

$$\int_X |f_n|\, d\mu = \int_{|f_n|\geq M_0} |f_n|\, d\mu + \int_{|f_n|<M_0} |f_n|\, d\mu \leq 1 + M_0\mu(X) < \infty.$$

This bound is uniform in n, so each f_n is absolutely integrable and $\sup_n \|f_n\|_{L^1(\mu)} \leq 1+M_0\mu(X) < +\infty$, which is (a).
Clause (c). This is where the finiteness of $\mu(X)$ does the real work, and it holds with no hypothesis at all beyond $\mu(X) < \infty$: for any $\delta > 0$ and any n, the integrand $|f_n|$ is bounded by δ on the set $\{|f_n| \leq \delta\}$, so

$$\int_{|f_n|\leq\delta} |f_n|\, d\mu \leq \delta\, \mu(\{|f_n| \leq \delta\}) \leq \delta\, \mu(X).$$

Taking the supremum over n gives $\sup_n \int_{|f_n|\leq\delta} |f_n|\, d\mu \leq \delta\mu(X) \to 0$ as $\delta \to 0$. This is (c). (Informally: on a finite measure space there is no room for escape to width infinity.)
All three clauses of Definition 1.5.11 hold, so (f_n) is uniformly integrable. This completes the equivalence.

Problem 1.5.12. (Uniform L^p bound on finite measure implies uniform integrability) Suppose that X have finite measure, let $1 < p < \infty$, and suppose that $f_n : X \to \mathbb{C}$ is a sequence of measurable functions such that $\sup_n \int_X |f_n|^p\, d\mu < \infty$. Show that the sequence f_n is uniformly integrable.

Solution. Set $A := \sup_n \int_X |f_n|^p\, d\mu < \infty$. Each $|f_n|^p$ is a nonnegative measurable function (the map $z \mapsto |z|^p$ is continuous on $\mathbb{C}$, so $|f_n|^p$ is measurable as a composition of a measurable function with a continuous one), so these integrals make sense.

Since $\mu(X) < \infty$, Exercise 1.5.11 reduces the problem to showing

$$\sup_n \int_{|f_n| \geq M} |f_n| \, d\mu \to 0 \qquad \text{as } M \to +\infty.$$

Fix $M > 0$ and fix n. At every point x of the set $\{|f_n| \geq M\}$ we may write

$$|f_n(x)| = |f_n(x)|^p \, |f_n(x)|^{1-p} \leq |f_n(x)|^p M^{1-p},$$

where the inequality uses $1 - p < 0$ (because $p > 1$) together with $|f_n(x)| \geq M > 0$, so that $t \mapsto t^{1-p}$ being decreasing on $(0, \infty)$ gives $|f_n(x)|^{1-p} \leq M^{1-p}$. (At a point where $|f_n(x)| = 0$ the set condition $|f_n(x)| \geq M > 0$ fails, so no division by zero occurs.)
Integrating this pointwise bound over $\{|f_n| \geq M\}$ and then enlarging the domain of integration to all of X (the integrand $|f_n|^p$ being nonnegative) yields

$$\int_{|f_n| \geq M} |f_n| \, d\mu \leq M^{1-p} \int_{|f_n| \geq M} |f_n|^p \, d\mu \leq M^{1-p} \int_X |f_n|^p \, d\mu \leq A M^{1-p}.$$

The bound AM^{1-p} does not depend on n, so

$$\sup_n \int_{|f_n| \geq M} |f_n| \, d\mu \leq \frac{A}{M^{p-1}} \to 0 \qquad \text{as } M \to +\infty,$$

because $p - 1 > 0$. By Exercise 1.5.11 the sequence (f_n) is uniformly integrable, as claimed.
(Two remarks. First, the argument genuinely needs $p > 1$: for $p = 1$ a uniform L^1 bound is not enough, as the sequence $f_n = n1_{[0,1/n]}$ on $[0, 1]$ shows, which has $\int |f_n| \, dm = 1$ for all n but $\int_{|f_n| \geq M} |f_n| \, dm = 1$ for every $n \geq M$. Second, the hypothesis $\mu(X) < \infty$ is used only through Exercise 1.5.11, i.e. to rule out escape to width infinity; the displayed estimate itself is valid on any measure space.)

Problem 1.5.13. Let $f_n : X \to \mathbb{C}$ be a uniformly integrable sequence of functions. Show that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\int_E |f_n| \, d\mu \leq \varepsilon$$

whenever $n \geq 1$ and E is a measurable set with $\mu(E) \leq \delta$.

Solution. Let $\varepsilon > 0$ be given. By clause (ii) of Definition 1.5.11 (no escape to vertical infinity),

$$\sup_n \int_{|f_n| \geq M} |f_n| \, d\mu \to 0 \qquad \text{as } M \to +\infty,$$

so we may fix a real number $M > 0$ with

$$\int_{|f_n| \geq M} |f_n| \, d\mu \leq \frac{\varepsilon}{2} \qquad \text{for every } n \geq 1.$$

Define

$$\delta := \frac{\varepsilon}{2M} > 0.$$

Now let $n \geq 1$ and let $E \in \mathcal{B}$ satisfy $\mu(E) \leq \delta$. Split E into the two measurable pieces $E \cap \{|f_n| \geq M\}$ and $E \cap \{|f_n| < M\}$; since $|f_n|$ is nonnegative and measurable, and each f_n is absolutely integrable (part of Definition 1.5.11), additivity of the integral over this partition gives

$$\int_E |f_n| \, d\mu = \int_{E \cap \{|f_n| \geq M\}} |f_n| \, d\mu + \int_{E \cap \{|f_n| < M\}} |f_n| \, d\mu.$$

The first term is bounded by enlarging the domain of integration from $E \cap \{|f_n| \geq M\}$ to all of $\{|f_n| \geq M\}$ (legitimate since the integrand is nonnegative), giving

$$\int_{E\cap\{|f_n|\geq M\}} |f_n|\,d\mu \leq \int_{|f_n|\geq M} |f_n|\,d\mu \leq \frac{\varepsilon}{2}$$

by the choice of M, uniformly in n.
On the second piece the integrand satisfies $|f_n| < M$ pointwise, so by monotonicity of the integral

$$\int_{E\cap\{|f_n|<M\}} |f_n|\,d\mu \leq M\,\mu(E\cap\{|f_n|<M\}) \leq M\mu(E) \leq M\delta = \frac{\varepsilon}{2}.$$

Adding the two bounds gives $\int_E |f_n|\,d\mu \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$, for every $n \geq 1$ and every measurable E with $\mu(E) \leq \delta$. Since $\varepsilon > 0$ was arbitrary, this is exactly the assertion. (Note that only clause (ii) of uniform integrability was used; clauses (i) and (iii) were not needed, and no finiteness assumption on $\mu(X)$ was required.)

Problem 1.5.14. This exercise is a partial converse to Exercise 1.5.13. Let X be a probability space, and let $f_n : X \to \mathbb{C}$ be a sequence of absolutely integrable functions with $\sup_n \|f_n\|_{L^1} < \infty$. Suppose that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\int_E |f_n|\,d\mu \leq \varepsilon$$

whenever $n \geq 1$ and E is a measurable set with $\mu(E) \leq \delta$. Show that the sequence f_n is uniformly integrable.

Solution. Since X is a probability space we have $\mu(X) = 1$; in particular $\mu(X) < \infty$. Put

$$A := \sup_n \|f_n\|_{L^1(\mu)} = \sup_n \int_X |f_n|\,d\mu < \infty.$$

We verify the three clauses of Definition 1.5.11.
Clause (i) (uniform bound on the L^1 norm). This is precisely the hypothesis $\sup_n \|f_n\|_{L^1(\mu)} = A < \infty$, and each f_n is absolutely integrable by assumption.
Clause (iii) (no escape to width infinity). As in Exercise 1.5.11, this is automatic on a finite measure space: for any $\delta > 0$ and any n, the integrand is at most δ on $\{|f_n| \leq \delta\}$, so

$$\int_{|f_n|\leq\delta} |f_n|\,d\mu \leq \delta\,\mu(X) = \delta,$$

and hence $\sup_n \int_{|f_n|\leq\delta} |f_n|\,d\mu \leq \delta \to 0$ as $\delta \to 0$.
Clause (ii) (no escape to vertical infinity). This is where the hypothesis is used. Let $\varepsilon > 0$, and let $\delta > 0$ be as provided by the hypothesis for this ε. By Markov's inequality (Exercise 1.4.36(vi)) applied to the nonnegative absolutely integrable function $|f_n|$, for every $M > 0$ and every n,

$$\mu(\{x \in X : |f_n(x)| \geq M\}) \leq \frac{1}{M}\int_X |f_n|\,d\mu \leq \frac{A}{M}.$$

Set $M_0 := (A+1)/\delta$. Then for every $M \geq M_0$ and every n,

$$\mu(\{|f_n| \geq M\}) \leq \frac{A}{M} \leq \frac{A\delta}{A+1} < \delta.$$

Thus for each such M and each n the measurable set $E := \{x \in X : |f_n(x)| \geq M\}$ satisfies $\mu(E) \leq \delta$, and the hypothesis (applied with this n and this E – the hypothesis is allowed to be used with a set depending on n, since it is asserted for all $n \geq 1$ and all measurable E of measure at most δ) gives

$$\int_{|f_n|\geq M} |f_n|\,d\mu \leq \varepsilon.$$

Taking the supremum over n, we conclude

$$\sup_n \int_{|f_n|\geq M} |f_n|\, d\mu \leq \varepsilon \qquad \text{for all } M \geq M_0.$$

As $\varepsilon > 0$ was arbitrary and for each ε we produced a threshold $M_0 = M_0(\varepsilon)$ beyond which the supremum is at most ε, this says exactly that

$$\sup_n \int_{|f_n|\geq M} |f_n|\, d\mu \to 0 \qquad \text{as } M \to +\infty,$$

which is clause (ii).
All three clauses hold, so (f_n) is uniformly integrable. (Equivalently, one could invoke Exercise 1.5.11 directly: on a finite measure space clause (ii) alone is enough, and clause (ii) is what the Markov inequality argument above supplies. The hypothesis $\sup_n \|f_n\|_{L^1} < \infty$ cannot be dropped: take $X = \{p\}$ a single point with $\mu(\{p\}) = 1$ and $f_n := n$. The only measurable set of measure at most δ, for any $\delta < 1$, is the empty set, so the ε-δ hypothesis holds vacuously; yet $\|f_n\|_{L^1(\mu)} = n$ is unbounded, and clause (i) of Definition 1.5.11 fails. It is exactly the uniform L^1 bound that lets Markov's inequality control $\mu(\{|f_n| \geq M\})$ uniformly in n.)

Problem 1.5.15. Give an example of a sequence f_n of uniformly integrable functions that converge pointwise almost everywhere to zero, but do not converge almost uniformly, in measure, or in L^1 norm.

Solution. Take $X = \mathbb{R}$ with Lebesgue measure m, and let

$$f_n := 1_{[n,n+1]}, \qquad n = 1, 2, 3, \ldots$$

This is the escape to horizontal infinity example (Example 1.5.2). The point of the exercise is that this sequence, unlike the escape to vertical infinity and escape to width infinity examples, is nevertheless uniformly integrable, so that Definition 1.5.11 is not by itself a substitute for domination in the dominated convergence theorem (Theorem 1.4.49).
Uniform integrability. We check the three conditions of Definition 1.5.11.
(i) Each f_n is the indicator of an interval of length 1, so $\|f_n\|_{L^1(m)} = \int_{\mathbb{R}} 1_{[n,n+1]}\, dm = m([n, n+1]) = 1$, and hence $\sup_n \|f_n\|_{L^1(m)} = 1 < +\infty$.
(ii) Since f_n only takes the values 0 and 1, for any $M > 1$ the set $\{x : |f_n(x)| \geq M\}$ is empty, and so $\int_{|f_n|\geq M} |f_n|\, dm = 0$ for every n. Thus $\sup_n \int_{|f_n|\geq M} |f_n|\, dm = 0$ for all $M > 1$, which certainly tends to 0 as $M \to +\infty$.
(iii) Similarly, for any $0 < \delta < 1$ the set $\{x : |f_n(x)| \leq \delta\}$ is exactly the set $\{x : f_n(x) = 0\}$, on which the integrand $|f_n|$ vanishes identically; hence $\int_{|f_n|\leq\delta} |f_n|\, dm = 0$ for every n. Thus $\sup_n \int_{|f_n|\leq\delta} |f_n|\, dm = 0$ for all $0 < \delta < 1$, which tends to 0 as $\delta \to 0$.
So the sequence (f_n) is uniformly integrable. (Equivalently, in the language of Remark 1.5.12 with $f_n = A_n 1_{E_n}$, $A_n = 1$ and $\mu(E_n) = 1$: the product $A_n\mu(E_n)$ stays bounded, and there is no subsequence along which $A_n \to \infty$ or $A_n \to 0$, so conditions (ii) and (iii) are vacuous.)
Pointwise convergence to zero. Fix $x \in \mathbb{R}$. For every $n > |x|$ we have $x \notin [n, n+1]$, so $f_n(x) = 0$. Hence $f_n(x) \to 0$ for every $x \in \mathbb{R}$; in particular $f_n \to 0$ pointwise, and therefore also pointwise almost everywhere.
Failure of L^1 convergence. As computed above, $\|f_n - 0\|_{L^1(m)} = 1$ for every n, which does not tend to 0.
Failure of convergence in measure. For $\varepsilon = 1/2$ we have $\{x : |f_n(x) - 0| \geq 1/2\} = [n, n+1]$, whose measure is 1 for every n. So $m(\{|f_n| \geq 1/2\}) \not\to 0$, and f_n does not converge to 0 in measure.
Failure of almost uniform convergence. By Exercise 1.5.2(vii), almost uniform convergence implies convergence in measure, so this already follows from the previous paragraph. For a direct argument: let E be any measurable set with $m(E) \leq 1/2$. For each n, the set $[n, n+1]\backslash E$ has measure at least $1 - 1/2 = 1/2 > 0$, hence is nonempty; picking x_n in it gives $f_n(x_n) = 1$. Therefore $\sup_{x\in\mathbb{R}\backslash E} |f_n(x) - 0| = 1$ for

every n, so f_n does not converge uniformly to 0 off E. Since no exceptional set of measure at most $1/2$ works, f_n does not converge almost uniformly to 0.
Finally, by Proposition 1.5.7 the limit is unique up to null sets in each of the seven modes, so there is no other candidate limit along which any of these three modes could hold: since $f_n \to 0$ pointwise almost everywhere, convergence of f_n almost uniformly, in measure, or in L^1 to any limit g would force $g = 0$ almost everywhere, contradicting the above.
(Note, consistently with the dominated convergence theorem, that this sequence is not dominated: any g with $g \geq |f_n|$ for all n almost everywhere satisfies $g \geq 1$ almost everywhere on $[1, +\infty)$, hence $\int_{\mathbb{R}} g\, dm = +\infty$ and g is not absolutely integrable.)

Problem 1.5.16. (Monotone convergence theorem) Suppose that $f_n : X \to [0, +\infty)$ are measurable, monotone non-decreasing in n, and are such that $\sup_n \int_X f_n\, d\mu < \infty$. Show that f_n converges in L^1 norm to $\sup_n f_n$. (Note that $\sup_n f_n$ can be infinite on a null set, but the definition of L^1 convergence can be easily modified to accomodate this.)

Solution. Write $f := \sup_n f_n$, a function from X to $[0, +\infty]$. Since the f_n are monotone non-decreasing in n, we in fact have $f(x) = \lim_{n\to\infty} f_n(x)$ for every $x \in X$, the limit being taken in $[0, +\infty]$; in particular f is measurable, being a pointwise limit of unsigned measurable functions (Exercise 1.4.29(vi)).
Set

$$A := \sup_n \int_X f_n\, d\mu < \infty.$$

By the monotone convergence theorem (Theorem 1.4.44),

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_X f_n\, d\mu = A < \infty,$$

the middle limit existing and equalling the supremum because $n \mapsto \int_X f_n\, d\mu$ is non-decreasing by monotonicity of the unsigned integral (Exercise 1.4.36(ii)).
In particular f is unsigned absolutely integrable. By Markov's inequality (Exercise 1.4.36(vi)), for every $\lambda \in (0, \infty)$ we have

$$\mu(\{x \in X : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_X f\, d\mu = \frac{A}{\lambda},$$

and since $\{x : f(x) = +\infty\} \subseteq \{x : f(x) \geq \lambda\}$ for every λ, letting $\lambda \to \infty$ gives $\mu(\{f = +\infty\}) = 0$. Thus f is finite almost everywhere, and the limit function $\sup_n f_n$ is (after modification on a null set, which affects none of the five measure-theoretic modes of convergence) an honest almost everywhere finite function; this is the modification of the definition of L^1 convergence alluded to in the statement. Concretely, we simply interpret $\|f - f_n\|_{L^1(\mu)}$ as $\int_X |f - f_n|\, d\mu$, where the integrand is an unsigned measurable function on X with values in $[0, +\infty]$.
Now fix n. Since $0 \leq f_n(x) < \infty$ and $f_n(x) \leq f(x)$ for every x, the difference $f - f_n : X \to [0, +\infty]$ is well defined and unsigned, with the convention $+\infty - c = +\infty$ for finite c (which is what occurs on the null set $\{f = +\infty\}$). It is measurable: for $m \geq n$ the function $f_m - f_n$ is a difference of finite-valued measurable functions, hence measurable, and it is unsigned and monotone non-decreasing in m with pointwise limit $f - f_n$, so Exercise 1.4.29(vi) applies. One has the pointwise identity

$$f_n + (f - f_n) = f \quad \text{everywhere on } X.$$

Applying additivity of the unsigned integral (Theorem 1.4.38) to this identity gives

$$\int_X f_n\, d\mu + \int_X (f - f_n)\, d\mu = \int_X f\, d\mu = A.$$

Since $\int_X f_n\, d\mu \leq A < \infty$ is finite, we may subtract it from both sides to obtain

$$\int_X (f - f_n)\, d\mu = A - \int_X f_n\, d\mu.$$

Finally, because $f \geq f_n$ pointwise, $|f - f_n| = f - f_n$, so

$$\|f - f_n\|_{L^1(\mu)} = \int_X |f - f_n| \, d\mu = A - \int_X f_n \, d\mu.$$

By the monotone convergence theorem computation above, $\int_X f_n \, d\mu \to A$ as $n \to \infty$, and therefore $\|f - f_n\|_{L^1(\mu)} \to 0$. That is, f_n converges in L^1 norm to $f = \sup_n f_n$, as required.
(The hypothesis $\sup_n \int_X f_n \, d\mu < \infty$ is essential: for $f_n := 1_{[0,n]}$ on $\mathbb{R}$ one has $f_n \uparrow 1_{[0,\infty)}$, but $\|f - f_n\|_{L^1} = +\infty$ for every n. The subtraction step above is exactly where finiteness is used.)

Problem 1.5.17. (Defect version of Fatou's lemma) Suppose that $f_n : X \to [0, +\infty)$ are measurable, are such that $\sup_n \int_X f_n \, d\mu < \infty$, and converge pointwise almost everywhere to some measurable limit $f : X \to [0, +\infty)$. Show that f_n converges in L^1 norm to f if and only if $\int_X f_n \, d\mu$ converges to $\int_X f \, d\mu$. Informally, we see that in the unsigned, bounded mass case, pointwise convergence implies L^1 norm convergence if and only if there is no loss of mass.

Solution. Write $A := \sup_n \int_X f_n \, d\mu < \infty$.
Preliminary: f is absolutely integrable. By Fatou's lemma (Corollary 1.4.47) applied to the sequence f_n, which converges to f pointwise almost everywhere,

$$\int_X f \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu \leq A < \infty.$$

(Strictly, Fatou's lemma is stated for the pointwise liminf; since $f_n \to f$ pointwise almost everywhere, $\liminf_n f_n = f$ almost everywhere, and by Exercise 1.4.36(i) the two have the same integral.) So f is unsigned absolutely integrable, and in particular all the integrals appearing below are finite.
The "only if" direction. Suppose $\|f_n - f\|_{L^1(\mu)} \to 0$. Since f_n and f are unsigned with finite integral, we have the pointwise bounds $f_n \leq f + |f_n - f|$ and $f \leq f_n + |f_n - f|$, whence by monotonicity (Exercise 1.4.36(ii)) and additivity (Theorem 1.4.38) of the unsigned integral,

$$\int_X f_n \, d\mu \leq \int_X f \, d\mu + \|f_n - f\|_{L^1(\mu)}, \qquad \int_X f \, d\mu \leq \int_X f_n \, d\mu + \|f_n - f\|_{L^1(\mu)}.$$

All quantities being finite, these rearrange to

$$\left| \int_X f_n \, d\mu - \int_X f \, d\mu \right| \leq \|f_n - f\|_{L^1(\mu)} \to 0,$$

so $\int_X f_n \, d\mu \to \int_X f \, d\mu$. (This is exactly the remark made after Remark 1.5.6, that L^1 convergence implies convergence of the integrals.)
The "if" direction. Suppose now that $\int_X f_n \, d\mu \to \int_X f \, d\mu$. Following the hint attached to Exercise 1.4.48, we consider the truncations

$$g_n := \min(f_n, f).$$

Each g_n is unsigned, finite-valued, and measurable, since its level sets $\{g_n > \lambda\} = \{f_n > \lambda\} \cap \{f > \lambda\}$ are measurable for every λ (Exercise 1.4.29(i)). We record the elementary pointwise identity, valid for any two real numbers $a, b \geq 0$:

$$|a - b| = a + b - 2\min(a, b),$$

which follows from $\max(a, b) + \min(a, b) = a + b$ and $|a - b| = \max(a, b) - \min(a, b)$. Applying it with $a = f_n(x)$, $b = f(x)$ gives, everywhere on X,

$$f_n + f = |f_n - f| + 2g_n.$$

All four functions here are unsigned and measurable, so by additivity of the unsigned integral (Theorem 1.4.38) together with homogeneity (Exercise 1.4.36(iii)),

$$\int_X f_n \, d\mu + \int_X f \, d\mu = \|f_n - f\|_{L^1(\mu)} + 2 \int_X g_n \, d\mu.$$

Every term on the left is finite, hence so is each term on the right, and we may rearrange:

$$\|f_n - f\|_{L^1(\mu)} = \int_X f_n \, d\mu + \int_X f \, d\mu - 2 \int_X g_n \, d\mu.$$

It therefore suffices to show that $\int_X g_n \, d\mu \to \int_X f \, d\mu$; for then the right-hand side converges to $\int_X f \, d\mu + \int_X f \, d\mu - 2 \int_X f \, d\mu = 0$, using the hypothesis $\int_X f_n \, d\mu \to \int_X f \, d\mu$.
To see this, apply the dominated convergence theorem (Theorem 1.4.49) to the sequence g_n. Indeed:

- $|g_n| = g_n \leq f$ pointwise for every n, and f is absolutely integrable by the preliminary step; so the sequence (g_n) is dominated by f.
- On the set of full measure where $f_n(x) \to f(x)$, continuity of min gives $g_n(x) = \min(f_n(x), f(x)) \to \min(f(x), f(x)) = f(x)$; so $g_n \to f$ pointwise almost everywhere.

Hence $\int_X g_n \, d\mu \to \int_X f \, d\mu$, and by the displayed identity $\|f_n - f\|_{L^1(\mu)} \to 0$, i.e. f_n converges to f in L^1 norm.
Remark. Rearranging the displayed identity gives

$$\int_X f_n \, d\mu - \int_X f \, d\mu - \|f_n - f\|_{L^1(\mu)} = 2 \left(\int_X g_n \, d\mu - \int_X f \, d\mu \right) \to 0,$$

which is precisely the conclusion of Exercise 1.4.48; the present exercise is the statement that, under the standing hypothesis $\sup_n \int_X f_n \, d\mu < \infty$, the defect $\|f_n - f\|_{L^1(\mu)}$ in Fatou's lemma vanishes exactly when no mass is lost in the limit. Loss of mass is illustrated by the escape to vertical infinity example $f_n := n 1_{[1/n, 2/n]}$ on $\mathbb{R}$ (Example 1.5.4): here $f_n \to 0 =: f$ pointwise, $\sup_n \int f_n \, dm = 1 < \infty$, but $\int f_n \, dm = 1 \not\to 0 = \int f \, dm$, and correspondingly $\|f_n - f\|_{L^1} = 1 \not\to 0$.

Problem 1.5.18. Suppose that $f_n : X \to \mathbb{C}$ are a dominated sequence of measurable functions, and let $f : X \to \mathbb{C}$ be another measurable function. Show that f_n converges pointwise almost everywhere to f if and only if f_n converges almost uniformly to f.

Solution. Recall the definition preceding Exercise 1.5.9: the sequence (f_n) is dominated means that there is an absolutely integrable $g : X \to \mathbb{C}$ with $|f_n(x)| \leq g(x)$ for all n and almost every x. Replacing g by $|g|$ we may and do assume $g : X \to [0, +\infty)$ is unsigned with $\int_X g \, d\mu < \infty$.
Almost uniform convergence implies pointwise almost everywhere convergence. This holds in complete generality (no domination needed) and is Exercise 1.5.2(iv); for completeness, here is the argument. For each $k \geq 1$ choose, by the definition of almost uniform convergence, a measurable set E_k with $\mu(E_k) \leq 1/k$ such that $f_n \to f$ uniformly on $X \setminus E_k$. Put $E := \bigcap_{k=1}^{\infty} E_k$; then $\mu(E) \leq \mu(E_k) \leq 1/k$ for every k, so $\mu(E) = 0$. If $x \notin E$, then $x \notin E_k$ for some k, and uniform (hence pointwise) convergence on $X \setminus E_k$ gives $f_n(x) \to f(x)$. Thus $f_n \to f$ pointwise outside the null set E.
Pointwise almost everywhere convergence implies almost uniform convergence. This is where domination is used; it is a version of Egorov's theorem (Exercise 1.4.31, Theorem 1.5.9) in which the finite measure hypothesis on X is replaced by domination.
Let Z be a null set outside of which all of the following hold simultaneously: $|f_n(x)| \leq g(x)$ for every n, and $f_n(x) \to f(x)$. (Such a Z exists, and may be taken measurable: each $\{x : |f_n(x)| > g(x)\}$ is measurable, being a level set of the measurable function $|f_n| - g$, and is null by the domination hypothesis; and the set $\{x : f_n(x) \not\to f(x)\} = \{x : \limsup_{n \to \infty} |f_n(x) - f(x)| > 0\}$ is measurable by Exercise 1.4.29, and is null by hypothesis. A countable union of measurable null sets is a measurable null set, so no completeness assumption on μ is needed here.) Note that for $x \notin Z$ we get $|f(x)| = \lim_n |f_n(x)| \leq g(x)$ as well, and hence

$$|f_n(x) - f(x)| \leq |f_n(x)| + |f(x)| \leq 2g(x) \qquad \text{for all } n \text{ and all } x \notin Z.$$

Call this bound $(*)$.
For $\eta > 0$ and $N \geq 1$ define the measurable sets

$$S_N^\eta := \bigcup_{n \geq N} \{x \in X \setminus Z : |f_n(x) - f(x)| > \eta\}.$$

For fixed η these are non-increasing in N. Two observations:
(a) S_1^η has finite measure. Indeed by $(*)$, if $x \in S_1^\eta$ then $2g(x) \geq |f_n(x) - f(x)| > \eta$ for some n, so $S_1^\eta \subseteq \{x : g(x) > \eta/2\}$, and Markov's inequality (Exercise 1.4.36(vi)) gives

$$\mu(S_1^\eta) \leq \mu(\{g > \eta/2\}) \leq \frac{2}{\eta} \int_X g \, d\mu < \infty.$$

(b) $\bigcap_{N=1}^\infty S_N^\eta = \emptyset$. Indeed if x lay in this intersection then $x \in X \setminus Z$ and $|f_n(x) - f(x)| > \eta$ for infinitely many n, contradicting $f_n(x) \to f(x)$.
Since the sets S_N^η are decreasing and the first one has finite measure, the downward monotone convergence property for sets (Exercise 1.4.23(iii)) applies and yields

$$\lim_{N\to\infty} \mu(S_N^\eta) = \mu\Big(\bigcap_{N\geq 1} S_N^\eta \Big) = 0.$$

(This is precisely the step at which domination substitutes for finiteness of $\mu(X)$; without some such hypothesis the downward convergence property fails, as the escape to horizontal infinity example of Exercise 1.5.15 shows.)
Now let $\varepsilon > 0$ be given. Using the $\varepsilon/2^k$ trick: for each $k \geq 1$, apply the above with $\eta = 1/k$ to find N_k such that

$$\mu\big(S_{N_k}^{1/k}\big) \leq \frac{\varepsilon}{2^k}.$$

Define the exceptional set

$$E := Z \cup \bigcup_{k=1}^\infty S_{N_k}^{1/k}.$$

Then E is measurable and, by countable subadditivity and $\mu(Z) = 0$,

$$\mu(E) \leq 0 + \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = \varepsilon.$$

We claim $f_n \to f$ uniformly on $X \setminus E$. Let $\eta > 0$, and choose k with $1/k \leq \eta$. If $x \in X \setminus E$ and $n \geq N_k$, then $x \notin Z$ and $x \notin S_{N_k}^{1/k}$; by the definition of $S_{N_k}^{1/k}$ (which is a union over all indices $n \geq N_k$) this forces

$$|f_n(x) - f(x)| \leq 1/k \leq \eta.$$

Thus $\sup_{x\in X\setminus E} |f_n(x) - f(x)| \leq \eta$ for all $n \geq N_k$, with N_k independent of x. As $\eta > 0$ was arbitrary, f_n converges to f uniformly on $X \setminus E$.
Since $\varepsilon > 0$ was arbitrary and $\mu(E) \leq \varepsilon$, this is exactly the assertion that f_n converges to f almost uniformly. This completes the proof of both directions.
Remark. It is worth recording exactly what this buys. For a dominated sequence, pointwise almost everywhere convergence and almost uniform convergence coincide (this exercise), and convergence in L^1 norm and convergence in measure coincide (Exercise 1.5.9). Moreover the first pair implies the second: given $f_n \to f$ pointwise almost everywhere, the bound $(*)$ shows the sequence $|f_n - f|$ is dominated by the absolutely integrable function $2g$ and tends to 0 pointwise almost everywhere, so the dominated convergence theorem (Theorem 1.4.49) gives $\|f_n - f\|_{L^1(\mu)} \to 0$. The implication does not reverse, even under domination: the typewriter sequence of Example 1.5.5 is dominated by the absolutely integrable function $1_{[0,1]}$ and converges to 0 in L^1 norm and in measure, but not pointwise almost everywhere and hence not almost uniformly. So domination merges these four modes into two equivalence classes, one strictly stronger than the other; it does not make them all equivalent.

Problem 1.5.19. Let X be a probability space (see Section 2.3). Given any real-valued measurable function $f : X \to \mathbb{R}$, we define the cumulative distribution function $F : \mathbb{R} \to [0, 1]$ of f to be the function $F(\lambda) := \mu(\{x \in X : f(x) \leq \lambda\})$. Given another sequence $f_n : X \to \mathbb{R}$ of real-valued

measurable functions, we say that f_n converges in distribution to f if the cumulative distribution function $F_n(\lambda)$ of f_n converges pointwise to the cumulative distribution function $F(\lambda)$ of f at all $\lambda \in \mathbb{R}$ for which F is continuous.

- (i) Show that if f_n converges to f in any of the seven senses discussed above (uniformly, essentially uniformly, almost uniformly, pointwise, pointwise almost everywhere, in L^1, or in measure), then it converges in distribution to f.
- (ii) Give an example in which f_n converges to f in distribution, but not in any of the above seven senses.
- (iii) Show that convergence in distribution is not linear, in the sense that if f_n converges to f in distribution, and g_n converges to g, then $f_n + g_n$ need not converge to $f + g$.
- (iv) Show that a sequence f_n can converge in distribution to two different limits f, g, which are not equal almost everywhere.

Solution. Throughout, $(X, \mathcal{B}, \mu)$ is a probability space, so $\mu(X) = 1 < \infty$; in particular the results of Section 1.5.3 for finite measure spaces are available.

(i) We first reduce all seven modes to convergence in measure, and then show that convergence in measure implies convergence in distribution.

Step 1: on a probability space, each of the seven modes implies convergence in measure. By Exercise 1.5.2, uniform convergence implies L^∞ convergence (part (ii)), L^∞ convergence implies almost uniform convergence (part (iii)), and almost uniform convergence implies convergence in measure (part (vii)). Also L^1 convergence implies convergence in measure (part (vi)), and pointwise convergence implies pointwise almost everywhere convergence (part (v)). Finally, since $\mu(X) = 1$ is finite, Egorov's theorem in the form of Theorem 1.5.9 says that pointwise almost everywhere convergence implies almost uniform convergence, hence convergence in measure. Chaining these, every one of the seven modes implies convergence in measure.

Step 2: convergence in measure implies convergence in distribution. Suppose $f_n \to f$ in measure, and let $\lambda \in \mathbb{R}$ be a point at which F is continuous. Fix $\eta > 0$.

If $f_n(x) \le \lambda$ and $|f_n(x) - f(x)| < \eta$, then $f(x) \le f_n(x) + \eta \le \lambda + \eta$. Hence

$$\{f_n \le \lambda\} \subseteq \{f \le \lambda + \eta\} \cup \{|f_n - f| \ge \eta\},$$

and so by monotonicity and finite subadditivity of μ,

$$F_n(\lambda) \le F(\lambda + \eta) + \mu(\{|f_n - f| \ge \eta\}).$$

Symmetrically, if $f(x) \le \lambda - \eta$ and $|f_n(x) - f(x)| < \eta$, then $f_n(x) < f(x) + \eta \le \lambda$, so

$$\{f \le \lambda - \eta\} \subseteq \{f_n \le \lambda\} \cup \{|f_n - f| \ge \eta\},$$

giving

$$F(\lambda - \eta) \le F_n(\lambda) + \mu(\{|f_n - f| \ge \eta\}).$$

By the definition of convergence in measure, $\mu(\{|f_n - f| \ge \eta\}) \to 0$ as $n \to \infty$. Letting $n \to \infty$ in the two displays therefore gives

$$F(\lambda - \eta) \le \liminf_{n\to\infty} F_n(\lambda) \le \limsup_{n\to\infty} F_n(\lambda) \le F(\lambda + \eta).$$

This holds for every $\eta > 0$. Since F is continuous at λ, letting $\eta \to 0^+$ makes both $F(\lambda - \eta)$ and $F(\lambda + \eta)$ converge to $F(\lambda)$, and we conclude

$$\lim_{n\to\infty} F_n(\lambda) = F(\lambda).$$

As λ was an arbitrary continuity point of F, f_n converges to f in distribution. Combining with Step 1 proves (i).

(ii) Take $X = [0, 1]$ with Lebesgue measure restricted to $[0, 1]$, which is a probability space. Let

$$f := 1_{[0,1/2)}, \qquad f_n := 1_{[1/2,1]} \quad \text{for every } n.$$

Both f and each f_n take only the values 0 and 1, and each takes the value 1 on a set of measure $1/2$ and the value 0 on a set of measure $1/2$. Hence their cumulative distribution functions coincide:

$$F(\lambda) = F_n(\lambda) = \begin{cases} 0, & \lambda < 0 \\ 1/2, & 0 \le \lambda < 1 \\ 1, & \lambda \ge 1. \end{cases}$$

So $F_n(\lambda) = F(\lambda)$ for every $\lambda \in \mathbb{R}$, and in particular $f_n \to f$ in distribution.
On the other hand $|f_n(x) - f(x)| = 1$ for every $x \in [0,1]$, since exactly one of $1_{[0,1/2)}(x)$, $1_{[1/2,1]}(x)$ equals 1. Therefore

$$\mu(\{x : |f_n(x) - f(x)| \ge 1/2\}) = \mu([0,1]) = 1 \not\to 0,$$

so f_n does not converge to f in measure. By Step 1 of part (i), every one of the seven modes of convergence implies convergence in measure on a probability space; hence f_n converges to f in none of the seven senses.
(iii) Keep $X = [0,1]$ with Lebesgue measure, and set

$$f := 1_{[0,1/2)}, \qquad g := 1_{[1/2,1]}, \qquad f_n := g, \qquad g_n := g \quad \text{for every } n.$$

By part (ii), $f_n \to f$ in distribution (the two have identical cumulative distribution functions). Also $g_n \to g$ in distribution trivially, since $g_n = g$ for all n so the distribution functions are literally equal. However $f_n + g_n = 2g = 2 \cdot 1_{[1/2,1]}$ for every n, while $f + g = 1_{[0,1/2)} + 1_{[1/2,1]} = 1$ identically on $[0,1]$ (the two indicators have disjoint supports whose union is all of $[0,1]$). The cumulative distribution functions are

$$F_{f+g}(\lambda) = \begin{cases} 0, & \lambda < 1 \\ 1, & \lambda \ge 1, \end{cases} \qquad F_{2g}(\lambda) = \begin{cases} 0, & \lambda < 0 \\ 1/2, & 0 \le \lambda < 2 \\ 1, & \lambda \ge 2. \end{cases}$$

The function F_{f+g} is continuous at $\lambda = 3/2$, and there $F_{f+g}(3/2) = 1$, whereas $F_{f_n+g_n}(3/2) = F_{2g}(3/2) = 1/2$ for every n. Thus $F_{f_n+g_n}(3/2) \to 1/2 \ne 1 = F_{f+g}(3/2)$, and $f_n + g_n$ does not converge in distribution to $f + g$. Hence convergence in distribution is not linear; contrast this with Exercise 1.5.1(ii), which shows that all seven of the other modes are linear.
(iv) Again on $X = [0,1]$ with Lebesgue measure, set

$$f_n := 1_{[0,1/2)} \quad \text{for every } n, \qquad f := 1_{[0,1/2)}, \qquad g := 1_{[1/2,1]}.$$

Since $f_n = f$ for all n, we have $F_n = F$ identically and $f_n \to f$ in distribution. But as computed in (ii), f and g have the same cumulative distribution function, so $F_n = F = F_g$ identically as well, and hence $f_n \to g$ in distribution too.
Yet f and g are not equal almost everywhere: $\{x \in [0,1] : f(x) \ne g(x)\} = [0,1]$, a set of measure 1. So convergence in distribution does not determine the limit even up to null sets; contrast this with Proposition 1.5.7, which shows that the seven other modes are mutually compatible in exactly that sense. (This is unsurprising: the cumulative distribution function only remembers the pushforward measure $f_*\mu$ on $\mathbb{R}$, and completely forgets how f is arranged over the underlying space X.)

2.6 §1.6 — Differentiation Theorems

Problem 1.6.1. If $F : [a,b] \to \mathbb{R}$ is everywhere differentiable, show that F is continuous and F' is measurable. If F is almost everywhere differentiable, show that the (almost everywhere defined) function F' is measurable (i.e. it is equal to an everywhere defined measurable function on $[a,b]$ outside of a null set), but give an example to demonstrate that F need not be continuous.

Solution. Throughout, $[a,b]$ is a compact interval of positive length, and differentiability at $x \in [a,b]$

means that the limit (1.19)

$$F'(x) = \lim_{y \to x;\, y \in [a,b] \setminus \{x\}} \frac{F(y) - F(x)}{y - x}$$

exists (and is a real number). Note that at the endpoints this is a one-sided limit.
Continuity in the everywhere differentiable case. Fix $x \in [a,b]$. For $y \in [a,b] \setminus \{x\}$ we may write

$$F(y) - F(x) = \frac{F(y) - F(x)}{y - x} \cdot (y - x).$$

As $y \to x$ the first factor converges to the real number $F'(x)$ and the second converges to 0, so the product converges to 0. Thus $F(y) \to F(x)$ as $y \to x$, i.e. F is continuous at x. As x was arbitrary, F is continuous on $[a,b]$. (Observe that this argument only used differentiability at the single point x; we record this for later use: *if F is differentiable at x, then F is continuous at x.*)
Measurability of F' in the everywhere differentiable case. Extend F to a function $\hat{F} : \mathbb{R} \to \mathbb{R}$ by setting $\hat{F}(x) := F(a)$ for $x < a$ and $\hat{F}(x) := F(b)$ for $x > b$. Since F is continuous on $[a,b]$, the extension $\hat{F}$ is continuous on all of $\mathbb{R}$, and hence Lebesgue measurable by Exercise 1.3.8(i). For each $n \geq 1$ define

$$g_n(x) := n\left(\hat{F}\left(x + \tfrac{1}{n}\right) - \hat{F}(x)\right).$$

Each g_n is continuous, hence measurable.
Let $x \in [a,b)$. For all n large enough that $\frac{1}{n} < b - x$ we have $x + \frac{1}{n} \in (x,b] \subset [a,b]$, and therefore

$$g_n(x) = \frac{F(x + \frac{1}{n}) - F(x)}{\frac{1}{n}}.$$

Since $x + \frac{1}{n} \to x$ with $x + \frac{1}{n} \in [a,b] \setminus \{x\}$, the limit (1.19) gives $g_n(x) \to F'(x)$.
Consequently the functions $g_n 1_{[a,b)}$, which are measurable as products of measurable functions (Exercise 1.3.8(vi), the indicator $1_{[a,b)}$ being measurable since $[a,b)$ is a Lebesgue measurable set), converge pointwise on all of $\mathbb{R}$ to the function $F'1_{[a,b)}$ (extended by zero outside $[a,b)$), so by Exercise 1.3.8(iv) the latter is measurable. Finally, the everywhere defined function on $[a,b]$ which equals F' on $[a,b)$ and equals $F'(b)$ at $x = b$ differs from the measurable function $F'1_{[a,b)}$ only at the single point b, which is a null set, and so is measurable by Exercise 1.3.8(iii). Hence F' is measurable on $[a,b]$.
The almost everywhere differentiable case. Now suppose $F : [a,b] \to \mathbb{R}$ is differentiable at every point of a set $E \subset [a,b]$ with $N := [a,b] \setminus E$ null. Null sets are Lebesgue measurable, so $E = [a,b] \setminus N$ is measurable.
Step 1: F is measurable. By the remark recorded above, F is continuous at every point $x \in E$ (as a function on $[a,b]$). Fix $\lambda \in \mathbb{R}$ and set

$$U_\lambda := \{x \in \mathbb{R} : \exists\, \delta > 0 \text{ such that } F(y) > \lambda \text{ for all } y \in [a,b] \text{ with } |y - x| < \delta\}.$$

The set U_λ is open: if $x \in U_\lambda$ with witness δ, then every x' with $|x' - x| < \delta/2$ lies in U_λ with witness $\delta/2$. Now if $x \in E$ and $F(x) > \lambda$, continuity of F at x produces such a δ, so $x \in U_\lambda$; conversely if $x \in E \cap U_\lambda$ then taking $y := x$ in the definition gives $F(x) > \lambda$. Hence

$$\{x \in E : F(x) > \lambda\} = E \cap U_\lambda,$$

which is measurable. Let $\tilde{F} : \mathbb{R} \to \mathbb{R}$ be F on E and 0 elsewhere; then $\{\tilde{F} > \lambda\}$ equals $E \cap U_\lambda$ when $\lambda \geq 0$ and equals $(E \cap U_\lambda) \cup (\mathbb{R} \setminus E)$ when $\lambda < 0$, and in either case is measurable. Since all superlevel sets of the real-valued $\tilde{F}$ are measurable, $\tilde{F}$ is measurable by Exercise 1.3.7(iv). Indeed, for open $U \subset \mathbb{C}$ we have $\tilde{F}^{-1}(U) = \tilde{F}^{-1}(U \cap \mathbb{R})$ (as $\tilde{F}$ is real-valued), and $U \cap \mathbb{R}$ is open in $\mathbb{R}$, hence a countable union of bounded open intervals (c,d), while

$$\tilde{F}^{-1}((c,d)) = \{\tilde{F} > c\} \setminus \bigcap_{n \geq 1} \left\{\tilde{F} > d - \tfrac{1}{n}\right\},$$

so $\tilde{F}^{-1}(U)$ is obtained from the sets $\{\tilde{F} > \lambda\}$ by countably many Boolean operations, and is measurable by Lemma 1.2.13. Now let $\hat{F} : \mathbb{R} \to \mathbb{R}$ equal F on $[a,b]$ and 0 outside $[a,b]$. Then $\hat{F}$ agrees with $\tilde{F}$

outside the null set N, so $\hat{F}$ is measurable by Exercise 1.3.8(iii). (Note that no continuity of F off E was needed: whatever F does on the null set N is invisible to measurability.)
Step 2: measurability of F'. Define, as before,

$$g_n(x) := n\left(\hat{F}\left(x+\tfrac{1}{n}\right) - \hat{F}(x)\right).$$

The translate $\hat{F}_n(x) := \hat{F}(x+\frac{1}{n})$ is measurable: for any open U one has $\hat{F}_n^{-1}(U) = \hat{F}^{-1}(U) - \frac{1}{n}$, which is Lebesgue measurable by translation invariance of Lebesgue measurability (Exercise 1.2.20), so the criterion of Exercise 1.3.7(iv) applies once more. Hence each g_n is measurable, being a linear combination of measurable functions (Exercise 1.3.8(vi)).
Let $x \in E \cap [a,b)$. For n large enough that $\frac{1}{n} < b-x$, the point $x+\frac{1}{n}$ lies in $[a,b]$, so $\hat{F}(x+\frac{1}{n}) = F(x+\frac{1}{n})$ and

$$g_n(x) = \frac{F(x+\frac{1}{n}) - F(x)}{\frac{1}{n}} \longrightarrow F'(x)$$

since F is differentiable at x. (It is irrelevant whether or not the points $x+\frac{1}{n}$ lie in E: differentiability at x is a statement about the values of F at *all* nearby points of $[a,b]$.)
Therefore the measurable functions $g_n 1_{E\cap[a,b)}$ converge pointwise everywhere on $\mathbb{R}$ to the function

$$G := F' 1_{E\cap[a,b)},$$

extended by 0 off $E \cap [a,b)$. By Exercise 1.3.8(iv), G is measurable. Since G is everywhere defined on $[a,b]$ and agrees with F' outside the null set $N \cup \{b\}$, the almost everywhere defined function F' is measurable in the required sense.
F need not be continuous. Take $[a,b] = [0,1]$ and

$$F(x) := \begin{cases} 1, & x = \frac{1}{2}, \\ 0, & x \neq \frac{1}{2}. \end{cases}$$

If $x \in [0,1]\setminus\{\frac{1}{2}\}$, then F vanishes identically on $[0,1]\cap(x-\delta, x+\delta)$ for $\delta := |x-\frac{1}{2}| > 0$, so all difference quotients at x with $|y-x| < \delta$ vanish and $F'(x) = 0$. Hence F is differentiable on $[0,1]\setminus\{\frac{1}{2}\}$, whose complement $\{\frac{1}{2}\}$ is null, so F is almost everywhere differentiable. But F is plainly discontinuous at $\frac{1}{2}$ (indeed $F(y) \to 0 \neq 1 = F(\frac{1}{2})$ as $y \to \frac{1}{2}$). Consistently with the above, $F' = 0$ off a null set, which is certainly measurable.

Problem 1.6.2. Give an example of a function $F : [a,b] \to \mathbb{R}$ which is everywhere differentiable, but not continuously differentiable. (Hint: choose an F that vanishes quickly at some point, say at the origin 0, but which also oscillates rapidly near that point.)

Solution. Take $[a,b] := [-1,1]$ and define

$$F(x) := \begin{cases} x^2 \sin(1/x), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

This vanishes to second order at the origin while oscillating with unbounded frequency there, as the hint suggests.
F is differentiable at every $x \neq 0$. On the open set $[-1,1]\setminus\{0\}$ the function F is a composition and product of the functions $x \mapsto x^2$, $x \mapsto 1/x$ and sin, all of which are differentiable there, so by the usual product and chain rules of single variable calculus

$$F'(x) = 2x\sin(1/x) - \cos(1/x) \qquad (x \neq 0).$$

(At the endpoints $x = \pm 1$ the limit (1.19) is one-sided, but the same formula holds, since the two-sided derivative of the smooth extension exists there.)

F is differentiable at 0, with $F'(0) = 0$. For $h \in [-1, 1] \setminus \{0\}$,

$$\left|\frac{F(h) - F(0)}{h}\right| = \left|\frac{h^2 \sin(1/h)}{h}\right| = |h|\,|\sin(1/h)| \leq |h|,$$

which tends to 0 as $h \to 0$. Hence the limit (1.19) exists at $x = 0$ and equals 0.
Thus F is everywhere differentiable on $[-1, 1]$.
F' is not continuous at 0. Consider the sequence $x_n := \frac{1}{2\pi n}$ for $n \geq 1$, which lies in $(0, 1]$ and tends to 0. Then $\sin(1/x_n) = \sin(2\pi n) = 0$ and $\cos(1/x_n) = \cos(2\pi n) = 1$, so

$$F'(x_n) = 2x_n \cdot 0 - 1 = -1$$

for every n. Hence $F'(x_n) \to -1$, whereas $F'(0) = 0$. So F' fails to be continuous at the origin, and F is not continuously differentiable.
(For contrast, note that F' is continuous at every other point of $[-1, 1]$, by the explicit formula above; the single point 0 is the sole obstruction. Note also that F' is bounded here, by $|F'| \leq 3$; one can make the derivative unbounded as well by taking $x^2 \sin(1/x^2)$ instead, whose derivative $2x \sin(1/x^2) - \frac{2}{x}\cos(1/x^2)$ is unbounded near the origin, showing that an everywhere differentiable function need not have a Riemann integrable derivative.)

Problem 1.6.3. Give an example to show that Rolle's theorem can fail if f is merely assumed to be almost everywhere differentiable, even if one adds the additional hypothesis that f is continuous. This example illustrates that everywhere differentiability is a significantly stronger property than almost everywhere differentiability. We will see further evidence of this fact later in these notes; there are many theorems that assert in their conclusion that a function is almost everywhere differentiable, but few that manage to conclude everywhere differentiability.

Solution. Take $[a, b] := [0, 1]$ and define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) := \left|x - \tfrac{1}{2}\right|.$$

f is continuous and satisfies the boundary hypothesis of Rolle's theorem. The absolute value function is continuous and $x \mapsto x - \frac{1}{2}$ is continuous, so f is continuous on $[0, 1]$. Moreover

$$f(0) = \left|-\tfrac{1}{2}\right| = \tfrac{1}{2} = \left|\tfrac{1}{2}\right| = f(1),$$

so $f(a) = f(b)$ as required in Theorem 1.6.2.
f is almost everywhere differentiable. If $x \in [0, \frac{1}{2})$, then $f(y) = \frac{1}{2} - y$ for all $y \in [0, 1]$ with $|y - x| < \frac{1}{2} - x$, so the difference quotients at x are identically -1 near x and $f'(x) = -1$. Similarly, if $x \in (\frac{1}{2}, 1]$ then $f(y) = y - \frac{1}{2}$ for y near x and $f'(x) = +1$. Thus f is differentiable at every point of $[0, 1] \setminus \{\frac{1}{2}\}$, and the exceptional set $\{\frac{1}{2}\}$ has Lebesgue measure zero; so f is almost everywhere differentiable.
(For completeness: f really is not differentiable at $\frac{1}{2}$, since the difference quotient $\frac{f(y) - f(\frac{1}{2})}{y - \frac{1}{2}} = \frac{|y - \frac{1}{2}|}{y - \frac{1}{2}}$ equals $+1$ for $y > \frac{1}{2}$ and -1 for $y < \frac{1}{2}$, so has no limit as $y \to \frac{1}{2}$.)
The conclusion of Rolle's theorem fails. The conclusion would assert the existence of some $x \in (0, 1)$ with $f'(x) = 0$. But at every point $x \in (0, 1)$ at which f is differentiable we computed $f'(x) = \pm 1 \neq 0$, and at the only remaining point $x = \frac{1}{2}$ the derivative does not exist at all. So there is no $x \in (0, 1)$ with $f'(x) = 0$, and Theorem 1.6.2 fails for this f.
Thus dropping "everywhere differentiable" to "continuous and almost everywhere differentiable" destroys Rolle's theorem, even though the exceptional set here is as small as it could possibly be (a single point). The reason the proof of Theorem 1.6.2 breaks down is that it locates an interior maximum x of f (or of $-f$) and then differentiates there; here the interior extremum is precisely the point $\frac{1}{2}$ at which differentiability was surrendered, so no information about f' can be extracted from it.

Problem 1.6.4. (Uniqueness of antiderivatives up to constants) Let $[a,b]$ be a compact interval of positive length, and let $F : [a,b] \to \mathbb{R}$ and $G : [a,b] \to \mathbb{R}$ be differentiable functions. Show that $F'(x) = G'(x)$ for every $x \in [a,b]$ if and only if $F(x) = G(x) + C$ for some constant $C \in \mathbb{R}$ and all $x \in [a,b]$.

Solution. The "if" direction. Suppose $F(x) = G(x) + C$ for all $x \in [a,b]$ and some constant C. Then for any $x \in [a,b]$ and $y \in [a,b] \setminus \{x\}$,

$$\frac{F(y) - F(x)}{y - x} = \frac{(G(y) + C) - (G(x) + C)}{y - x} = \frac{G(y) - G(x)}{y - x},$$

so the two difference quotients are identical functions of y; since the limit (1.19) of the right-hand side exists and equals $G'(x)$, the same holds for the left-hand side, giving $F'(x) = G'(x)$. (This direction needs no mean value theorem.)
The "only if" direction. Suppose now $F'(x) = G'(x)$ for every $x \in [a,b]$. Put $H := F - G$, so $H : [a,b] \to \mathbb{R}$. For $x \in [a,b]$ and $y \in [a,b] \setminus \{x\}$ we have

$$\frac{H(y) - H(x)}{y - x} = \frac{F(y) - F(x)}{y - x} - \frac{G(y) - G(x)}{y - x},$$

and both terms on the right converge as $y \to x$ (to $F'(x)$ and $G'(x)$ respectively), so by the limit laws H is differentiable at x with

$$H'(x) = F'(x) - G'(x) = 0.$$

Thus H is differentiable on $[a,b]$ with identically vanishing derivative.
Set $C := H(a)$. We claim $H(x) = C$ for every $x \in [a,b]$. This is trivial for $x = a$. So fix $x \in (a,b]$ and consider the compact interval $[a,x]$, which has positive length since $x > a$. The restriction $H|_{[a,x]}$ is differentiable at every point of $[a,x]$: indeed, for $z \in [a,x]$ the limit

$$\lim_{y \to z;\, y \in [a,x] \setminus \{z\}} \frac{H(y) - H(z)}{y - z}$$

is a limit of the same difference quotient along a subset of $[a,b] \setminus \{z\}$, so it exists and equals $H'(z) = 0$ because the corresponding limit over the larger set $[a,b] \setminus \{z\}$ exists and equals 0.
Therefore the mean value theorem (Corollary 1.6.5) applies to $H|_{[a,x]}$ on $[a,x]$, and yields a point $\xi \in (a,x)$ with

$$H'(\xi) = \frac{H(x) - H(a)}{x - a}.$$

The left-hand side is 0, and $x - a \neq 0$, so $H(x) = H(a) = C$.
Hence $H \equiv C$ on $[a,b]$, i.e. $F(x) = G(x) + C$ for all $x \in [a,b]$, as required.
Remark. The hypothesis that the domain is an interval is essential: on the disconnected domain $[0,1] \cup [2,3]$ the functions $F := 1_{[2,3]}$ and $G := 0$ have equal (vanishing) derivatives everywhere but differ by a non-constant function. It is precisely the mean value theorem, which needs an interval to run through, that rules this out above.

Problem 1.6.5. Let $f : \mathbb{R} \to \mathbb{C}$ be an absolutely integrable function, and let $F : \mathbb{R} \to \mathbb{C}$ be the definite integral $F(x) := \int_{[-\infty,x]} f(t)\, dt$. Show that F is continuous.

Solution. First note that F is well defined: for each $x \in \mathbb{R}$ the function $f 1_{[-\infty,x]}$ is measurable (a product of measurable functions, Exercise 1.3.8(vi), the indicator of the measurable set $[-\infty,x] \cap \mathbb{R} = (-\infty,x]$ being measurable), and it obeys $|f 1_{[-\infty,x]}| \leq |f|$ pointwise, so it is absolutely integrable since f is.
Since $\mathbb{R}$ is a metric space, it suffices to prove sequential continuity: we fix $x \in \mathbb{R}$ and an arbitrary sequence $(x_n)_{n \geq 1}$ of reals with $x_n \to x$, and show that $F(x_n) \to F(x)$.

By linearity of the Lebesgue integral on $L^1(\mathbb{R} \to \mathbb{C})$ (Exercise 1.3.19),

$$F(x_n) - F(x) = \int_{\mathbb{R}} f(t) \left(1_{[-\infty,x_n]}(t) - 1_{[-\infty,x]}(t)\right)\ dt = \int_{\mathbb{R}} h_n(t)\ dt,$$

where

$$h_n := f \cdot \left(1_{[-\infty,x_n]} - 1_{[-\infty,x]}\right).$$

Each h_n is measurable, being a product of measurable functions.
Domination. The bracket takes values in $\{-1, 0, 1\}$, so $|h_n(t)| \leq |f(t)|$ for every $t \in \mathbb{R}$ and every n, and $|f|$ is absolutely integrable by hypothesis.
Pointwise convergence. Let $t \in \mathbb{R}$ with $t \neq x$. If $t < x$, then since $x_n \to x$ we have $x_n > t$ for all sufficiently large n, so $1_{[-\infty,x_n]}(t) = 1 = 1_{[-\infty,x]}(t)$ and hence $h_n(t) = 0$ for all large n. If $t > x$, then $x_n < t$ for all sufficiently large n, so both indicators vanish at t and again $h_n(t) = 0$ for all large n. Thus $h_n \to 0$ pointwise on $\mathbb{R} \setminus \{x\}$, i.e. pointwise almost everywhere (the exceptional set $\{x\}$ being a null set).
Conclusion. By the dominated convergence theorem (Theorem 1.4.49), applied on the measure space $\mathbb{R}$ with Lebesgue measure and with dominating function $|f|$,

$$F(x_n) - F(x) = \int_{\mathbb{R}} h_n(t)\ dt \longrightarrow \int_{\mathbb{R}} 0\ dt = 0.$$

Hence $F(x_n) \to F(x)$. As the sequence $x_n \to x$ was arbitrary, F is continuous at x; and as $x \in \mathbb{R}$ was arbitrary, F is continuous on $\mathbb{R}$.
Remark. Equivalently, for $x < y$ one has $|F(y) - F(x)| \leq \int_{(x,y]} |f(t)|\ dt$ by the triangle inequality (Lemma 1.3.19), and the above dominated convergence argument is exactly the statement that this quantity tends to 0 as $y - x \to 0$. Note that no continuity of f was used, and indeed f may be wildly discontinuous; integration always smooths by one degree of regularity.

Problem 1.6.6. Show that Theorem 1.6.11 follows from Theorem 1.6.12.

Solution. Recall the two statements. Theorem 1.6.11 asserts that if $f : \mathbb{R} \to \mathbb{C}$ is absolutely integrable and $F(x) := \int_{[-\infty,x]} f(t)\ dt$, then F is continuous and almost everywhere differentiable, with $F'(x) = f(x)$ for almost every x. Theorem 1.6.12 asserts that for such an f,

$$\lim_{h \to 0^+} \frac{1}{h} \int_{[x,x+h]} f(t)\ dt = f(x) \tag{1.20}$$

for almost every $x \in \mathbb{R}$, and

$$\lim_{h \to 0^+} \frac{1}{h} \int_{[x-h,x]} f(t)\ dt = f(x) \tag{1.21}$$

for almost every $x \in \mathbb{R}$.
Continuity. This is exactly Exercise 1.6.5, and needs nothing from Theorem 1.6.12.
Identification of the difference quotients. Let $x \in \mathbb{R}$ and $h > 0$. Since $1_{[-\infty,x+h]} - 1_{[-\infty,x]} = 1_{(x,x+h]}$ pointwise, linearity of the integral gives

$$F(x+h) - F(x) = \int_{\mathbb{R}} f\, 1_{(x,x+h]} = \int_{(x,x+h]} f(t)\ dt.$$

The functions $f1_{(x,x+h]}$ and $f1_{[x,x+h]}$ agree outside the null set $\{x\}$, hence have the same integral, so

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_{[x,x+h]} f(t)\ dt. \tag{*}$$

Likewise $1_{[-\infty,x]} - 1_{[-\infty,x-h]} = 1_{(x-h,x]}$, so $F(x) - F(x-h) = \int_{(x-h,x]} f = \int_{[x-h,x]} f$ (again the two integrands differ only on the null set $\{x-h\}$), and hence

$$\frac{F(x-h) - F(x)}{(x-h) - x} = \frac{F(x) - F(x-h)}{h} = \frac{1}{h} \int_{[x-h,x]} f(t)\ dt. \tag{**}$$

Assembling the two one-sided limits. Let $A \subset \mathbb{R}$ be the set of x for which (1.20) holds and $B \subset \mathbb{R}$ the set for which (1.21) holds. By Theorem 1.6.12 both $\mathbb{R} \setminus A$ and $\mathbb{R} \setminus B$ are null, so their union $\mathbb{R} \setminus (A \cap B)$ is null (a finite union of null sets is null); thus $A \cap B$ has full measure.
Fix $x \in A \cap B$ and let $\varepsilon > 0$. By (1.20) there is $\delta_1 > 0$ such that

$$\left| \frac{1}{h} \int_{[x,x+h]} f(t)\ dt - f(x) \right| \leq \varepsilon \quad \text{whenever } 0 < h < \delta_1,$$

and by (1.21) there is $\delta_2 > 0$ such that

$$\left| \frac{1}{h} \int_{[x-h,x]} f(t)\ dt - f(x) \right| \leq \varepsilon \quad \text{whenever } 0 < h < \delta_2.$$

Set $\delta := \min(\delta_1, \delta_2) > 0$ and let $y \in \mathbb{R} \setminus \{x\}$ satisfy $|y - x| < \delta$.
If $y > x$, write $h := y - x \in (0, \delta)$; by $(*)$,

$$\left| \frac{F(y) - F(x)}{y - x} - f(x) \right| = \left| \frac{1}{h} \int_{[x,x+h]} f(t)\ dt - f(x) \right| \leq \varepsilon.$$

If $y < x$, write $h := x - y \in (0, \delta)$, so $y = x - h$; by $(**)$,

$$\left| \frac{F(y) - F(x)}{y - x} - f(x) \right| = \left| \frac{1}{h} \int_{[x-h,x]} f(t)\ dt - f(x) \right| \leq \varepsilon.$$

In either case the difference quotient lies within ε of $f(x)$. Since $\varepsilon > 0$ was arbitrary, the limit (1.19) exists at x and

$$F'(x) = \lim_{y \to x;\, y \neq x} \frac{F(y) - F(x)}{y - x} = f(x).$$

Thus F is differentiable at every point of the full-measure set $A \cap B$, with derivative f there. In other words F is almost everywhere differentiable with $F'(x) = f(x)$ for almost every $x \in \mathbb{R}$, which together with the continuity established above is precisely Theorem 1.6.11.
Remark. Both halves of Theorem 1.6.12 are genuinely needed: (1.20) alone gives only the right-hand derivative of F, and a function can have a right derivative everywhere without being differentiable (for instance $x \mapsto |x|$ at the origin).

Problem 1.6.7. Let $f : \mathbb{R}^d \to \mathbb{C}$, $g : \mathbb{R}^d \to \mathbb{C}$ be Lebesgue measurable functions such that f is absolutely integrable and g is essentially bounded (i.e. bounded outside of a null set). Show that the convolution $f * g : \mathbb{R}^d \to \mathbb{C}$ defined by the formula

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

is well-defined (in the sense that the integrand on the right-hand side is absolutely integrable) and that $f * g$ is a bounded, continuous function.

Solution. Since g is essentially bounded, we may fix $M \geq 0$ and a null set $N \subset \mathbb{R}^d$ such that

$$|g(z)| \leq M \quad \text{for all } z \in \mathbb{R}^d \setminus N.$$

Write $\|f\|_{L^1} := \int_{\mathbb{R}^d} |f(y)|\ dy < \infty$.
We record once and for all the affine invariance we shall use. For each fixed $x \in \mathbb{R}^d$, the map $T_x(y) := x - y$ is the composition of the invertible linear map $R(y) := -y$ (of determinant $(-1)^d$, of absolute value 1) with the translation $y \mapsto y + x$; note T_x is its own inverse.
At the level of sets, Exercise 1.2.20 (translation invariance) and Exercise 1.2.21 (linear change of variables) show that $T_x(E) = x - E$ is Lebesgue measurable with $m(x - E) = |\det R|\, m(E) = m(E)$

whenever E is Lebesgue measurable; in particular T_x maps null sets to null sets. Consequently, if F is a Lebesgue measurable function then so is $F \circ T_x$, since $(F \circ T_x)^{-1}(U) = T_x(F^{-1}(U))$ for every open U, and we may then apply Exercise 1.3.7(iv). At the level of integrals, Exercise 1.3.15 (translation invariance) and Exercise 1.3.16 (linear change of variables, with $|\det R| = 1$), together with Exercise 1.3.20 which extends both to complex-valued absolutely integrable functions, give

$$\int_{\mathbb{R}^d} F(T_x(y))\ dy = \int_{\mathbb{R}^d} F(z)\ dz \tag{$\dagger$}$$

for any unsigned measurable or absolutely integrable F.

Well-definedness. Fix $x \in \mathbb{R}^d$. By the above, $y \mapsto g(x-y) = g(T_x(y))$ is measurable, hence so is the product $y \mapsto f(y)g(x-y)$ by Exercise 1.3.8(vi). Moreover, the set $T_x(N) = x - N$ is null, and for every $y \notin x - N$ we have $x - y \notin N$ and therefore $|g(x-y)| \leq M$. Consequently

$$|f(y)g(x-y)| \leq M|f(y)| \quad \text{for almost every } y \in \mathbb{R}^d,$$

and since the unsigned Lebesgue integral is monotone with respect to almost everywhere domination and is homogeneous (Exercise 1.3.10(ii),(iii)),

$$\int_{\mathbb{R}^d} |f(y)g(x-y)|\ dy \leq M \int_{\mathbb{R}^d} |f(y)|\ dy = M\|f\|_{L^1} < \infty.$$

Thus the integrand is absolutely integrable and $f * g(x)$ is a well-defined complex number.

Boundedness. By the triangle inequality for the Lebesgue integral (Lemma 1.3.19) and the bound just obtained,

$$|f * g(x)| \leq \int_{\mathbb{R}^d} |f(y)g(x-y)|\ dy \leq M\|f\|_{L^1}$$

for every $x \in \mathbb{R}^d$. So $f * g$ is bounded, with $\sup_x |f * g(x)| \leq M\|f\|_{L^1}$.

Rewriting the convolution. Applying ($\dagger$) to the absolutely integrable function $F(z) := f(x-z)g(z)$ (measurable and absolutely integrable by the same reasoning as above, with the roles of the two factors exchanged), and noting $F(T_x(y)) = f(x-(x-y))g(x-y) = f(y)g(x-y)$, we obtain

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

Continuity. Let $x, x' \in \mathbb{R}^d$. Both $z \mapsto f(x-z)g(z)$ and $z \mapsto f(x'-z)g(z)$ are absolutely integrable, so by linearity of the integral,

$$f * g(x) - f * g(x') = \int_{\mathbb{R}^d} \big(f(x-z) - f(x'-z)\big)\, g(z)\ dz,$$

and hence, by the triangle inequality (Lemma 1.3.19) together with $|g(z)| \leq M$ for almost every z,

$$|f * g(x) - f * g(x')| \leq \int_{\mathbb{R}^d} |f(x-z) - f(x'-z)|\, |g(z)|\ dz \leq M \int_{\mathbb{R}^d} |f(x-z) - f(x'-z)|\ dz.$$

Now apply ($\dagger$) once more, substituting $w := x - z$ (that is, integrating the unsigned function $w \mapsto |f(w) - f(x'-x+w)|$ against the map T_x). Since $x' - z = x' - x + w = w - h$ with

$$h := x - x',$$

we get

$$\int_{\mathbb{R}^d} |f(x-z) - f(x'-z)|\ dz = \int_{\mathbb{R}^d} |f(w) - f(w-h)|\ dw = \int_{\mathbb{R}^d} |f(w) - f_h(w)|\ dw,$$

where $f_h(w) := f(w-h)$ is the shifted function of Proposition 1.6.13. Therefore

$$|f * g(x) - f * g(x')| \leq M \int_{\mathbb{R}^d} |f_h(w) - f(w)|\ dw, \qquad h = x - x'.$$

By Proposition 1.6.13 (translation is continuous in L^1), the right-hand side tends to 0 as $h \to 0$. Explicitly, given $\varepsilon > 0$ there exists $\delta > 0$ such that $\int_{\mathbb{R}^d} |f_h - f| \leq \varepsilon/(M+1)$ whenever $|h| \leq \delta$; and then

$$|f * g(x) - f * g(x')| \leq \frac{M}{M+1}\varepsilon \leq \varepsilon \quad \text{whenever } |x - x'| \leq \delta.$$

Since δ depends only on ε (and on f, M), this shows that $f * g$ is in fact *uniformly* continuous on $\mathbb{R}^d$, and in particular continuous.

Remark. Note the division of labour: the essential boundedness of g supplied the quantitative estimate, while the absolute integrability of f supplied the regularity, via the density argument hidden inside Proposition 1.6.13. Neither factor was assumed continuous, yet the convolution is uniformly continuous – an instance of the smoothing phenomenon discussed after the exercise, and the basis for the Steinhaus theorem in Exercise 1.6.8.

Problem 1.6.8. (Steinhaus theorem) Let $E \subset \mathbb{R}^d$ be a Lebesgue measurable set of positive measure. Show that the set $E - E := \{x - y : x, y \in E\}$ contains an open neighbourhood of the origin. (Hint: reduce to the case when E is bounded, and then apply the previous exercise to the convolution $1_E * 1_{-E}$, where $-E := \{-y : y \in E\}$.)

Solution. Reduction to the bounded case. Since $E = \bigcup_{N=1}^{\infty}(E \cap B(0,N))$ is an increasing union, upward monotone convergence of Lebesgue measure gives $m(E \cap B(0,N)) \to m(E) > 0$, so there is an N with $m(E \cap B(0,N)) > 0$. Writing $E' := E \cap B(0,N)$, we have $E' - E' \subseteq E - E$, so it suffices to prove the claim for E'. Thus we may and do assume that E is measurable, bounded, and $0 < m(E) < \infty$.

The convolution. Let $f := 1_E$ and $g := 1_{-E}$. The set $-E$ is measurable (reflection $y \mapsto -y$ is a linear bijection, and Lebesgue measurability is preserved under it), so g is measurable, and g is bounded by 1, hence essentially bounded. Also f is measurable with $\int_{\mathbb{R}^d} |f| = m(E) < \infty$, so f is absolutely integrable. By Exercise 1.6.7 the convolution

$$f * g(x) = \int_{\mathbb{R}^d} 1_E(y) 1_{-E}(x - y)\, dy$$

is well-defined and is a bounded continuous function on $\mathbb{R}^d$.

Let us evaluate it. For fixed x, the condition $x - y \in -E$ is equivalent to $y - x \in E$, i.e. to $y \in E + x$. Hence the integrand is $1_{E \cap (E+x)}(y)$ and

$$f * g(x) = m\big(E \cap (E + x)\big).$$

Conclusion. At the origin, $f * g(0) = m(E) > 0$. Since $f * g$ is continuous, there is $\delta > 0$ such that $f * g(x) > 0$ for all $|x| < \delta$. Fix such an x. Then $m(E \cap (E + x)) > 0$, so in particular $E \cap (E + x)$ is non-empty: there is a point z with $z \in E$ and $z \in E + x$, i.e. $z = w + x$ for some $w \in E$. Therefore

$$x = z - w \in E - E.$$

Thus $B(0,\delta) \subseteq E - E$, and $E - E$ contains an open neighbourhood of the origin. (For the original unreduced set this gives $B(0,\delta) \subseteq E' - E' \subseteq E - E$ as well.) $\square$

Problem 1.6.9. A homomorphism $f : \mathbb{R}^d \to \mathbb{C}$ is a map with the property that $f(x+y) = f(x)+f(y)$ for all $x, y \in \mathbb{R}^d$.

- (i) Show that all measurable homomorphisms are continuous. (Hint: for any disk D centered at the origin in the complex plane, show that $f^{-1}(z + D)$ has positive measure for at least one $z \in \mathbb{C}$, and then use the Steinhaus theorem from the previous exercise.)
- (ii) Show that f is a measurable homomorphism if and only if it takes the form $f(x_1, \ldots, x_d) = x_1 z_1 + \cdots + x_d z_d$ for all $x_1, \ldots, x_d \in \mathbb{R}$ and some complex coefficients $z_1, \ldots, z_d$. (Hint: first establish this for rational $x_1, \ldots, x_d$, and then use the previous part of this exercise.)

- (iii) (For readers familiar with Zorn's lemma, see §2.4 of *An epsilon of room, Vol. I*) Show that there exist homomorphisms $f : \mathbb{R}^d \to \mathbb{C}$ which are not of the form in the previous exercise. (Hint: view $\mathbb{R}^d$ (or $\mathbb{C}$) as a vector space over the rationals $\mathbb{Q}$, and use the fact (from Zorn's lemma) that every vector space - even an infinite-dimensional one - has at least one basis.) This gives an alternate construction of a non-measurable set to that given in previous notes.

Solution. Preliminaries on homomorphisms. Let f be any homomorphism. Setting $x = y = 0$ gives $f(0) = 2f(0)$, so $f(0) = 0$; then $0 = f(x + (-x)) = f(x) + f(-x)$ gives $f(-x) = -f(x)$. An easy induction gives $f(nx) = nf(x)$ for all positive integers n, and combined with $f(-x) = -f(x)$ we get $f(nx) = nf(x)$ for all $n \in \mathbb{Z}$. If $q = m/n$ with $m \in \mathbb{Z}$, $n \geq 1$, then $nf(qx) = f(nqx) = f(mx) = mf(x)$, so

$$f(qx) = qf(x) \qquad \text{for all } q \in \mathbb{Q},\ x \in \mathbb{R}^d.$$

(i) Suppose f is a measurable homomorphism, i.e. $f^{-1}(U)$ is Lebesgue measurable for every open $U \subseteq \mathbb{C}$.
First observe that it suffices to prove continuity at the origin. Indeed, for any x_0 and x we have $f(x) - f(x_0) = f(x - x_0)$, so if $f(h) \to f(0) = 0$ as $h \to 0$ then $f(x) \to f(x_0)$ as $x \to x_0$.
Fix $\varepsilon > 0$ and let $D := \{w \in \mathbb{C} : |w| < \varepsilon\}$. Choose a countable set $S \subseteq \mathbb{C}$ whose translates cover the plane, e.g. $S := \{\frac{\varepsilon}{2}(j + ik) : j, k \in \mathbb{Z}\}$; then every $w \in \mathbb{C}$ lies within distance $\frac{\varepsilon}{2}\sqrt{2} < \varepsilon$ of some point of S, so $\mathbb{C} = \bigcup_{z \in S}(z + D)$. Consequently

$$\mathbb{R}^d = \bigcup_{z \in S} f^{-1}(z + D),$$

a countable union of measurable sets (each $z + D$ is open, and f is measurable). If every $f^{-1}(z + D)$ were null, countable subadditivity would force $m(\mathbb{R}^d) = 0$, which is false. Hence there is $z \in S$ with

$$m\big(A\big) > 0, \qquad A := f^{-1}(z + D).$$

By the Steinhaus theorem (Exercise 1.6.8), $A - A$ contains a ball $B(0, \delta)$ for some $\delta > 0$. Let $|x| < \delta$ and write $x = a - a'$ with $a, a' \in A$. Then

$$f(x) = f(a) - f(a') \in (z + D) - (z + D) = D - D,$$

so $|f(x)| < 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary (apply the above with $\varepsilon/2$ in place of ε to get $|f(x)| < \varepsilon$ for $|x| < \delta$), f is continuous at the origin, and hence continuous everywhere.
(ii) If $f(x_1, \ldots, x_d) = x_1 z_1 + \cdots + x_d z_d$, then f is clearly additive, and it is continuous, hence measurable (preimages of open sets are open, so Borel, so Lebesgue measurable).
Conversely, let f be a measurable homomorphism, and set $z_j := f(e_j)$ where $e_1, \ldots, e_d$ is the standard basis of $\mathbb{R}^d$. For rationals $q_1, \ldots, q_d$, additivity and the rational homogeneity established above give

$$f(q_1, \ldots, q_d) = f\Big(\sum_{j=1}^{d} q_j e_j\Big) = \sum_{j=1}^{d} f(q_j e_j) = \sum_{j=1}^{d} q_j z_j.$$

Thus f agrees with the map $g(x_1, \ldots, x_d) := \sum_j x_j z_j$ on $\mathbb{Q}^d$. By part (i), f is continuous, and g is manifestly continuous; two continuous functions agreeing on the dense set $\mathbb{Q}^d$ agree on all of $\mathbb{R}^d$. Hence $f = g$, as required.
(iii) Regard $\mathbb{R}^d$ as a vector space over the field $\mathbb{Q}$. By Zorn's lemma this vector space has a basis B (a Hamel basis): every $x \in \mathbb{R}^d$ has a unique representation $x = \sum_{b \in B} q_b(x)\, b$ with $q_b(x) \in \mathbb{Q}$ and only finitely many $q_b(x)$ non-zero. Fix any $b_0 \in B$ (note $B \neq \emptyset$ since $\mathbb{R}^d \neq \{0\}$) and define

$$f(x) := q_{b_0}(x) \in \mathbb{Q} \subseteq \mathbb{C}.$$

Uniqueness of the representation makes each coordinate functional q_b additive (indeed $\mathbb{Q}$-linear), so f is a homomorphism from $\mathbb{R}^d$ to $\mathbb{C}$.

Now f is not of the form $x \mapsto \sum_j x_j z_j$. Indeed, the range of f is contained in $\mathbb{Q}$, hence is countable; on the other hand, if f had the stated form with some $z_j \neq 0$, then $t \mapsto f(te_j) = tz_j$ would already take uncountably many values as t ranges over $\mathbb{R}$, and if all $z_j = 0$ then $f \equiv 0$, contradicting $f(b_0) = 1$. So no such representation exists.
By part (ii), such an f cannot be measurable: if it were, it would have the linear form just excluded. By the definition of measurability, this means there is some open set $U \subseteq \mathbb{C}$ whose preimage $f^{-1}(U)$ is not Lebesgue measurable, giving an alternate construction of a non-measurable set. □

Problem 1.6.10. Show that any open subset U of $\mathbb{R}$ can be written as the union of at most countably many disjoint non-empty open intervals, whose endpoints lie outside of U. (Hint: first show that every x in U is contained in a maximal open subinterval (a, b) of U, and that these maximal open subintervals are disjoint, with each such interval containing at least one rational number.)

Solution. We may assume $U \neq \emptyset$ (otherwise take the empty family).
Construction of the maximal intervals. For each $x \in U$ set

$$a_x := \inf\{a \in [-\infty, x) : (a, x] \subseteq U\}, \qquad b_x := \sup\{b \in (x, +\infty] : [x, b) \subseteq U\},$$

and $I_x := (a_x, b_x)$, an open interval (possibly unbounded) containing x. The two sets over which we take the infimum and supremum are non-empty: U is open, so $(x - \varepsilon, x + \varepsilon) \subseteq U$ for some $\varepsilon > 0$, whence $a_x \leq x - \varepsilon < x < x + \varepsilon \leq b_x$.
$I_x \subseteq U$. Let $y \in (a_x, x]$. By definition of the infimum there is $a < y$ with $(a, x] \subseteq U$, and $y \in (a, x]$, so $y \in U$. Symmetrically every $y \in [x, b_x)$ lies in U. Hence $I_x \subseteq U$.
Maximality. Suppose $J = (\alpha, \beta)$ is any open interval with $x \in J \subseteq U$. For every $a \in (\alpha, x)$ we have $(a, x] \subseteq J \subseteq U$, so $a_x \leq a$; letting $a \downarrow \alpha$ gives $a_x \leq \alpha$. Symmetrically $b_x \geq \beta$. Hence $J \subseteq I_x$: the interval I_x is the largest open interval containing x and contained in U.
The endpoints lie outside U. Suppose, for contradiction, that a_x is finite and $a_x \in U$. As U is open there is $\varepsilon > 0$ with $(a_x - \varepsilon, a_x + \varepsilon) \subseteq U$. Combining this with $(a_x, x] \subseteq I_x \cup \{x\} \subseteq U$ we obtain $(a_x - \varepsilon, x] \subseteq U$, so $a_x \leq a_x - \varepsilon$, absurd. Hence $a_x \notin U$; symmetrically $b_x \notin U$. (If an endpoint is $\pm\infty$ the requirement is vacuous, since $\pm\infty \notin U \subseteq \mathbb{R}$.)
Disjointness. Suppose $x, y \in U$ and $I_x \cap I_y \neq \emptyset$. Then $I_x \cup I_y$ is a union of two overlapping intervals, hence itself an open interval; it contains x and is contained in U, so by maximality $I_x \cup I_y \subseteq I_x$, i.e. $I_y \subseteq I_x$. By symmetry $I_x \subseteq I_y$, so $I_x = I_y$. Thus any two members of the family $\mathcal{I} := \{I_x : x \in U\}$ are either equal or disjoint; listing the distinct members of $\mathcal{I}$ gives a pairwise disjoint family of non-empty open intervals.
Countability. Each $I \in \mathcal{I}$ is a non-empty open interval, hence contains a rational number; choose one, $q(I) \in I \cap \mathbb{Q}$. If $I \neq I'$ are distinct members of $\mathcal{I}$ they are disjoint, so $q(I) \neq q(I')$. Thus $I \mapsto q(I)$ is an injection of $\mathcal{I}$ into $\mathbb{Q}$, and $\mathcal{I}$ is at most countable.
The union. Since $x \in I_x \subseteq U$ for every $x \in U$, we get $U = \bigcup_{x \in U} I_x = \bigcup_{I \in \mathcal{I}} I$.
So U is the union of an at most countable, pairwise disjoint family of non-empty open intervals whose endpoints lie outside U. □

Problem 1.6.11. (Two-sided Hardy-Littlewood maximal inequality) Let $f : \mathbb{R} \to \mathbb{C}$ be an absolutely integrable function, and let $\lambda > 0$. Show that

$$m\Big(\Big\{x \in \mathbb{R} : \sup_{x \in I} \frac{1}{|I|} \int_I |f(t)|\, dt \geq \lambda\Big\}\Big) \leq \frac{2}{\lambda} \int_{\mathbb{R}} |f(t)|\, dt,$$

where the supremum ranges over all intervals I of positive length that contain x.

Solution. Write

$$Mf(x) := \sup_{x \in I} \frac{1}{|I|} \int_I |f(t)|\, dt, \qquad M^+ f(x) := \sup_{h > 0} \frac{1}{h} \int_{[x, x+h]} |f(t)|\, dt,$$

$$M^- f(x) := \sup_{h>0} \frac{1}{h} \int_{[x-h,x]} |f(t)|\, dt.$$

Step 0: the one-sided maximal functions are lower semicontinuous. Let $H(x) := \int_{[-\infty,x]} |f(t)|\, dt$, which is finite and continuous by Exercise 1.6.5. Since a point is null, $\frac{1}{h}\int_{[x,x+h]} |f| = \frac{H(x+h)-H(x)}{h}$, which for each fixed $h > 0$ is a continuous function of x; likewise $\frac{1}{h}\int_{[x-h,x]} |f| = \frac{H(x)-H(x-h)}{h}$. Hence $M^+ f$ and $M^- f$ are suprema of families of continuous functions, so they are lower semicontinuous and the sets $\{M^\pm f > \mu\}$ are open, in particular Lebesgue measurable. (The same remark applies to the sets $\{M^\pm f \geq \mu\} = \bigcap_{n>1/\mu} \{M^\pm f > \mu - \frac{1}{n}\}$ for $\mu > 0$.)
Step 1: the one-sided bounds. Lemma 1.6.16 states that $m(\{M^+ f \geq \mu\}) \leq \frac{1}{\mu}\int_{\mathbb{R}} |f|$ for every $\mu > 0$. For M^-, let $\tilde{f}(x) := f(-x)$, which is measurable and absolutely integrable with $\int_{\mathbb{R}} |\tilde{f}| = \int_{\mathbb{R}} |f|$ by the reflection invariance of the Lebesgue integral. Substituting $t = -s$,

$$\frac{1}{h} \int_{[x-h,x]} |f(t)|\, dt = \frac{1}{h} \int_{[-x,-x+h]} |\tilde{f}(s)|\, ds,$$

so $M^- f(x) = M^+ \tilde{f}(-x)$ and hence $\{M^- f \geq \mu\} = -\{M^+ \tilde{f} \geq \mu\}$. Since Lebesgue measure is invariant under $x \mapsto -x$, Lemma 1.6.16 applied to $\tilde{f}$ gives

$$m(\{M^- f \geq \mu\}) = m(\{M^+ \tilde{f} \geq \mu\}) \leq \frac{1}{\mu} \int_{\mathbb{R}} |f|.$$

Step 2: splitting an interval at x. Fix $\mu > 0$ and suppose $Mf(x) > \mu$. Then there is an interval I of positive length with $x \in I$ and $\int_I |f| > \mu |I|$. Such an I is necessarily bounded: if $|I| = \infty$ then $\mu|I| = \infty$, whereas $\int_I |f| \leq \int_{\mathbb{R}} |f| < \infty$. So let $c < d$ be the (finite) endpoints of I, so that $c \leq x \leq d$ and $|I| = d - c$. Since the single point x is null,

$$\int_{[c,x]} |f| + \int_{[x,d]} |f| = \int_{[c,d]} |f| \geq \int_I |f| > \mu\,(x - c) + \mu\,(d - x).$$

If $c < x < d$, at least one of the two summands on the left must exceed the corresponding summand on the right; taking $h := x - c > 0$ in the first case gives $M^- f(x) > \mu$, and $h := d - x > 0$ in the second case gives $M^+ f(x) > \mu$. If $x = c$ then the first left summand vanishes and we get $\int_{[x,d]} |f| > \mu(d - x)$ with $d - x = |I| > 0$, so $M^+ f(x) > \mu$; symmetrically $x = d$ gives $M^- f(x) > \mu$. In every case

$$\{Mf > \mu\} \subseteq \{M^+ f > \mu\} \cup \{M^- f > \mu\}.$$

Step 3: measurability. The set $\{Mf > \mu\}$ is in fact open. Indeed, with I, c, d as in Step 2, put $I_\varepsilon := (c - \varepsilon, d + \varepsilon)$ for $\varepsilon > 0$. Then $\int_{I_\varepsilon} |f| \geq \int_I |f| > \mu|I|$, and since $\mu|I| < \int_I |f|$ strictly, for ε small enough we still have $\int_{I_\varepsilon} |f| > \mu(|I| + 2\varepsilon) = \mu|I_\varepsilon|$. As I_ε is an open interval of positive length containing x, every $y \in I_\varepsilon$ satisfies $Mf(y) \geq \frac{1}{|I_\varepsilon|}\int_{I_\varepsilon} |f| > \mu$. Thus $\{Mf > \mu\}$ contains a neighbourhood of each of its points, so it is open, hence measurable; and $\{Mf \geq \lambda\} = \bigcap_{n>1/\lambda} \{Mf > \lambda - \frac{1}{n}\}$ is measurable too.
Step 4: conclusion. Let $0 < \mu < \lambda$. By Steps 2 and 1 together with subadditivity,

$$m(\{Mf \geq \lambda\}) \leq m(\{Mf > \mu\}) \leq m(\{M^+ f > \mu\}) + m(\{M^- f > \mu\}) \leq \frac{2}{\mu} \int_{\mathbb{R}} |f(t)|\, dt,$$

using $\{M^\pm f > \mu\} \subseteq \{M^\pm f \geq \mu\}$. Letting $\mu \uparrow \lambda$ gives

$$m(\{Mf \geq \lambda\}) \leq \frac{2}{\lambda} \int_{\mathbb{R}} |f(t)|\, dt,$$

as required. □

Problem 1.6.12. (Rising sun inequality) Let $f : \mathbb{R} \to \mathbb{R}$ be an absolutely integrable function, and let $f^* : \mathbb{R} \to \mathbb{R}$ be the one-sided signed Hardy-Littlewood maximal function

$$f^*(x) := \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} f(t)\, dt.$$

Establish the rising sun inequality

$$\lambda\, m(\{f^*(x) > \lambda\}) \leq \int_{x:f^*(x)>\lambda} f(x)\, dx$$

for all real λ (note here that we permit λ to be zero or negative), and show that this inequality implies Lemma 1.6.16. (Hint: First do the $\lambda = 0$ case, by invoking the rising sun lemma.) See [Ta2009, §2.9] for some further discussion of inequalities of this type, and applications to ergodic theory (and in particular the maximal ergodic theorem).

Solution. Setup. Let $F(x) := \int_{[-\infty,x]} f(t)\, dt$, which is finite for every x (as f is absolutely integrable) and continuous by Exercise 1.6.5. Since a single point is null, $\int_{[x,x+h]} f = F(x+h) - F(x)$ for $h > 0$. Also, by dominated convergence (Theorem 1.4.49),

$$\lim_{x\to-\infty} F(x) = 0, \qquad \lim_{x\to+\infty} F(x) = \int_{\mathbb{R}} f =: L.$$

Fix $\lambda \in \mathbb{R}$ and set $G(x) := F(x) - \lambda x$, a continuous function on $\mathbb{R}$. For $h > 0$ we have $\frac{1}{h}\int_{[x,x+h]} f > \lambda$ if and only if $G(x+h) > G(x)$. Since a supremum exceeds λ exactly when one of the terms does,

$$E_\lambda := \{x : f^*(x) > \lambda\} = \{x \in \mathbb{R} : G(y) > G(x) \text{ for some } y > x\}.$$

This is precisely the "shadow" set of the rising sun lemma, now formed on all of $\mathbb{R}$.
Step 1: E_λ is open, and its components. If $x \in E_\lambda$, pick $y > x$ with $G(y) > G(x)$. By continuity of G at x there is a neighbourhood $V \subseteq (-\infty, y)$ of x on which $G < G(y)$; every $x' \in V$ then satisfies $y > x'$ and $G(y) > G(x')$, so $V \subseteq E_\lambda$. Thus E_λ is open, hence measurable, and by Exercise 1.6.10 we may write

$$E_\lambda = \bigsqcup_n I_n, \qquad I_n = (a_n, b_n)$$

as an at most countable disjoint union of non-empty open intervals (possibly unbounded) whose finite endpoints lie outside E_λ.
Step 2: the rising sun estimate on a component with finite right endpoint. Suppose $I = (a, b)$ is a component with $b < \infty$. We claim $G(x) \leq G(b)$ for every $x \in (a, b)$.
Suppose not, so $G(x) > G(b)$ for some $x \in (a, b)$. Since $b \notin E_\lambda$, we have $G(y) \leq G(b)$ for all $y > b$. Let $A := \{s \in [x, b] : G(s) \geq G(x)\}$. This is a closed bounded set containing x but not b (as $G(b) < G(x)$), so $s^* := \sup A$ belongs to A and lies in $[x, b)$. Now $s^* \in (a, b) \subseteq E_\lambda$, so there is $y > s^*$ with $G(y) > G(s^*) \geq G(x) > G(b)$. Then $y \neq b$, and $y > b$ is impossible since that would force $G(y) \leq G(b)$. Hence $s^* < y < b$, and $G(y) > G(x)$ gives $y \in A$, contradicting $s^* = \sup A$. This proves the claim.
Step 3: components with right endpoint $+\infty$. Suppose $I = (a, \infty)$ is a component and that $G(y) \to \ell$ as $y \to +\infty$, where $\ell \in [-\infty, \infty)$. We claim $G(x) \leq \ell$ for all $x \in (a, \infty)$. Indeed, if $G(x) > \ell$ for some such x, choose c with $\ell < c < G(x)$; then $A := \{y \geq x : G(y) \geq c\}$ is closed, contains x, and is bounded above (since $G(y) < c$ for all large y), so $s^* := \max A$ exists. As $s^* \in (a, \infty) \subseteq E_\lambda$ there is $y > s^*$ with $G(y) > G(s^*) \geq c$, so $y \in A$, contradicting maximality.
Step 4: the case $\lambda > 0$. Here $G(y) = F(y) - \lambda y \to -\infty$ as $y \to +\infty$ (since F is bounded). By Step 3 (with $\ell = -\infty$), no component can have right endpoint $+\infty$: the claim $G(x) \leq -\infty$ is absurd. Also no component can have left endpoint $-\infty$: if $(-\infty, b)$ were a component with b finite, Step 2 would give $G(x) \leq G(b)$ for all $x < b$, contradicting $G(x) = F(x) - \lambda x \to +\infty$ as $x \to -\infty$. Hence every $I_n = (a_n, b_n)$ is bounded, and by Step 2 plus the continuity of G (let $x \downarrow a_n$) we get $G(b_n) \geq G(a_n)$,

i.e.

$$\int_{I_n} f = F(b_n) - F(a_n) \geq \lambda(b_n - a_n) = \lambda\, m(I_n) \geq 0.$$

Summing over n and using $\sum_n \int_{I_n} |f| \leq \int_{\mathbb{R}} |f| < \infty$ (the I_n are disjoint), all series converge absolutely and countable additivity of the integral gives

$$\int_{E_\lambda} f = \sum_n \int_{I_n} f \geq \lambda \sum_n m(I_n) = \lambda\, m(E_\lambda),$$

which is the desired inequality. (In particular $\lambda\, m(E_\lambda) \leq \int_{\mathbb{R}} |f| < \infty$, so $m(E_\lambda) < \infty$ automatically.)
Step 5: the case $\lambda = 0$. Now $G = F$, and both limits $F(-\infty) = 0$, $F(+\infty) = L$ exist and are finite. There are four kinds of components, and in each we show $\int_{I_n} f \geq 0$:

- $I_n = (a_n, b_n)$ bounded: by Step 2, $F(x) \leq F(b_n)$ on (a_n, b_n); letting $x \downarrow a_n$ and using continuity, $F(a_n) \leq F(b_n)$, so $\int_{I_n} f = F(b_n) - F(a_n) \geq 0$.
- $I_n = (a_n, \infty)$ with a_n finite: by Step 3 with $\ell = L$, $F(x) \leq L$ for $x > a_n$; letting $x \downarrow a_n$ gives $F(a_n) \leq L$, so $\int_{I_n} f = L - F(a_n) \geq 0$.
- $I_n = (-\infty, b_n)$ with b_n finite: by Step 2, $F(x) \leq F(b_n)$ for all $x < b_n$; letting $x \to -\infty$ gives $0 \leq F(b_n)$, so $\int_{I_n} f = F(b_n) - 0 \geq 0$.
- $I_n = \mathbb{R}$: by Step 3, $F(x) \leq L$ for all x; letting $x \to -\infty$ gives $0 \leq L = \int_{I_n} f$.

Summing (again the series converges absolutely), $\int_{E_0} f = \sum_n \int_{I_n} f \geq 0 = 0 \cdot m(E_0)$, which is the inequality for $\lambda = 0$.
Step 6: the case $\lambda < 0$. If $m(E_\lambda) = \infty$ then the left-hand side is $-\infty$ while the right-hand side is a real number (bounded in absolute value by $\int_{\mathbb{R}} |f|$), so the inequality is trivially true. If $m(E_\lambda) < \infty$, then every component has finite length, hence finite endpoints, and Step 2 applies exactly as in Step 4: $G(b_n) \geq G(a_n)$, i.e. $\int_{I_n} f - \lambda m(I_n) \geq 0$. Summing the (non-negative) terms and using $\sum_n m(I_n) = m(E_\lambda) < \infty$ together with absolute convergence of $\sum_n \int_{I_n} f$, we get $\int_{E_\lambda} f \geq \lambda\, m(E_\lambda)$.
This establishes the rising sun inequality for all real λ.
Step 7: deduction of Lemma 1.6.16. Let $f : \mathbb{R} \to \mathbb{C}$ be absolutely integrable and $\lambda > 0$. Apply the rising sun inequality to the real-valued absolutely integrable function $g := |f|$, whose one-sided signed maximal function is

$$g^*(x) = \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} |f(t)|\, dt.$$

For any $0 < \mu < \lambda$, the inequality (Step 4) gives

$$\mu\, m(\{g^* > \mu\}) \leq \int_{\{g^* > \mu\}} |f| \leq \int_{\mathbb{R}} |f(t)|\, dt,$$

so $m(\{g^* > \mu\}) \leq \frac{1}{\mu} \int_{\mathbb{R}} |f|$. Since $\{g^* \geq \lambda\} \subseteq \{g^* > \mu\}$, monotonicity gives $m(\{g^* \geq \lambda\}) \leq \frac{1}{\mu} \int_{\mathbb{R}} |f|$ for every $\mu < \lambda$; letting $\mu \uparrow \lambda$ yields

$$m\Big(\Big\{x \in \mathbb{R} : \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} |f(t)|\, dt \geq \lambda\Big\}\Big) \leq \frac{1}{\lambda} \int_{\mathbb{R}} |f(t)|\, dt,$$

which is exactly Lemma 1.6.16. □

Problem 1.6.13. Show that the left and right-hand sides in Lemma 1.6.16 are in fact equal. (Hint: one may first wish to try this in the case when f has compact support, in which case one can apply the rising sun lemma to a sufficiently large interval containing the support of f.)
[Erratum: in the published errata for the text, "Lemma 1.6.16" here should read "Exercise 1.6.12", and the hypothesis $\lambda > 0$ should be added. So the exercise to be solved is: for absolutely integrable $f : \mathbb{R} \to \mathbb{R}$ and $\lambda > 0$, show that

$$\lambda\, m(\{f^*(x) > \lambda\}) = \int_{x : f^*(x) > \lambda} f(x)\, dx.$$

]

Solution. Why the correction is needed. As literally printed, the assertion is false. Take $f = 1_{[0,1]}$ and $\lambda = 2$. Since $|f| \le 1$, every average $\frac{1}{h}\int_{[x,x+h]} |f|$ is at most 1, so the set on the left-hand side of Lemma 1.6.16 is empty and the left-hand side is 0; the right-hand side is $\frac{1}{2}\int_{\mathbb{R}} |f| = \frac{1}{2}$. So we prove the corrected statement, which is the assertion that the rising sun inequality of Exercise 1.6.12 is an identity when $\lambda > 0$.
Setup. Let $f : \mathbb{R} \to \mathbb{R}$ be absolutely integrable and $\lambda > 0$. As in Exercise 1.6.12, put

$$F(x) := \int_{[-\infty,x]} f(t)\,dt, \qquad G(x) := F(x) - \lambda x,$$

so F is finite and continuous by Exercise 1.6.5, hence so is G, and

$$E_\lambda := \{x : f^*(x) > \lambda\} = \{x \in \mathbb{R} : G(y) > G(x) \text{ for some } y > x\}.$$

By Step 1 of Exercise 1.6.12 this set is open, and by Exercise 1.6.10 we may write $E_\lambda = \bigsqcup_n I_n$ with $I_n = (a_n, b_n)$ an at most countable family of disjoint non-empty open intervals whose finite endpoints lie outside E_λ. By Step 4 of Exercise 1.6.12 (using $\lambda > 0$), every I_n is a bounded interval, so a_n and b_n are both finite and both lie outside E_λ; moreover $m(E_\lambda) \le \frac{1}{\lambda}\int_{\mathbb{R}} |f| < \infty$.
The key point: equality on each component. Fix n.
One inequality is the rising sun estimate already proved: by Step 2 of Exercise 1.6.12, $G(x) \le G(b_n)$ for every $x \in (a_n, b_n)$, and letting $x \downarrow a_n$ and using continuity of G gives

$$G(a_n) \le G(b_n).$$

For the reverse inequality we use that the *left* endpoint also lies outside E_λ. Since $a_n \notin E_\lambda$, there is no $y > a_n$ with $G(y) > G(a_n)$; that is,

$$G(y) \le G(a_n) \qquad \text{for all } y > a_n.$$

Taking $y = b_n$ gives $G(b_n) \le G(a_n)$.
Therefore $G(a_n) = G(b_n)$, i.e. $F(b_n) - F(a_n) = \lambda(b_n - a_n)$, i.e.

$$\int_{I_n} f(x)\,dx = \lambda\, m(I_n).$$

(This is the precise sense in which the rising sun lemma "loses nothing": on each shadow interval the function G returns exactly to its starting height.)
Summation. The intervals I_n are disjoint, and $\sum_n \int_{I_n} |f| \le \int_{\mathbb{R}} |f| < \infty$, so $\sum_n \int_{I_n} f$ converges absolutely and equals $\int_{E_\lambda} f$ by countable additivity of the integral. Likewise $\sum_n m(I_n) = m(E_\lambda) < \infty$. Summing the identities above,

$$\int_{x : f^*(x) > \lambda} f(x)\,dx = \sum_n \int_{I_n} f = \lambda \sum_n m(I_n) = \lambda\, m(\{f^*(x) > \lambda\}),$$

which is the desired equality.
Remark (the non-strict version, and what survives of the printed claim). The same identity holds with $>$ replaced by $\ge$: since $\{f^* \ge \lambda\} = \bigcap_{0<\mu<\lambda} \{f^* > \mu\}$ is a decreasing intersection of sets of finite measure, downward monotone convergence for the measure and for the integral of the absolutely integrable f lets us pass to the limit $\mu \uparrow \lambda$ in $\mu\, m(\{f^* > \mu\}) = \int_{\{f^* > \mu\}} f$, giving $\lambda\, m(\{f^* \ge \lambda\}) = \int_{\{f^* \ge \lambda\}} f$. Applying this to $|f|$ in place of f shows that the correct sharp form of Lemma 1.6.16 is

$$m\Big(\Big\{x : \sup_{h>0} \tfrac{1}{h}\int_{[x,x+h]} |f| \ \ge \lambda\Big\}\Big) = \frac{1}{\lambda}\int_E |f(t)|\,dt,$$

where E is the set appearing on the left; the printed right-hand side $\frac{1}{\lambda}\int_{\mathbb{R}} |f|$ is larger in general, and equal precisely when f vanishes almost everywhere outside E. □

Problem 1.6.14. Call a function $f : \mathbb{R}^d \to \mathbb{C}$ locally integrable if, for every $x \in \mathbb{R}^d$, there exists an open neighbourhood of x on which f is absolutely integrable.

- (i) Show that f is locally integrable if and only if $\int_{B(0,r)} |f(x)|\,dx < \infty$ for all $r > 0$.
- (ii) Show that Theorem 1.6.19 implies a generalisation of itself in which the condition of absolute integrability of f is weakened to local integrability.

Solution. (i) Suppose first that $\int_{B(0,r)} |f| < \infty$ for every $r > 0$; here (and throughout) absolute integrability of f on a set includes the assertion that f restricted to that set is measurable. Given $x \in \mathbb{R}^d$, take $r := |x|+1$; then $B(0,r)$ is an open neighbourhood of x on which f is absolutely integrable. So f is locally integrable.
Conversely, suppose f is locally integrable, and fix $r > 0$. For each x in the closed ball $\overline{B}(0,r)$ choose an open neighbourhood U_x of x with $\int_{U_x} |f| < \infty$, and choose $\rho_x > 0$ with $B(x,\rho_x) \subseteq U_x$. The balls $\{B(x,\rho_x)\}_{x\in\overline{B}(0,r)}$ form an open cover of the set $\overline{B}(0,r)$, which is closed and bounded, hence compact by the Heine-Borel theorem. Extract a finite subcover $B(x_1,\rho_{x_1}),\dots,B(x_k,\rho_{x_k})$. Then

$$B(0,r) \subseteq \overline{B}(0,r) \subseteq \bigcup_{j=1}^{k} B(x_j,\rho_{x_j}) \subseteq \bigcup_{j=1}^{k} U_{x_j},$$

so f is measurable on $B(0,r)$ (a finite union of sets on which it is measurable), and by monotonicity and finite subadditivity of the integral of the non-negative function $|f|$,

$$\int_{B(0,r)} |f| \le \sum_{j=1}^{k} \int_{U_{x_j}} |f| < \infty.$$

(Letting r run through the positive integers also shows that f is measurable on all of $\mathbb{R}^d = \bigcup_n B(0,n)$, a countable union.)
(ii) The generalisation is: *if $f : \mathbb{R}^d \to \mathbb{C}$ is locally integrable, then for almost every $x \in \mathbb{R}^d$ one has*

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\,dy = 0$$

and consequently $\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\,dy = f(x)$.
To deduce this from Theorem 1.6.19, for each integer $n \ge 1$ set

$$f_n := f\,1_{B(0,n)}.$$

By part (i), f_n is measurable and $\int_{\mathbb{R}^d} |f_n| = \int_{B(0,n)} |f| < \infty$, so f_n is absolutely integrable. Theorem 1.6.19 applies to f_n: there is a null set $N_n \subseteq \mathbb{R}^d$ such that every $x \notin N_n$ is a Lebesgue point of f_n, i.e.

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f_n(y) - f_n(x)|\,dy = 0.$$

Let $N := \bigcup_{n\ge 1} N_n$; as a countable union of null sets, N is null. Let $x \notin N$ and pick an integer $n > |x|$. Then $f_n(x) = f(x)$, and for every radius r with $0 < r < n - |x|$ we have $B(x,r) \subseteq B(0,n)$, so $f_n = f$ on $B(x,r)$ and therefore

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\,dy = \frac{1}{m(B(x,r))} \int_{B(x,r)} |f_n(y) - f_n(x)|\,dy.$$

Since $x \notin N_n$, the right-hand side tends to 0 as $r \to 0$; hence so does the left-hand side. Thus every $x \notin N$ is a Lebesgue point of f, and the first conclusion holds almost everywhere.
The second conclusion follows from the first exactly as in the text: by the triangle inequality,

$$\Big|\frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\,dy - f(x)\Big| = \Big|\frac{1}{m(B(x,r))} \int_{B(x,r)} \big(f(y)-f(x)\big)\,dy\Big| \le \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)-f(x)|\,dy,$$

which tends to 0 for almost every x. (Note all the integrals appearing here are finite: for r small, $B(x,r)$ is contained in some $B(0,n)$, on which f is absolutely integrable by part (i).) $\square$

Problem 1.6.15. For each $h > 0$, let E_h be a subset of $B(0,h)$ with the property that $m(E_h) \geq c\, m(B(0,h))$ for some $c > 0$ independent of h. Show that if $f : \mathbb{R}^d \to \mathbb{C}$ is locally integrable, and x is a Lebesgue point of f, then

$$\lim_{h\to 0} \frac{1}{m(E_h)} \int_{x+E_h} f(y)\, dy = f(x).$$

Conclude that Theorem 1.6.19 implies Theorem 1.6.12.

Solution. Throughout, the sets E_h are understood to be Lebesgue measurable (otherwise $m(E_h)$ is not defined), and x is a Lebesgue point of f, i.e.

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0.$$

First note that the quantities involved are all finite and the denominators are nonzero. Indeed $m(B(0,h)) = \omega_d h^d$ with $\omega_d := m(B(0,1)) \in (0, +\infty)$ (the unit ball contains a small box and is contained in a large one), so $m(E_h) \geq c\,\omega_d h^d > 0$. Also, by translation invariance of Lebesgue measure, $m(B(x,h)) = m(B(0,h))$. Since f is locally integrable, $\int_{B(x,h)} |f(y)|\, dy < \infty$ for every $h > 0$ by Exercise 1.6.14(i) (applied after translating), so f is absolutely integrable on $x + E_h \subset B(x,h)$ and all the integrals below are well defined and finite.

Now fix $h > 0$. Since $x + E_h$ has measure $m(E_h)$ (translation invariance), we may write $f(x) = \frac{1}{m(E_h)} \int_{x+E_h} f(x)\, dy$, and hence by the triangle inequality for integrals

$$\left| \frac{1}{m(E_h)} \int_{x+E_h} f(y)\, dy - f(x) \right| = \left| \frac{1}{m(E_h)} \int_{x+E_h} \big(f(y) - f(x)\big) dy \right| \leq \frac{1}{m(E_h)} \int_{x+E_h} |f(y) - f(x)|\, dy.$$

The integrand is unsigned, and $x + E_h \subset x + B(0,h) = B(x,h)$, so by monotonicity of the unsigned integral we may enlarge the domain of integration:

$$\frac{1}{m(E_h)} \int_{x+E_h} |f(y) - f(x)|\, dy \leq \frac{1}{m(E_h)} \int_{B(x,h)} |f(y) - f(x)|\, dy.$$

Finally, using the hypothesis $m(E_h) \geq c\, m(B(0,h)) = c\, m(B(x,h))$ we obtain

$$\left| \frac{1}{m(E_h)} \int_{x+E_h} f(y)\, dy - f(x) \right| \leq \frac{1}{c} \cdot \frac{1}{m(B(x,h))} \int_{B(x,h)} |f(y) - f(x)|\, dy.$$

Since c is independent of h and x is a Lebesgue point of f, the right-hand side tends to 0 as $h \to 0$. This proves the claimed limit.

Now we deduce Theorem 1.6.12 from Theorem 1.6.19. Work in dimension $d = 1$ and let $f : \mathbb{R} \to \mathbb{C}$ be absolutely integrable (hence locally integrable). By Theorem 1.6.19, almost every $x \in \mathbb{R}$ is a Lebesgue point of f; fix such an x.

For the right averages, take $E_h := [0,h)$ for $h > 0$. This is a measurable subset of $B(0,h) = (-h,h)$, and $m(E_h) = h = \frac{1}{2} m(B(0,h))$, so the hypothesis holds with $c = 1/2$. Since $x + E_h = [x, x+h)$ differs from $[x, x+h]$ by the single point $x+h$, which is null, we have $\int_{x+E_h} f = \int_{[x,x+h]} f$, and the result just proved gives

$$\lim_{h\to 0^+} \frac{1}{h} \int_{[x,x+h]} f(t)\, dt = f(x),$$

which is (1.20).

For the left averages, take instead $E_h := (-h, 0]$, again a measurable subset of $(-h,h)$ with $m(E_h) = h = \frac{1}{2} m(B(0,h))$, so $c = 1/2$ works. Now $x + E_h = (x-h, x]$, which differs from $[x-h, x]$ by the null set $\{x-h\}$, so the same result gives

$$\lim_{h\to 0^+} \frac{1}{h} \int_{[x-h,x]} f(t)\, dt = f(x),$$

which is (1.21). Since both hold at every Lebesgue point, and almost every point is a Lebesgue point, Theorem 1.6.12 follows.

Problem 1.6.16. Show that Theorem 1.6.19 holds whenever f is continuous.

Solution. Let $f : \mathbb{R}^d \to \mathbb{C}$ be continuous and absolutely integrable. We show the stronger statement that every point $x \in \mathbb{R}^d$ is a Lebesgue point of f; the exceptional set is then empty, which is certainly null.
Fix $x \in \mathbb{R}^d$ and let $\varepsilon > 0$. By continuity of f at x, there is $\delta > 0$ such that $|f(y) - f(x)| \le \varepsilon$ whenever $|y - x| < \delta$.
Let $0 < r < \delta$. The function $y \mapsto |f(y) - f(x)|$ is continuous, hence Lebesgue measurable, and it is bounded by ε on $B(x, r)$. Also $m(B(x, r)) = \omega_d r^d \in (0, +\infty)$, where $\omega_d := m(B(0, 1))$ is finite and positive since the unit ball contains a nondegenerate box and is contained in a box. Hence by monotonicity of the unsigned integral,

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy \le \frac{1}{m(B(x,r))} \int_{B(x,r)} \varepsilon\, dy = \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this shows

$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0,$$

which is (1.26); so x is a Lebesgue point of f.
The second conclusion of Theorem 1.6.19 follows from the first exactly as in the text: by the triangle inequality,

$$\left| \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dy - f(x) \right| = \left| \frac{1}{m(B(x,r))} \int_{B(x,r)} \big(f(y) - f(x)\big) dy \right| \le \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy,$$

which tends to 0 as $r \to 0$. Thus both conclusions of Theorem 1.6.19 hold at every point when f is continuous.
(Note that absolute integrability of f was not needed for this argument; local integrability, which continuity already supplies since a continuous function is bounded on each ball, suffices.)

Problem 1.6.17. Use the density argument to show that Theorem 1.6.20 implies Theorem 1.6.19.

Solution. Let $f : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable. For an absolutely integrable $F : \mathbb{R}^d \to \mathbb{C}$ define the Hardy-Littlewood maximal function

$$MF(x) := \sup_{r > 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |F(y)|\, dy \in [0, +\infty],$$

and define the fluctuation

$$\Lambda F(x) := \limsup_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |F(y) - F(x)|\, dy \in [0, +\infty].$$

Theorem 1.6.20 asserts that $m(\{MF \ge \lambda\}) \le \frac{C_d}{\lambda} \|F\|_{L^1}$ for all $\lambda > 0$, where $\|F\|_{L^1} := \int_{\mathbb{R}^d} |F(t)|\, dt$. Conclusion (1.26) of Theorem 1.6.19 is exactly the assertion that $\Lambda f(x) = 0$ for almost every x, and, as shown in the text just before Exercise 1.6.15, the second conclusion of Theorem 1.6.19 follows from the first by the triangle inequality. So it suffices to prove $\Lambda f = 0$ almost everywhere.
Step 1: a sublinearity estimate. Let $g : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable and put $h := f - g$. For every x and every $r > 0$, the pointwise triangle inequality $|f(y) - f(x)| \le |g(y) - g(x)| + |h(y) - h(x)|$ and the triangle inequality for integrals give

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy \le \frac{1}{m(B(x,r))} \int_{B(x,r)} |g(y) - g(x)|\, dy + \frac{1}{m(B(x,r))} \int_{B(x,r)} |h(y) - h(x)|\, dy,$$

and moreover

$$\frac{1}{m(B(x,r))}\int_{B(x,r)}|h(y)-h(x)|\,dy \leq \frac{1}{m(B(x,r))}\int_{B(x,r)}|h(y)|\,dy + |h(x)| \leq Mh(x)+|h(x)|.$$

Taking $\limsup_{r\to 0}$ and using subadditivity of the lim sup we obtain the pointwise bound

$$\Lambda f(x) \leq \Lambda g(x) + Mh(x) + |h(x)| \qquad \text{for all } x \in \mathbb{R}^d.$$

Step 2: the dense subclass. Let $\varepsilon > 0$. By Littlewood's second principle (Theorem 1.3.20(iii)) there is a continuous, compactly supported $g : \mathbb{R}^d \to \mathbb{C}$ with $\|f-g\|_{L^1} \leq \varepsilon$. Such a g is continuous, so by Exercise 1.6.16 every point is a Lebesgue point of g, i.e. $\Lambda g \equiv 0$. With $h := f - g$ (absolutely integrable, with $\|h\|_{L^1} \leq \varepsilon$), Step 1 becomes

$$\Lambda f(x) \leq Mh(x) + |h(x)| \qquad \text{for all } x \in \mathbb{R}^d.$$

Step 3: the quantitative estimate. Fix $\lambda > 0$. By the previous display,

$$\{x : \Lambda f(x) > \lambda\} \subset \{x : Mh(x) \geq \lambda/2\} \cup \{x : |h(x)| \geq \lambda/2\}.$$

By Theorem 1.6.20, $m(\{Mh \geq \lambda/2\}) \leq \frac{2C_d}{\lambda}\|h\|_{L^1} \leq \frac{2C_d}{\lambda}\varepsilon$. By Markov's inequality (Lemma 1.3.15) applied to the unsigned measurable function $|h|$, $m(\{|h| \geq \lambda/2\}) \leq \frac{2}{\lambda}\|h\|_{L^1} \leq \frac{2}{\lambda}\varepsilon$. By monotonicity and finite subadditivity of Lebesgue outer measure,

$$m^*(\{x : \Lambda f(x) > \lambda\}) \leq \frac{2(C_d+1)}{\lambda}\,\varepsilon.$$

(We use outer measure here because we have not yet checked that the set on the left is measurable; this will follow in a moment.) The left-hand side does not depend on ε, and $\varepsilon > 0$ was arbitrary, so

$$m^*(\{x : \Lambda f(x) > \lambda\}) = 0,$$

and in particular $\{\Lambda f > \lambda\}$ is a null set, hence Lebesgue measurable with measure zero.
Step 4: conclusion. Since

$$\{x : \Lambda f(x) > 0\} = \bigcup_{n=1}^{\infty}\{x : \Lambda f(x) > 1/n\}$$

is a countable union of null sets, it is null by countable subadditivity. Hence $\Lambda f(x) = 0$ for almost every x; that is, for almost every x,

$$\lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)}|f(y)-f(x)|\,dy = 0,$$

(the lim sup of a nonnegative quantity being zero forces the limit to exist and equal zero). This is (1.26), and as noted the second conclusion of Theorem 1.6.19 follows. Combined with Exercise 1.6.14(ii), the same conclusion extends to merely locally integrable f.

Problem 1.6.18. Technically speaking, the above algorithmic argument was not phrased in the standard language of formal mathematical deduction, because in that language, any mathematical object (such as the natural number m) can only be defined once, and not redefined multiple times as is done in most algorithms. Rewrite the above argument in a way that avoids redefining any variable. (Hint: introduce a "time" variable t, and recursively construct families $B'_{1,t}, \ldots, B'_{m_t,t}$ of balls that represent the outcome of the above algorithm after t iterations (or t_* iterations, if the algorithm halted at some previous time $t_* < t$). For this particular algorithm, there are also more ad hoc approaches that exploit the relatively simple nature of the algorithm to allow for a less notationally complicated construction.) More generally, it is possible to use this time parameter

trick to convert any construction involving a provably terminating algorithm into a construction that does not redefine any variable. (It is however dangerous to work with any algorithm that has an infinite run time, unless one has a suitably strong convergence result for the algorithm that allows one to take limits, either in the classical sense or in the more general sense of jumping to limit ordinals; in the latter case, one needs to use transfinite induction in order to ensure that the use of such algorithms is rigorous; see §2.4 of An epsilon of room, Vol. I.)

Solution. We give the ad hoc version suggested at the end of the hint: instead of recording the whole family $B'_{1,t}, \dots, B'_{m_t,t}$ at each time t, we record only the single new index chosen at time t. This is enough because the family produced by the algorithm at time t is always an initial segment of one fixed sequence of indices.

Setup. Let $B_1, \dots, B_n$ be open balls in $\mathbb{R}^d$, say $B_i = B(c_i, r_i)$ with $r_i > 0$ (balls of radius 0 are empty and may be discarded at the outset without affecting either the disjointness conclusion or the covering conclusion). Write $[n] := \{1, \dots, n\}$.

The recursion. We define, by recursion on $t \in \{0, 1, 2, \dots\}$, a natural number t_* and a finite sequence of indices $j_1, j_2, \dots, j_{t_*} \in [n]$. Precisely, suppose that for some $t \geq 0$ the distinct indices $j_1, \dots, j_t$ have already been defined (for $t = 0$ this is the empty list, so no assumption is being made). Define

$$A_t := \{i \in [n] : B_i \cap B_{j_s} = \emptyset \text{ for all } 1 \leq s \leq t\}$$

(so $A_0 = [n]$). Two cases:

If $A_t = \emptyset$, we declare $t_* := t$ and the recursion stops: no further j's are defined.

If $A_t \neq \emptyset$, let $\rho_t := \max_{i \in A_t} r_i$ (a maximum over a nonempty finite set, hence attained), and define

$$j_{t+1} := \min\{i \in A_t : r_i = \rho_t\},$$

the least index attaining the maximal radius. (Taking the least such index is a definite prescription; it is what replaces the informal instruction to break the tie arbitrarily, and it also avoids any appeal to the axiom of choice.) The recursion then continues with $t + 1$ in place of t.

Each of the objects A_t, ρ_t, j_{t+1}, t_* is defined exactly once, in terms of previously defined objects only; this is a legitimate definition by recursion, and no variable is ever redefined.

Termination. If j_{t+1} is defined then $j_{t+1} \in A_t$, so $B_{j_{t+1}}$ is disjoint from $B_{j_1}, \dots, B_{j_t}$; since $B_{j_{t+1}}$ is nonempty it is not disjoint from itself, and therefore $j_{t+1} \notin \{j_1, \dots, j_t\}$. Hence $j_1, \dots, j_t$ are always distinct elements of $[n]$, which forces $t \leq n$. Consequently the recursion cannot run past time n: there is a least $t \leq n$ with $A_t = \emptyset$, and t_* is that value. Set

$$m := t_*, \qquad B'_s := B_{j_s} \quad (1 \leq s \leq m),$$

so that $B'_1, \dots, B'_m$ is a subcollection of $B_1, \dots, B_n$. (If one wishes to match the hint's notation exactly, put $m_t := \min(t, t_*)$ and $B'_{s,t} := B_{j_s}$ for $1 \leq s \leq m_t$; the families of the hint are then precisely the initial segments of the single sequence just constructed, and $B'_{s,t} = B'_s$ for all $t \geq t_*$.)

Disjointness. Let $1 \leq s < s' \leq m$. Then $j_{s'} \in A_{s'-1} \subset A_s$ (the sets A_t are decreasing in t), so by the definition of A_s we have $B_{j_{s'}} \cap B_{j_s} = \emptyset$. Thus $B'_1, \dots, B'_m$ are pairwise disjoint.

Two properties of the selection. Fix $i \in [n]$. Since $A_{t_*} = A_m = \emptyset$, we have $i \notin A_m$, i.e. B_i meets at least one of $B'_1, \dots, B'_m$. Let s be the least index with $B_i \cap B'_s \neq \emptyset$. Then B_i is disjoint from $B_{j_1}, \dots, B_{j_{s-1}}$, i.e. $i \in A_{s-1}$. Since j_s was chosen to maximise the radius over A_{s-1}, we conclude

$$r_i \leq r_{j_s}.$$

Also, picking any $z \in B_i \cap B'_s$, the triangle inequality gives $|c_i - c_{j_s}| \leq |c_i - z| + |z - c_{j_s}| < r_i + r_{j_s} \leq 2r_{j_s}$, so

$$c_i \in 2B'_s.$$

The covering. With s as above, let $y \in B_i$. Then, using $z \in B_i \cap B'_s$ again,

$$|y - c_{j_s}| \leq |y - c_i| + |c_i - z| + |z - c_{j_s}| < r_i + r_i + r_{j_s} \leq 3r_{j_s},$$

so $B_i \subset 3B'_s$. As $i \in [n]$ was arbitrary,

$$\bigcup_{i=1}^{n} B_i \subset \bigcup_{j=1}^{m} 3B'_j,$$

which is (1.27). Finite subadditivity and $m(3B'_j) = 3^d m(B'_j)$ then give $m(\bigcup_{i=1}^n B_i) \le 3^d \sum_{j=1}^m m(B'_j)$, completing the proof of Lemma 1.6.22 in a form in which no variable is ever redefined.
(The general recipe implicit in the hint is the same: given a provably terminating algorithm, index the state of the algorithm by a time parameter t, define the state at time $t+1$ from the state at time t by recursion, define t_* to be the first time the halting condition holds, and take the output to be the state at time t_*. Termination of the algorithm is exactly the statement that such a t_* exists. Here the state at time t was the tuple $(j_1, \dots, j_t)$, which is determined by its entries, so only the entries had to be named.)

Problem 1.6.19. Improve the constant 3^d in the Hardy-Littlewood maximal inequality to 2^d. (Hint: observe that with the construction used to prove the Vitali covering lemma, the centres of the balls B_i are contained in $\bigcup_{j=1}^m 2B'_j$ and not just in $\bigcup_{j=1}^m 3B'_j$. To exploit this observation one may need to first create an epsilon of room, as the centers are not by themselves sufficient to cover the required set.)

Solution. We show that Theorem 1.6.20 holds with $C_d = 2^d$.
The refined covering lemma. We first record the extra information contained in the proof of Lemma 1.6.22, as isolated in the solution to Exercise 1.6.18. Let $B_1, \dots, B_n$ be open balls, $B_i = B(c_i, r_i)$ with $r_i > 0$, and let $B'_1, \dots, B'_m$ (with $B'_s = B(c'_s, r'_s)$) be the disjoint subcollection produced by the greedy selection. Then:
For each i there is an index $s = s(i)$ such that $r_i \le r'_s$ and $|c_i - c'_s| < r_i + r'_s \le 2r'_s$.
Indeed, take s to be the first selected ball that B_i meets; as shown in the proof of Lemma 1.6.22, B_i was still available when B'_s was chosen, so $r_i \le r'_s$ by maximality, and if $z \in B_i \cap B'_s$ then $|c_i - c'_s| \le |c_i - z| + |z - c'_s| < r_i + r'_s$.
Consequently, for any $\varepsilon > 0$ we get the stronger conclusion

$$B(c_i, \varepsilon r_i) \subset B\big(c'_s, 2r'_s + \varepsilon r_i\big) \subset (2+\varepsilon)B'_s,$$

using $r_i \le r'_s$ in the last inclusion. In words: after shrinking each ball by the factor ε about its own centre, the shrunken balls are covered by the $(2+\varepsilon)$-dilates of the selected balls. This shrinking is the epsilon of room of the hint: the centres alone form a finite set and cannot cover anything of positive measure, but the tiny balls around the centres can.
Proof of the maximal inequality with constant 2^d. Fix an absolutely integrable $f : \mathbb{R}^d \to \mathbb{C}$ and $\lambda > 0$, and write $Mf(x) := \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)|\, dy$ and $\|f\|_{L^1} := \int_{\mathbb{R}^d} |f(t)|\, dt$. As in the text, it suffices to prove

$$m(\{x \in \mathbb{R}^d : Mf(x) > \lambda\}) \le \frac{2^d}{\lambda} \|f\|_{L^1},$$

since the non-strict version follows by applying this with λ replaced by $\lambda' < \lambda$ (note $\{Mf \ge \lambda\} \subset \{Mf > \lambda'\}$) and then letting $\lambda' \to \lambda^-$.
The set $\{Mf > \lambda\}$ is open, hence Lebesgue measurable. Indeed, if $Mf(x) > \lambda$, choose $r > 0$ with $\int_{B(x,r)} |f| > \lambda\, m(B(x,r)) = \lambda \omega_d r^d$ (where $\omega_d := m(B(0,1))$), and then, by continuity of $\rho \mapsto \rho^d$, choose $\rho > r$ for which $\lambda \omega_d \rho^d < \int_{B(x,r)} |f|$ still holds; for any y with $|y - x| < \rho - r$ we have $B(x,r) \subset B(y,\rho)$ and hence $\int_{B(y,\rho)} |f| \ge \int_{B(x,r)} |f| > \lambda \omega_d \rho^d = \lambda\, m(B(y,\rho))$, so $Mf(y) > \lambda$. Therefore, by inner regularity, it suffices to show

$$m(K) \le \frac{2^d}{\lambda} \|f\|_{L^1}$$

for every compact $K \subset \{Mf > \lambda\}$.

Fix such a K and fix $\varepsilon > 0$. For every $x \in K$ we have $Mf(x) > \lambda$, so there is a radius $r_x > 0$ with

$$\int_{B(x,r_x)} |f(y)|\,dy > \lambda\, m(B(x,r_x)).$$

The shrunken balls $B(x, \varepsilon r_x)$, $x \in K$, are open and cover K (each x lies in its own). By compactness there are $x_1, \dots, x_n \in K$ with

$$K \subset \bigcup_{i=1}^{n} B(x_i, \varepsilon r_i), \qquad r_i := r_{x_i}.$$

Now apply the greedy selection of Lemma 1.6.22 to the full balls $B_i := B(x_i, r_i)$, $i = 1, \dots, n$, obtaining a disjoint subcollection $B'_1, \dots, B'_m$. By the refined covering lemma above, for each i there is $s(i)$ with $B(x_i, \varepsilon r_i) \subset (2+\varepsilon) B'_{s(i)}$, and therefore

$$K \subset \bigcup_{i=1}^{n} B(x_i, \varepsilon r_i) \subset \bigcup_{j=1}^{m} (2+\varepsilon) B'_j.$$

Each B'_j is one of the B_i, so it satisfies $\lambda\, m(B'_j) < \int_{B'_j} |f(y)|\,dy$. Using finite subadditivity, the scaling identity $m(cB) = c^d m(B)$, and the disjointness of the B'_j (so that $\sum_j \int_{B'_j} |f| = \int_{\bigcup_j B'_j} |f| \le \|f\|_{L^1}$), we obtain

$$m(K) \le \sum_{j=1}^{m} m\big((2+\varepsilon)B'_j\big) = (2+\varepsilon)^d \sum_{j=1}^{m} m(B'_j) \le \frac{(2+\varepsilon)^d}{\lambda} \sum_{j=1}^{m} \int_{B'_j} |f(y)|\,dy \le \frac{(2+\varepsilon)^d}{\lambda} \|f\|_{L^1}.$$

The left-hand side does not depend on ε, so letting $\varepsilon \to 0^+$ gives

$$m(K) \le \frac{2^d}{\lambda} \|f\|_{L^1},$$

as required. Taking the supremum over compact $K \subset \{Mf > \lambda\}$ and then removing the strictness in λ as described above yields Theorem 1.6.20 with $C_d = 2^d$.

Problem 1.6.20. (Dyadic maximal inequality) If $f : \mathbb{R}^d \to \mathbb{C}$ is an absolutely integrable function, establish the dyadic Hardy-Littlewood maximal inequality

$$m\left(\left\{x \in \mathbb{R}^d : \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\,dy \ge \lambda\right\}\right) \le \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(t)|\,dt$$

where the supremum ranges over all dyadic cubes Q that contain x. (Hint: the nesting property of dyadic cubes will be useful when it comes to the covering lemma stage of the argument, much as it was in Exercise 1.1.14.)

Solution. Recall from Exercise 1.1.14 that a dyadic cube is a half-open box

$$Q = \left[\frac{i_1}{2^n}, \frac{i_1+1}{2^n}\right) \times \cdots \times \left[\frac{i_d}{2^n}, \frac{i_d+1}{2^n}\right)$$

for integers $n, i_1, \dots, i_d$; its measure is $|Q| = 2^{-dn}$. Write $\mathcal{D}$ for the (countable) collection of all dyadic cubes, $\|f\|_{L^1} := \int_{\mathbb{R}^d} |f(t)|\,dt$, and

$$M_{\mathcal{D}} f(x) := \sup_{Q \in \mathcal{D},\, x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\,dy.$$

This is well defined: for each n the dyadic cubes of sidelength 2^{-n} partition $\mathbb{R}^d$, so every x lies in exactly one of them, and the supremum is over a nonempty countable set.

Step 0: the nesting property. If $Q, R \in \mathcal{D}$ and $Q \cap R \neq \emptyset$, then $Q \subset R$ or $R \subset Q$. To see this, let Q have sidelength 2^{-n} and R sidelength $2^{-n'}$, and suppose without loss of generality $n \geq n'$, so Q is the smaller cube. It suffices to show that Q is contained in some dyadic cube of sidelength $2^{-n'}$: for then that cube meets R, and two dyadic cubes of the same sidelength are either equal or disjoint, so that cube is R, giving $Q \subset R$. Coordinatewise, with $k := n - n' \geq 0$ and $i = q2^k + s$ the division of i by 2^k (so $q \in \mathbb{Z}$, $0 \leq s < 2^k$), we have

$$\frac{q}{2^{n'}} = \frac{q2^k}{2^n} \leq \frac{i}{2^n} \quad \text{and} \quad \frac{i+1}{2^n} = \frac{q2^k + s + 1}{2^n} \leq \frac{q2^k + 2^k}{2^n} = \frac{q+1}{2^{n'}},$$

so $[i2^{-n}, (i+1)2^{-n}) \subset [q2^{-n'}, (q+1)2^{-n'})$. Doing this in each coordinate exhibits the required containing cube.

Step 1: the strict-inequality estimate. Fix $\mu > 0$ and let

$$\mathcal{Q}_\mu := \left\{ Q \in \mathcal{D} : \frac{1}{|Q|} \int_Q |f(y)|\, dy > \mu \right\}, \qquad \Omega_\mu := \bigcup_{Q \in \mathcal{Q}_\mu} Q.$$

Since $\mathcal{D}$ is countable, Ω_μ is a countable union of boxes, hence Lebesgue measurable.

Every $Q \in \mathcal{Q}_\mu$ satisfies

$$|Q| < \frac{1}{\mu} \int_Q |f(y)|\, dy \leq \frac{1}{\mu} \|f\|_{L^1} < \infty,$$

so the cubes in $\mathcal{Q}_\mu$ have uniformly bounded measure. This is the substitute for a covering lemma. Given $Q \in \mathcal{Q}_\mu$, the dyadic cubes containing Q form an increasing chain $Q = Q^{(0)} \subset Q^{(1)} \subset Q^{(2)} \subset \cdots$ with $|Q^{(k)}| = 2^{dk}|Q| \to \infty$ (each dyadic cube of sidelength 2^{-n} is contained in exactly one dyadic cube of sidelength 2^{-n+1}, by Step 0, and by Step 0 again these are the only dyadic cubes containing Q). Only finitely many members of this chain can lie in $\mathcal{Q}_\mu$, so there is a largest one, Q^*; it belongs to $\mathcal{Q}_\mu$, contains Q, and is maximal in $\mathcal{Q}_\mu$ with respect to inclusion, since any element of $\mathcal{Q}_\mu$ strictly containing Q^* would contain Q and hence be a later member of the chain.

Let $\mathcal{M}_\mu$ be the set of maximal elements of $\mathcal{Q}_\mu$. By the previous paragraph, $\Omega_\mu = \bigcup_{Q \in \mathcal{M}_\mu} Q$. Distinct elements of $\mathcal{M}_\mu$ are disjoint: if $Q_1, Q_2 \in \mathcal{M}_\mu$ meet, then by Step 0 one contains the other, and maximality forces $Q_1 = Q_2$. Since $\mathcal{M}_\mu \subset \mathcal{D}$ is countable, countable additivity gives

$$m(\Omega_\mu) = \sum_{Q \in \mathcal{M}_\mu} |Q| < \frac{1}{\mu} \sum_{Q \in \mathcal{M}_\mu} \int_Q |f(y)|\, dy = \frac{1}{\mu} \int_{\Omega_\mu} |f(y)|\, dy \leq \frac{1}{\mu} \|f\|_{L^1},$$

where we used disjointness (and countable additivity of the integral over a disjoint countable union) in the middle equality. (If $\mathcal{M}_\mu = \emptyset$ then $\Omega_\mu = \emptyset$ and the bound is trivial.)

Step 2: passing to the non-strict inequality. Let $\lambda > 0$ and set $E_\lambda := \{x : M_{\mathcal{D}} f(x) \geq \lambda\}$. If $M_{\mathcal{D}} f(x) \geq \lambda$ and $0 < \mu < \lambda$, then the supremum defining $M_{\mathcal{D}} f(x)$ exceeds μ, so some dyadic cube containing x lies in $\mathcal{Q}_\mu$; that is, $x \in \Omega_\mu$. Conversely if $x \in \Omega_\mu$ for every $\mu < \lambda$ then $M_{\mathcal{D}} f(x) \geq \mu$ for every $\mu < \lambda$, so $M_{\mathcal{D}} f(x) \geq \lambda$. Hence, taking $\mu = \lambda(1 - 1/n)$,

$$E_\lambda = \bigcap_{n=2}^{\infty} \Omega_{\lambda(1-1/n)},$$

which is measurable, being a countable intersection of measurable sets. By monotonicity and Step 1, for every $n \geq 2$,

$$m(E_\lambda) \leq m\big(\Omega_{\lambda(1-1/n)}\big) \leq \frac{1}{\lambda(1-1/n)} \|f\|_{L^1}.$$

Letting $n \to \infty$ gives

$$m\left(\left\{ x \in \mathbb{R}^d : \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\, dy \geq \lambda \right\} \right) \leq \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(t)|\, dt,$$

as required. Note that the constant here is 1, independent of the dimension d; the nesting of the dyadic grid replaces the Vitali covering lemma and costs nothing.

Problem 1.6.21. (Besicovich covering lemma in one dimension) Let $I_1, \dots, I_n$ be a finite family of open intervals in $\mathbb{R}$ (not necessarily disjoint). Show that there exists a subfamily $I'_1, \dots, I'_m$ of intervals such that
(i) $\bigcup_{i=1}^n I_i = \bigcup_{j=1}^m I'_j$; and
(ii) Each point $x \in \mathbb{R}$ is contained in at most two of the I'_j.
(Hint: First refine the family of intervals so that no interval I_i is contained in the union of the other intervals. At that point, show that it is no longer possible for a point to be contained in three of the intervals.) There is a variant of this lemma that holds in higher dimensions, known as the Besicovitch covering lemma.

Solution. Step 1: the refinement. Run the following finite pruning procedure. Start with the family $\mathcal{F}_0 := (I_1, \dots, I_n)$. Given a family $\mathcal{F}_t$, if some member I of $\mathcal{F}_t$ satisfies

$$I \subset \bigcup \{J : J \in \mathcal{F}_t,\ J \neq I\}$$

(where the members are compared as members of the list, so that a repeated interval counts as two distinct members), let $\mathcal{F}_{t+1}$ be $\mathcal{F}_t$ with that member deleted; otherwise stop. Deleting such a member does not change the union of the family, since the deleted interval was already covered by the others. Each step reduces the number of members by one, so the procedure stops after at most n steps. Let $I'_1, \dots, I'_m$ be the members of the final family; it is a subfamily of $I_1, \dots, I_n$, it satisfies (i), and it has the property
$(*)$ no I'_j is contained in $\bigcup_{k \neq j} I'_k$.
In particular each I'_j is nonempty (the empty interval is contained in any union, including the empty one), the I'_j are pairwise distinct as sets (if $I'_j = I'_k$ with $j \neq k$ then $I'_j \subset \bigcup_{l \neq j} I'_l$, contradicting $(*)$), and for each j there is a witness point $z_j \in I'_j$ lying in no other I'_k.
Step 2: no point lies in three of the I'_j. Suppose for contradiction that some $x \in \mathbb{R}$ lies in $I'_{j_1} \cap I'_{j_2} \cap I'_{j_3}$ with j_1, j_2, j_3 distinct. Write $I'_{j_l} = (a_l, b_l)$ for $l = 1, 2, 3$; these are nonempty open intervals with

$$a_l < x < b_l \qquad (l = 1, 2, 3),$$

where we allow $a_l = -\infty$ and $b_l = +\infty$.
Choose $u \in \{1, 2, 3\}$ with $a_u = \min(a_1, a_2, a_3)$ and $v \in \{1, 2, 3\}$ with $b_v = \max(b_1, b_2, b_3)$ (ties broken arbitrarily).
Case 1: $u = v$. Then for each $l \neq u$ we have $a_u \leq a_l$ and $b_l \leq b_u$, so $I'_{j_l} = (a_l, b_l) \subset (a_u, b_u) = I'_{j_u}$. Picking any such l (there are two of them, and $j_l \neq j_u$), we get $I'_{j_l} \subset \bigcup_{k \neq j_l} I'_k$, contradicting $(*)$.
Case 2: $u \neq v$. Let r be the remaining index in $\{1, 2, 3\}$, so $r \neq u$ and $r \neq v$. Since both (a_u, b_u) and (a_v, b_v) contain x, their union is again an interval, namely

$$I'_{j_u} \cup I'_{j_v} = (\min(a_u, a_v), \max(b_u, b_v)) = (a_u, b_v),$$

using that a_u is the least of the three left endpoints and b_v the greatest of the three right endpoints. But then, since $a_u \leq a_r$ and $b_r \leq b_v$,

$$I'_{j_r} = (a_r, b_r) \subset (a_u, b_v) = I'_{j_u} \cup I'_{j_v} \subset \bigcup_{k \neq j_r} I'_k,$$

again contradicting $(*)$.
Both cases are impossible, so no point of $\mathbb{R}$ lies in three of the intervals $I'_1, \dots, I'_m$. That is, every $x \in \mathbb{R}$ lies in at most two of them, which is (ii). Together with Step 1 this proves the lemma.
Remark. Conclusion (ii) says exactly that $\sum_{j=1}^m 1_{I'_j} \leq 2 \cdot 1_{\bigcup_{i=1}^n I_i}$ pointwise on $\mathbb{R}$. Integrating this against any measure μ on the Borel σ-algebra (the indicator functions here are Borel measurable, and the sum is finite, so no convergence theorem is needed) gives

$$\sum_{j=1}^m \mu(I'_j) \leq 2\, \mu\left(\bigcup_{i=1}^n I_i \right).$$

This is what makes the lemma a substitute for Lemma 1.6.22 in the argument of Exercise 1.6.22. The essential difference from Lemma 1.6.22 is that here the union of the original family is preserved exactly, and no dilation of the selected intervals is needed; the price is that the selected intervals overlap, but only with multiplicity two. Lemma 1.6.22 is unusable in that setting because it would require an estimate of the form $\mu(3I) \le C\mu(I)$, which a general Borel measure μ need not satisfy.

Problem 1.6.22. Let μ be a Borel measure (i.e. a countably additive measure on the Borel σ-algebra) on $\mathbb{R}$, such that $0 < \mu(I) < \infty$ for every interval I of positive length. Assume that μ is inner regular, in the sense that $\mu(E) = \sup_{K \subset E,\ \text{compact}} \mu(K)$ for every Borel measurable set E. (As it turns out, from the theory of Radon measures, all locally finite Borel measures have this property, but we will not prove this here; see Section 1.10 of An epsilon of room, Vol. I.) Establish the Hardy-Littlewood maximal inequality

$$\mu\Big(\Big\{x \in \mathbb{R} : \sup_{x \in I} \frac{1}{\mu(I)} \int_I |f(y)|\, d\mu(y) \ge \lambda\Big\}\Big) \le \frac{2}{\lambda} \int_{\mathbb{R}} |f(y)|\, d\mu(y)$$

for any absolutely integrable function $f \in L^1(\mu)$, where the supremum ranges over all open intervals I that contain x. Note that this essentially generalises Exercise 1.6.11, in which μ is replaced by Lebesgue measure. (Hint: Repeat the proof of the usual Hardy-Littlewood maximal inequality, but use the Besicovich covering lemma in place of the Vitali-type covering lemma. Why do we need the former lemma here instead of the latter?)

Solution. Throughout, $f \in L^1(\mu)$ is fixed, and for an open interval I we write

$$A_I := \frac{1}{\mu(I)} \int_I |f(y)|\, d\mu(y).$$

This is well defined: a nonempty open interval has positive length, so $0 < \mu(I) < \infty$ by hypothesis. Let

$$E_\lambda := \Big\{x \in \mathbb{R} : \sup_{I \ni x} A_I \ge \lambda\Big\},$$

the supremum being over open intervals I containing x.

Step 1: E_λ is Borel. For $\nu \in \mathbb{R}$ put

$$U_\nu := \bigcup \{I : I \text{ an open interval with } \textstyle\int_I |f|\, d\mu > \nu\, \mu(I)\}.$$

Being a union of open sets, U_ν is open, hence Borel. By construction $x \in U_\nu$ if and only if there is an open interval $I \ni x$ with $A_I > \nu$, i.e. if and only if $\sup_{I \ni x} A_I > \nu$. Consequently

$$E_\lambda = \bigcap_{n \ge 1} U_{\lambda - 1/n},$$

since $\sup_{I \ni x} A_I \ge \lambda$ holds precisely when $\sup_{I \ni x} A_I > \lambda - 1/n$ for every n. Thus E_λ is a countable intersection of open sets and in particular Borel measurable, so $\mu(E_\lambda)$ makes sense.

Step 2: the estimate on each U_ν. We claim that for every $\nu > 0$,

$$\mu(U_\nu) \le \frac{2}{\nu} \int_{\mathbb{R}} |f(y)|\, d\mu(y).$$

By inner regularity it suffices to show $\mu(K) \le \frac{2}{\nu} \int_{\mathbb{R}} |f|\, d\mu$ for every compact $K \subset U_\nu$.
Fix such a K. For each $x \in K$ there is, by definition of U_ν, an open interval $I_x \ni x$ with

$$\int_{I_x} |f(y)|\, d\mu(y) > \nu\, \mu(I_x).$$

The family $\{I_x\}_{x \in K}$ is an open cover of the compact set K, so by the Heine-Borel theorem there are finitely many $I_1, \dots, I_n$ from this family with $K \subset \bigcup_{i=1}^n I_i$. By the Besicovitch covering lemma in one

dimension (Exercise 1.6.21) we may pass to a subfamily $I'_1, \dots, I'_m$ with

$$\bigcup_{j=1}^{m} I'_j = \bigcup_{i=1}^{n} I_i \supset K$$

and such that every point of $\mathbb{R}$ lies in at most two of the I'_j; equivalently, $\sum_{j=1}^{m} 1_{I'_j} \le 2$ pointwise. Therefore, using finite subadditivity of μ and then the defining property of each I'_j,

$$\begin{aligned}\mu(K) \;\le\; \mu\Big(\bigcup_{j=1}^{m} I'_j\Big) \;\le\; \sum_{j=1}^{m} \mu(I'_j) \;&<\; \frac{1}{\nu}\sum_{j=1}^{m}\int_{I'_j} |f(y)|\,d\mu(y) \\ &=\; \frac{1}{\nu}\int_{\mathbb{R}} \Big(\sum_{j=1}^{m} 1_{I'_j}(y)\Big)|f(y)|\,d\mu(y) \;\le\; \frac{2}{\nu}\int_{\mathbb{R}} |f(y)|\,d\mu(y),\end{aligned}$$

where the interchange of the finite sum with the integral is just linearity of the integral. Taking the supremum over all compact $K \subset U_\nu$ and using inner regularity gives the claim.

Step 3: conclusion. Fix $\lambda > 0$. For every n with $1/n < \lambda$ we have $E_\lambda \subset U_{\lambda - 1/n}$ by Step 1, so by Step 2 and monotonicity,

$$\mu(E_\lambda) \le \mu(U_{\lambda-1/n}) \le \frac{2}{\lambda - 1/n}\int_{\mathbb{R}} |f(y)|\,d\mu(y).$$

Letting $n \to \infty$ yields

$$\mu(E_\lambda) \le \frac{2}{\lambda}\int_{\mathbb{R}} |f(y)|\,d\mu(y),$$

which is the desired inequality.

Why Besicovitch and not the Vitali-type lemma. The Vitali-type covering lemma used to prove Theorem 1.6.20 produces a **disjoint** subfamily $B'_1, \dots, B'_m$ whose threefold dilates $3B'_j$ cover the original union; the passage from $\sum_j \mu(3B'_j)$ back to $\sum_j \mu(B'_j)$ requires the doubling property $m(3B) = 3^d m(B)$ of Lebesgue measure. A general Borel measure μ as above need not be doubling: there need be no constant C with $\mu(3I) \le C\mu(I)$ for all intervals I. For instance $d\mu = e^{-y^2}\,dy$ is a finite Borel measure that is positive and finite on every interval of positive length and is inner regular, yet taking $I = (a, a+1)$ with $a \to +\infty$ one has $\mu(I) \le e^{-a^2}$ while $3I = (a-1, a+2) \supset (a-1, a-\frac{1}{2})$ gives $\mu(3I) \ge \frac{1}{2}e^{-(a-1/2)^2} = \frac{1}{2}e^{1/4}\,e^{-a^2}e^{a-1/2}$, so $\mu(3I)/\mu(I) \to \infty$. The Besicovitch lemma avoids dilations altogether: it keeps the **original** intervals and only asserts bounded overlap, and bounded overlap costs a factor 2 for **any** measure, since $\sum_j 1_{I'_j} \le 2$ is a pointwise statement about sets. This is exactly why the constant 2 here is independent of μ.

Problem 1.6.23. (Cousin's theorem) Prove Cousin's theorem: given any function $\delta : [a,b] \to (0, +\infty)$ on a compact interval $[a,b]$ of positive length, there exists a partition $a = t_0 < t_1 < \cdots < t_k = b$ with $k \ge 1$, together with real numbers $t_j^* \in [t_{j-1}, t_j]$ for each $1 \le j \le k$ and $t_j - t_{j-1} \le \delta(t_j^*)$. (Hint: use the Heine-Borel theorem, which asserts that any open cover of $[a,b]$ has a finite subcover, followed by the Besicovitch covering lemma.) This theorem is useful in a variety of applications related to the second fundamental theorem of calculus, as we shall see below. The positive function δ is known as a gauge function.

Solution. Call a closed subinterval $[c,d] \subset [a,b]$ with $c < d$ **good** (for the gauge δ) if there is a partition $c = t_0 < t_1 < \cdots < t_k = d$ with $k \ge 1$ and tags $t_j^* \in [t_{j-1}, t_j]$ satisfying $t_j - t_{j-1} \le \delta(t_j^*)$ for all $1 \le j \le k$. We must show that $[a,b]$ is good.

Concatenation. If $c < e < d$ and both $[c,e]$ and $[e,d]$ are good, then $[c,d]$ is good: simply list the partition points of $[c,e]$ followed by those of $[e,d]$ (the point e occurring once), and keep the corresponding tags. Every constraint $t_j - t_{j-1} \le \delta(t_j^*)$ is inherited from one of the two halves. Contrapositively:

Halving principle: if $[c,d]$ is not good, then at least one of its two halves $[c, \frac{c+d}{2}]$, $[\frac{c+d}{2}, d]$ is

not good.
The bisection. Suppose for contradiction that $[a,b]$ is not good. Set $[c_0,d_0] := [a,b]$, and recursively, given a non-good interval $[c_i,d_i]$ of length $(b-a)2^{-i} > 0$, use the halving principle to choose one of its two halves, called $[c_{i+1},d_{i+1}]$, which is again not good; it has length $(b-a)2^{-(i+1)}$. This produces a nested decreasing sequence

$$[a,b] = [c_0,d_0] \supset [c_1,d_1] \supset [c_2,d_2] \supset \cdots, \qquad d_i - c_i = (b-a)2^{-i},$$

of non-good intervals.
The c_i form a non-decreasing sequence bounded above by b, and the d_i a non-increasing sequence bounded below by a, with $d_i - c_i \to 0$; hence by the nested interval property (a form of the completeness of $\mathbb{R}$, equivalently of the compactness of $[a,b]$) there is a point

$$x^* \in \bigcap_{i \geq 0} [c_i,d_i].$$

In particular $x^* \in [a,b]$, so $\delta(x^*) > 0$ is defined and strictly positive.
The contradiction. Choose i so large that $(b-a)2^{-i} \leq \delta(x^*)$; this is possible since $2^{-i} \to 0$ and $\delta(x^*) > 0$. For this i consider the partition of $[c_i,d_i]$ into the single subinterval ($k = 1$ in the notation of the statement) $t_0 := c_i < t_1 := d_i$, tagged by $t_1^* := x^*$. This is legitimate: $c_i < d_i$, the tag satisfies $x^* \in [c_i,d_i] = [t_0,t_1]$, and

$$t_1 - t_0 = d_i - c_i = (b-a)2^{-i} \leq \delta(x^*) = \delta(t_1^*).$$

Hence $[c_i,d_i]$ is good, contradicting its construction. Therefore $[a,b]$ is good, which is the assertion of Cousin's theorem. □
Remark (the route suggested by the hint). One can instead argue by compactness directly: the intervals $J_x := (x - \delta(x)/2,\ x + \delta(x)/2)$, $x \in [a,b]$, form an open cover of $[a,b]$, so by Heine-Borel finitely many of them cover $[a,b]$, and Exercise 1.6.21 lets one thin this finite family down to one with the same union in which every point lies in at most two intervals; one then chooses the partition points inside the successive overlaps. The bisection argument above is a cleaner packaging of the same use of compactness, so we have given it instead. A third variant is to let S be the set of $x \in (a,b]$ for which $[a,x]$ is good, note that S is nonempty (any $x \in (a, \min(b, a+\delta(a))]$ lies in S, tagging $[a,x]$ by a), and check that $c := \sup S$ lies in S and equals b, using the gauge value $\delta(c)$ in both steps.

Problem 1.6.24. If $E \subset \mathbb{R}^d$ is Lebesgue measurable, show that almost every point in E is a point of density for E, and almost every point in the complement of E is not a point of density for E.

Solution. Recall the definition given just before the exercise: $x \in \mathbb{R}^d$ is a **point of density** for E if

$$\frac{m(E \cap B(x,r))}{m(B(x,r))} \to 1 \quad \text{as } r \to 0.$$

Note $m(B(x,r)) = \omega_d r^d > 0$ for $r > 0$, where $\omega_d := m(B(0,1))$, so the ratio is always defined.
Step 1: the bounded case. Suppose first that E is a measurable set of finite measure, so that $f := 1_E$ is absolutely integrable. By the Lebesgue differentiation theorem in general dimension (Theorem 1.6.19), for almost every $x \in \mathbb{R}^d$ we have

$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} 1_E(y)\, dy = 1_E(x),$$

that is,

$$\lim_{r \to 0} \frac{m(E \cap B(x,r))}{m(B(x,r))} = 1_E(x).$$

Let N be the null set off which this holds. If $x \in E \setminus N$ then the limit is 1, so x is a point of density for E; if $x \in E^c \setminus N$ then the limit exists and equals $0 \neq 1$, so x is **not** a point of density for E. Thus the conclusion holds for E of finite measure, with exceptional set N.

Step 2: localisation. For a general measurable $E \subset \mathbb{R}^d$ and each integer $n \geq 2$ set

$$E_n := E \cap B(0,n),$$

a measurable set of finite measure. If $x \in B(0,n-1)$ and $0 < r < 1$ then $B(x,r) \subset B(0,n)$, and hence

$$E \cap B(x,r) = E \cap B(0,n) \cap B(x,r) = E_n \cap B(x,r).$$

Since the definition of a point of density only involves the ratios for arbitrarily small r, we conclude: **For $x \in B(0,n-1)$, x is a point of density for E if and only if it is a point of density for E_n.**
Step 3: conclusion. Apply Step 1 to E_n: there is a null set $N_n \subset \mathbb{R}^d$ such that every $x \in E_n \setminus N_n$ is a point of density for E_n and every $x \notin E_n$ with $x \notin N_n$ is not a point of density for E_n. Let

$$N := \bigcup_{n \geq 2} N_n,$$

which is a countable union of null sets and hence null (by countable subadditivity of Lebesgue outer measure).
Let $x \in E \setminus N$. Choose $n \geq 2$ with $x \in B(0,n-1)$. Then $x \in E \cap B(0,n-1) \subset E_n$ and $x \notin N_n$, so x is a point of density for E_n, and hence, by Step 2, a point of density for E. Thus every point of E outside the null set N is a point of density for E; that is, almost every point of E is a point of density for E.
Now let $x \in E^c \setminus N$, and again choose $n \geq 2$ with $x \in B(0,n-1)$. Since $E_n \subset E$ we have $x \notin E_n$, and $x \notin N_n$, so x is not a point of density for E_n; by Step 2 again, x is not a point of density for E. Thus almost every point of the complement of E fails to be a point of density for E. □
Remark. Steps 2 and 3 can be replaced by a single appeal to Exercise 1.6.14(ii): 1_E is locally integrable for any measurable E (indeed $\int_{B(0,r)} 1_E \leq m(B(0,r)) < \infty$), so Theorem 1.6.19 applies to $f = 1_E$ directly and Step 1 goes through verbatim for arbitrary measurable E.

Problem 1.6.25. Let $E \subset \mathbb{R}^d$ be a measurable set of positive measure, and let $\varepsilon > 0$.
(i) Using Exercise 1.6.15 and Exercise 1.6.24, show that there exists a cube $Q \subset \mathbb{R}^d$ of positive sidelength such that $m(E \cap Q) > (1-\varepsilon)m(Q)$.
(ii) Give an alternate proof of the above claim that avoids the Lebesgue differentiation theorem. (Hint: reduce to the case when E is bounded, then approximate E by an almost disjoint union of cubes.)
(iii) Use the above result to give an alternate proof of the Steinhaus theorem (Exercise 1.6.8).
Of course, one can replace cubes here by other comparable shapes, such as balls. (Indeed, a good principle to adopt in analysis is that cubes and balls are "equivalent up to constants", in that a cube of some sidelength can be contained in a ball of comparable radius, and vice versa. This type of mental equivalence is analogous to, though not identical with, the famous dictum that a topologist cannot distinguish a doughnut from a coffee cup.)

Solution. We may assume $0 < \varepsilon < 1$, since the statement for a smaller ε implies it for a larger one.
(i). Write $f := 1_E$; as noted in Exercise 1.6.24 this is locally integrable, since $\int_{B(0,r)} 1_E \leq m(B(0,r)) < \infty$ for every $r > 0$. By Theorem 1.6.19 (extended to locally integrable functions via Exercise 1.6.14(ii)) almost every point of $\mathbb{R}^d$ is a Lebesgue point of f. Since $m(E) > 0$, the set of Lebesgue points of f lying in E is nonempty; fix such a point x, so that $f(x) = 1_E(x) = 1$ and

$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |1_E(y) - 1|\, dy = 0.$$

For $h > 0$ let

$$R_h := \left(-\tfrac{h}{\sqrt{d}}, \tfrac{h}{\sqrt{d}} \right)^d$$

be the open cube of sidelength $2h/\sqrt{d}$ centred at the origin. If $y \in R_h$ then $|y|^2 = \sum_{i=1}^d y_i^2 <$

$d \cdot (h/\sqrt{d})^2 = h^2$, so $R_h \subset B(0,h)$, as Exercise 1.6.15 requires. Moreover

$$m(R_h) = \Big(\frac{2h}{\sqrt{d}}\Big)^d = \frac{(2/\sqrt{d})^d}{\omega_d}\, m(B(0,h)), \qquad \omega_d := m(B(0,1)),$$

so $m(R_h) \geq c\, m(B(0,h))$ with $c := (2/\sqrt{d})^d/\omega_d > 0$ independent of h. Thus the family $(R_h)_{h>0}$ satisfies the hypotheses of Exercise 1.6.15, and that exercise gives

$$\lim_{h\to 0} \frac{1}{m(R_h)} \int_{x+R_h} 1_E(y)\, dy = 1_E(x) = 1,$$

that is, $m(E \cap (x+R_h))/m(x+R_h) \to 1$ as $h \to 0$ (using translation invariance $m(x+R_h) = m(R_h)$). Choosing h small enough that this ratio exceeds $1-\varepsilon$ and setting $Q := x + R_h$, a cube of positive sidelength $2h/\sqrt{d}$, we obtain $m(E \cap Q) > (1-\varepsilon)m(Q)$ as required. (If one insists on a closed cube, take the closure of Q; this changes neither $m(Q)$ nor $m(E \cap Q)$, since the boundary of a cube is null.)
Two remarks on the ingredients. The only role of $m(E) > 0$ is to guarantee that E contains at least one Lebesgue point of 1_E; equivalently one may quote Exercise 1.6.24 as the exercise suggests, which says that almost every point of E is a point of density for E, and pick any such $x \in E$ (a set of positive measure remains nonempty after deleting a null set). The two are the same thing here: for $x \in E$ one has $|1_E(y) - 1_E(x)| = 1_{E^c}(y)$, so

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |1_E(y) - 1_E(x)|\, dy = 1 - \frac{m(E \cap B(x,r))}{m(B(x,r))},$$

and hence x is a Lebesgue point of 1_E precisely when it is a point of density for E. Exercise 1.6.15 then transfers the conclusion from balls to the inscribed cubes $x + R_h$.
(ii). First we reduce to E of finite measure. Since $m(E) > 0$ and $E = \bigcup_n (E \cap B(0,n))$ is an increasing union, upward monotone convergence for Lebesgue measure gives $m(E \cap B(0,n)) > 0$ for some n. Replacing E by $E' := E \cap B(0,n)$ only decreases the quantity $m(E \cap Q)$, so a cube Q with $m(E' \cap Q) > (1-\varepsilon)m(Q)$ works for E as well. Hence we may assume $0 < m(E) < \infty$.
Since $\varepsilon \in (0,1)$ and $0 < m(E) < \infty$, we have $m(E) < m(E)/(1-\varepsilon)$, so we may pick $\eta > 0$ with $m(E) + \eta < m(E)/(1-\varepsilon)$. By the outer regularity built into the definition of Lebesgue measurability, there is an open set $U \supset E$ with $m(U \setminus E) \leq \eta$, whence

$$m(U) \leq m(E) + \eta < \frac{m(E)}{1-\varepsilon}.$$

By Lemma 1.2.11 we may write $U = \bigcup_{n=1}^{\infty} Q_n$ as a countable union of almost disjoint closed cubes Q_n, each of positive sidelength, and by Lemma 1.2.9,

$$m(U) = \sum_{n=1}^{\infty} m(Q_n).$$

On the other hand $E \subset U = \bigcup_n Q_n$, so $E = \bigcup_n (E \cap Q_n)$ and countable subadditivity gives

$$m(E) \leq \sum_{n=1}^{\infty} m(E \cap Q_n).$$

Suppose, for contradiction, that $m(E \cap Q_n) \leq (1-\varepsilon)m(Q_n)$ for every n. Then

$$m(E) \leq \sum_{n=1}^{\infty} (1-\varepsilon)m(Q_n) = (1-\varepsilon)m(U) < (1-\varepsilon) \cdot \frac{m(E)}{1-\varepsilon} = m(E),$$

a contradiction (all quantities are finite since $m(E) < \infty$ forces $m(U) < \infty$). Hence some Q_n satisfies $m(E \cap Q_n) > (1-\varepsilon)m(Q_n)$, and Q_n is a cube of positive sidelength. No use was made of the Lebesgue differentiation theorem.

(iii) (Steinhaus theorem). Let $E \subset \mathbb{R}^d$ be measurable with $m(E) > 0$; we show $E - E$ contains an open neighbourhood of the origin.
Apply (i) or (ii) with $\varepsilon := 1/4$: there is a cube Q, of sidelength $\ell > 0$, with

$$m(F) > \tfrac{3}{4}\ell^d, \qquad F := E \cap Q.$$

Set $\eta := 1 - 2^{-1/d} \in (0,1)$, so that $(1-\eta)^d = 1/2$. We claim that

$$\{v \in \mathbb{R}^d : |v_i| < \eta\ell \text{ for all } i\} \subset E - E,$$

which is an open neighbourhood of 0 and so finishes the proof.
Let v satisfy $|v_i| < \eta\ell$ for all i, and suppose for contradiction that $F \cap (F+v) = \emptyset$. Both F and $F+v$ are measurable subsets of $Q \cup (Q+v)$, so by disjointness and translation invariance,

$$2\,m(F) = m(F) + m(F+v) = m\big(F \cup (F+v)\big) \le m\big(Q \cup (Q+v)\big) = 2m(Q) - m\big(Q \cap (Q+v)\big).$$

Writing $Q = \prod_{i=1}^d [a_i, a_i+\ell]$, the set $Q \cap (Q+v)$ is the box $\prod_{i=1}^d [\max(a_i, a_i+v_i), \min(a_i+\ell, a_i+\ell+v_i)]$, of sidelengths $\ell - |v_i| > 0$; hence

$$m\big(Q \cap (Q+v)\big) = \prod_{i=1}^d (\ell - |v_i|) \ge (1-\eta)^d \ell^d = \tfrac{1}{2}\ell^d.$$

Combining, and using $m(Q) = \ell^d$,

$$2 \cdot \tfrac{3}{4}\ell^d < 2m(F) \le 2\ell^d - \tfrac{1}{2}\ell^d = \tfrac{3}{2}\ell^d,$$

i.e. $\frac{3}{2}\ell^d < \frac{3}{2}\ell^d$, absurd. Therefore $F \cap (F+v) \neq \emptyset$: there are $y, z \in F \subset E$ with $y = z + v$, so $v = y - z \in E - E$. This proves the claim, and hence the Steinhaus theorem. □

Problem 1.6.26.
(i) Give an example of a compact set $K \subset \mathbb{R}$ of positive measure such that $m(K \cap I) < |I|$ for every interval I of positive length. (Hint: first construct an open dense subset of $[0,1]$ of measure strictly less than 1.)
(ii) Give an example of a measurable set $E \subset \mathbb{R}$ such that $0 < m(E \cap I) < |I|$ for every interval I of positive length. (Hint: first work in a bounded interval, such as $(-1, 2)$. The complement of the set K in the first example is the union of at most countably many open intervals, thanks to Exercise 1.6.10. Now fill in these open intervals and iterate.)

Solution. **(i).** Enumerate the rationals in $[0,1]$ as $q_1, q_2, q_3, \ldots$ and set

$$U := \bigcup_{n=1}^{\infty} \big(q_n - 2^{-n-3},\ q_n + 2^{-n-3}\big), \qquad K := [0,1] \setminus U.$$

Then U is open, the n-th generating interval having length 2^{-n-2}, so by countable subadditivity

$$m(U) \le \sum_{n=1}^{\infty} 2^{-n-2} = \tfrac{1}{4}.$$

The set $U \cap [0,1]$ is an open subset of $[0,1]$ containing every rational of $[0,1]$, hence dense in $[0,1]$, and of measure at most $1/4 < 1$ (this is the set the hint asks for). The set $K = [0,1] \cap U^c$ is closed and bounded, hence compact, and since $[0,1] \subset K \cup U$,

$$m(K) \ge m([0,1]) - m(U) \ge 1 - \tfrac{1}{4} = \tfrac{3}{4} > 0.$$

Now let I be any interval of positive length. If $|I| = \infty$ then $m(K \cap I) \le m(K) \le 1 < \infty = |I|$, so assume $|I| < \infty$.

Case 1: $I \cap (0,1) = \emptyset$. Then I lies in $(-\infty, 0]$ or in $[1, \infty)$, so $K \cap I \subset \{0, 1\}$ and $m(K \cap I) = 0 < |I|$.
Case 2: $I \cap (0,1) \neq \emptyset$. The interior I° is a nonempty open interval, and $I^\circ \cap (0,1)$ is nonempty: indeed, if $x \in I \cap (0,1)$ lies in I° we are done, and if x is an endpoint of I then points of I° sufficiently close to x still lie in $(0,1)$. Thus $V := I^\circ \cap (0,1)$ is a nonempty open set, so it contains some rational q_n together with a small interval around q_n; intersecting with the n-th generating interval of U we see that $V \cap U$ contains a nonempty open interval and therefore $m(I \cap U) \geq m(V \cap U) > 0$. Since $K \cap I \subset I \setminus U$,

$$m(K \cap I) \leq |I| - m(I \cap U) < |I|.$$

Thus K is a compact set of positive measure with $m(K \cap I) < |I|$ for every interval I of positive length. (Note K is nowhere dense: it is closed and contains no rational number, hence contains no interval of positive length, so its interior is empty.)
(ii). We construct E as a countable disjoint union of affine copies of the set K from part (i), placed so as to meet every interval, while leaving room for a second family of copies in the complement.
Building blocks. For real $\alpha < \beta$ let $\varphi_{\alpha,\beta}(t) := \alpha + (\beta - \alpha)t$ and put $K_{[\alpha,\beta]} := \varphi_{\alpha,\beta}(K) \subset [\alpha, \beta]$. Since $\varphi_{\alpha,\beta}$ is an increasing affine homeomorphism of $\mathbb{R}$, the set $K_{[\alpha,\beta]}$ is compact, has empty interior (so is nowhere dense), and by the scaling behaviour of Lebesgue measure

$$m\big(K_{[\alpha,\beta]}\big) = (\beta - \alpha)\, m(K) \in \big(0,\ \beta - \alpha\big].$$

Recursive construction. Let $J_1, J_2, J_3, \ldots$ enumerate the countably many open intervals (p, q) with $p < q$ rational. We construct compact nowhere dense sets $A_1, B_1, A_2, B_2, \ldots$, pairwise disjoint, with

$$A_n \cup B_n \subset J_n, \qquad m(A_n) > 0, \quad m(B_n) > 0, \quad m(A_n) \leq 2^{-n}.$$

Suppose $A_1, B_1, \ldots, A_{n-1}, B_{n-1}$ have been chosen, and let $S := \bigcup_{i<n}(A_i \cup B_i)$. This is a finite union of closed sets with empty interior, hence itself closed with empty interior. (If F, G are closed with empty interior and $V \subset F \cup G$ is open and nonempty, then $V \setminus F$ is open and contained in G, so it is empty, whence $V \subset F$, contradicting $F^\circ = \emptyset$.) Consequently $J_n \setminus S$ is a nonempty open set, so it contains an open interval (α, β) with $\alpha < \beta$. Choose two disjoint closed subintervals $[\alpha_1, \beta_1], [\alpha_2, \beta_2] \subset (\alpha, \beta)$ with $\beta_1 - \alpha_1 \leq 2^{-n}$, and set

$$A_n := K_{[\alpha_1,\beta_1]}, \qquad B_n := K_{[\alpha_2,\beta_2]}.$$

These are compact, nowhere dense, of positive measure, contained in J_n, disjoint from each other and from S, and $m(A_n) \leq \beta_1 - \alpha_1 \leq 2^{-n}$. By construction each later pair A_m, B_m $(m > n)$ is disjoint from A_n and B_n, so the whole family $\{A_n\} \cup \{B_n\}$ is pairwise disjoint.
The set. Put

$$E := \bigcup_{n=1}^{\infty} A_n.$$

This is a countable union of compact sets, hence Borel, hence Lebesgue measurable, and

$$m(E) \leq \sum_{n=1}^{\infty} m(A_n) \leq \sum_{n=1}^{\infty} 2^{-n} = 1 < \infty.$$

Also $E \cap B_n = \emptyset$ for every n, by the pairwise disjointness above.
Verification. Let I be an interval of positive length. Its interior I° is a nonempty open interval, so we may pick rationals $p < q$ with $(p, q) \subset I^\circ \subset I$; say $(p, q) = J_n$. Then $A_n \subset J_n \subset I$ and $B_n \subset J_n \subset I$.
Lower bound: $m(E \cap I) \geq m(A_n) > 0$.
Upper bound: if $|I| = \infty$ then $m(E \cap I) \leq m(E) \leq 1 < \infty = |I|$. If $|I| < \infty$, then $E \cap I$ and $I \setminus E$ partition I, so

$$m(E \cap I) = |I| - m(I \setminus E) \leq |I| - m(B_n) < |I|,$$

using $B_n \subset I \setminus E$ and $m(B_n) > 0$.
Hence $0 < m(E \cap I) < |I|$ for every interval I of positive length. $\square$
Remark. The construction is exactly the "fill in the complementary intervals and iterate" procedure of the hint, organised so that the bookkeeping is done once: instead of iterating through the complementary intervals of a single K (which by Exercise 1.6.10 form a countable disjoint family of open

intervals), we run through a countable base $\{J_n\}$ of $\mathbb{R}$ and insert two fat nowhere dense sets into each stage, one to be kept and one to be discarded. Nowhere density of the pieces is what guarantees there is always room left at the next stage.

Problem 1.6.27. (Approximations to the identity) Define a good kernel to be a measurable function $P : \mathbb{R}^d \to \mathbb{R}^+$ which is non-negative, radial (which means that there is a function $\tilde{P} : [0, +\infty) \to \mathbb{R}^+$ such that $P(x) = \tilde{P}(|x|)$), radially non-increasing (so that $\tilde{P}$ is a non-increasing function), and has total mass $\int_{\mathbb{R}^d} P(x)\,dx$ equal to 1. The functions $P_t(x) := \frac{1}{t^d} P(\frac{x}{t})$ for $t > 0$ are then said to be a good family of approximations to the identity. [Different texts have slightly different notions of what a good kernel is; the "right" class of kernels to consider depends to some extent on what type of convergence results one is interested in (e.g. almost everywhere convergence, convergence in L^1 or L^∞ norm, etc.), and on what hypotheses one wishes to place on the original function f.]

(i) Show that the heat kernels $P_t(x) := \frac{1}{(4\pi t^2)^{d/2}} e^{-|x|^2/4t^2}$ and Poisson kernels $P_t(x) := c_d \frac{t}{(t^2+|x|^2)^{(d+1)/2}}$ are good families of approximations to the identity, if the constant $c_d > 0$ is chosen correctly (in fact one has $c_d = \Gamma((d+1)/2)/\pi^{(d+1)/2}$, but you are not required to establish this). [Note that we have modified the usual formulation of the heat kernel by replacing t with t^2 in order to make it conform to the notational conventions used in this exercise.]

(ii) Show that if P is a good kernel, then

$$c_d < \sum_{n=-\infty}^{\infty} 2^{dn} \tilde{P}(2^n) \leq C_d$$

for some constants $0 < c_d < C_d$ depending only on d. (Hint: compare P with such "horizontal wedding cake" functions as $\sum_{n=-\infty}^{\infty} 1_{2^{n-1}<|x|\leq 2^n} \tilde{P}(2^n)$.)

(iii) Establish the quantitative upper bound

$$\Big| \int_{\mathbb{R}^d} f(y) P_t(x-y)\,dy \Big| \leq C'_d \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\,dy$$

for any absolutely integrable function f and some constant $C'_d > 0$ depending only on d.

(iv) Show that if $f : \mathbb{R}^d \to \mathbb{C}$ is absolutely integrable and x is a Lebesgue point of f, then the convolution

$$f * P_t(x) := \int_{\mathbb{R}^d} f(y) P_t(x-y)\,dy$$

converges to $f(x)$ as $t \to 0$. (Hint: split $f(y)$ as the sum of $f(x)$ and $f(y) - f(x)$.) In particular, $f * P_t$ converges pointwise almost everywhere to f.

Solution. Throughout write $\omega_d := m(B(0,1))$, so that $m(B(x,r)) = \omega_d r^d$, and set

$$A_n := \{x \in \mathbb{R}^d : 2^{n-1} < |x| \leq 2^n\} \qquad (n \in \mathbb{Z}),$$

so that the A_n are disjoint measurable sets with $\bigcup_{n\in\mathbb{Z}} A_n = \mathbb{R}^d \setminus \{0\}$ and

$$m(A_n) = \omega_d\big(2^{nd} - 2^{(n-1)d}\big) = \omega_d(1 - 2^{-d})\,2^{nd}.$$

We record two facts used repeatedly. First, if P is a good kernel then so is each P_t: it is non-negative and measurable, it is radial with profile $\widetilde{P}_t(r) = t^{-d}\tilde{P}(r/t)$, which is again non-increasing, and by the scaling law for Lebesgue measure,

$$\int_{\mathbb{R}^d} P_t(x)\,dx = \int_{\mathbb{R}^d} t^{-d} P(x/t)\,dx = \int_{\mathbb{R}^d} P(u)\,du = 1.$$

Second, P is bounded: $\tilde{P}$ is non-increasing and finite-valued, so $P(x) = \tilde{P}(|x|) \leq \tilde{P}(0) < \infty$.

(i) The heat kernels. Here $P(x) = P_1(x) = (4\pi)^{-d/2} e^{-|x|^2/4}$, and indeed $t^{-d}P(x/t) = t^{-d}(4\pi)^{-d/2}e^{-|x|^2/4t^2} = (4\pi t^2)^{-d/2}e^{-|x|^2/4t^2} = P_t(x)$, so the given family is of the required form. P

is continuous (hence measurable), strictly positive, and radial with $\tilde{P}(r) = (4\pi)^{-d/2}e^{-r^2/4}$, which is non-increasing on $[0,\infty)$. For the total mass, substitute $x = 2u$ (a dilation, under which Lebesgue measure scales by 2^d) and use the standard Gaussian integral $\int_{\mathbb{R}^d} e^{-|u|^2}du = \pi^{d/2}$ (which follows from the one-dimensional case by Tonelli):

$$\int_{\mathbb{R}^d}(4\pi)^{-d/2}e^{-|x|^2/4}\,dx = (4\pi)^{-d/2}2^d\int_{\mathbb{R}^d}e^{-|u|^2}\,du = (4\pi)^{-d/2}2^d\pi^{d/2} = (4\pi)^{-d/2}(4\pi)^{d/2} = 1.$$

So the heat kernels form a good family.

The Poisson kernels. Take $P(x) := c_d(1+|x|^2)^{-(d+1)/2}$; then

$$t^{-d}P(x/t) = \frac{c_d\,t^{-d}}{(1+|x|^2/t^2)^{(d+1)/2}} = \frac{c_d\,t^{-d}t^{d+1}}{(t^2+|x|^2)^{(d+1)/2}} = c_d\frac{t}{(t^2+|x|^2)^{(d+1)/2}},$$

which is the stated Poisson kernel. P is continuous, positive, radial with $\tilde{P}(r) = c_d(1+r^2)^{-(d+1)/2}$ non-increasing. It remains to see that $M := \int_{\mathbb{R}^d}(1+|x|^2)^{-(d+1)/2}dx$ is finite and positive, for then the choice $c_d := 1/M$ normalises the mass to 1. Positivity is clear. For finiteness, split into $B(0,1)$ and the annuli A_n, $n \geq 1$: on $B(0,1)$ the integrand is at most 1, contributing at most ω_d; and on A_n with $n \geq 1$ we have $(1+|x|^2)^{-(d+1)/2} \leq |x|^{-(d+1)} \leq 2^{-(n-1)(d+1)}$, so

$$M \leq \omega_d + \sum_{n\geq 1}2^{-(n-1)(d+1)}\omega_d 2^{nd} = \omega_d + \omega_d 2^{d+1}\sum_{n\geq 1}2^{-n} < \infty.$$

So the Poisson kernels form a good family as well.

(ii). Let P be a good kernel. Since $\tilde{P}$ is non-increasing, for $x \in A_n$ we have $2^{n-1} < |x| \leq 2^n$ and therefore

$$\tilde{P}(2^n) \leq P(x) \leq \tilde{P}(2^{n-1}).$$

The single point 0 is null, so by countable additivity of the integral of a non-negative measurable function over the disjoint decomposition $\mathbb{R}^d \setminus \{0\} = \bigsqcup_n A_n$,

$$1 = \int_{\mathbb{R}^d}P = \sum_{n\in\mathbb{Z}}\int_{A_n}P.$$

The lower bound $P \geq \tilde{P}(2^n)$ on A_n gives

$$1 \geq \sum_n \tilde{P}(2^n)m(A_n) = \omega_d(1-2^{-d})\sum_n 2^{dn}\tilde{P}(2^n),$$

i.e.

$$\sum_{n\in\mathbb{Z}}2^{dn}\tilde{P}(2^n) \leq \frac{1}{\omega_d(1-2^{-d})} = \frac{2^d}{\omega_d(2^d-1)} =: C_d.$$

The upper bound $P \leq \tilde{P}(2^{n-1})$ on A_n gives, after reindexing $m := n-1$ (so $2^{nd} = 2^d2^{md}$),

$$1 \leq \sum_n \tilde{P}(2^{n-1})m(A_n) = \omega_d(1-2^{-d})\sum_n 2^{nd}\tilde{P}(2^{n-1}) = \omega_d(2^d-1)\sum_m 2^{md}\tilde{P}(2^m),$$

i.e.

$$\sum_{n\in\mathbb{Z}}2^{dn}\tilde{P}(2^n) \geq \frac{1}{\omega_d(2^d-1)} > \frac{1}{2\omega_d(2^d-1)} =: c_d.$$

Both c_d and C_d depend only on d, and $0 < c_d < C_d$, as required. (The strict inequality on the left was obtained by simply halving the natural constant.)

(iii). Write

$$Mf(x) := \sup_{r>0}\frac{1}{m(B(x,r))}\int_{B(x,r)}|f(y)|\,dy$$

for the Hardy-Littlewood maximal function appearing on the right-hand side. We prove the bound for an arbitrary good kernel Q in place of P_t; since each P_t is a good kernel (see above) and the constants c_d, C_d of part (ii) depend only on d, the desired estimate follows on taking $Q := P_t$.
Decomposing $\{y : y \neq x\}$ into the annuli $x - A_n$ (i.e. $2^{n-1} < |x-y| \le 2^n$), using $Q(x-y) = \tilde{Q}(|x-y|) \le \tilde{Q}(2^{n-1})$ there, and enlarging the region of integration to the closed ball of radius 2^n about x (whose measure and integral agree with those of $B(x, 2^n)$, the sphere being null):

$$\begin{aligned}\int_{\mathbb{R}^d} |f(y)|\, Q(x-y)\, dy &= \sum_{n\in\mathbb{Z}} \int_{2^{n-1}<|x-y|\le 2^n} |f(y)| Q(x-y)\, dy \\ &\le \sum_{n\in\mathbb{Z}} \tilde{Q}(2^{n-1}) \int_{B(x,2^n)} |f(y)|\, dy \\ &\le \sum_{n\in\mathbb{Z}} \tilde{Q}(2^{n-1})\, \omega_d 2^{nd}\, Mf(x) \\ &= \omega_d 2^d\, Mf(x) \sum_{n\in\mathbb{Z}} 2^{(n-1)d} \tilde{Q}(2^{n-1}) \;\le\; \omega_d 2^d C_d\, Mf(x),\end{aligned}$$

using part (ii) in the last step. Since $\omega_d C_d = 2^d/(2^d-1)$, we may take

$$C_d' := \frac{4^d}{2^d - 1},$$

and then

$$\Big| \int_{\mathbb{R}^d} f(y) P_t(x-y)\, dy \Big| \le \int_{\mathbb{R}^d} |f(y)| P_t(x-y)\, dy \le C_d'\, Mf(x),$$

as claimed. In particular the convolution $f * P_t(x)$ is absolutely convergent whenever $Mf(x) < \infty$; it is in any case absolutely convergent for every x, since P_t is bounded and f is absolutely integrable.
(iv). We first record a decay lemma.
Lemma. If P is a good kernel then $s^d \tilde{P}(s) \to 0$ as $s \to +\infty$.
Proof. By (ii) the series $\sum_n 2^{dn} \tilde{P}(2^n)$ converges (all terms being non-negative), so its terms tend to 0; in particular $2^{dn}\tilde{P}(2^n) \to 0$ as $n \to +\infty$. Given $s \ge 1$, pick n with $2^{n-1} < s \le 2^n$; since $\tilde{P}$ is non-increasing, $\tilde{P}(s) \le \tilde{P}(2^{n-1})$ and so

$$s^d \tilde{P}(s) \le 2^{nd} \tilde{P}(2^{n-1}) = 2^d \cdot 2^{(n-1)d} \tilde{P}(2^{n-1}) \to 0$$

as $s \to \infty$ (note $n \to \infty$ as $s \to \infty$). □
Now let f be absolutely integrable and let x be a Lebesgue point of f, i.e.

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0.$$

Since $\int P_t = 1$ we may write, as suggested by the hint,

$$f * P_t(x) - f(x) = \int_{\mathbb{R}^d} \big(f(y) - f(x)\big) P_t(x-y)\, dy,$$

so that

$$|f * P_t(x) - f(x)| \le \int_{\mathbb{R}^d} |f(y) - f(x)|\, P_t(x-y)\, dy.$$

Fix $\varepsilon > 0$. By the Lebesgue point property, choose $r_0 > 0$ such that

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy \le \varepsilon \qquad \text{for all } 0 < r \le r_0,$$

and let $n_0 \in \mathbb{Z}$ be the largest integer with $2^{n_0} \le r_0$; put $r_1 := 2^{n_0} \le r_0$. Split the integral at radius r_1.

Near part. Exactly as in (iii), but summing only over $n \le n_0$ (these annuli cover $0 < |x-y| \le r_1$) and using $2^n \le r_0$ so that the averages are at most ε:

$$\int_{|x-y|\le r_1} |f(y)-f(x)|P_t(x-y)\,dy \le \sum_{n\le n_0} \widetilde{P}_t(2^{n-1}) \int_{B(x,2^n)} |f(y)-f(x)|\,dy$$
$$\le \sum_{n\le n_0} \widetilde{P}_t(2^{n-1})\,\omega_d 2^{nd}\,\varepsilon \;\le\; \omega_d 2^d C_d\,\varepsilon \;=\; C'_d\,\varepsilon,$$

where we used part (ii) applied to the good kernel P_t. Crucially this bound is uniform in t.
Far part. On $|x-y| > r_1$ we have $P_t(x-y) \le \widetilde{P}_t(r_1) = t^{-d}\tilde{P}(r_1/t)$. Split further:

$$\int_{|x-y|>r_1} |f(y)-f(x)|P_t(x-y)\,dy \le \widetilde{P}_t(r_1) \int_{\mathbb{R}^d} |f(y)|\,dy \;+\; |f(x)| \int_{|z|>r_1} P_t(z)\,dz.$$

For the first term, by the Lemma with $s := r_1/t \to +\infty$ as $t \to 0$,

$$\widetilde{P}_t(r_1) = t^{-d}\tilde{P}(r_1/t) = \frac{1}{r_1^d}\, s^d \tilde{P}(s) \longrightarrow 0 \qquad (t \to 0),$$

and $\int |f| < \infty$, so this term tends to 0. For the second term, by the scaling $z = tu$,

$$\int_{|z|>r_1} P_t(z)\,dz = \int_{|u|>r_1/t} P(u)\,du \longrightarrow 0 \qquad (t \to 0),$$

since $\int_{\mathbb{R}^d} P = 1 < \infty$ and the sets $\{|u| > R\}$ decrease to the null set $\emptyset$ as $R \to \infty$ (downward monotone convergence for the finite measure $A \mapsto \int_A P$).
Conclusion. Combining, for every $\varepsilon > 0$,

$$\limsup_{t\to 0} |f * P_t(x) - f(x)| \le C'_d\,\varepsilon.$$

Letting $\varepsilon \to 0$ gives $f * P_t(x) \to f(x)$ as $t \to 0$.
Finally, by Theorem 1.6.19 almost every $x \in \mathbb{R}^d$ is a Lebesgue point of the absolutely integrable function f; hence $f * P_t \to f$ pointwise almost everywhere as $t \to 0$. □

Problem 1.6.28. (Weierstrass function) Let $F : \mathbb{R} \to \mathbb{R}$ be the function

$$F(x) := \sum_{n=1}^{\infty} 4^{-n} \sin(8^n \pi x).$$

- **(i)** Show that F is well-defined (in the sense that the series is absolutely convergent) and that F is a bounded continuous function.
- **(ii)** Show that for every 8-dyadic interval $[\frac{j}{8^n}, \frac{j+1}{8^n}]$ with $n \ge 1$, one has $|F(\frac{j+1}{8^n}) - F(\frac{j}{8^n})| \ge c4^{-n}$ for some absolute constant $c > 0$.
- **(iii)** Show that F is not differentiable at any point $x \in \mathbb{R}$. (Hint: argue by contradiction and use the previous part of this exercise.) Note that it is not enough to formally differentiate the series term by term and observe that the resulting series is divergent - why not?

Solution. **(i).** For every x and every $n \ge 1$ we have $|4^{-n}\sin(8^n\pi x)| \le 4^{-n}$, and $\sum_{n\ge 1} 4^{-n} = \frac{1}{3} < \infty$. Hence the series defining $F(x)$ is absolutely convergent for each x, so F is well defined, and

$$|F(x)| \le \sum_{n=1}^{\infty} 4^{-n} = \tfrac{1}{3}$$

for all x, so F is bounded. Moreover the bound $\sup_x |4^{-n}\sin(8^n\pi x)| \le 4^{-n}$ with $\sum 4^{-n} < \infty$ is exactly the hypothesis of the Weierstrass M-test, so the partial sums $F_N(x) := \sum_{n=1}^{N} 4^{-n}\sin(8^n\pi x)$ converge

to F **uniformly** on $\mathbb{R}$. Each F_N is continuous (a finite sum of continuous functions), and a uniform limit of continuous functions is continuous; hence F is continuous.

(ii) The statement as printed is false, and we prove a corrected version. Observe first that if $x = j/8^n$ with $j \in \mathbb{Z}$ and $n \geq 1$, then for every $m \geq n$ the quantity $8^m x = 8^{m-n} j$ is an integer, so $\sin(8^m \pi x) = 0$. Consequently

$$F\Big(\frac{j}{8^n}\Big) = \sum_{m=1}^{n-1} 4^{-m} \sin\Big(\frac{\pi j}{8^{n-m}}\Big) = 4^{-n} \sum_{k=1}^{n-1} 4^k \sin\Big(\frac{\pi j}{8^k}\Big),$$

(the substitution being $k = n - m$), and hence

$$F\Big(\frac{j+1}{8^n}\Big) - F\Big(\frac{j}{8^n}\Big) = 4^{-n} G_n(j), \qquad G_n(j) := \sum_{k=1}^{n-1} 4^k \Big[\sin\Big(\frac{\pi(j+1)}{8^k}\Big) - \sin\Big(\frac{\pi j}{8^k}\Big)\Big].$$

For $n = 1$ the sum defining $G_1(j)$ is empty, so $G_1(j) = 0$ for every j: the function F vanishes identically on $\frac{1}{8}\mathbb{Z}$. Thus for $n = 1$ the asserted inequality reads $0 \geq c/4$ with $c > 0$, which is false. The failure at $n = 1$ is already a complete refutation of the printed statement, which quantifies over all $n \geq 1$. Nor does the statement look repairable by discarding finitely many n: since $j \mapsto G_n(j)$ has period $2 \cdot 8^{n-1}$, the quantity $\min_j |G_n(j)|$ is computable by a finite search, and numerically it comes out as 0.304, 0.0877, 0.00644, 0.00113, 6.4×10^{-5}, 4.4×10^{-6} for $n = 2, \ldots, 7$ (rounded), i.e. it appears to decay geometrically rather than staying bounded below. The obstruction is structural: at the 8-adic rationals the **high** frequencies $m \geq n$, which are the ones carrying the large oscillation at scale 8^{-n}, all vanish simultaneously, leaving only the smooth low-frequency tail, and that tail may nearly cancel. What survives, and what part (iii) actually needs, is the following statement: **at every point, and at every scale 8^{-n}, F oscillates by at least a constant multiple of 4^{-n}.** We phrase it using a second-order difference, which annihilates the smooth part far more efficiently than a first difference does.

Lemma. For $x \in \mathbb{R}$, $n \geq 1$ and $\theta \in \{1, 2\}$ set $h := \theta 8^{-n}$ and

$$D_\theta(x, n) := F(x + h) - 2F\big(x + \tfrac{h}{2}\big) + F(x).$$

Then, choosing $\theta := 1$ if $|\cos(8^n \pi x)| \geq 2^{-1/2}$ and $\theta := 2$ otherwise, one has

$$|D_\theta(x, n)| \geq \tfrac{6}{5}\, 4^{-n}.$$

Proof of the Lemma. Fix x, n, θ and write $\phi_m := 8^m \pi h = \theta\, 8^{m-n} \pi$ and $\alpha_m := 8^m \pi x$. Since the series for F converges absolutely, all three of $F(x + h), F(x + \frac{h}{2}), F(x)$ are given by absolutely convergent series and we may combine them term by term. By the sum-to-product identity $\sin(u + \phi) + \sin u = 2 \sin(u + \frac{\phi}{2}) \cos\frac{\phi}{2}$,

$$\sin(\alpha_m + \phi_m) - 2 \sin\big(\alpha_m + \tfrac{\phi_m}{2}\big) + \sin \alpha_m = 2 \sin\big(\alpha_m + \tfrac{\phi_m}{2}\big)\Big(\cos\tfrac{\phi_m}{2} - 1\Big),$$

so that

$$D_\theta(x, n) = \sum_{m=1}^{\infty} 4^{-m} \cdot 2 \sin\big(\alpha_m + \tfrac{\phi_m}{2}\big)\Big(\cos\tfrac{\phi_m}{2} - 1\Big).$$

High frequencies $m > n$ vanish. Here $\frac{\phi_m}{2} = \frac{\theta}{2} 8^{m-n} \pi$. If $\theta = 2$ this is $8^{m-n}\pi$ with 8^{m-n} even, and if $\theta = 1$ it is $\frac{8^{m-n}}{2}\pi$ with $\frac{8^{m-n}}{2} = 4 \cdot 8^{m-n-1}$ again an even integer. In both cases $\cos\frac{\phi_m}{2} = 1$ and the term is 0.

The critical frequency $m = n$. If $\theta = 1$ then $\frac{\phi_n}{2} = \frac{\pi}{2}$ and $\cos\frac{\phi_n}{2} - 1 = -1$, so the $m = n$ term equals

$$4^{-n} \cdot 2 \sin\big(8^n \pi x + \tfrac{\pi}{2}\big) \cdot (-1) = -2 \cdot 4^{-n} \cos(8^n \pi x).$$

If $\theta = 2$ then $\frac{\phi_n}{2} = \pi$ and $\cos\frac{\phi_n}{2} - 1 = -2$, so the $m = n$ term equals

$$4^{-n} \cdot 2 \sin\big(8^n \pi x + \pi\big) \cdot (-2) = 4 \cdot 4^{-n} \sin(8^n \pi x).$$

Low frequencies $m < n$ are small. Using $|\sin| \le 1$ and $|\cos u - 1| \le \frac{u^2}{2}$ with $u = \frac{\phi_m}{2} = \frac{\theta}{2} 8^{m-n}\pi$,

$$\Big|4^{-m} \cdot 2\sin\big(\alpha_m + \tfrac{\phi_m}{2}\big)\big(\cos\tfrac{\phi_m}{2} - 1\big)\Big| \le 4^{-m}\Big(\tfrac{\theta}{2}8^{m-n}\pi\Big)^2 = \frac{\theta^2\pi^2}{4}\,16^m\,8^{-2n},$$

and since $\sum_{m=1}^{n-1} 16^m < \frac{16^n}{15}$ and $16^n 8^{-2n} = (16/64)^n = 4^{-n}$, the total contribution of the frequencies $m < n$ is at most

$$\frac{\theta^2\pi^2}{60}\,4^{-n}.$$

Combining. By the triangle inequality, if $\theta = 1$,

$$|D_1(x,n)| \ge 2\cdot 4^{-n}|\cos(8^n\pi x)| - \frac{\pi^2}{60}4^{-n} \ge \Big(\frac{2}{\sqrt{2}} - \frac{\pi^2}{60}\Big)4^{-n} > 1.24\cdot 4^{-n},$$

using $|\cos(8^n\pi x)| \ge 2^{-1/2}$ in this case; and if $\theta = 2$, then $|\cos(8^n\pi x)| < 2^{-1/2}$ forces $|\sin(8^n\pi x)| > 2^{-1/2}$, so

$$|D_2(x,n)| \ge 4\cdot 4^{-n}|\sin(8^n\pi x)| - \frac{4\pi^2}{60}4^{-n} \ge \Big(\frac{4}{\sqrt{2}} - \frac{\pi^2}{15}\Big)4^{-n} > 2.17\cdot 4^{-n}.$$

In either case $|D_\theta(x,n)| \ge \frac{6}{5}4^{-n}$, proving the Lemma. $\square$

Since $|D_\theta(x,n)| \le |F(x+h)-F(x+\frac{h}{2})| + |F(x+\frac{h}{2})-F(x)|$, the Lemma yields the following statement in the spirit of the printed part (ii): **for every $x \in \mathbb{R}$ and $n \ge 1$ there is an interval $[u,v] \subset [x, x+2\cdot 8^{-n}]$ of length $\frac{1}{2}8^{-n}$ or 8^{-n} with $|F(v)-F(u)| \ge \frac{3}{5}4^{-n}$.**

(iii). Suppose for contradiction that F is differentiable at some $x \in \mathbb{R}$. Then, writing $\rho(h) := F(x+h) - F(x) - hF'(x) = o(h)$ as $h \to 0$, we have for $h \ne 0$

$$F(x+h) - 2F\big(x+\tfrac{h}{2}\big) + F(x) = \big(hF'(x) + \rho(h)\big) - 2\big(\tfrac{h}{2}F'(x) + \rho(\tfrac{h}{2})\big) = \rho(h) - 2\rho(\tfrac{h}{2}),$$

and since $\rho(h) = o(h)$ and $\rho(h/2) = o(h/2) = o(h)$ as $h \to 0$, we conclude

$$F(x+h) - 2F\big(x+\tfrac{h}{2}\big) + F(x) = o(h) \qquad (h \to 0).$$

Now for each $n \ge 1$ apply the Lemma of part (ii) at this x, obtaining $\theta_n \in \{1,2\}$ and $h_n := \theta_n 8^{-n}$ with

$$\big|F(x+h_n) - 2F\big(x+\tfrac{h_n}{2}\big) + F(x)\big| \ge \tfrac{6}{5}4^{-n}.$$

Since $0 < h_n \le 2\cdot 8^{-n}$, we have $h_n \to 0$, while

$$\frac{\big|F(x+h_n) - 2F(x+\frac{h_n}{2}) + F(x)\big|}{h_n} \ge \frac{\frac{6}{5}4^{-n}}{2\cdot 8^{-n}} = \frac{3}{5}\cdot 2^n \longrightarrow \infty.$$

This contradicts the fact that the second difference is $o(h_n)$. Hence F is not differentiable at x; as x was arbitrary, F is nowhere differentiable. $\square$

Why formal term-by-term differentiation proves nothing. Differentiating the series termwise gives $\pi\sum_{n\ge1} 2^n\cos(8^n\pi x)$, whose terms do not tend to 0 for most x, so the series diverges. But divergence of the differentiated series is simply irrelevant to the differentiability of the sum: differentiation does not commute with infinite summation in the absence of a hypothesis such as uniform convergence of the differentiated series, and a series of smooth functions can perfectly well converge to a smooth function while its formal derivative series diverges. For an explicit example, set

$$u_n(x) := \frac{\sin(n!\,x)}{n} - \frac{\sin((n+1)!\,x)}{n+1} \qquad (n \ge 1),$$

so that the partial sums telescope: $\sum_{n=1}^{N} u_n(x) = \sin x - \frac{\sin((N+1)!\,x)}{N+1} \to \sin x$ for every x, since the remainder is bounded by $\frac{1}{N+1}$. Thus $\sum_{n\ge1} u_n = \sin$, which is differentiable everywhere. Yet at $x = 0$,

$$\sum_{n=1}^{N} u_n'(0) = \sum_{n=1}^{N}\Big(\frac{n!}{n} - \frac{(n+1)!}{n+1}\Big) = \sum_{n=1}^{N}\big((n-1)! - n!\big) = 1 - N! \to -\infty,$$

so the termwise differentiated series diverges at a point of differentiability of the sum. This is why the genuine argument in (iii) works with difference quotients of F itself rather than with the formal derivative series.

Problem 1.6.29. Show that every monotone function is measurable.

Solution. Let $F : \mathbb{R} \to \mathbb{R}$ be monotone. We may assume F is monotone non-decreasing: if F is non-increasing then $-F$ is non-decreasing, and once $-F$ is known to be measurable so is $F = -(-F)$ (multiplication by the constant -1 preserves measurability, by Exercise 1.3.8(v) applied to the continuous map $\varphi(z) := -z$).
So suppose F is non-decreasing. By Exercise 1.3.7(iv), F is measurable as soon as $F^{-1}(V)$ is Lebesgue measurable for every open $V \subseteq \mathbb{C}$. Since F takes real values, $F^{-1}(V) = F^{-1}(V \cap \mathbb{R})$, and $V \cap \mathbb{R}$ is an open subset of $\mathbb{R}$; conversely every open $U \subseteq \mathbb{R}$ arises this way (take $V := \{z \in \mathbb{C} : \mathrm{Re}(z) \in U,\ |\mathrm{Im}(z)| < 1\}$). So it suffices to show that $F^{-1}(U)$ is Lebesgue measurable for every open $U \subseteq \mathbb{R}$. Since every open subset of $\mathbb{R}$ is an at most countable union of open intervals (Exercise 1.6.10), and the Lebesgue measurable sets form a σ-algebra, it suffices to show that $F^{-1}((\alpha, \beta))$ is measurable for all $-\infty \leq \alpha < \beta \leq +\infty$. Writing $F^{-1}((\alpha, \beta)) = S_\alpha \setminus S_{\beta^-}$ where

$$S_\lambda := \{x \in \mathbb{R} : F(x) > \lambda\}, \qquad S_{\beta^-} := \{x \in \mathbb{R} : F(x) \geq \beta\} = \bigcap_{k=1}^{\infty} S_{\beta - 1/k}$$

(and with the conventions $S_{-\infty} = \mathbb{R}$ and $S_{+\infty^-} = \emptyset$), it suffices to prove that each S_λ, $\lambda \in \mathbb{R}$, is measurable.
Fix $\lambda \in \mathbb{R}$. The key point is that S_λ is an **up-set**: if $x \in S_\lambda$ and $y > x$, then by monotonicity $F(y) \geq F(x) > \lambda$, so $y \in S_\lambda$. Consequently, if S_λ is empty we are done ($\emptyset$ is measurable); otherwise set

$$c := \inf S_\lambda \in [-\infty, +\infty).$$

If $c = -\infty$, then every real y exceeds some element of S_λ, so the up-set property gives $S_\lambda = \mathbb{R}$, which is measurable. If $c \in \mathbb{R}$, then every $y > c$ exceeds some element of S_λ (by definition of the infimum), so $y \in S_\lambda$; and no $y < c$ lies in S_λ. Hence

$$(c, +\infty) \subseteq S_\lambda \subseteq [c, +\infty),$$

so S_λ is either $(c, +\infty)$ or $[c, +\infty)$ according as $F(c) \leq \lambda$ or $F(c) > \lambda$. In every case S_λ is an interval, hence Lebesgue measurable (indeed Borel measurable).
Therefore $F^{-1}(U)$ is measurable for every open $U \subseteq \mathbb{R}$, and F is measurable.
Two remarks. First, the argument shows more: the sets S_λ are intervals, hence Borel, so every monotone function is in fact Borel measurable. Second, the same proof works verbatim for a monotone function defined only on an interval $[a, b]$ (the sets S_λ are then relative up-sets of $[a, b]$, hence subintervals of $[a, b]$); and for a monotone function taking values in $[0, +\infty]$ one argues directly from Lemma 1.3.9(v), since the sets $\{x : F(x) > \lambda\}$ have just been shown to be intervals.

Problem 1.6.30. If F is monotone, show that the four Dini derivatives of F are measurable. (Hint: the main difficulty is to reformulate the derivatives so that h ranges over a countable set rather than an uncountable one.)

Solution. Recall the four Dini derivatives of F at x:

$$D^+F(x) := \limsup_{h \to 0^+} \frac{F(x+h) - F(x)}{h}, \quad D_+F(x) := \liminf_{h \to 0^+} \frac{F(x+h) - F(x)}{h},$$
$$D^-F(x) := \limsup_{h \to 0^-} \frac{F(x+h) - F(x)}{h}, \quad D_-F(x) := \liminf_{h \to 0^-} \frac{F(x+h) - F(x)}{h}.$$

Reductions. If F is monotone non-increasing then $-F$ is non-decreasing, and $D^+F = -D_+(-F)$, $D_+F = -D^+(-F)$, and similarly on the left. Negation preserves measurability here: if G takes values in $[0, +\infty]$ and is unsigned measurable, then for every $\lambda \in \mathbb{R}$ the set $\{x : -G(x) > \lambda\} = \{x : G(x) < -\lambda\}$ is measurable by Lemma 1.3.9(vii) (for $\lambda \geq 0$ the set is empty), so $-G$ is a measurable extended-real-valued function. Hence it suffices to treat the case where F is monotone non-decreasing. Next, put

$\tilde{F}(x) := -F(-x)$, which is again monotone non-decreasing. For $h > 0$,

$$\frac{\tilde{F}(-x+h) - \tilde{F}(-x)}{h} = \frac{-F(x-h) + F(x)}{h} = \frac{F(x+(-h)) - F(x)}{-h},$$

so letting $h \to 0^+$ we get $D^- F(x) = D^+ \tilde{F}(-x)$ and $D_- F(x) = D_+ \tilde{F}(-x)$. Since $y \mapsto -y$ is a measure-preserving linear change of variables, measurability of $D^+ \tilde{F}$ and $D_+ \tilde{F}$ implies measurability of $D^- F$ and $D_- F$. So it suffices to prove that $D^+ F$ and $D_+ F$ are measurable for F monotone non-decreasing. Note that in this case all four Dini derivatives take values in $[0, +\infty]$, so they are unsigned functions and measurability is in the sense of Lemma 1.3.9.

Notation. For $x \in \mathbb{R}$ and $h > 0$ write

$$\Delta(x, h) := \frac{F(x+h) - F(x)}{h} \in [0, +\infty),$$

which is non-negative because F is non-decreasing. For $\delta > 0$ set

$$S_\delta(x) := \sup_{0<h<\delta} \Delta(x, h), \qquad I_\delta(x) := \inf_{0<h<\delta} \Delta(x, h),$$

and let $S_\delta^{\mathbb{Q}}, I_\delta^{\mathbb{Q}}$ denote the same expressions with h restricted to the rationals in $(0, \delta)$. Since $\delta \mapsto S_\delta(x)$ is non-decreasing and $\delta \mapsto I_\delta(x)$ is non-increasing,

$$D^+ F(x) = \lim_{\delta \to 0^+} S_\delta(x) = \inf_{n \geq 1} S_{1/n}(x), \qquad D_+ F(x) = \lim_{\delta \to 0^+} I_\delta(x) = \sup_{n \geq 1} I_{1/n}(x).$$

The countability step. We claim that for every $\delta > 0$ and every x,

$$S_\delta(x) = S_\delta^{\mathbb{Q}}(x) \qquad \text{and} \qquad I_\delta(x) = I_\delta^{\mathbb{Q}}(x).$$

The inequalities $S_\delta^{\mathbb{Q}} \leq S_\delta$ and $I_\delta^{\mathbb{Q}} \geq I_\delta$ are trivial, since the rational h are among all h.

For the reverse inequality in the supremum case, fix $0 < h < \delta$ and $\varepsilon > 0$, and choose a rational h' with

$$h < h' < \min\big(\delta,\ (1+\varepsilon)h\big)$$

(possible since $\delta > h$ and $(1+\varepsilon)h > h$, and the rationals are dense). As F is non-decreasing and $x + h' > x + h$, we have $F(x+h') \geq F(x+h)$, and hence, using $F(x+h) - F(x) \geq 0$,

$$\Delta(x, h') = \frac{F(x+h') - F(x)}{h'} \geq \frac{F(x+h) - F(x)}{h'} \geq \frac{F(x+h) - F(x)}{(1+\varepsilon)h} = \frac{\Delta(x, h)}{1+\varepsilon}.$$

Taking the supremum over rational $h' \in (0, \delta)$ on the left and then over all $h \in (0, \delta)$ on the right gives $S_\delta^{\mathbb{Q}}(x) \geq S_\delta(x)/(1+\varepsilon)$; letting $\varepsilon \to 0$ (which is legitimate also when $S_\delta(x) = +\infty$) yields $S_\delta^{\mathbb{Q}}(x) \geq S_\delta(x)$, as required.

For the infimum case, fix $0 < h < \delta$ and $\varepsilon > 0$, and choose a rational h' with

$$\frac{h}{1+\varepsilon} < h' < h.$$

Then $x + h' < x + h$, so $F(x+h') \leq F(x+h)$, and hence

$$\Delta(x, h') = \frac{F(x+h') - F(x)}{h'} \leq \frac{F(x+h) - F(x)}{h'} \leq (1+\varepsilon)\,\frac{F(x+h) - F(x)}{h} = (1+\varepsilon)\Delta(x, h).$$

Taking infima as before gives $I_\delta^{\mathbb{Q}}(x) \leq (1+\varepsilon) I_\delta(x)$ for every $\varepsilon > 0$, hence $I_\delta^{\mathbb{Q}}(x) \leq I_\delta(x)$. This proves the claim.

Conclusion. By Exercise 1.6.29, F is measurable, and for each fixed $h > 0$ the translate $x \mapsto F(x+h)$ is measurable as well (translation of a measurable function is measurable, since translation preserves Lebesgue measurability of sets). Hence for each fixed $h > 0$ the function $x \mapsto \Delta(x, h) = h^{-1}\big(F(x+h) - F(x)\big)$ is measurable by Exercise 1.3.8(vi); it takes values in $[0, +\infty)$, so by the compatibility of

complex-valued and unsigned measurability noted after Exercise 1.3.7 it is unsigned measurable. By the claim, for each $n \geq 1$,

$$S_{1/n} = \sup_{h \in \mathbb{Q} \cap (0,1/n)} \Delta(\cdot, h), \qquad I_{1/n} = \inf_{h \in \mathbb{Q} \cap (0,1/n)} \Delta(\cdot, h)$$

are suprema and infima of **countably many** unsigned measurable functions, hence unsigned measurable by Exercise 1.3.3(iii). Applying Exercise 1.3.3(iii) once more,

$$D^+F = \inf_{n \geq 1} S_{1/n}, \qquad D_+F = \sup_{n \geq 1} I_{1/n}$$

are unsigned measurable. By the reductions at the start, all four Dini derivatives of a monotone F are measurable.

(The hypothesis of monotonicity was genuinely used twice: once to know that F itself is measurable, and once — crucially — to compare $\Delta(x, h)$ with $\Delta(x, h')$ for nearby h', which is what lets us replace the uncountable supremum/infimum by a countable one. For a general measurable F no such comparison is available, and indeed the individual functions $\Delta(\cdot, h)$ give no control on the uncountable supremum.)

Problem 1.6.31. Prove Lemma 1.6.26 in the discontinuous case. (Hint: the rising sun lemma is no longer available, but one can use either the Vitali-type covering lemma (which will give $C = 3$) or the Besicovitch lemma (which will give $C = 2$), by modifying the proof of Theorem 1.6.20.)

Solution. We must show: there is an absolute constant $C > 0$ such that for every monotone non-decreasing $F : [a, b] \to \mathbb{R}$ (no continuity assumed) and every $\lambda > 0$,

$$m\big(\{x \in [a, b] : D^+F(x) \geq \lambda\}\big) \leq C\, \frac{F(b) - F(a)}{\lambda},$$

and likewise for the other three Dini derivatives. We give the argument with $C = 3$ using the Vitali-type covering lemma, and then indicate the refinement to $C = 2$ using the one-dimensional Besicovitch lemma (Exercise 1.6.21).

Preliminary reductions. Extend F to all of $\mathbb{R}$ by setting $F(x) := F(a)$ for $x < a$ and $F(x) := F(b)$ for $x > b$; the extension is monotone non-decreasing on $\mathbb{R}$ and agrees with F on $[a, b]$, so the Dini derivatives at points of the open interval (a, b) are unchanged. By Exercise 1.6.30 the function D^+F is measurable, so all the sets below are measurable. The two endpoints a, b form a null set, so it suffices to bound the measure of subsets of (a, b).

Also, it suffices to prove the strict-inequality bound

$$m\big(\{x \in (a, b) : D^+F(x) > \lambda\}\big) \leq 3\, \frac{F(b) - F(a)}{\lambda} \qquad (\lambda > 0). \tag{$*$}$$

Indeed, granting $(*)$, for any $0 < \lambda' < \lambda$ we have $\{x \in (a, b) : D^+F(x) \geq \lambda\} \subseteq \{x \in (a, b) : D^+F(x) > \lambda'\}$, whence $m(\{D^+F \geq \lambda\} \cap (a, b)) \leq 3(F(b) - F(a))/\lambda'$; letting $\lambda' \uparrow \lambda$ gives the non-strict bound with the constant 3.

A covering lemma for intervals. In dimension $d = 1$ an open ball is exactly a bounded open interval, so the Vitali-type covering lemma (Lemma 1.6.22) is available verbatim; for the reader's convenience we recall its greedy proof in this notation. For a bounded interval I with centre c and length $|I|$, write $3I$ for the closed interval with centre c and length $3|I|$.

Claim: if $I_1, \ldots, I_n$ are bounded intervals in $\mathbb{R}$, then there is a subcollection $I_{j_1}, \ldots, I_{j_m}$ of pairwise disjoint intervals with $\bigcup_{i=1}^n I_i \subseteq \bigcup_{k=1}^m 3I_{j_k}$.

Proof of the claim: run the greedy algorithm of Lemma 1.6.22: choose I_{j_1} of maximal length among $I_1, \ldots, I_n$; discard every I_i meeting I_{j_1}; from the survivors choose I_{j_2} of maximal length; and so on until nothing is left. The selected intervals are pairwise disjoint by construction, and every I_i meets some selected I_{j_k} with $|I_{j_k}| \geq |I_i|$ (namely the first selected interval that caused I_i to be discarded, or I_i itself if I_i was selected). It remains to note that if two intervals I, J meet and $|I| \leq |J|$, then $I \subseteq 3J$: writing c, d for the centres and $\rho := |J|/2 \geq \sigma := |I|/2$ for the half-lengths, $I \cap J \neq \emptyset$ forces

$|d-c| \le \rho+\sigma \le 2\rho$, so any $y \in I$ satisfies $|y-c| \le |y-d|+|d-c| \le \sigma+2\rho \le 3\rho$, i.e. $y \in 3J$. This proves the claim. $\square$

Proof of $(*)$. Let $E := \{x \in (a,b) : D^+F(x) > \lambda\}$; this is measurable, so by inner regularity of Lebesgue measure (the same step used in the proof of Theorem 1.6.20) it suffices to show

$$m(K) \le 3\,\frac{F(b)-F(a)}{\lambda}$$

for every compact $K \subseteq E$.

Fix such a K; we may assume $K \neq \emptyset$, as otherwise there is nothing to prove. For each $x \in K$ we have $D^+F(x) > \lambda$, so by definition of the lim sup there exist arbitrarily small $h > 0$ with

$$\frac{F(x+h)-F(x)}{h} > \lambda;$$

since $x < b$, we may choose such an $h = h_x$ with $0 < h_x < b-x$. Set

$$I_x := [x,\ x+h_x] \subseteq [a,b], \qquad \text{so that} \qquad F(x+h_x)-F(x) > \lambda\,|I_x|.$$

The intervals $\{I_x\}_{x\in K}$ cover K (each contains its own left endpoint x). They are not open, so we cannot immediately invoke compactness; but each I_x is contained in the open interval $I_x' := (x-\eta_x, x+h_x)$, and since the inequality $F(x+h_x)-F(x) > \lambda h_x$ is strict we may pick $\eta_x > 0$ so small that

$$F(x+h_x)-F(x) > \lambda\,(h_x+\eta_x) = \lambda\,|I_x'|,$$

and also so small that $x-\eta_x > a-1$. Since F is non-decreasing and $x-\eta_x < x$, we have $F(x-\eta_x) \le F(x)$, so writing $I_x' = (\alpha_x, \beta_x)$ we get

$$F(\beta_x)-F(\alpha_x)\ \ge\ F(x+h_x)-F(x)\ >\ \lambda\,|I_x'|. \tag{$\dagger$}$$

The I_x' are open and cover the compact set K, so finitely many of them, say $J_1,\dots,J_n$, cover K. Applying the Claim to $J_1,\dots,J_n$ we obtain pairwise disjoint $J_1',\dots,J_m'$ among them with $\bigcup_i J_i \subseteq \bigcup_k 3J_k'$. Writing $J_k' = (\alpha_k,\beta_k)$, finite subadditivity and $(\dagger)$ give

$$m(K) \le m\Big(\bigcup_{k=1}^{m} 3J_k'\Big) \le \sum_{k=1}^{m} 3|J_k'| < \frac{3}{\lambda}\sum_{k=1}^{m}\big(F(\beta_k)-F(\alpha_k)\big).$$

Finally, the intervals J_k' are pairwise disjoint, so after relabelling we may assume $\alpha_1 < \beta_1 \le \alpha_2 < \beta_2 \le \cdots \le \alpha_m < \beta_m$. All the α_k, β_k lie in $[a-1, b]$, where the extended F is constant equal to $F(a)$ on $[a-1, a]$; hence, by the monotonicity of F, the sum telescopes into

$$\sum_{k=1}^{m}\big(F(\beta_k)-F(\alpha_k)\big) \le F(\beta_m)-F(\alpha_1) \le F(b)-F(a-1) = F(b)-F(a).$$

Therefore $m(K) \le 3(F(b)-F(a))/\lambda$, and $(*)$ follows. This establishes the lemma for D^+F with $C = 3$.

The other three Dini derivatives. Since $D_+F(x) \le D^+F(x)$ pointwise, the bound for D_+F follows immediately from the bound for D^+F. For the left-hand derivatives, apply what we have proved to $\tilde F(x) := -F(-x)$ on $[-b,-a]$: this is again monotone non-decreasing, its total increment is $\tilde F(-a) - \tilde F(-b) = F(b)-F(a)$, and, as computed in Exercise 1.6.30, $D^-F(x) = D^+\tilde F(-x)$ and $D_-F(x) = D_+\tilde F(-x)$. Since $y \mapsto -y$ preserves Lebesgue measure, the bounds transfer verbatim. This proves Lemma 1.6.26 in the discontinuous case with $C = 3$.

Improvement to $C = 2$ via Besicovitch. Retain the open intervals I_x' satisfying $(\dagger)$ and the finite subcover $J_1,\dots,J_n$ of K. By the one-dimensional Besicovitch covering lemma (Exercise 1.6.21) there is a subfamily $J_1',\dots,J_m'$ with $\bigcup_k J_k' = \bigcup_i J_i \supseteq K$ and with every point of $\mathbb{R}$ lying in at most two of the J_k'. We claim this family splits into two pairwise disjoint subfamilies. Order the $J_k' = (\alpha_k,\beta_k)$ so that $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_m$, and colour them greedily in this order with two colours. When we come to colour J_k', the intervals among $J_1',\dots,J_{k-1}'$ that meet J_k' are exactly those J_j' with $\alpha_j \le \alpha_k < \beta_j$;

there are finitely many of them, so for $\varepsilon > 0$ small enough the point $\alpha_k + \varepsilon$ lies in J_k' and in every one of them. Multiplicity at most two at that point forces there to be at most one such J_j', so at most one colour is forbidden and a colour is available. Two intervals of the same colour never meet, so each colour class is a pairwise disjoint family. Applying the telescoping estimate above separately to each of these two disjoint families (each of them being a family of disjoint subintervals of $[a-1, b]$),

$$\sum_{k=1}^{m} \big(F(\beta_k) - F(\alpha_k)\big) \le 2\big(F(b) - F(a)\big),$$

while (†) and countable (here finite) subadditivity give

$$m(K) \le \sum_{k=1}^{m} |J_k'| < \frac{1}{\lambda} \sum_{k=1}^{m} \big(F(\beta_k) - F(\alpha_k)\big) \le \frac{2}{\lambda}\big(F(b) - F(a)\big).$$

This yields the lemma with $C = 2$.
(As Remark 1.6.29 explains, the loss of the constant C here is what prevents one from running the argument of Lemma 1.6.28 in the discontinuous case: with a constant C in play one can no longer stop D^+F from being up to C times as large as D_-F. This is why the proof of Theorem 1.6.25 must instead route through the jump-function decomposition.)

Problem 1.6.32. Show that the decomposition of a bounded monotone non-decreasing function F into continuous F_c and jump components F_{pp} given by the above lemma is unique.

Solution. Suppose $F = F_c + F_{pp} = G_c + G_{pp}$, where F_c, G_c are continuous monotone non-decreasing and F_{pp}, G_{pp} are jump functions in the sense of Definition 1.6.30. We must show $F_c = G_c$ and $F_{pp} = G_{pp}$.
Step 0: normalising a jump function. Let $J = \sum_n c_n J_n$ be a jump function, with J_n the basic jump function with point of discontinuity x_n and fraction θ_n, with $c_n > 0$ and $\sum_n c_n < \infty$. First, we may assume the points x_n are distinct: if a subfamily S of indices shares a common point x_0, then $\sum_{n\in S} c_n J_n = c\, J'$ where $c := \sum_{n\in S} c_n$ and J' is the basic jump function with point x_0 and fraction $\theta := \big(\sum_{n\in S} c_n\theta_n\big)/c \in [0,1]$, as one checks by evaluating at $x < x_0$, $x = x_0$ and $x > x_0$. So from now on the x_n are distinct.
Step 1: a jump function vanishes at $-\infty$. Given $\varepsilon > 0$, pick N with $\sum_{n>N} c_n < \varepsilon$. For $x < \min_{n\le N} x_n$ we have $J_n(x) = 0$ for all $n \le N$, so $0 \le J(x) \le \sum_{n>N} c_n < \varepsilon$. Hence

$$\lim_{x\to-\infty} J(x) = 0.$$

Step 2: increments of a jump function are determined by its jump data. For $a < b$ we compute, using the absolute convergence of $\sum_n c_n$ to interchange the (finite) difference with the sum,

$$J(b) - J(a) = \sum_n c_n\big(J_n(b) - J_n(a)\big),$$

where, directly from Definition 1.6.30,

$$J_n(b) - J_n(a) = \begin{cases} 1 & \text{if } a < x_n < b, \\ 1-\theta_n & \text{if } x_n = a, \\ \theta_n & \text{if } x_n = b, \\ 0 & \text{if } x_n < a \text{ or } x_n > b. \end{cases}$$

Thus

$$J(b) - J(a) = \sum_{n : x_n \in (a,b)} c_n \;+\; (1-\theta_m) c_m \mathbf{1}_{a\in\{x_n\}} \;+\; \theta_k c_k \mathbf{1}_{b\in\{x_n\}}, \tag{1}$$

where m (resp. k) is the index with $x_m = a$ (resp. $x_k = b$), when such an index exists. So J is determined, up to the additive constant $J(-\infty)$ — which is 0 by Step 1 — by the family of pairs (c_n, θ_n) attached to the points x_n.

Step 3: the jump data of J is read off from one-sided limits. Fix $x \in \mathbb{R}$ and write $J_-(x) := \lim_{y\to x^-} J(y)$, $J_+(x) := \lim_{y\to x^+} J(y)$ (these exist by monotonicity). We claim

$$J_+(x) - J_-(x) = c_k, \qquad J(x) - J_-(x) = \theta_k c_k \tag{2}$$

if $x = x_k$ for some (necessarily unique) k, and $J_+(x) - J_-(x) = J(x) - J_-(x) = 0$ otherwise. Indeed, take $a < x$ in (1): the sum $\sum_{n:x_n\in[a,x)} c_n$ tends to 0 as $a \uparrow x$, since given $\varepsilon > 0$ we may pick N with $\sum_{n>N} c_n < \varepsilon$ and then take a close enough to x that no x_n with $n \le N$ lies in $[a, x)$; hence the first two terms of (1) tend to 0, and (1) gives $J(x) - J_-(x) = \theta_k c_k$ (interpreted as 0 when x is not one of the x_n). The computation of $J_+(x) - J(x) = (1 - \theta_k)c_k$ is identical: take $a := x$ and $b > x$ in (1) and let $b \downarrow x$, noting that both the first and the third terms of (1) are bounded by $\sum_{n:x_n\in(x,b]} c_n$, which tends to 0 by the same tail argument. Adding the two identities gives (2).
Step 4: the jump data of F_{pp} is determined by F. Since F_c is continuous, $F_-(x) = F_c(x) + (F_{pp})_-(x)$ and $F_+(x) = F_c(x) + (F_{pp})_+(x)$ for every x, and $F(x) = F_c(x) + F_{pp}(x)$. Subtracting,

$$F_+(x) - F_-(x) = (F_{pp})_+(x) - (F_{pp})_-(x), \qquad F(x) - F_-(x) = F_{pp}(x) - (F_{pp})_-(x).$$

By Step 3, this says: the set of discontinuity points of F_{pp} is exactly the set A of discontinuity points of F, and for $x \in A$ the corresponding coefficient and fraction are

$$c_x = F_+(x) - F_-(x), \qquad \theta_x = \frac{F(x) - F_-(x)}{F_+(x) - F_-(x)},$$

quantities depending on F alone. Exactly the same computation applies to the second decomposition, so G_{pp} has the very same jump data $\{(x, c_x, \theta_x)\}_{x\in A}$ as F_{pp}.
Step 5: conclusion. Since F_{pp} and G_{pp} have identical jump data, formula (1) gives $F_{pp}(b) - F_{pp}(a) = G_{pp}(b) - G_{pp}(a)$ for all $a < b$, i.e. $F_{pp} - G_{pp}$ is a constant κ. Letting $a \to -\infty$ and using Step 1 for both functions gives $\kappa = 0 - 0 = 0$. Hence $F_{pp} = G_{pp}$, and therefore $F_c = F - F_{pp} = F - G_{pp} = G_c$ as well.
(The normalisation $J(-\infty) = 0$, which Step 1 shows is automatic for every jump function in the sense of Definition 1.6.30, is what makes the decomposition unique on the nose rather than merely up to an additive constant: if constants were admitted as jump functions, then $(F_c + \kappa, F_{pp} - \kappa)$ would be another decomposition for every $\kappa \in \mathbb{R}$.)

Problem 1.6.33. Find a suitable generalisation of the notion of a jump function that allows one to extend the above decomposition to unbounded monotone functions, and then prove this extension. (Hint: the notion to shoot for here is that of a "locally jump function".)

Solution. The obstruction to Lemma 1.6.31(iii) for unbounded F is that the jumps c_x of F need no longer be globally summable (consider the floor function $F(x) := \lfloor x \rfloor$, which jumps by 1 at every integer, so that $\sum_{x\in A} c_x = +\infty$ and no jump function in the sense of Definition 1.6.30 can have these jumps); they are, however, always summable over bounded sets, as the estimate below shows. This suggests the following definition, together with a renormalisation of the basic jump functions to the left of the origin so that the defining series converges pointwise.
Definition (locally jump function). For a basic jump function J with point of discontinuity x_0 and fraction θ, define its **recentred** version

$$\tilde{J} := J - \mathbf{1}_{x_0<0}, \qquad \text{i.e.} \qquad \tilde{J}(x) = \begin{cases} J(x) & \text{if } x_0 \ge 0, \\ J(x) - 1 & \text{if } x_0 < 0. \end{cases}$$

A **locally jump function** is a function of the form $F = \sum_n c_n \tilde{J}_n$, where n ranges over an at most countable index set, the J_n are basic jump functions with pairwise distinct points of discontinuity x_n and fractions θ_n, the c_n are positive reals, and the summability is only local:

$$\sum_{n\,:\,|x_n|\le R} c_n < \infty \qquad \text{for every } R > 0.$$

Basic properties. Fix $R > 0$ and let $x \in [-R, R]$. If $x_n > R$ then $x < x_n$ and $x_n \geq 0$, so $\tilde{J}_n(x) = J_n(x) = 0$; if $x_n < -R$ then $x > x_n$ and $x_n < 0$, so $\tilde{J}_n(x) = 1 - 1 = 0$. Hence on $[-R, R]$ the series reduces to $\sum_{|x_n| \leq R} c_n \tilde{J}_n$, which converges absolutely and uniformly there because $|\tilde{J}_n| \leq 1$ and $\sum_{|x_n| \leq R} c_n < \infty$. In particular F is well defined; being a pointwise limit of non-decreasing functions (each $\tilde{J}_n$ is J_n shifted by a constant, hence monotone non-decreasing), F is monotone non-decreasing. Every jump function differs from a locally jump function by an additive constant: after the merging normalisation of Step 0 of Exercise 1.6.32 its points of discontinuity are distinct, the c_n are globally summable (so certainly locally so), and $\sum_n c_n J_n = \sum_n c_n \tilde{J}_n + \sum_{n: x_n < 0} c_n$. So this is a genuine generalisation, apart from the change of normalisation point from $-\infty$ to 0^-.

The increment formula (1) of Exercise 1.6.32 continues to hold verbatim, since $\tilde{J}_n$ and J_n differ by a constant: for $a < b$,

$$F(b) - F(a) = \sum_{n\,:\,x_n \in (a,b)} c_n \;+\; (1 - \theta_m) c_m \mathbf{1}_{a \in \{x_n\}} \;+\; \theta_k c_k \mathbf{1}_{b \in \{x_n\}}, \tag{1'}$$

(all sums finite, by local summability). The argument of Step 3 of Exercise 1.6.32 also goes through: to compute the one-sided limits at a point x, fix $R > |x| + 1$ and discard the terms with $|x_n| > R$, which do not contribute on $[-R, R]$; the remaining family has $\sum_{|x_n| \leq R} c_n < \infty$, so the tail argument of Step 3 applies verbatim to it. This gives, for every x,

$$F_+(x) - F_-(x) = c_k, \qquad F(x) - F_-(x) = \theta_k c_k \tag{2'}$$

when $x = x_k$, and $F_+(x) = F(x) = F_-(x)$ otherwise. Finally, the **normalisation**

$$\lim_{x \to 0^-} F(x) = 0 \tag{3'}$$

holds: for $x < 0$ only the terms with $x_n < 0$ and $x \leq x_n$ contribute, and $|F(x)| \leq \sum_{x_n \in [x, 0)} c_n$, which tends to 0 as $x \uparrow 0$ by the tail argument of Step 3 of Exercise 1.6.32 (choose N with $\sum_{|x_n| \leq 1, n > N} c_n < \varepsilon$, then take x close enough to 0 that no x_n with $n \leq N$ lies in $[x, 0)$). This replaces the normalisation $J(-\infty) = 0$ of ordinary jump functions, which is unavailable here.

Theorem (decomposition, unbounded case). Every monotone non-decreasing $F : \mathbb{R} \to \mathbb{R}$ can be written as $F = F_c + F_{pp}$, where F_c is continuous and monotone non-decreasing and F_{pp} is a locally jump function; and this decomposition is unique.

Proof of existence. Parts (i) and (ii) of Lemma 1.6.31 did not use boundedness, so the set A of discontinuities of F is at most countable and every $x \in A$ is a jump discontinuity, with $F_-(x) < F(x) \leq F_+(x)$ or $F_-(x) \leq F(x) < F_+(x)$. As in the proof of that lemma set, for $x \in A$,

$$c_x := F_+(x) - F_-(x) > 0, \qquad \theta_x := \frac{F(x) - F_-(x)}{F_+(x) - F_-(x)} \in [0, 1],$$

and let J_x be the basic jump function with point of discontinuity x and fraction θ_x.

Local summability: fix $R > 0$. For $x \in A \cap [-R, R]$ the intervals $(F_-(x), F_+(x))$ have length c_x, are pairwise disjoint (by monotonicity of F), and are all contained in the bounded interval $[F(-R - 1), F(R + 1)]$, since $F(-R - 1) \leq F_-(x)$ and $F_+(x) \leq F(R + 1)$ for $|x| \leq R$. By countable additivity,

$$\sum_{x \in A \cap [-R, R]} c_x \leq F(R + 1) - F(-R - 1) < \infty.$$

Hence $F_{pp} := \sum_{x \in A} c_x \tilde{J}_x$ is a locally jump function. (This is precisely where boundedness of F was used in Lemma 1.6.31 to obtain a **global** bound, and precisely what the local definition relaxes.)

$F_c := F - F_{pp}$ is continuous: by (2') the function F_{pp} has, at every point x, the one-sided limits

$$(F_{pp})_+(x) = (F_{pp})_-(x) + c_x, \qquad F_{pp}(x) = (F_{pp})_-(x) + \theta_x c_x$$

(with $c_x := 0$ for $x \notin A$), which is exactly the same relation satisfied by F by the definition of c_x, θ_x. Since one-sided limits are additive, $(F_c)_-(x) = F_-(x) - (F_{pp})_-(x)$ and $(F_c)_+(x) = F_+(x) - (F_{pp})_+(x)$, so

$$(F_c)_-(x) = F_-(x) - (F_{pp})_-(x) = \big(F(x) - \theta_x c_x\big) - \big(F_{pp}(x) - \theta_x c_x\big) = F_c(x),$$

and likewise $(F_c)_+(x) = F_c(x)$. Thus F_c is continuous at every x.
F_c is monotone non-decreasing: we must show $F_{pp}(b) - F_{pp}(a) \leq F(b) - F(a)$ for $a < b$. By (1'), the left-hand side equals

$$\sum_{x \in A \cap (a,b)} c_x + (1-\theta_a)c_a \, \mathbf{1}_{a \in A} + \theta_b c_b \, \mathbf{1}_{b \in A}.$$

Now c_x is the length of $(F_-(x), F_+(x))$ for $x \in A \cap (a,b)$; $(1-\theta_a)c_a = F_+(a) - F(a)$ is the length of $(F(a), F_+(a))$; and $\theta_b c_b = F(b) - F_-(b)$ is the length of $(F_-(b), F(b))$. By the monotonicity of F all of these intervals are pairwise disjoint and contained in $(F(a), F(b))$, so by countable additivity the sum is at most $F(b) - F(a)$, as required. This proves existence.
Proof of uniqueness. Suppose $F = F_c + F_{pp} = G_c + G_{pp}$ with F_c, G_c continuous and F_{pp}, G_{pp} locally jump functions. Exactly as in Step 4 of Exercise 1.6.32 (which used only continuity of the continuous part and (2')), both F_{pp} and G_{pp} have the same jump data, namely the discontinuity set A of F with coefficients $c_x = F_+(x) - F_-(x)$ and fractions $\theta_x = (F(x) - F_-(x))/c_x$. By the increment formula (1'), $F_{pp}(b) - F_{pp}(a) = G_{pp}(b) - G_{pp}(a)$ for all $a < b$, so $F_{pp} - G_{pp}$ is a constant; letting $x \to 0^-$ and using the normalisation (3') for both, that constant is 0. Hence $F_{pp} = G_{pp}$ and $F_c = G_c$. $\square$
Consistency with Lemma 1.6.31: if F is bounded, then taking $R \to \infty$ in the summability estimate above gives $\sum_{x \in A} c_x < \infty$, so F_{pp} differs from the honest jump function $\sum_{x \in A} c_x J_x$ by the constant $\sum_{x \in A, x < 0} c_x$, and one recovers the bounded decomposition of Lemma 1.6.31 (with the constant absorbed into F_c).

Problem 1.6.34. If $F : \mathbb{R} \to \mathbb{R}$ is a monotone function, show that $\|F\|_{TV([a,b])} = |F(b) - F(a)|$ for any interval $[a,b]$, and that F has bounded variation on $\mathbb{R}$ if and only if it is bounded.

Solution. Throughout we use Definition 1.6.33 in the form

$$\|F\|_{TV([a,b])} = \sup_{a \leq x_0 < x_1 < \cdots < x_n \leq b} \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})|,$$

the supremum being over all finite increasing sequences in $[a,b]$ with $n \geq 0$ (the empty sum, for $n = 0$, being 0).
Reduction to the non-decreasing case. If F is monotone non-increasing then $-F$ is monotone non-decreasing, and since $|(-F)(x_i) - (-F)(x_{i-1})| = |F(x_i) - F(x_{i-1})|$ for every partition, we have $\| -F\|_{TV(I)} = \|F\|_{TV(I)}$ for every I; also $|(-F)(b) - (-F)(a)| = |F(b) - F(a)|$, and $-F$ is bounded exactly when F is. So we may and do assume F is monotone non-decreasing.
The identity on $[a,b]$. If $a = b$ then the only admissible sequences have $n = 0$, so both sides are 0. Assume $a < b$. Let $a \leq x_0 < x_1 < \cdots < x_n \leq b$ be any admissible sequence. Since F is non-decreasing, each increment $F(x_i) - F(x_{i-1})$ is non-negative, so the absolute values may be dropped and the sum telescopes:

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| = \sum_{i=1}^{n} \big(F(x_i) - F(x_{i-1})\big) = F(x_n) - F(x_0) \leq F(b) - F(a),$$

the last step again by monotonicity ($x_0 \geq a$ and $x_n \leq b$). Taking the supremum gives $\|F\|_{TV([a,b])} \leq F(b) - F(a)$. Conversely, the admissible sequence $x_0 := a < x_1 := b$ contributes exactly $F(b) - F(a)$, so $\|F\|_{TV([a,b])} \geq F(b) - F(a)$. Hence

$$\|F\|_{TV([a,b])} = F(b) - F(a) = |F(b) - F(a)|,$$

the last equality because $F(b) \geq F(a)$.
Bounded variation on $\mathbb{R}$ versus boundedness. Every finite increasing sequence $x_0 < \cdots < x_n$ of reals lies in $[-N, N]$ for N large, and conversely every sequence in $[-N, N]$ is an admissible sequence for $\mathbb{R}$; hence, as recorded after Definition 1.6.33,

$$\|F\|_{TV(\mathbb{R})} = \sup_{N \geq 1} \|F\|_{TV([-N,N])} = \sup_{N \geq 1} \big(F(N) - F(-N)\big),$$

using the identity just proved. Note that $N \mapsto F(N) - F(-N)$ is non-decreasing, so this supremum is also the limit as $N \to \infty$.
If F is bounded, say $|F| \leq M$, then $F(N) - F(-N) \leq 2M$ for every N, so $\|F\|_{TV(\mathbb{R})} \leq 2M < \infty$ and F has bounded variation.
Conversely, suppose $L := \|F\|_{TV(\mathbb{R})} < \infty$. Let $x \in \mathbb{R}$ be arbitrary and choose $N \geq |x|$. If $x \geq 0$, then monotonicity gives $F(0) \leq F(x) \leq F(N)$ and $F(-N) \leq F(0)$, so

$$F(0) \leq F(x) \leq F(N) \leq F(-N) + L \leq F(0) + L.$$

If $x \leq 0$, then similarly $F(-N) \leq F(x) \leq F(0)$ and $F(N) \geq F(0)$, so

$$F(0) \geq F(x) \geq F(-N) \geq F(N) - L \geq F(0) - L.$$

In both cases $|F(x) - F(0)| \leq L$, so F is bounded (by $|F(0)| + L$). This completes the proof.

Problem 1.6.35. For any functions $F, G : \mathbb{R} \to \mathbb{R}$, establish the triangle property $\|F + G\|_{TV(\mathbb{R})} \leq \|F\|_{TV(\mathbb{R})} + \|G\|_{TV(\mathbb{R})}$ and the homogeneity property $\|cF\|_{TV(\mathbb{R})} = |c|\|F\|_{TV(\mathbb{R})}$ for any $c \in \mathbb{R}$. Also show that $\|F\|_{TV} = 0$ if and only if F is constant.

Solution. All three claims follow from the fact that the quantities being supremised already obey the corresponding properties partition by partition. Throughout, fix an arbitrary admissible finite increasing sequence $x_0 < x_1 < \cdots < x_n$ of reals ($n \geq 0$), and write

$$V(H; x_0, \ldots, x_n) := \sum_{i=1}^{n} |H(x_i) - H(x_{i-1})| \in [0, +\infty)$$

for a function $H : \mathbb{R} \to \mathbb{R}$, so that $\|H\|_{TV(\mathbb{R})} = \sup V(H; x_0, \ldots, x_n)$, the supremum being over all such sequences, and taking values in $[0, +\infty]$.
Triangle inequality. By the triangle inequality in $\mathbb{R}$ applied termwise,

$$V(F+G; x_0, \ldots, x_n) = \sum_{i=1}^{n} \big|\big(F(x_i) - F(x_{i-1})\big) + \big(G(x_i) - G(x_{i-1})\big)\big| \leq V(F; x_0, \ldots, x_n) + V(G; x_0, \ldots, x_n),$$

and the right-hand side is at most $\|F\|_{TV(\mathbb{R})} + \|G\|_{TV(\mathbb{R})}$ (an inequality in $[0, +\infty]$, valid also when one of the two norms is infinite). Taking the supremum over all admissible sequences on the left gives

$$\|F + G\|_{TV(\mathbb{R})} \leq \|F\|_{TV(\mathbb{R})} + \|G\|_{TV(\mathbb{R})}.$$

Homogeneity. For any $c \in \mathbb{R}$,

$$V(cF; x_0, \ldots, x_n) = \sum_{i=1}^{n} |c|\,|F(x_i) - F(x_{i-1})| = |c|\,V(F; x_0, \ldots, x_n).$$

If $c \neq 0$, then multiplication by the positive constant $|c|$ is an increasing bijection of $[0, +\infty]$ and so commutes with suprema; taking suprema over all admissible sequences therefore gives $\|cF\|_{TV(\mathbb{R})} = |c|\,\|F\|_{TV(\mathbb{R})}$, including the case where the right-hand side is $+\infty$. If $c = 0$ then cF is the zero function, all of whose sums V vanish, so $\|cF\|_{TV(\mathbb{R})} = 0 = |c|\,\|F\|_{TV(\mathbb{R})}$ under the convention $0 \cdot (+\infty) = 0$ used throughout the text.
Vanishing of the norm. If F is constant then every increment $F(x_i) - F(x_{i-1})$ vanishes, so $V(F; x_0, \ldots, x_n) = 0$ for every admissible sequence and hence $\|F\|_{TV(\mathbb{R})} = 0$. Conversely, suppose $\|F\|_{TV(\mathbb{R})} = 0$, and let $x < y$ be arbitrary reals. Applying the definition to the admissible sequence $x_0 := x < x_1 := y$ gives

$$|F(y) - F(x)| = V(F; x, y) \leq \|F\|_{TV(\mathbb{R})} = 0,$$

so $F(y) = F(x)$. As $x < y$ were arbitrary, F is constant.
(Thus $\|\cdot\|_{TV(\mathbb{R})}$ is a seminorm on the space of functions $F : \mathbb{R} \to \mathbb{R}$ which vanishes exactly on the constants; it becomes a genuine norm on the quotient by the constants, or equivalently on the space of bounded-variation functions normalised by $\lim_{x \to -\infty} F(x) = 0$, which exists by Exercise 1.6.37(i).)

Problem 1.6.36. If $F : \mathbb{R} \to \mathbb{R}$ is a function, show that $\|F\|_{TV([a,b])} + \|F\|_{TV([b,c])} = \|F\|_{TV([a,c])}$ whenever $a \leq b \leq c$.

Solution. Throughout, a *partition of* $[a,b]$ means a finite increasing sequence $a \leq x_0 < x_1 < \cdots < x_n \leq b$ with $n \geq 0$; we identify it with the finite set $P = \{x_0, \ldots, x_n\} \subseteq [a,b]$, and write

$$S(P) := \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})|$$

for the associated sum (interpreted as 0 when $n = 0$, i.e. when P is a singleton or empty). With this notation Definition 1.6.33 reads $\|F\|_{TV([a,b])} = \sup_{P \subseteq [a,b]} S(P)$, the supremum being over all finite subsets P of $[a,b]$, and taking values in $[0, +\infty]$.
Step 1: refinement increases S. If $P \subseteq P'$ are finite subsets of $\mathbb{R}$, then $S(P) \leq S(P')$.
It suffices to treat the case where $P' = P \cup \{y\}$ consists of one extra point $y \notin P$, the general case following by induction on $|P' \setminus P|$. Write $P = \{x_0 < \cdots < x_n\}$. If $y < x_0$ or $y > x_n$, then $S(P')$ is $S(P)$ plus the single extra nonnegative term $|F(x_0) - F(y)|$ resp. $|F(y) - F(x_n)|$, so $S(P') \geq S(P)$. Otherwise $x_{i-1} < y < x_i$ for exactly one i, and $S(P')$ is obtained from $S(P)$ by replacing the term $|F(x_i) - F(x_{i-1})|$ by $|F(y) - F(x_{i-1})| + |F(x_i) - F(y)|$, which is at least as large by the triangle inequality. (If $P = \emptyset$ the claim is trivial.)
Step 2: sums split at a common point. Suppose R is a finite set with $b \in R$, and put $R_1 := R \cap (-\infty, b]$ and $R_2 := R \cap [b, +\infty)$, so that $R = R_1 \cup R_2$ and $R_1 \cap R_2 = \{b\}$. Then

$$S(R) = S(R_1) + S(R_2).$$

Indeed, listing R in increasing order as $z_0 < \cdots < z_m$, the point b occurs as some z_k; every consecutive pair (z_{j-1}, z_j) has either $j \leq k$ (both endpoints in R_1) or $j > k$ (both endpoints in R_2), and these two families of consecutive pairs are exactly the consecutive pairs of R_1 and of R_2 respectively. Summing the corresponding terms gives the identity.
Step 3: the inequality $\|F\|_{TV([a,b])} + \|F\|_{TV([b,c])} \leq \|F\|_{TV([a,c])}$. Let $P \subseteq [a,b]$ and $Q \subseteq [b,c]$ be finite. Put $R := P \cup Q \cup \{b\} \subseteq [a,c]$. By Step 1, $S(P) \leq S(P \cup \{b\})$ and $S(Q) \leq S(Q \cup \{b\})$, and by Step 2 applied to R (noting $R \cap (-\infty, b] = P \cup \{b\}$ and $R \cap [b, +\infty) = Q \cup \{b\}$, since $P \subseteq [a,b]$ and $Q \subseteq [b,c]$),

$$S(P) + S(Q) \leq S(P \cup \{b\}) + S(Q \cup \{b\}) = S(R) \leq \|F\|_{TV([a,c])}.$$

Taking the supremum first over P and then over Q gives $\|F\|_{TV([a,b])} + \|F\|_{TV([b,c])} \leq \|F\|_{TV([a,c])}$. (If one of the two total variations on the left is infinite, the displayed inequality shows the right-hand side is infinite too, so the inequality still holds in $[0, +\infty]$.)
Step 4: the reverse inequality. Let $Z \subseteq [a,c]$ be finite. By Step 1, $S(Z) \leq S(Z \cup \{b\})$, and by Step 2 applied to $R := Z \cup \{b\}$,

$$S(Z) \leq S(R_1) + S(R_2) \leq \|F\|_{TV([a,b])} + \|F\|_{TV([b,c])},$$

since $R_1 = R \cap (-\infty, b] \subseteq [a,b]$ and $R_2 = R \cap [b, +\infty) \subseteq [b,c]$. Taking the supremum over Z gives $\|F\|_{TV([a,c])} \leq \|F\|_{TV([a,b])} + \|F\|_{TV([b,c])}$.
Combining Steps 3 and 4 yields the claimed identity in $[0, +\infty]$. (The degenerate cases are consistent: if $a = b$ then every finite subset of $[a,b]$ is a singleton or empty, so $\|F\|_{TV([a,a])} = 0$ and the identity is trivial.)

Problem 1.6.37. (i) Show that every function $f : \mathbb{R} \to \mathbb{R}$ of bounded variation is bounded, and that the limits $\lim_{x \to +\infty} f(x)$ and $\lim_{x \to -\infty} f(x)$ are well-defined.
(ii) Give a counterexample of a bounded, continuous, compactly supported function f that is not of bounded variation.

Solution. As in Exercise 1.6.36 we write $S(P) := \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$ for a finite set $P = \{x_0 < \cdots < x_n\}$ of reals, so that $\|f\|_{TV(\mathbb{R})} = \sup_P S(P)$ by Definition 1.6.33.
Part (i). Suppose $V := \|f\|_{TV(\mathbb{R})} < \infty$. For any $x \neq 0$ the two-point set $P = \{0, x\}$ (in increasing order) is admissible, and $S(P) = |f(x) - f(0)|$; hence $|f(x) - f(0)| \leq V$ for all x, and so

$$|f(x)| \leq |f(0)| + V \qquad \text{for all } x \in \mathbb{R},$$

i.e. f is bounded.
Now we show $\lim_{x\to+\infty} f(x)$ exists. Since f is bounded, the quantities

$$L := \limsup_{x\to+\infty} f(x), \qquad \ell := \liminf_{x\to+\infty} f(x)$$

are finite real numbers with $\ell \leq L$. Suppose for contradiction that $\ell < L$, and set $\varepsilon := (L - \ell)/3 > 0$. By the definition of lim sup, for every M there exists $x > M$ with $f(x) > L - \varepsilon$; by the definition of lim inf, for every M there exists $y > M$ with $f(y) < \ell + \varepsilon$.
Fix $N \geq 1$. Construct recursively real numbers $z_1 < z_2 < \cdots < z_{2N}$ as follows: having chosen $z_1 < \cdots < z_{j-1}$ (with $z_0 := 0$ say), choose $z_j > z_{j-1}$ with $f(z_j) > L - \varepsilon$ if j is odd, and $f(z_j) < \ell + \varepsilon$ if j is even; this is possible by the previous paragraph. Then for each $j = 2, \ldots, 2N$ we have

$$|f(z_j) - f(z_{j-1})| \geq (L - \varepsilon) - (\ell + \varepsilon) = L - \ell - 2\varepsilon = \varepsilon,$$

and therefore $S(\{z_1, \ldots, z_{2N}\}) \geq (2N - 1)\varepsilon$. Since N was arbitrary this forces $\|f\|_{TV(\mathbb{R})} = +\infty$, a contradiction. Hence $\ell = L$, and since this common value is finite, $\lim_{x\to+\infty} f(x)$ exists (and is finite).
The argument at $-\infty$ is identical: with $L' := \limsup_{x\to-\infty} f(x)$ and $\ell' := \liminf_{x\to-\infty} f(x)$ finite, if $\ell' < L'$ we build a *decreasing* sequence $z_1 > z_2 > \cdots > z_{2N}$ tending to $-\infty$ on which f alternates between values $> L' - \varepsilon$ and values $< \ell' + \varepsilon$; reading it in increasing order gives an admissible partition with $S \geq (2N - 1)\varepsilon$, again contradicting bounded variation. So $\lim_{x\to-\infty} f(x)$ exists as well.
Part (ii). Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) := \begin{cases} x \sin(\pi/x), & 0 < x \leq 1, \\ 0, & x \leq 0 \text{ or } x \geq 1. \end{cases}$$

This is well-defined at $x = 1$, since $1 \cdot \sin(\pi) = 0$.
The function f is continuous: on $(0, 1)$ it is a composition and product of continuous functions; at $x = 1$ both formulas give 0 and $x \sin(\pi/x) \to \sin \pi = 0$ as $x \to 1^-$; at $x = 0$ we have $|f(x)| \leq |x| \to 0 = f(0)$ for $0 < x \leq 1$, and $f \equiv 0$ to the left; and f is identically 0 on the open sets $(-\infty, 0)$ and $(1, \infty)$. It is bounded, since $|f(x)| \leq |x| \leq 1$ everywhere, and it is supported in $[0, 1]$, hence compactly supported.
Now set $x_k := \dfrac{2}{2k+1}$ for $k = 1, 2, \ldots$, so that $0 < x_k \leq 2/3 < 1$ and $x_{k+1} < x_k$. We compute

$$\sin(\pi/x_k) = \sin\left(\frac{(2k+1)\pi}{2}\right) = (-1)^k,$$

so $f(x_k) = (-1)^k \frac{2}{2k+1}$. In particular consecutive values $f(x_k)$ and $f(x_{k+1})$ have opposite signs, whence

$$|f(x_{k+1}) - f(x_k)| = \frac{2}{2k+3} + \frac{2}{2k+1} \geq \frac{2}{2k+3}.$$

Fix N and take the partition $x_{N+1} < x_N < \cdots < x_1$ (these are $N + 1$ points in increasing order). Its sum is

$$S = \sum_{k=1}^{N} |f(x_{k+1}) - f(x_k)| \geq \sum_{k=1}^{N} \frac{2}{2k+3},$$

which diverges to $+\infty$ as $N \to \infty$ by comparison with the harmonic series. Hence $\|f\|_{TV(\mathbb{R})} = +\infty$, so this bounded, continuous, compactly supported f is not of bounded variation.

Problem 1.6.38. Let $f : \mathbb{R} \to \mathbb{R}$ be an absolutely integrable function, and let $F : \mathbb{R} \to \mathbb{R}$ be the indefinite integral $F(x) := \int_{[-\infty,x]} f(y)\,dy$. Show that F is of bounded variation, and that $\|F\|_{TV(\mathbb{R})} = \|f\|_{L^1(\mathbb{R})}$. (Hint: the upper bound $\|F\|_{TV(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})}$ is relatively easy to establish. To obtain the lower bound, use the density argument.)

Solution. For $g \in L^1(\mathbb{R})$ write $I_g(x) := \int_{(-\infty,x]} g(y)\,dy$ for its indefinite integral (the set $\{-\infty\}$ carries no measure, so this is the same as the F of the statement), and set

$$T(g) := \|I_g\|_{TV(\mathbb{R})} \in [0, +\infty].$$

We must show $T(f) = \|f\|_{L^1(\mathbb{R})}$ for every $f \in L^1(\mathbb{R})$. Note first that I_g is well defined and real-valued since g is absolutely integrable, and that for $u < v$ we have, by additivity of the integral over the disjoint decomposition $(-\infty, v] = (-\infty, u] \sqcup (u, v]$,

$$I_g(v) - I_g(u) = \int_{(u,v]} g(y)\,dy.$$

The upper bound. Let $x_0 < x_1 < \cdots < x_n$ be any finite increasing sequence. By the display above and the triangle inequality for the integral (Lemma 1.3.19),

$$\sum_{i=1}^{n} |I_g(x_i) - I_g(x_{i-1})| = \sum_{i=1}^{n} \left| \int_{(x_{i-1},x_i]} g \right| \leq \sum_{i=1}^{n} \int_{(x_{i-1},x_i]} |g|.$$

The intervals $(x_{i-1}, x_i]$, $i = 1, \ldots, n$, are pairwise disjoint subsets of $\mathbb{R}$, so by additivity and monotonicity of the unsigned integral the right-hand side is at most $\int_{\mathbb{R}} |g| = \|g\|_{L^1(\mathbb{R})}$. Taking the supremum over all partitions gives

$$T(g) \leq \|g\|_{L^1(\mathbb{R})} \qquad \text{for all } g \in L^1(\mathbb{R}). \tag{1}$$

In particular $T(f) \leq \|f\|_{L^1(\mathbb{R})} < \infty$, so $F = I_f$ is of bounded variation.
A stability estimate. Let $g, h \in L^1(\mathbb{R})$. By linearity of the integral (Exercise 1.3.19) we have $I_h = I_g + I_{h-g}$ pointwise, so the triangle inequality for total variation (Exercise 1.6.35) gives

$$T(h) = \|I_g + I_{h-g}\|_{TV(\mathbb{R})} \leq T(g) + T(h-g) \leq T(g) + \|h-g\|_{L^1(\mathbb{R})},$$

the last step by (1). Rearranging (all three quantities being finite, by (1)),

$$T(g) \geq T(h) - \|h-g\|_{L^1(\mathbb{R})}. \tag{2}$$

The dense subclass: step functions. Let g be a step function on $\mathbb{R}$, that is (in the terminology preceding Theorem 1.3.20) a finite linear combination of indicator functions of boxes; in dimension $d = 1$ a box is a bounded interval. The finitely many endpoints of these intervals cut $\mathbb{R}$ into finitely many points and finitely many open intervals, on each of which g is constant; moreover g vanishes outside a bounded set. Discarding the pieces on which g vanishes and the finitely many cut points (a set of measure zero, which affects neither $\|g\|_{L^1}$ nor I_g), we may write

$$g = \sum_{j=1}^{k} c_j 1_{[a_j,b_j)} \quad \text{almost everywhere,}$$

with $c_j \neq 0$ and $a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_k < b_k$. Then

$$\|g\|_{L^1(\mathbb{R})} = \sum_{j=1}^{k} |c_j|(b_j - a_j).$$

Consider the finite increasing sequence obtained by listing the points $a_1, b_1, a_2, b_2, \ldots, a_k, b_k$ in increasing order (deleting repetitions when $b_j = a_{j+1}$). For each j we have $I_g(b_j) - I_g(a_j) = \int_{(a_j,b_j]} g =$

$c_j(b_j - a_j)$, while for the intervening gaps $I_g(a_{j+1}) - I_g(b_j) = \int_{(b_j, a_{j+1}]} g = 0$. Hence the sum of absolute consecutive differences along this partition is exactly $\sum_{j=1}^{k} |c_j|(b_j - a_j) = \|g\|_{L^1(\mathbb{R})}$, and therefore

$$T(g) \geq \|g\|_{L^1(\mathbb{R})}.$$

Combined with (1), $T(g) = \|g\|_{L^1(\mathbb{R})}$ for every step function g.
The density argument. Let $f \in L^1(\mathbb{R})$ and $\varepsilon > 0$. By Theorem 1.3.20(ii) there is a step function g with $\|f - g\|_{L^1(\mathbb{R})} \leq \varepsilon$. By (2) with $h = g$, and then by the previous paragraph and the triangle inequality in L^1,

$$T(f) \geq T(g) - \|g - f\|_{L^1(\mathbb{R})} = \|g\|_{L^1(\mathbb{R})} - \|g - f\|_{L^1(\mathbb{R})} \geq \|f\|_{L^1(\mathbb{R})} - 2\varepsilon.$$

Letting $\varepsilon \to 0$ gives $T(f) \geq \|f\|_{L^1(\mathbb{R})}$, and with (1) we conclude

$$\|F\|_{TV(\mathbb{R})} = T(f) = \|f\|_{L^1(\mathbb{R})},$$

as required.

Problem 1.6.39. Let $F : \mathbb{R} \to \mathbb{R}$ be of bounded variation. Define the positive variation F^+ by (1.30), i.e.

$$F^+(x) := \sup_{x_0 < \cdots < x_n \leq x} \sum_{i=1}^{n} \max(F(x_i) - F(x_{i-1}), 0),$$

and the negative variation F^- by

$$F^-(x) := \sup_{x_0 < \cdots < x_n \leq x} \sum_{i=1}^{n} \max(-F(x_i) + F(x_{i-1}), 0).$$

Establish the identities

$$F(x) = F(-\infty) + F^+(x) - F^-(x),$$

$$\|F\|_{TV([a,b])} = F^+(b) - F^+(a) + F^-(b) - F^-(a),$$

and

$$\|F\|_{TV} = F^+(+\infty) + F^-(+\infty)$$

for every interval $[a, b]$, where $F(-\infty) := \lim_{x \to -\infty} F(x)$, $F^+(+\infty) := \lim_{x \to +\infty} F^+(x)$, and $F^-(+\infty) := \lim_{x \to +\infty} F^-(x)$. (Hint: The main difficulty comes from the fact that a partition $x_0 < \cdots < x_n \leq x$ that is good for F^+ need not be good for F^-, and vice versa. However, this can be fixed by taking a good partition for F^+ and a good partition for F^- and combining them together into a common refinement.)

Solution. Notation. For a finite set $P = \{x_0 < x_1 < \cdots < x_n\} \subset \mathbb{R}$ put

$$S^+(P) := \sum_{i=1}^{n} \max(F(x_i) - F(x_{i-1}), 0), \quad S^-(P) := \sum_{i=1}^{n} \max(F(x_{i-1}) - F(x_i), 0),$$

$$S(P) := \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})|,$$

all of them 0 when $n = 0$. Since $\max(t, 0) + \max(-t, 0) = |t|$ and $\max(t, 0) - \max(-t, 0) = t$, we have for every such P the two basic identities

$$S^+(P) + S^-(P) = S(P), \qquad S^+(P) - S^-(P) = F(x_n) - F(x_0), \tag{a}$$

the second by telescoping. With this notation, $F^+(x) = \sup\{S^+(P) : P \subset (-\infty, x] \text{ finite}\}$ and $F^-(x) = \sup\{S^-(P) : P \subset (-\infty, x] \text{ finite}\}$. Observe also that $S^-(P)$ is the S^+-sum of the function $-F$; thus

$$F^- = (-F)^+, \tag{b}$$

and since $\| - F\|_{TV} = \|F\|_{TV}$ by Exercise 1.6.35, every statement we prove for the positive variation of an arbitrary bounded variation function applies verbatim to F^- upon replacing F by $-F$. We will use this repeatedly.

Since $0 \le S^{\pm}(P) \le S(P) \le \|F\|_{TV(\mathbb{R})} < \infty$, the functions F^+ and F^- take values in $[0, \|F\|_{TV}]$; in particular they are finite. They are non-decreasing, since enlarging x enlarges the family of admissible P.

Lemma A (refinement). If $P \subseteq P'$ are finite subsets of $\mathbb{R}$, then $S^+(P) \le S^+(P')$ and $S^-(P) \le S^-(P')$.

It suffices to insert one point $y \notin P$, and by (b) it suffices to treat S^+. Write $P = \{x_0 < \cdots < x_n\}$. If y lies outside $[x_0, x_n]$ then $S^+(P')$ equals $S^+(P)$ plus one extra nonnegative term. If $x_{i-1} < y < x_i$, then the term $\max(F(x_i) - F(x_{i-1}), 0)$ of $S^+(P)$ is replaced in $S^+(P')$ by $\max(F(y) - F(x_{i-1}), 0) + \max(F(x_i) - F(y), 0)$, which is at least as large because $\max(s + t, 0) \le \max(s, 0) + \max(t, 0)$ applied to $s = F(y) - F(x_{i-1})$, $t = F(x_i) - F(y)$.

Lemma B (splitting). If R is finite and $c \in R$, then with $R_1 := R \cap (-\infty, c]$ and $R_2 := R \cap [c, +\infty)$ we have $S^{\pm}(R) = S^{\pm}(R_1) + S^{\pm}(R_2)$.

Indeed, listing R in increasing order, the point c occurs in the list, and every consecutive pair of R lies entirely in R_1 or entirely in R_2; these pairs are exactly the consecutive pairs of R_1 together with those of R_2.

Lemma C (common refinement). Let $J \subseteq \mathbb{R}$ be an interval (possibly unbounded). Then

$$\sup_{P \subset J \text{ finite}} S(P) = \sup_{P \subset J} S^+(P) + \sup_{P \subset J} S^-(P).$$

Abbreviate $A := \sup_{P \subset J} S^+(P)$ and $B := \sup_{P \subset J} S^-(P)$; both are finite, being at most $\|F\|_{TV}$. The inequality $\le$ is immediate from $S = S^+ + S^-$, which gives $S(P) \le A + B$ for every finite $P \subset J$. For $\ge$, let $\varepsilon > 0$ and choose finite $P^+, P^- \subset J$ with $S^+(P^+) \ge A - \varepsilon$ and $S^-(P^-) \ge B - \varepsilon$. Let $P^* := P^+ \cup P^- \subset J$ be their common refinement. By Lemma A,

$$S(P^*) = S^+(P^*) + S^-(P^*) \ge S^+(P^+) + S^-(P^-) \ge A + B - 2\varepsilon,$$

and letting $\varepsilon \to 0$ gives the claim. This is exactly the manoeuvre suggested in the hint.

Identity 1: $F(x) = F(-\infty) + F^+(x) - F^-(x)$.

First, $F(-\infty)$ exists (and is finite) by Exercise 1.6.37(i), since F has bounded variation. Fix $x \in \mathbb{R}$ and let $\varepsilon > 0$. Choose finite $P^+, P^- \subset (-\infty, x]$ with $S^+(P^+) \ge F^+(x) - \varepsilon$ and $S^-(P^-) \ge F^-(x) - \varepsilon$. Choose also $M > 0$ so large that

$$|F(y) - F(-\infty)| \le \varepsilon \qquad \text{whenever } y \le -M,$$

and pick a real z with $z \le -M$, $z < x$, and $z < y$ for every $y \in P^+ \cup P^-$; this is possible because $P^+ \cup P^-$ is a finite (possibly empty) set. Set

$$P := P^+ \cup P^- \cup \{z, x\} \subset (-\infty, x],$$

so that $\min P = z$ and $\max P = x$. By Lemma A, $S^+(P) \ge S^+(P^+) \ge F^+(x) - \varepsilon$, and also $S^+(P) \le F^+(x)$ since $P \subset (-\infty, x]$; hence $|S^+(P) - F^+(x)| \le \varepsilon$, and likewise $|S^-(P) - F^-(x)| \le \varepsilon$. By the second identity in (a),

$$F(x) - F(z) = S^+(P) - S^-(P),$$

so

$$|(F(x) - F(z)) - (F^+(x) - F^-(x))| \le 2\varepsilon.$$

Since $|F(z) - F(-\infty)| \le \varepsilon$, the triangle inequality gives

$$|(F(x) - F(-\infty)) - (F^+(x) - F^-(x))| \le 3\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $F(x) - F(-\infty) = F^+(x) - F^-(x)$, which is the first identity.

Identity 2: $\|F\|_{TV([a,b])} = (F^+(b) - F^+(a)) + (F^-(b) - F^-(a))$.

For an interval J write $V^+(J) := \sup\{S^+(P) : P \subset J \text{ finite}\}$ and similarly $V^-(J)$. We claim

$$V^+([a, b]) = F^+(b) - F^+(a). \tag{c}$$

Proof that $V^+([a,b]) \le F^+(b) - F^+(a)$: let $P \subset [a,b]$ be finite; we may assume $P \neq \emptyset$ and, by Lemma A, that $a \in P$ (adjoining a keeps P inside $[a,b]$ and does not decrease S^+). Given $\varepsilon > 0$, choose finite $Q \subset (-\infty, a]$ with $S^+(Q) \ge F^+(a) - \varepsilon$; again by Lemma A we may assume $a \in Q$. Then $R := Q \cup P \subset (-\infty, b]$ satisfies $R \cap (-\infty, a] = Q$ and $R \cap [a, +\infty) = P$, so by Lemma B,

$$F^+(b) \ge S^+(R) = S^+(Q) + S^+(P) \ge F^+(a) - \varepsilon + S^+(P).$$

Letting $\varepsilon \to 0$ and taking the supremum over P gives $F^+(b) - F^+(a) \ge V^+([a,b])$ (all quantities are finite).
Proof that $V^+([a,b]) \ge F^+(b) - F^+(a)$: let $\varepsilon > 0$ and choose finite $R \subset (-\infty, b]$ with $S^+(R) \ge F^+(b) - \varepsilon$; by Lemma A we may assume $a \in R$ (adjoining a keeps R inside $(-\infty, b]$, as $a \le b$). Splitting at a via Lemma B, with $R_1 := R \cap (-\infty, a]$ and $R_2 := R \cap [a, b]$,

$$F^+(b) - \varepsilon \le S^+(R) = S^+(R_1) + S^+(R_2) \le F^+(a) + V^+([a,b]).$$

Letting $\varepsilon \to 0$ proves (c). Applying (c) to $-F$ and using (b) gives $V^-([a,b]) = F^-(b) - F^-(a)$.
Now by Lemma C with $J = [a,b]$ and Definition 1.6.33,

$$\|F\|_{TV([a,b])} = \sup_{P \subset [a,b]} S(P) = V^+([a,b]) + V^-([a,b]) = (F^+(b) - F^+(a)) + (F^-(b) - F^-(a)),$$

which is the second identity.
Identity 3: $\|F\|_{TV} = F^+(+\infty) + F^-(+\infty)$.
Write $G := F^+ + F^-$, a bounded non-decreasing function, so that the limits $G(+\infty) = F^+(+\infty) + F^-(+\infty)$ and $G(-\infty) := \lim_{x \to -\infty} G(x)$ exist. By Identity 2, $\|F\|_{TV([-N,N])} = G(N) - G(-N)$ for every N, and since $\|F\|_{TV(\mathbb{R})} = \sup_N \|F\|_{TV([-N,N])}$ (as noted after Definition 1.6.33), letting $N \to \infty$ gives

$$\|F\|_{TV(\mathbb{R})} = G(+\infty) - G(-\infty).$$

So it remains to prove $G(-\infty) = 0$, i.e. $F^+(-\infty) = F^-(-\infty) = 0$; by (b) it suffices to treat F^+.
Suppose, for contradiction, that $c := F^+(-\infty) = \lim_{x \to -\infty} F^+(x) > 0$ (the limit exists by monotonicity and boundedness of F^+, and is ≥ 0). Then $F^+(x) \ge c$ for every x. We construct finite sets $P_1, P_2, \ldots$ recursively, as follows. Set $m_0 := 0$. Given m_{k-1}, use $F^+(m_{k-1}) \ge c$ to select a finite set $P_k \subset (-\infty, m_{k-1}]$ with $m_{k-1} \in P_k$ (adjoining m_{k-1} only increases S^+, by Lemma A) and

$$S^+(P_k) \ge c/2,$$

and then set $m_k := \min P_k = \min(P_1 \cup \cdots \cup P_k)$. Then $P_1 \cup \cdots \cup P_k$ is a finite set with

$$S^+(P_1 \cup \cdots \cup P_k) \ge \sum_{j=1}^{k} S^+(P_j) \ge kc/2.$$

Indeed, this follows by induction on k: the point m_{k-1} lies in $P_1 \cup \cdots \cup P_k$, and splitting that set at m_{k-1} by Lemma B gives $S^+(P_1 \cup \cdots \cup P_k) = S^+(P_k) + S^+(P_1 \cup \cdots \cup P_{k-1})$, since $P_k \subset (-\infty, m_{k-1}]$ while $P_1 \cup \cdots \cup P_{k-1} \subset [m_{k-1}, +\infty)$. Letting $k \to \infty$ contradicts $S^+ \le \|F\|_{TV} < \infty$. Hence $F^+(-\infty) = 0$, and symmetrically $F^-(-\infty) = 0$, so $G(-\infty) = 0$ and

$$\|F\|_{TV} = G(+\infty) = F^+(+\infty) + F^-(+\infty),$$

completing the proof.

Problem 1.6.40. Call a function locally of bounded variation if it is of bounded variation on every compact interval $[a,b]$. Show that every function that is locally of bounded variation is differentiable almost everywhere.

Solution. Let $F : \mathbb{R} \to \mathbb{R}$ be locally of bounded variation, so that $\|F\|_{TV([a,b])} < \infty$ for every compact interval $[a, b]$. The point is to localise F so as to apply Corollary 1.6.35 (the BV differentiation theorem: every bounded variation function is differentiable almost everywhere), which requires bounded variation on all of $\mathbb{R}$.
For each $n \geq 1$ define the truncation map $\pi_n(x) := \max(-n, \min(n, x))$, and set

$$F_n := F \circ \pi_n, \qquad \text{i.e.} \quad F_n(x) = \begin{cases} F(-n), & x < -n, \\ F(x), & -n \leq x \leq n, \\ F(n), & x > n. \end{cases}$$

Claim: $\|F_n\|_{TV(\mathbb{R})} = \|F\|_{TV([-n,n])} < \infty$.
For the inequality $\geq$, note that π_n is the identity on $[-n, n]$, so every finite increasing sequence in $[-n, n]$ produces the same sum for F_n as for F; taking suprema gives $\|F_n\|_{TV(\mathbb{R})} \geq \|F\|_{TV([-n,n])}$. For the inequality $\leq$, let $x_0 < x_1 < \cdots < x_m$ be any finite increasing sequence of reals. Since π_n is non-decreasing, the numbers $\pi_n(x_0) \leq \pi_n(x_1) \leq \cdots \leq \pi_n(x_m)$ form a non-decreasing sequence in $[-n, n]$. Any index i with $\pi_n(x_i) = \pi_n(x_{i-1})$ contributes a zero term to $\sum_{i=1}^m |F_n(x_i) - F_n(x_{i-1})|$; deleting those indices, the remaining terms are precisely the consecutive differences along the strictly increasing list $y_0 < y_1 < \cdots < y_k$ of the distinct values taken by π_n on $\{x_0, \ldots, x_m\}$. Hence

$$\sum_{i=1}^{m} |F_n(x_i) - F_n(x_{i-1})| = \sum_{j=1}^{k} |F(y_j) - F(y_{j-1})| \leq \|F\|_{TV([-n,n])},$$

since $y_0 < \cdots < y_k$ all lie in $[-n, n]$. Taking the supremum proves the claim.
Thus each F_n is a function of bounded variation on $\mathbb{R}$, so by Corollary 1.6.35 there is a null set $E_n \subset \mathbb{R}$ such that F_n is differentiable at every point of $\mathbb{R} \setminus E_n$. Since F_n and F agree on the open interval $(-n, n)$, and differentiability at a point only depends on the values of the function in an arbitrarily small neighbourhood of that point, F is differentiable at every point of $(-n, n) \setminus E_n$.
Finally let $E := \bigcup_{n \geq 1} E_n$. This is a countable union of null sets, hence null (by countable subadditivity of Lebesgue outer measure). If $x \in \mathbb{R} \setminus E$, choose $n > |x|$; then $x \in (-n, n) \setminus E_n$, so F is differentiable at x. Hence F is differentiable outside the null set E, i.e. almost everywhere.

Problem 1.6.41. (Lipschitz differentiation theorem, one-dimensional case) A function $f : \mathbb{R} \to \mathbb{R}$ is said to be Lipschitz continuous if there exists a constant $C > 0$ such that $|f(x) - f(y)| \leq C|x - y|$ for all $x, y \in \mathbb{R}$; the smallest C with this property is known as the Lipschitz constant of f. Show that every Lipschitz continuous function F is locally of bounded variation, and hence differentiable almost everywhere. Furthermore, show that the derivative F', when it exists, is bounded in magnitude by the Lipschitz constant of F.

Solution. Let $F : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous, and let

$$\mathcal{C} := \{C > 0 : |F(x) - F(y)| \leq C|x - y| \text{ for all } x, y \in \mathbb{R}\},$$

which is non-empty by hypothesis. We first check that the Lipschitz constant $C_F := \inf \mathcal{C}$ really is the smallest element of $\mathcal{C}$ (rather than merely an infimum): if $C_k \in \mathcal{C}$ with $C_k \to C_F$, then for fixed x, y we have $|F(x) - F(y)| \leq C_k|x - y|$ for each k, and letting $k \to \infty$ gives $|F(x) - F(y)| \leq C_F|x - y|$. So C_F itself is a Lipschitz constant for F (if $C_F = 0$ then F is constant, and every statement below is trivial).
Locally of bounded variation. Fix a compact interval $[a, b]$ and let $a \leq x_0 < x_1 < \cdots < x_n \leq b$ be any finite increasing sequence. Then

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| \leq \sum_{i=1}^{n} C_F(x_i - x_{i-1}) = C_F(x_n - x_0) \leq C_F(b - a),$$

where we used that the increments $x_i - x_{i-1}$ are positive and telescope. Taking the supremum over all

such sequences gives, by Definition 1.6.33,

$$\|F\|_{TV([a,b])} \leq C_F(b-a) < \infty.$$

Hence F is of bounded variation on every compact interval, i.e. locally of bounded variation. By Exercise 1.6.40, F is therefore differentiable almost everywhere.
Bound on the derivative. Suppose F is differentiable at x. For every $h \neq 0$ the Lipschitz bound gives

$$\left|\frac{F(x+h)-F(x)}{h}\right| \leq \frac{C_F|h|}{|h|} = C_F.$$

Letting $h \to 0$, and using that the absolute value function is continuous, we get

$$|F'(x)| = \lim_{h\to 0}\left|\frac{F(x+h)-F(x)}{h}\right| \leq C_F.$$

Thus F', wherever it exists, is bounded in magnitude by the Lipschitz constant of F.

Problem 1.6.42. A function $f : \mathbb{R} \to \mathbb{R}$ is said to be convex if one has $f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$ for all $x < y$ and $0 < t < 1$. Show that if f is convex, then it is continuous and almost everywhere differentiable, and its derivative f' is equal almost everywhere to a monotone non-decreasing function, and so is itself almost everywhere differentiable. (Hint: Drawing the graph of f, together with a number of chords and tangent lines, is likely to be very helpful in providing visual intuition.) Thus we see that in some sense, convex functions are "almost everywhere twice differentiable". Similar claims also hold for concave functions, of course.

Solution. For $u \neq v$ write

$$s(u,v) := \frac{f(v)-f(u)}{v-u}$$

for the slope of the chord of the graph of f joining $(u, f(u))$ to $(v, f(v))$; note $s(u,v) = s(v,u)$.
Step 1: the three-chord inequality. If $x < y < z$ then

$$s(x,y) \leq s(x,z) \leq s(y,z).$$

Indeed, write $y = (1-t)x+tz$ with $t := \frac{y-x}{z-x} \in (0,1)$, so that $y-x = t(z-x)$ and $z-y = (1-t)(z-x)$. Convexity gives $f(y) \leq (1-t)f(x)+tf(z)$. Subtracting $f(x)$ yields $f(y)-f(x) \leq t(f(z)-f(x))$; dividing by $y-x = t(z-x) > 0$ gives $s(x,y) \leq s(x,z)$. Subtracting the convexity inequality from $f(z)$ instead yields $f(z)-f(y) \geq f(z)-(1-t)f(x)-tf(z) = (1-t)(f(z)-f(x))$; dividing by $z-y = (1-t)(z-x) > 0$ gives $s(y,z) \geq s(x,z)$.
Step 2: f is locally Lipschitz, hence continuous. Fix a compact interval $[b,c]$ with $b < c$, and pick $a < b$ and $d > c$. Let $b \leq u < v \leq c$. First, $s(a,b) \leq s(u,v)$: if $b < u$, Step 1 for $a < b < u$ gives $s(a,b) \leq s(a,u)$, while if $b = u$ this is an identity; and Step 1 for $a < u < v$ gives $s(a,u) \leq s(a,v) \leq s(u,v)$. Second, $s(u,v) \leq s(c,d)$: Step 1 for $u < v < d$ gives $s(u,v) \leq s(u,d) \leq s(v,d)$, and if $v < c$ then Step 1 for $v < c < d$ gives $s(v,d) \leq s(c,d)$, while if $v = c$ this is an identity. Hence with

$$K := \max\big(|s(a,b)|, |s(c,d)|\big)$$

we have $|s(u,v)| \leq K$, that is,

$$|f(v)-f(u)| \leq K|v-u| \qquad \text{for all } u,v \in [b,c].$$

So f is Lipschitz on every compact interval; in particular it is continuous at every point of $\mathbb{R}$ (every point lies in the interior of such an interval).
Step 3: f is differentiable almost everywhere. From Step 2, for any compact interval $[b,c]$ and any finite increasing sequence $b \leq x_0 < \cdots < x_n \leq c$ we have $\sum_{i=1}^n |f(x_i)-f(x_{i-1})| \leq K\sum_{i=1}^n (x_i - x_{i-1}) \leq K(c-b)$, exactly as in the proof of Exercise 1.6.41. Hence $\|f\|_{TV([b,c])} \leq K(c-b) < \infty$ for every

compact interval, so f is locally of bounded variation, and by Exercise 1.6.40 it is differentiable almost everywhere.
Step 4: the right derivative exists everywhere and is monotone non-decreasing. Fix $x \in \mathbb{R}$. For $h > 0$ the map $h \mapsto s(x, x+h)$ is non-decreasing: if $0 < h < h'$, then $x < x+h < x+h'$ and Step 1 gives $s(x, x+h) \leq s(x, x+h')$. It is also bounded below, since Step 1 applied to $x-1 < x < x+h$ gives $s(x-1, x) \leq s(x, x+h)$. A non-decreasing function of h bounded below therefore has a finite limit as $h \to 0^+$, namely

$$D^+f(x) := \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} = \inf_{h>0} s(x, x+h) \in \mathbb{R}.$$

So $D^+f : \mathbb{R} \to \mathbb{R}$ is defined at every point.
D^+f is non-decreasing. Let $x < y$. Taking $h := y - x$ in the infimum gives $D^+f(x) \leq s(x, y)$. On the other hand, for any $k > 0$ Step 1 applied to $x < y < y + k$ gives $s(x, y) \leq s(y, y+k)$, and taking the infimum over $k > 0$ gives $s(x, y) \leq D^+f(y)$. Hence

$$D^+f(x) \leq s(x, y) \leq D^+f(y),$$

so D^+f is monotone non-decreasing on $\mathbb{R}$.
Step 5: conclusion. Let E be the (full measure, by Step 3) set of points where f is differentiable. If $x \in E$ then the two-sided limit $\lim_{h \to 0} \frac{f(x+h)-f(x)}{h}$ exists and equals $f'(x)$; in particular the one-sided limit through $h > 0$ exists and equals $f'(x)$, so

$$f'(x) = D^+f(x) \qquad \text{for every } x \in E.$$

Since $\mathbb{R} \setminus E$ is null, f' agrees almost everywhere with the monotone non-decreasing function D^+f, as claimed.
Finally, by Theorem 1.6.25 (the monotone differentiation theorem) every monotone function is differentiable almost everywhere; applying this to D^+f shows that D^+f, and hence the almost everywhere defined derivative f' (up to the null set on which they may differ), is itself differentiable almost everywhere. This is the precise sense in which a convex function is "almost everywhere twice differentiable". For a concave function g the function $-g$ is convex, so all of the above applies to $-g$; in particular g is continuous, almost everywhere differentiable, and g' agrees almost everywhere with the monotone non-increasing function $-D^+(-g)$, which is again almost everywhere differentiable.

Problem 1.6.43. Show that any function of bounded variation has an (almost everywhere defined) derivative that is absolutely integrable.

Solution. Let $F : \mathbb{R} \to \mathbb{R}$ have bounded variation, so that $\|F\|_{TV(\mathbb{R})} < \infty$. We must show that F is differentiable almost everywhere, that F' (defined a.e.) is measurable, and that $\int_{\mathbb{R}} |F'(x)|\, dx < \infty$. In fact we will prove the sharp-looking bound

$$\int_{\mathbb{R}} |F'(x)|\, dx \leq \|F\|_{TV(\mathbb{R})}.$$

Step 1: a monotone decomposition. Let F^+ and F^- be the positive and negative variations of F, as in (1.30) and Exercise 1.6.39. Both are monotone non-decreasing functions taking values in $[0, \|F\|_{TV(\mathbb{R})}]$, hence are bounded, and by Exercise 1.6.39 we have the decomposition

$$F(x) = F(-\infty) + F^+(x) - F^-(x) \qquad (x \in \mathbb{R}),$$

together with the identity $\|F\|_{TV(\mathbb{R})} = F^+(+\infty) + F^-(+\infty)$, where $F^{\pm}(+\infty) := \lim_{x \to +\infty} F^{\pm}(x)$ (these limits exist by monotonicity and boundedness). In particular $F^+(+\infty)$ and $F^-(+\infty)$ are finite. (Alternatively one may simply invoke Proposition 1.6.34 to write F as a difference of two bounded monotone functions; the explicit decomposition above is what gives the clean constant.)
Step 2: almost everywhere differentiability and measurability. By the monotone differentiation theorem (Theorem 1.6.25), each of F^+ and F^- is differentiable outside a null set; let N be the union of these

two null sets, so $m(N) = 0$. For $x \notin N$ the function F, being an affine combination of F^+ and F^- plus a constant, is differentiable at x with

$$F'(x) = (F^+)'(x) - (F^-)'(x).$$

Thus F is differentiable almost everywhere; this is of course just Corollary 1.6.35. Moreover F^+ and F^- are monotone, hence measurable by Exercise 1.6.29, and as noted in the discussion preceding Proposition 1.6.37 (an easy modification of Exercise 1.6.1) the almost everywhere defined derivative of a monotone function is measurable. Hence $(F^+)'$ and $(F^-)'$ are measurable, and so is F'.
Step 3: integrability. Fix $M > 0$. The restriction of F^+ to $[-M, M]$ is monotone non-decreasing, and its derivative agrees almost everywhere on $[-M, M]$ with $(F^+)'$ (the two endpoints form a null set). Hence Proposition 1.6.37 applies and gives

$$\int_{[-M,M]} (F^+)'(x)\,dx \leq F^+(M) - F^+(-M) \leq F^+(+\infty),$$

where the last inequality uses that F^+ is non-decreasing and non-negative with limit $F^+(+\infty)$ at $+\infty$. Since F^+ is monotone non-decreasing, $(F^+)' \geq 0$ wherever it is defined, so the unsigned measurable functions $(F^+)'\mathbf{1}_{[-M,M]}$ increase pointwise almost everywhere to $(F^+)'$ as $M \to \infty$ along the integers. By the monotone convergence theorem for unsigned functions (Theorem 1.4.44),

$$\int_{\mathbb{R}} (F^+)'(x)\,dx = \lim_{M\to\infty} \int_{[-M,M]} (F^+)'(x)\,dx \leq F^+(+\infty) < \infty.$$

The identical argument gives $\int_{\mathbb{R}} (F^-)'(x)\,dx \leq F^-(+\infty) < \infty$.
Finally, almost everywhere we have $|F'| \leq (F^+)' + (F^-)'$ by the triangle inequality, so by monotonicity and additivity of the unsigned integral,

$$\int_{\mathbb{R}} |F'(x)|\,dx \leq \int_{\mathbb{R}} (F^+)'(x)\,dx + \int_{\mathbb{R}} (F^-)'(x)\,dx \leq F^+(+\infty) + F^-(+\infty) = \|F\|_{TV(\mathbb{R})} < \infty.$$

Thus F' is an almost everywhere defined measurable function which is absolutely integrable, as claimed.
Remark. If F is merely locally of bounded variation in the sense of Exercise 1.6.40, the same argument applied on each compact interval shows that F is differentiable almost everywhere and that F' is locally absolutely integrable, with $\int_{[a,b]} |F'| \leq \|F\|_{TV([a,b])}$; global absolute integrability may of course fail, as the example $F(x) = x$ shows.

Problem 1.6.44. (Second fundamental theorem for Lipschitz functions) Let $F : [a, b] \to \mathbb{R}$ be Lipschitz continuous. Show that

$$\int_{[a,b]} F'(x)\,dx = F(b) - F(a).$$

(Hint: Argue as in the proof of Proposition 1.6.37, but use the dominated convergence theorem (Theorem 1.4.49) in place of Fatou's lemma (Corollary 1.4.47).)

Solution. We may assume $a < b$, the case $a = b$ being trivial. Let C be a Lipschitz constant for F, so $|F(x) - F(y)| \leq C|x - y|$ for all $x, y \in [a, b]$.
Step 1: extension. As in the proof of Proposition 1.6.37 it is convenient to extend F to all of $\mathbb{R}$, by declaring $F(x) := F(b)$ for $x > b$ and $F(x) := F(a)$ for $x < a$. The extended function is still Lipschitz with the same constant C. Indeed, writing $\pi(x) := \min(\max(x, a), b)$ for the (1-Lipschitz) retraction of $\mathbb{R}$ onto $[a, b]$, the extension is exactly $F \circ \pi$, and hence

$$|F(\pi(x)) - F(\pi(y))| \leq C|\pi(x) - \pi(y)| \leq C|x - y|.$$

By Exercise 1.6.41 the extended F is differentiable almost everywhere, with $|F'| \leq C$ wherever the

derivative exists; and $F' = 0$ outside $[a,b]$ since F is constant on each of $(-\infty, a)$ and $(b, +\infty)$. Note also that the derivative of the extension agrees with the derivative of the original F at every interior point of $[a,b]$ where the latter exists, and $\{a,b\}$ is null, so the integral $\int_{[a,b]} F'$ is unambiguous.
Step 2: Newton quotients and dominated convergence. For $n \geq 1$ define

$$f_n(x) := \frac{F(x+1/n) - F(x)}{1/n} = n\big(F(x+1/n) - F(x)\big).$$

Each f_n is measurable (indeed continuous, since F is continuous), and by the Lipschitz bound

$$|f_n(x)| = n\,|F(x+1/n) - F(x)| \leq n \cdot C \cdot \tfrac{1}{n} = C$$

for every $x \in \mathbb{R}$ and every n. Since F is differentiable almost everywhere, $f_n \to F'$ pointwise almost everywhere on $[a,b]$.
The constant function C is absolutely integrable on the bounded interval $[a,b]$ (its integral is $C(b-a) < \infty$), so the dominated convergence theorem (Theorem 1.4.49) applies on $[a,b]$ and yields

$$\int_{[a,b]} F'(x)\,dx = \lim_{n\to\infty} \int_{[a,b]} n\big(F(x+1/n) - F(x)\big)\,dx.$$

This is the only place where the Lipschitz hypothesis is used, and it is precisely what upgrades the inequality of Proposition 1.6.37 (obtained there from Fatou's lemma, Corollary 1.4.47) to an equality.
Step 3: computing the limit. Fix n large enough that $1/n \leq b - a$. By translation invariance of the Lebesgue integral,

$$\int_{[a,b]} F(x+1/n)\,dx = \int_{[a+1/n,\, b+1/n]} F(y)\,dy,$$

so that

$$\int_{[a,b]} n\big(F(x+1/n) - F(x)\big)\,dx = n\left(\int_{[a+1/n,\, b+1/n]} F(y)\,dy - \int_{[a,b]} F(y)\,dy\right).$$

Since $1/n \leq b-a$, the intervals $[a+1/n, b+1/n]$ and $[a,b]$ overlap in $[a+1/n, b]$, and up to null sets

$$[a+1/n, b+1/n] = [a+1/n, b] \cup [b, b+1/n], \qquad [a,b] = [a, a+1/n] \cup [a+1/n, b].$$

The common piece cancels, and we are left with

$$\int_{[a,b]} n\big(F(x+1/n) - F(x)\big)\,dx = n\int_{[b,\, b+1/n]} F(y)\,dy - n\int_{[a,\, a+1/n]} F(y)\,dy.$$

(All the integrals here are finite, since F is bounded on bounded sets.)
For the first term, $F(y) = F(b)$ identically on $[b, b+1/n]$ by construction of the extension, so

$$n\int_{[b,\, b+1/n]} F(y)\,dy = n \cdot \tfrac{1}{n} F(b) = F(b)$$

exactly, for every n. For the second term we use continuity of F at a: for $y \in [a, a+1/n]$ we have $|F(y) - F(a)| \leq C/n$, whence

$$\left| n\int_{[a,\, a+1/n]} F(y)\,dy - F(a) \right| = \left| n\int_{[a,\, a+1/n]} (F(y) - F(a))\,dy \right| \leq n \cdot \tfrac{1}{n} \cdot \tfrac{C}{n} = \tfrac{C}{n} \xrightarrow[n\to\infty]{} 0.$$

Combining, the limit in Step 2 equals $F(b) - F(a)$, and therefore

$$\int_{[a,b]} F'(x)\,dx = F(b) - F(a),$$

as required.

Problem 1.6.45. (Integration by parts formula) Let $F, G : [a,b] \to \mathbb{R}$ be Lipschitz continuous functions. Show that

$$\int_{[a,b]} F'(x)G(x)\,dx = F(b)G(b) - F(a)G(a) - \int_{[a,b]} F(x)G'(x)\,dx.$$

(Hint: first show that the product of two Lipschitz continuous functions on $[a,b]$ is again Lipschitz continuous.)

Solution. Let C_F, C_G be Lipschitz constants for F and G respectively.
Step 1: the product is Lipschitz. A Lipschitz function on the compact interval $[a,b]$ is bounded: for $x \in [a,b]$ we have $|F(x)| \le |F(a)| + C_F(b-a) =: M_F$, and similarly $|G(x)| \le M_G$ for a finite constant M_G. (Boundedness is where compactness of the interval is used; on $\mathbb{R}$ the product of two Lipschitz functions need not be Lipschitz, as $x \mapsto x^2$ shows.) Set $H := FG$. For $x, y \in [a,b]$,

$$\begin{aligned} |H(x) - H(y)| &= |F(x)G(x) - F(y)G(y)| \\ &\le |F(x)|\,|G(x) - G(y)| + |G(y)|\,|F(x) - F(y)| \\ &\le (M_F C_G + M_G C_F)\,|x - y|, \end{aligned}$$

so H is Lipschitz continuous on $[a,b]$ with constant $C_H := M_F C_G + M_G C_F$.
Step 2: the product rule almost everywhere. By Exercise 1.6.41, F and G are each differentiable outside a null set; let $N \subset [a,b]$ be the union of these two null sets together with the endpoints $\{a,b\}$, so $m(N) = 0$. Fix $x \in [a,b] \setminus N$. Then for $h \neq 0$ small,

$$\frac{H(x+h) - H(x)}{h} = F(x+h)\,\frac{G(x+h) - G(x)}{h} + G(x)\,\frac{F(x+h) - F(x)}{h}.$$

As $h \to 0$ we have $F(x+h) \to F(x)$ (Lipschitz functions are continuous), and the two difference quotients converge to $G'(x)$ and $F'(x)$ respectively. Hence H is differentiable at x with

$$H'(x) = F'(x)G(x) + F(x)G'(x).$$

So this identity holds for almost every $x \in [a,b]$.
Step 3: integrability of the two terms. The functions F, G are continuous, hence measurable; F', G' are almost everywhere defined and measurable (being almost everywhere pointwise limits of the continuous difference quotients $n(F(\cdot + 1/n) - F(\cdot))$, and likewise for G). Therefore $F'G$ and FG' are measurable. Moreover, by Exercise 1.6.41 again, $|F'| \le C_F$ and $|G'| \le C_G$ wherever defined, so almost everywhere on $[a,b]$

$$|F'(x)G(x)| \le C_F M_G, \qquad |F(x)G'(x)| \le M_F C_G.$$

Since $[a,b]$ has finite measure, both $F'G$ and FG' are absolutely integrable on $[a,b]$, and in particular the two integrals in the statement are well defined and finite.
Step 4: conclusion. Since $H = FG$ is Lipschitz on $[a,b]$ by Step 1, Exercise 1.6.44 applies to H and gives

$$\int_{[a,b]} H'(x)\,dx = H(b) - H(a) = F(b)G(b) - F(a)G(a).$$

On the other hand, by Step 2 the integrand H' agrees almost everywhere with $F'G + FG'$, and by Step 3 both summands are absolutely integrable, so by linearity of the absolutely convergent integral,

$$\int_{[a,b]} H'(x)\,dx = \int_{[a,b]} F'(x)G(x)\,dx + \int_{[a,b]} F(x)G'(x)\,dx.$$

Comparing the two displays and rearranging,

$$\int_{[a,b]} F'(x)G(x)\,dx = F(b)G(b) - F(a)G(a) - \int_{[a,b]} F(x)G'(x)\,dx,$$

which is the desired integration by parts formula.

Problem 1.6.46. Show that if F is a jump function, then F' vanishes almost everywhere. (Hint: use the density argument, starting from piecewise constant jump functions and using Proposition 1.6.37 as the quantitative estimate.)

Solution. Recall Definition 1.6.30: a jump function is a function of the form

$$F = \sum_n c_n J_n,$$

where n ranges over an at most countable index set, each J_n is a basic jump function (so J_n takes the value 0 to the left of its point of discontinuity x_n, the value 1 to the right, and some value $\theta_n \in [0,1]$ at x_n), the c_n are positive reals, and $\sum_n c_n < \infty$. We may assume the index set is $\{1,2,3,\dots\}$; if it is finite then F is a piecewise constant jump function and the conclusion is immediate (see Step 1 below), so assume it is infinite.

Since each J_n is monotone non-decreasing with values in $[0,1]$ and the c_n are positive and summable, F is monotone non-decreasing with values in $[0, \sum_n c_n]$; in particular F is a bounded monotone function. By the monotone differentiation theorem (Theorem 1.6.25), F is differentiable almost everywhere, and $F' \geq 0$ wherever it is defined. Our task is to show $F' = 0$ almost everywhere.

Step 1: the dense subclass. For each $N \geq 1$ let

$$F_N := \sum_{n=1}^{N} c_n J_n, \qquad R_N := F - F_N = \sum_{n>N} c_n J_n.$$

Then F_N is a piecewise constant jump function: it is constant on each of the finitely many open intervals into which the points $x_1, \dots, x_N$ divide $\mathbb{R}$. Consequently $F_N'(x)$ exists and equals 0 for every x outside the finite set $\{x_1, \dots, x_N\}$. This is the trivial case of the claim, and it is the dense subclass from which the density argument starts: as noted in the text following Definition 1.6.30, $F_N \to F$ uniformly, since

$$\sup_{x \in \mathbb{R}} |F(x) - F_N(x)| = \sup_x R_N(x) \leq \varepsilon_N := \sum_{n>N} c_n,$$

and $\varepsilon_N \to 0$ by the absolute convergence of $\sum_n c_n$.

Step 2: the quantitative estimate. The remainder $R_N = \sum_{n>N} c_n J_n$ is itself a jump function, hence monotone non-decreasing, with values in $[0, \varepsilon_N]$. Fix any compact interval $[a,b]$. The restriction of R_N to $[a,b]$ is monotone non-decreasing, so Proposition 1.6.37 applies and yields

$$\int_{[a,b]} R_N'(x)\, dx \leq R_N(b) - R_N(a) \leq \varepsilon_N,$$

the last inequality because R_N takes values in the interval $[0, \varepsilon_N]$. (Here R_N' exists almost everywhere, again by Theorem 1.6.25 applied to the bounded monotone function R_N.) This is the quantitative estimate the hint refers to: it is small precisely because R_N has small total oscillation, and it holds with no loss of constants.

Step 3: transferring the estimate to F. Fix N. If $x \notin \{x_1, \dots, x_N\}$, then F_N is constant on a neighbourhood of x, so for h small enough $F(x+h) - F(x) = R_N(x+h) - R_N(x)$; hence F is differentiable at x if and only if R_N is, and in that case $F'(x) = R_N'(x)$. Since $\{x_1, \dots, x_N\}$ is finite and therefore null, we conclude

$$F'(x) = R_N'(x) \quad \text{for almost every } x \in \mathbb{R}.$$

Combining with Step 2, for every compact interval $[a,b]$ and every $N \geq 1$,

$$\int_{[a,b]} F'(x)\, dx = \int_{[a,b]} R_N'(x)\, dx \leq \varepsilon_N.$$

Step 4: conclusion. The left-hand side of the last display does not depend on N, while $\varepsilon_N \to 0$ as $N \to \infty$. Therefore

$$\int_{[a,b]} F'(x)\, dx = 0$$

for every compact interval $[a,b]$. Since $F' \geq 0$ almost everywhere (as F is monotone non-decreasing) and F' is measurable, the vanishing property of the unsigned integral (Exercise 1.4.36(viii)) gives $F'(x) = 0$ for almost every $x \in [a,b]$. Taking $[a,b] = [-M, M]$ for $M = 1, 2, 3, \ldots$ and using that a countable union of null sets is null, we conclude that $F'(x) = 0$ for almost every $x \in \mathbb{R}$, as claimed.

Remark. The point of the exercise is that a jump function, though it may be discontinuous on a dense set (for instance $\sum_{n=1}^{\infty} 2^{-n} \mathbf{1}_{[q_n, +\infty)}$ for an enumeration $q_1, q_2, \ldots$ of the rationals), carries all of its variation on a countable, hence null, set; the classical derivative cannot see this variation, and so the second fundamental theorem of calculus fails as badly as possible for such F: one has $\int_{[a,b]} F' = 0$ while $F(b) - F(a)$ can be any positive number.

Problem 1.6.47. (Cantor function) Define the functions $F_0, F_1, F_2, \cdots : [0,1] \to \mathbb{R}$ recursively as follows:

1. Set $F_0(x) := x$ for all $x \in [0,1]$.
2. For each $n = 1, 2, \ldots$ in turn, define

$$F_n(x) := \begin{cases} \frac{1}{2} F_{n-1}(3x) & \text{if } x \in [0, 1/3]; \\ \frac{1}{2} & \text{if } x \in (1/3, 2/3); \\ \frac{1}{2} + \frac{1}{2} F_{n-1}(3x-2) & \text{if } x \in [2/3, 1]. \end{cases}$$

(i) Graph F_0, F_1, F_2, and F_3 (preferably on a single graph).

(ii) Show that for each $n = 0, 1, \ldots$, F_n is a continuous monotone non-decreasing function with $F_n(0) = 0$ and $F_n(1) = 1$. (Hint: induct on n.)

(iii) Show that for each $n = 0, 1, \ldots$, one has $|F_{n+1}(x) - F_n(x)| \leq 2^{-n}$ for each $x \in [0,1]$. Conclude that the F_n converge uniformly to a limit $F : [0,1] \to \mathbb{R}$. This limit is known as the Cantor function.

(iv) Show that the Cantor function F is continuous and monotone non-decreasing, with $F(0) = 0$ and $F(1) = 1$.

(v) Show that if $x \in [0,1]$ lies outside the middle thirds Cantor set (Exercise 1.2.9), then F is constant in a neighbourhood of x, and in particular $F'(x) = 0$. Conclude that $\int_{[0,1]} F'(x)\, dx = 0 \neq 1 = F(1) - F(0)$, so that the second fundamental theorem of calculus fails for this function.

(vi) Show that $F\left(\sum_{n=1}^{\infty} a_n 3^{-n}\right) = \sum_{n=1}^{\infty} \frac{a_n}{2} 2^{-n}$ for any digits $a_1, a_2, \cdots \in \{0, 2\}$. Thus the Cantor function, in some sense, converts base three expansions to base two expansions.

(vii) Let $I = \left[\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n}\right]$ be one of the intervals used in the nth cover I_n of C (see Exercise 1.2.9), thus $n \geq 0$ and $a_1, \ldots, a_n \in \{0, 2\}$. Show that I is an interval of length 3^{-n}, but $F(I)$ is an interval of length 2^{-n}.

(viii) Show that F is not differentiable at any element of the Cantor set C.

Solution. Throughout, $C := \bigcap_{n \geq 0} I_n$ denotes the middle thirds Cantor set of Exercise 1.2.9, where I_n is the union of the 2^n closed intervals $\left[\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n}\right]$ with $a_1, \ldots, a_n \in \{0, 2\}$. Recall from that exercise that $m(C) = 0$.

Part (i). A verbal description of the graphs (which is what one draws). Each F_n is piecewise linear. Unwinding the recursion, F_n has slope $(3/2)^n$ on each of the 2^n closed intervals making up I_n, and is constant on each of the $2^n - 1$ open middle-third intervals removed at levels $1, \ldots, n$. (The total rise is $2^n \cdot (3/2)^n \cdot 3^{-n} = 1$, consistent with $F_n(0) = 0$, $F_n(1) = 1$.) Explicitly:

- F_0 is the diagonal segment from $(0,0)$ to $(1,1)$.
- F_1 rises with slope $3/2$ on $[0, 1/3]$, is constant $= 1/2$ on $[1/3, 2/3]$, then rises with slope $3/2$ on $[2/3, 1]$.
- F_2 has slope $9/4$ on each of $[0, 1/9], [2/9, 1/3], [2/3, 7/9], [8/9, 1]$, and takes the constant values $1/4, 1/2, 3/4$ on the removed intervals $(1/9, 2/9), (1/3, 2/3), (7/9, 8/9)$.
- F_3 has slope $27/8$ on each of the eight intervals of I_3, and takes the constant values $1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8$ on the seven removed intervals $(1/27, 2/27), (1/9, 2/9), (7/27, 8/27), (1/3, 2/3), (19/27, 20/27), (7/9, 8/9), (25/27, 26/27)$, in that left-to-right order.

Drawn on one axis, the graphs form the familiar sequence of staircases converging to the devil's stair-

case.
Part (ii). We induct on n. For $n = 0$ the function $F_0(x) = x$ is continuous, strictly increasing, and $F_0(0) = 0$, $F_0(1) = 1$.
Suppose the claim holds for F_{n-1} ($n \geq 1$). Consider the three pieces of the definition of F_n.
On $[0, 1/3]$, the map $x \mapsto 3x$ is an increasing bijection onto $[0, 1]$, so $x \mapsto \frac{1}{2}F_{n-1}(3x)$ is continuous and non-decreasing there, with value $\frac{1}{2}F_{n-1}(0) = 0$ at $x = 0$ and value $\frac{1}{2}F_{n-1}(1) = \frac{1}{2}$ at $x = 1/3$; its range is contained in $[0, 1/2]$.
On $(1/3, 2/3)$, F_n is the constant $1/2$.
On $[2/3, 1]$, the map $x \mapsto 3x - 2$ is an increasing bijection onto $[0, 1]$, so $x \mapsto \frac{1}{2} + \frac{1}{2}F_{n-1}(3x - 2)$ is continuous and non-decreasing, with value $\frac{1}{2}$ at $x = 2/3$ and value 1 at $x = 1$; its range is contained in $[1/2, 1]$.
The three pieces agree at the junctions: the first piece has value $1/2$ at $x = 1/3$ and the middle piece is identically $1/2$, so F_n is continuous at $1/3$; the middle piece has left limit $1/2$ at $x = 2/3$ and the third piece has value $1/2$ at $2/3$, so F_n is continuous at $2/3$. Hence F_n is continuous on $[0, 1]$. Since each piece is non-decreasing, and the ranges on the three pieces are ordered as $[0, 1/2]$, $\{1/2\}$, $[1/2, 1]$ from left to right, F_n is non-decreasing on all of $[0, 1]$. Finally $F_n(0) = \frac{1}{2}F_{n-1}(0) = 0$ and $F_n(1) = \frac{1}{2} + \frac{1}{2}F_{n-1}(1) = 1$. This closes the induction. Note in passing that every F_n takes values in $[0, 1]$.
Part (iii). We show $\sup_{x\in[0,1]} |F_{n+1}(x) - F_n(x)| \leq 2^{-n}$ by induction on n.
For $n = 0$: both F_1 and F_0 take values in $[0, 1]$ by (ii), so $|F_1 - F_0| \leq 1 = 2^0$.
Suppose the bound holds for $n - 1$, i.e. $\sup_x |F_n(x) - F_{n-1}(x)| \leq 2^{-(n-1)}$, where $n \geq 1$. Both F_{n+1} and F_n are given by the recursion (applied to F_n and F_{n-1} respectively). For $x \in [0, 1/3]$,

$$|F_{n+1}(x) - F_n(x)| = \tfrac{1}{2}|F_n(3x) - F_{n-1}(3x)| \leq \tfrac{1}{2} \cdot 2^{-(n-1)} = 2^{-n}.$$

For $x \in (1/3, 2/3)$ both values equal $1/2$, so the difference is 0. For $x \in [2/3, 1]$,

$$|F_{n+1}(x) - F_n(x)| = \tfrac{1}{2}|F_n(3x - 2) - F_{n-1}(3x - 2)| \leq 2^{-n}.$$

This closes the induction.
Consequently, for $m > n \geq 0$ and all x,

$$|F_m(x) - F_n(x)| \leq \sum_{k=n}^{m-1} |F_{k+1}(x) - F_k(x)| \leq \sum_{k=n}^{\infty} 2^{-k} = 2^{-n+1},$$

so (F_n) is uniformly Cauchy on $[0, 1]$. Since $\mathbb{R}$ is complete, F_n converges pointwise to some $F : [0, 1] \to \mathbb{R}$, and letting $m \to \infty$ in the display gives $\sup_x |F(x) - F_n(x)| \leq 2^{-n+1}$; that is, $F_n \to F$ uniformly. This F is the Cantor function.
Part (iv). A uniform limit of continuous functions is continuous, so F is continuous. If $0 \leq x \leq y \leq 1$ then $F_n(x) \leq F_n(y)$ for every n by (ii), and passing to the limit gives $F(x) \leq F(y)$; so F is monotone non-decreasing. Finally $F(0) = \lim_n F_n(0) = 0$ and $F(1) = \lim_n F_n(1) = 1$. Also $0 \leq F \leq 1$ throughout.
Passing to the limit in the recursion (which is legitimate pointwise), F satisfies the functional equations

$$F(x) = \tfrac{1}{2}F(3x) \ \ (x \in [0, 1/3]), \qquad F(x) = \tfrac{1}{2} \ \ (x \in [1/3, 2/3]), \qquad F(x) = \tfrac{1}{2} + \tfrac{1}{2}F(3x - 2) \ \ (x \in [2/3, 1]).$$

(For the middle equation, note $F_n \equiv 1/2$ on $(1/3, 2/3)$ for every $n \geq 1$, so $F \equiv 1/2$ there, and continuity extends this to the closed interval.) These equations will be used repeatedly below.
Part (v). We first show that F is constant on each of the open intervals removed in the construction of C, by induction on the level at which the interval is removed.
At level 1 the removed interval is $(1/3, 2/3)$, on which $F \equiv 1/2$ by the functional equation. Suppose every interval removed at level n carries a constant value of F, and let J be an interval removed at level $n + 1$. By construction J lies either in $[0, 1/3]$ or in $[2/3, 1]$, and in the first case $J = \frac{1}{3}J'$ for some interval J' removed at level n, while in the second $J = \frac{2+J'}{3}$ for such a J'. In the first case $F(x) = \frac{1}{2}F(3x)$ for $x \in J$ and $3x$ ranges over J', on which F is constant by hypothesis; hence F is constant on J. The second case is identical using $F(x) = \frac{1}{2} + \frac{1}{2}F(3x - 2)$. This closes the induction.

Now let $x \in [0,1] \setminus C$. Since $C = \bigcap_n I_n$, there is some n with $x \notin I_n$, i.e. x lies in one of the open intervals removed at level at most n. That interval is an open neighbourhood of x in $[0,1]$ on which F is constant, so F is constant near x; in particular, if $x \in (0,1)$ then $F'(x)$ exists and equals 0.
Since $m(C) = 0$ by Exercise 1.2.9, we conclude that $F'(x)$ exists and equals 0 for almost every $x \in [0,1]$. Hence F' is (almost everywhere) the zero function, and

$$\int_{[0,1]} F'(x)\,dx = 0.$$

But $F(1) - F(0) = 1 - 0 = 1 \neq 0$. So the second fundamental theorem of calculus fails for the Cantor function, even though F is continuous and monotone. (Note this is consistent with Proposition 1.6.37, which only asserts the inequality $\int_{[0,1]} F' \leq F(1) - F(0)$.)
Part (vi). Let $a_1, a_2, \cdots \in \{0,2\}$ and put $x := \sum_{n=1}^{\infty} a_n 3^{-n} \in [0,1]$. Write $\sigma(a) := \sum_{n=1}^{\infty} a_{n+1} 3^{-n}$ for the shifted number.
Suppose first $a_1 = 0$. Then $x = \sum_{n\geq 2} a_n 3^{-n} = \frac{1}{3}\sigma(a)$, and $0 \leq x \leq \frac{1}{3}\sum_{n\geq 1} 2 \cdot 3^{-n} = \frac{1}{3}$, so $x \in [0, 1/3]$ and $3x = \sigma(a)$. The functional equation gives

$$F(x) = \tfrac{1}{2}F(\sigma(a)) = \tfrac{a_1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2}F(\sigma(a)),$$

the last equality because $a_1 = 0$.
Suppose instead $a_1 = 2$. Then $x = \frac{2}{3} + \frac{1}{3}\sigma(a) \in [2/3, 1]$ and $3x - 2 = \sigma(a)$, so

$$F(x) = \tfrac{1}{2} + \tfrac{1}{2}F(\sigma(a)) = \tfrac{a_1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2}F(\sigma(a)).$$

In both cases $F(x) = \frac{a_1}{2}2^{-1} + \frac{1}{2}F(\sigma(a))$. Iterating this identity N times (each iteration shifting the digit sequence by one) gives

$$F(x) = \sum_{n=1}^{N} \frac{a_n}{2} 2^{-n} + 2^{-N} F(\sigma^N(a)), \qquad \sigma^N(a) := \sum_{n=1}^{\infty} a_{n+N} 3^{-n} \in [0,1].$$

Since $0 \leq F \leq 1$ by (iv), the error term satisfies $0 \leq 2^{-N}F(\sigma^N(a)) \leq 2^{-N} \to 0$. Letting $N \to \infty$,

$$F\left(\sum_{n=1}^{\infty} a_n 3^{-n}\right) = \sum_{n=1}^{\infty} \frac{a_n}{2} 2^{-n},$$

as claimed. (Since $a_n/2 \in \{0,1\}$, the right-hand side is precisely the binary number with digits $a_n/2$: the Cantor function reads a base-three expansion using only the digits 0 and 2, halves each digit, and reinterprets the result in base two.)
Part (vii). Let $n \geq 0$, $a_1, \ldots, a_n \in \{0,2\}$, and let $\alpha := \sum_{i=1}^{n} a_i 3^{-i}$, $\beta := \alpha + 3^{-n}$, so $I = [\alpha, \beta]$ has length 3^{-n}; that part is immediate.
Both endpoints admit ternary expansions using only the digits 0 and 2. Indeed $\alpha = \sum_{i=1}^{\infty} a_i 3^{-i}$ with $a_i := 0$ for $i > n$, while

$$\beta = \sum_{i=1}^{n} a_i 3^{-i} + 3^{-n} = \sum_{i=1}^{n} a_i 3^{-i} + \sum_{i>n} 2 \cdot 3^{-i},$$

using the geometric series $\sum_{i>n} 2 \cdot 3^{-i} = 2 \cdot \frac{3^{-(n+1)}}{1-1/3} = 3^{-n}$. So $\beta = \sum_{i=1}^{\infty} b_i 3^{-i}$ with $b_i = a_i$ for $i \leq n$ and $b_i = 2$ for $i > n$.
Applying part (vi) to each,

$$F(\alpha) = \sum_{i=1}^{n} \frac{a_i}{2} 2^{-i}, \qquad F(\beta) = \sum_{i=1}^{n} \frac{a_i}{2} 2^{-i} + \sum_{i>n} 2^{-i} = F(\alpha) + 2^{-n},$$

since $\sum_{i>n} 2^{-i} = 2^{-n}$. Hence $F(\beta) - F(\alpha) = 2^{-n}$.
Finally, F is continuous and non-decreasing on the interval I, so its image $F(I)$ is exactly the interval $[F(\alpha), F(\beta)]$ (continuity and the intermediate value theorem give surjectivity onto this interval; monotonicity gives the containment). Its length is 2^{-n}.

Part (viii). Let $x \in C$. For each $n \geq 0$, x belongs to I_n, so x lies in one of its constituent intervals $I^{(n)} = [\alpha_n, \beta_n]$, of length $\beta_n - \alpha_n = 3^{-n}$ and with $F(\beta_n) - F(\alpha_n) = 2^{-n}$ by part (vii).
We claim that for each n there is a point $y_n \in \{\alpha_n, \beta_n\}$ with $y_n \neq x$ and

$$\frac{F(y_n) - F(x)}{y_n - x} \;\geq\; \left(\frac{3}{2}\right)^n.$$

Indeed, if $x = \alpha_n$, take $y_n := \beta_n$: then the quotient is exactly $2^{-n}/3^{-n} = (3/2)^n$. Symmetrically if $x = \beta_n$, take $y_n := \alpha_n$; since F is non-decreasing the quotient is again $(F(\beta_n) - F(\alpha_n))/(\beta_n - \alpha_n) = (3/2)^n$. Otherwise $\alpha_n < x < \beta_n$, and writing

$$p_1 := F(x) - F(\alpha_n), \quad p_2 := F(\beta_n) - F(x), \quad q_1 := x - \alpha_n, \quad q_2 := \beta_n - x,$$

we have $p_1, p_2 \geq 0$ (monotonicity), $q_1, q_2 > 0$, $p_1 + p_2 = 2^{-n}$ and $q_1 + q_2 = 3^{-n}$. If both $p_1/q_1 < (3/2)^n$ and $p_2/q_2 < (3/2)^n$ held, then $p_1 + p_2 < (3/2)^n(q_1 + q_2) = (3/2)^n 3^{-n} = 2^{-n}$, a contradiction. So at least one of the two ratios is $\geq (3/2)^n$; take y_n to be the corresponding endpoint, noting that for $y_n = \alpha_n$ the difference quotient $\frac{F(\alpha_n) - F(x)}{\alpha_n - x}$ equals p_1/q_1, and for $y_n = \beta_n$ it equals p_2/q_2.
Now $|y_n - x| \leq 3^{-n} \to 0$, so $y_n \to x$ with $y_n \neq x$, while the difference quotients $\frac{F(y_n) - F(x)}{y_n - x} \geq (3/2)^n \to +\infty$. If F were differentiable at x, then every sequence of difference quotients at x along points tending to x would converge to the finite number $F'(x)$; this contradicts the unboundedness just established. Hence F is not differentiable at any $x \in C$.
Combining with part (v): the Cantor function is differentiable with derivative 0 at every point of $[0, 1] \setminus C$, and fails to be differentiable at every point of C. All of its variation is concentrated on the null set C, which is exactly why the classical derivative cannot detect it.

Problem 1.6.48.

(i) Show that every absolutely continuous function is uniformly continuous and therefore continuous.

(ii) Show that every absolutely continuous function is of bounded variation on every compact interval $[a, b]$. (Hint: first show this is true for any sufficiently small interval.) In particular (by Exercise 1.6.40), absolutely continuous functions are differentiable almost everywhere.

(iii) Show that every Lipschitz continuous function is absolutely continuous.

(iv) Show that the function $x \mapsto \sqrt{x}$ is absolutely continuous, but not Lipschitz continuous, on the interval $[0, 1]$.

(v) Show that the Cantor function from Exercise 1.6.47 is continuous, monotone, and uniformly continuous, but not absolutely continuous, on $[0, 1]$.

(vi) If $f : \mathbb{R} \to \mathbb{R}$ is absolutely integrable, show that the indefinite integral $F(x) := \int_{[-\infty, x]} f(y)\, dy$ is absolutely continuous, and that F is differentiable almost everywhere with $F'(x) = f(x)$ for almost every x.

(vii) Show that the sum or product of two absolutely continuous functions on an interval $[a, b]$ remains absolutely continuous. What happens if we work on $\mathbb{R}$ instead of on $[a, b]$?

Solution. Throughout we use Definition 1.6.39: F is absolutely continuous if for every $\varepsilon > 0$ there is $\delta > 0$ such that $\sum_{j=1}^n |F(b_j) - F(a_j)| \leq \varepsilon$ whenever $(a_1, b_1), \ldots, (a_n, b_n)$ are disjoint intervals in the domain of F with $\sum_{j=1}^n (b_j - a_j) \leq \delta$.
Part (i). Let F be absolutely continuous and let $\varepsilon > 0$. Take $\delta > 0$ as in the definition. Applying the definition with the single interval $n = 1$, $(a_1, b_1) = (a, b)$, we get $|F(b) - F(a)| \leq \varepsilon$ whenever (a, b) is an interval of length at most δ. This is precisely the definition of uniform continuity recalled in the text just before Definition 1.6.39, and uniform continuity trivially implies continuity.
Part (ii). Let F be absolutely continuous and let $[a, b]$ be a compact interval in its domain. Apply the definition of absolute continuity with $\varepsilon = 1$ to obtain $\delta > 0$ such that $\sum_j |F(b_j) - F(a_j)| \leq 1$ whenever the (a_j, b_j) are disjoint intervals of total length at most δ.
We first show that $\|F\|_{TV([c,d])} \leq 1$ for every subinterval $[c, d] \subset [a, b]$ with $d - c \leq \delta$. Indeed,

let $c \le x_0 < x_1 < \cdots < x_n \le d$ be any finite increasing sequence in $[c,d]$. The open intervals $(x_0,x_1),(x_1,x_2),\dots,(x_{n-1},x_n)$ are disjoint and have total length $x_n - x_0 \le d - c \le \delta$, so by the choice of δ,

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| \le 1.$$

Taking the supremum over all such sequences gives $\|F\|_{TV([c,d])} \le 1$, as claimed. This is the sufficiently small interval step of the hint. (Note how the definition of absolute continuity is used with n arbitrary, which is exactly what makes it strong enough to bound a total variation.)
Now partition $[a,b]$ into $N := \lceil (b-a)/\delta \rceil$ consecutive closed subintervals $[c_{k-1},c_k]$, $k = 1,\dots,N$, each of length at most δ, where $a = c_0 < c_1 < \cdots < c_N = b$. By repeated application of Exercise 1.6.36,

$$\|F\|_{TV([a,b])} = \sum_{k=1}^{N} \|F\|_{TV([c_{k-1},c_k])} \le N < \infty.$$

So F has bounded variation on every compact interval, i.e. F is locally of bounded variation. By Exercise 1.6.40, F is therefore differentiable almost everywhere.
Part (iii). Let F be Lipschitz continuous with constant C. If $C = 0$ then F is constant and the claim is trivial, so assume $C > 0$. Given $\varepsilon > 0$, set $\delta := \varepsilon/C$. If $(a_1,b_1),\dots,(a_n,b_n)$ are disjoint intervals with $\sum_j (b_j - a_j) \le \delta$, then

$$\sum_{j=1}^{n} |F(b_j) - F(a_j)| \le \sum_{j=1}^{n} C(b_j - a_j) \le C\delta = \varepsilon.$$

Hence F is absolutely continuous. (Note that the disjointness of the intervals was not even needed here; it is needed in general.)
Part (iv). Write $F(x) := \sqrt{x}$ on $[0,1]$.
F is not Lipschitz: for $x \in (0,1]$ we have $\frac{|F(x)-F(0)|}{|x-0|} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}}$, which tends to $+\infty$ as $x \to 0^+$. So no finite Lipschitz constant can exist.
F is absolutely continuous. The key elementary fact is that the increments of $\sqrt{\cdot}$ decrease as one moves to the right: for $a \ge 0$ and $h > 0$,

$$\sqrt{a+h} - \sqrt{a} = \frac{h}{\sqrt{a+h} + \sqrt{a}},$$

which is a non-increasing function of a (the denominator increases with a). Now let $(a_1,b_1),\dots,(a_n,b_n)$ be disjoint subintervals of $[0,1]$; relabelling, we may assume $a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_n < b_n$. Put $h_j := b_j - a_j$ and $s_j := h_1 + \cdots + h_j$ (with $s_0 := 0$), and let $L := s_n = \sum_j (b_j - a_j)$. Since the intervals $(a_1,b_1),\dots,(a_{j-1},b_{j-1})$ all lie in $[0,a_j]$ and are disjoint, we have $a_j \ge s_{j-1}$. Therefore, by the displayed monotonicity,

$$\sqrt{b_j} - \sqrt{a_j} = \sqrt{a_j + h_j} - \sqrt{a_j} \le \sqrt{s_{j-1} + h_j} - \sqrt{s_{j-1}} = \sqrt{s_j} - \sqrt{s_{j-1}}.$$

Summing over j, the right-hand side telescopes:

$$\sum_{j=1}^{n} |F(b_j) - F(a_j)| = \sum_{j=1}^{n} (\sqrt{b_j} - \sqrt{a_j}) \le \sqrt{s_n} - \sqrt{s_0} = \sqrt{L}.$$

So given $\varepsilon > 0$ we may take $\delta := \varepsilon^2$: if $L \le \delta$ then the sum above is at most $\sqrt{\delta} = \varepsilon$. Hence $\sqrt{\cdot}$ is absolutely continuous on $[0,1]$. (Alternatively, one can deduce this from part (vi), since $\sqrt{x} = \int_{[0,x]} \frac{1}{2\sqrt{y}}\,dy$ and $y \mapsto \frac{1}{2\sqrt{y}} \mathbf{1}_{(0,1)}(y)$ is absolutely integrable, with integral 1.)
Part (v). Let F be the Cantor function of Exercise 1.6.47. By Exercise 1.6.47(iv), F is continuous and monotone non-decreasing on $[0,1]$. Being continuous on the compact interval $[0,1]$, it is uniformly continuous (this is the Heine-Cantor theorem; alternatively it follows from the fact that a continuous function on a compact set is uniformly continuous).

F is not absolutely continuous. We show that the definition fails for $\varepsilon = 1/2$, i.e. that no $\delta > 0$ works. Let $\delta > 0$ be arbitrary and choose n so large that $(2/3)^n \le \delta$. Let $[\alpha_1, \beta_1], \dots, [\alpha_{2^n}, \beta_{2^n}]$ be the 2^n pairwise disjoint closed intervals whose union is I_n, the nth cover of the Cantor set from Exercise 1.2.9; each has length 3^{-n}, so the corresponding open intervals (α_k, β_k) are disjoint with total length

$$\sum_{k=1}^{2^n} (\beta_k - \alpha_k) = 2^n \cdot 3^{-n} = (2/3)^n \le \delta.$$

On the other hand, by Exercise 1.6.47(vii) each of these intervals satisfies $F(\beta_k) - F(\alpha_k) = 2^{-n}$, so

$$\sum_{k=1}^{2^n} |F(\beta_k) - F(\alpha_k)| = 2^n \cdot 2^{-n} = 1 > \tfrac{1}{2}.$$

Since δ was arbitrary, F is not absolutely continuous. (This is of course consistent with Exercise 1.6.47(v): were F absolutely continuous, Theorem 1.6.40 would force $F(1) - F(0) = \int_{[0,1]} F' = 0$.)
Part (vi). Let $f : \mathbb{R} \to \mathbb{R}$ be absolutely integrable and $F(x) := \int_{[-\infty, x]} f(y)\, dy$.
Absolute continuity. The single function f is uniformly integrable by Exercise 1.5.10, so by Exercise 1.5.13 there exists, for each $\varepsilon > 0$, a $\delta > 0$ such that

$$\int_E |f(y)|\, dy \le \varepsilon \quad \text{whenever } E \subset \mathbb{R} \text{ is measurable with } m(E) \le \delta.$$

Let $(a_1, b_1), \dots, (a_n, b_n)$ be disjoint intervals with $\sum_j (b_j - a_j) \le \delta$, and set $E := \bigcup_{j=1}^n (a_j, b_j)$, so that $m(E) = \sum_j (b_j - a_j) \le \delta$ by disjointness. For each j, additivity of the integral gives $F(b_j) - F(a_j) = \int_{(a_j, b_j]} f(y)\, dy$, whence (the single point b_j being null)

$$\sum_{j=1}^n |F(b_j) - F(a_j)| \le \sum_{j=1}^n \int_{(a_j, b_j)} |f(y)|\, dy = \int_E |f(y)|\, dy \le \varepsilon,$$

using disjointness again in the middle equality. Hence F is absolutely continuous.
Almost everywhere differentiability with $F' = f$. This is exactly the one-dimensional Lebesgue differentiation theorem, Theorem 1.6.11, applied to the absolutely integrable function f: it asserts that F is continuous, differentiable almost everywhere, and $F'(x) = f(x)$ for almost every $x \in \mathbb{R}$.
Part (vii). Let $F, G : [a, b] \to \mathbb{R}$ be absolutely continuous.
Sum. Let $\varepsilon > 0$. Choose $\delta_F > 0$ for F with tolerance $\varepsilon/2$ and $\delta_G > 0$ for G with tolerance $\varepsilon/2$, and set $\delta := \min(\delta_F, \delta_G)$. If the (a_j, b_j) are disjoint with total length at most δ, then by the triangle inequality

$$\sum_j |(F+G)(b_j) - (F+G)(a_j)| \le \sum_j |F(b_j) - F(a_j)| + \sum_j |G(b_j) - G(a_j)| \le \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

So $F + G$ is absolutely continuous. (The same computation works verbatim on $\mathbb{R}$.)
Product. By part (i), F and G are continuous on the compact interval $[a, b]$, hence bounded there; let $M \ge 1$ be a common bound, $|F| \le M$ and $|G| \le M$ on $[a, b]$. For any $a_j < b_j$,

$$|F(b_j)G(b_j) - F(a_j)G(a_j)| \le |F(b_j)|\,|G(b_j) - G(a_j)| + |G(a_j)|\,|F(b_j) - F(a_j)|,$$

so

$$\sum_j |(FG)(b_j) - (FG)(a_j)| \le M \left(\sum_j |G(b_j) - G(a_j)| + \sum_j |F(b_j) - F(a_j)| \right).$$

Given $\varepsilon > 0$, choose $\delta > 0$ small enough that each of the two sums on the right is at most $\varepsilon/(2M)$ whenever the disjoint intervals have total length at most δ (possible by the absolute continuity of F and of G, taking the smaller of the two resulting deltas). Then the left side is at most ε, and FG is absolutely continuous.

What happens on $\mathbb{R}$. The sum of two absolutely continuous functions on $\mathbb{R}$ is still absolutely continuous, by the argument above. The product, however, need not be: the only place compactness was used is in the boundedness of F and G, and absolutely continuous functions on $\mathbb{R}$ need not be bounded. For a counterexample take $F(x) = G(x) = x$, which is Lipschitz with constant 1 and hence absolutely continuous on $\mathbb{R}$ by part (iii). Their product is $x \mapsto x^2$, which is not uniformly continuous on $\mathbb{R}$ (for instance $|(x + \frac{1}{x})^2 - x^2| = 2 + \frac{1}{x^2} > 2$ while the two points are at distance $\frac{1}{x} \to 0$), so by part (i) it is not absolutely continuous on $\mathbb{R}$. Thus the product statement genuinely requires a bounded interval, or at least the boundedness of the two factors: if F, G are absolutely continuous and bounded on $\mathbb{R}$, the displayed argument shows FG is absolutely continuous on $\mathbb{R}$ as well.

Problem 1.6.49.
(i) Show that absolutely continuous functions map null sets to null sets, i.e. if $F : \mathbb{R} \to \mathbb{R}$ is absolutely continuous and E is a null set then $F(E) := \{F(x) : x \in E\}$ is also a null set.
(ii) Show that the Cantor function does not have this property.

Solution. Part (i). Let $F : \mathbb{R} \to \mathbb{R}$ be absolutely continuous and let $E \subset \mathbb{R}$ be null. We must show $m^*(F(E)) = 0$; since a set of outer measure zero is automatically Lebesgue measurable and null, this suffices (and it is the right formulation, since $F(E)$ need not obviously be measurable a priori).
Reduction to a bounded set. Write $E = \bigcup_{k\in\mathbb{Z}} E_k$ with $E_k := E \cap [k, k+1]$; each E_k is null, and $F(E) = \bigcup_{k\in\mathbb{Z}} F(E_k)$. Since a countable union of null sets is null, it suffices to show $m^*(F(E_k)) = 0$ for each k. So we may and do assume $E \subset [a, b]$ for some compact interval $[a, b]$.
The main estimate. Let $\varepsilon > 0$. By absolute continuity (Definition 1.6.39) choose $\delta > 0$ such that

$$\sum_{j=1}^{n} |F(\beta_j) - F(\alpha_j)| \le \varepsilon$$

whenever $(\alpha_1, \beta_1), \dots, (\alpha_n, \beta_n)$ are finitely many disjoint intervals with $\sum_{j=1}^{n}(\beta_j - \alpha_j) \le \delta$.
Since $m(E) = 0$, by the definition of Lebesgue outer measure (equivalently, by outer regularity) there is an open set $U \supset E$ with $m(U) < \delta$; replacing U by $U \cap (a - 1, b + 1)$, which still contains E and is still open with measure less than δ, we may assume U is bounded. Every open subset of $\mathbb{R}$ is an at most countable union of disjoint open intervals, so write

$$U = \bigcup_j (a_j, b_j), \qquad \sum_j (b_j - a_j) = m(U) < \delta,$$

the union being over an at most countable index set, which we enumerate as $j = 1, 2, 3, \dots$ (padding with empty intervals if the collection is finite), and the intervals being pairwise disjoint and bounded. For each j put

$$M_j := \sup_{x\in(a_j,b_j)} F(x), \qquad m_j := \inf_{x\in(a_j,b_j)} F(x).$$

These are finite: F is continuous by Exercise 1.6.48(i), and $[a_j, b_j]$ is compact, so F is bounded on it. By construction $F((a_j, b_j)) \subset [m_j, M_j]$, hence

$$m^*\big(F((a_j, b_j))\big) \le M_j - m_j.$$

We claim $\sum_j (M_j - m_j) \le \varepsilon$. Fix $\eta > 0$ and a positive integer N. For each $j \le N$, the definitions of supremum and infimum let us pick $c_j, d_j \in (a_j, b_j)$ with

$$F(c_j) < m_j + \tfrac{\eta}{2} 2^{-j}, \qquad F(d_j) > M_j - \tfrac{\eta}{2} 2^{-j},$$

so that

$$|F(d_j) - F(c_j)| \;\ge\; F(d_j) - F(c_j) \;>\; (M_j - m_j) - \eta 2^{-j}.$$

Let $S := \{j \le N : c_j \ne d_j\}$. For $j \le N$ with $j \notin S$ we have $c_j = d_j$, so the last display forces $M_j - m_j < \eta 2^{-j}$ outright. For $j \in S$ let I_j be the nondegenerate open interval with endpoints c_j and

d_j; then $I_j \subset (a_j, b_j)$, so the I_j ($j \in S$) are finitely many pairwise disjoint intervals (the (a_j, b_j) being pairwise disjoint) of total length at most $\sum_j (b_j - a_j) < \delta$. This is a legitimate input to the definition of absolute continuity, which is stated for finite collections; hence by the choice of δ,

$$\sum_{j \in S} |F(d_j) - F(c_j)| \le \varepsilon.$$

Combining the three previous displays,

$$\begin{aligned}\sum_{j \le N} (M_j - m_j) &= \sum_{j \in S} (M_j - m_j) + \sum_{j \le N,\, j \notin S} (M_j - m_j) \\ &< \sum_{j \in S} \big(|F(d_j) - F(c_j)| + \eta 2^{-j}\big) + \sum_{j \le N,\, j \notin S} \eta 2^{-j} \\ &\le \varepsilon + \eta \sum_{j \ge 1} 2^{-j} = \varepsilon + \eta.\end{aligned}$$

Letting $\eta \to 0$ gives $\sum_{j \le N} (M_j - m_j) \le \varepsilon$ for every N, and letting $N \to \infty$ gives $\sum_j (M_j - m_j) \le \varepsilon$, as claimed.

Conclusion. Since $E \subset U$, we have $F(E) \subset F(U) = \bigcup_j F((a_j, b_j))$, so by countable subadditivity of outer measure,

$$m^*(F(E)) \le \sum_j m^*\big(F((a_j, b_j))\big) \le \sum_j (M_j - m_j) \le \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $m^*(F(E)) = 0$, i.e. $F(E)$ is null.

Part (ii). Let $F : [0,1] \to \mathbb{R}$ be the Cantor function of Exercise 1.6.47 (extend it to $\mathbb{R}$ by $F(x) := 0$ for $x < 0$ and $F(x) := 1$ for $x > 1$ if one insists on a function on all of $\mathbb{R}$; this changes nothing below). Let C be the middle thirds Cantor set. By Exercise 1.2.9, $m(C) = 0$, so C is a null set. We claim that $F(C) = [0,1]$, which has measure 1; hence F does not map null sets to null sets.

First, $F(C) \subset [0,1]$ since $0 \le F \le 1$ by Exercise 1.6.47(iv).

Conversely, let $y \in [0,1]$ be arbitrary and choose a binary expansion $y = \sum_{n=1}^{\infty} b_n 2^{-n}$ with digits $b_n \in \{0,1\}$ (every $y \in [0,1]$ admits one). Set $a_n := 2b_n \in \{0,2\}$ and

$$x := \sum_{n=1}^{\infty} a_n 3^{-n}.$$

Then $x \in C$: indeed, for each $n \ge 0$ we may write $x = \sum_{i=1}^{n} a_i 3^{-i} + t_n$ with $t_n = \sum_{i>n} a_i 3^{-i} \in \big[0, \sum_{i>n} 2 \cdot 3^{-i}\big] = [0, 3^{-n}]$, so x lies in the interval $\big[\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n}\big]$, which is one of the constituent intervals of the nth cover I_n from Exercise 1.2.9. Thus $x \in \bigcap_{n \ge 0} I_n = C$.

By Exercise 1.6.47(vi),

$$F(x) = \sum_{n=1}^{\infty} \frac{a_n}{2} 2^{-n} = \sum_{n=1}^{\infty} b_n 2^{-n} = y.$$

Hence $y \in F(C)$, and therefore $[0,1] \subset F(C)$.

So $F(C) = [0,1]$ and $m(F(C)) = 1 \ne 0$, even though $m(C) = 0$. The Cantor function therefore does not map null sets to null sets. Combined with part (i), this gives a second proof that the Cantor function is not absolutely continuous, complementing the direct argument of Exercise 1.6.48(v).

Problem 1.6.50. Show that a function $F : [a,b] \to \mathbb{R}$ is absolutely continuous if and only if it takes the form

$$F(x) = \int_{[a,x]} f(y)\, dy + C$$

for some absolutely integrable $f : [a,b] \to \mathbb{R}$ and a constant C.

Solution. We prove the two implications separately.
The "if" direction. Suppose $f : [a,b] \to \mathbb{R}$ is absolutely integrable, $C \in \mathbb{R}$, and $F(x) := \int_{[a,x]} f(y)\,dy + C$ for all $x \in [a,b]$. Note first that F is well defined, since $|\int_{[a,x]} f| \le \int_{[a,b]} |f| < \infty$.
By Exercise 1.5.10(i), the constant sequence $f_n := f$ is uniformly integrable. Hence by Exercise 1.5.13, for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\int_E |f(y)|\,dy \le \varepsilon$$

whenever $E \subset [a,b]$ is measurable with $m(E) \le \delta$.
Let $\varepsilon > 0$ and let $\delta > 0$ be as above. Let $(a_1,b_1),\dots,(a_n,b_n)$ be a finite collection of disjoint intervals in $[a,b]$ of total length $\sum_{j=1}^n (b_j - a_j) \le \delta$. For each j we have, by additivity of the integral over the disjoint decomposition $[a,b_j] = [a,a_j] \cup (a_j,b_j]$,

$$F(b_j) - F(a_j) = \int_{(a_j,b_j]} f(y)\,dy = \int_{(a_j,b_j)} f(y)\,dy,$$

the last equality because a single point is a null set. Therefore, writing $E := \bigcup_{j=1}^n (a_j,b_j)$, which is a measurable set of measure $\sum_{j=1}^n (b_j - a_j) \le \delta$, we get

$$\sum_{j=1}^n |F(b_j) - F(a_j)| \le \sum_{j=1}^n \int_{(a_j,b_j)} |f(y)|\,dy = \int_E |f(y)|\,dy \le \varepsilon,$$

using countable (here finite) additivity and the disjointness of the intervals. Since $\varepsilon > 0$ was arbitrary, F is absolutely continuous on $[a,b]$. (This is the same computation as in Exercise 1.6.48(vi), applied to f extended by zero to all of $\mathbb{R}$.)
The "only if" direction. Suppose now that $F : [a,b] \to \mathbb{R}$ is absolutely continuous. By Exercise 1.6.48(ii), F is of bounded variation on $[a,b]$; hence by the BV differentiation theorem (Corollary 1.6.35) F is differentiable almost everywhere on $[a,b]$, and by Exercise 1.6.43 the (almost everywhere defined) derivative F' is absolutely integrable on $[a,b]$. Define $f : [a,b] \to \mathbb{R}$ by setting $f(x) := F'(x)$ at every point where F is differentiable and $f(x) := 0$ at the remaining points; this is an everywhere defined, absolutely integrable function (it is measurable by Exercise 1.6.1, and agrees with F' outside a null set, so its integral is unaffected).
Now fix $x \in [a,b]$. If $x = a$ the identity below is trivial, so assume $x > a$. The restriction $F|_{[a,x]}$ is again absolutely continuous (any finite collection of disjoint intervals inside $[a,x]$ is in particular a finite collection of disjoint intervals inside $[a,b]$, so the same δ works). By the second fundamental theorem for absolutely continuous functions (Theorem 1.6.40) applied on $[a,x]$,

$$\int_{[a,x]} f(y)\,dy = \int_{[a,x]} F'(y)\,dy = F(x) - F(a).$$

Thus $F(x) = \int_{[a,x]} f(y)\,dy + C$ with $C := F(a)$, for every $x \in [a,b]$, as required.
Combining the two directions, F is absolutely continuous if and only if it is an indefinite Lebesgue integral of an absolutely integrable function plus a constant. Moreover the representation is essentially unique: setting $x = a$ gives $C = F(a)$, and by Exercise 1.6.48(vi) (applied to f extended by zero to $\mathbb{R}$, so that $F(x) - C$ becomes the indefinite integral $\int_{[-\infty,x]} f$) the function F is differentiable almost everywhere with $F'(x) = f(x)$ for almost every x.

Problem 1.6.51. (Compatibility of the strong and weak derivatives in the absolutely continuous case) Let $F : [a,b] \to \mathbb{R}$ be an absolutely continuous function, and let $\varphi : [a,b] \to \mathbb{R}$ be a continuously differentiable function supported in a compact subset of (a,b). Show that

$$\int_{[a,b]} F'\varphi(x)\,dx = -\int_{[a,b]} F\varphi'(x)\,dx.$$

Solution. Write $K \subset (a,b)$ for the compact set outside of which φ vanishes. Since $a \notin K$ and $b \notin K$, we record at the outset the only consequence of the support hypothesis that we shall need:

$$\varphi(a) = \varphi(b) = 0.$$

Step 1: all four functions in sight are absolutely integrable. Since F is absolutely continuous it is of bounded variation on $[a,b]$ (Exercise 1.6.48(ii)), hence differentiable almost everywhere (Corollary 1.6.35) with F' absolutely integrable on $[a,b]$ (Exercise 1.6.43). Also F is continuous (Exercise 1.6.48(i)), hence bounded on the compact interval $[a,b]$, say $|F| \le M_F$. The functions φ, φ' are continuous on $[a,b]$, hence bounded there, say $|\varphi| \le M_\varphi$ and $|\varphi'| \le M_{\varphi'}$, and measurable. Consequently

$$\int_{[a,b]} |F'\varphi| \le M_\varphi \int_{[a,b]} |F'| < \infty, \qquad \int_{[a,b]} |F\varphi'| \le M_F M_{\varphi'}(b-a) < \infty,$$

so both integrals in the statement are well defined and finite.

Step 2: $F\varphi$ is absolutely continuous. Since φ is continuously differentiable on the compact interval $[a,b]$, its derivative is bounded by $M_{\varphi'}$, so by the mean value theorem $|\varphi(y)-\varphi(x)| \le M_{\varphi'}|y-x|$ for all $x, y \in [a,b]$; that is, φ is Lipschitz continuous, and hence absolutely continuous by Exercise 1.6.48(iii). By Exercise 1.6.48(vii), the product $F\varphi$ of two absolutely continuous functions on $[a,b]$ is absolutely continuous on $[a,b]$.

Step 3: the product rule holds almost everywhere. Let $x \in [a,b]$ be any point at which F is differentiable; by Step 1 this excludes only a null set. Since φ is differentiable everywhere, the usual product rule (valid whenever both factors are differentiable at the point in question, and proved by writing

$$\frac{F(y)\varphi(y) - F(x)\varphi(x)}{y-x} = \frac{F(y)-F(x)}{y-x}\varphi(y) + F(x)\frac{\varphi(y)-\varphi(x)}{y-x}$$

and letting $y \to x$, using the continuity of φ at x) gives that $F\varphi$ is differentiable at x with

$$(F\varphi)'(x) = F'(x)\varphi(x) + F(x)\varphi'(x).$$

Thus $(F\varphi)' = F'\varphi + F\varphi'$ almost everywhere on $[a,b]$.

Step 4: apply the second fundamental theorem. By Step 2 we may apply Theorem 1.6.40 to the absolutely continuous function $F\varphi$:

$$\int_{[a,b]} (F\varphi)'(x)\,dx = F(b)\varphi(b) - F(a)\varphi(a) = 0,$$

since $\varphi(a) = \varphi(b) = 0$. On the other hand, by Step 3 the integrand agrees almost everywhere with $F'\varphi + F\varphi'$, and both of these summands are absolutely integrable by Step 1, so by linearity of the Lebesgue integral (and the fact that altering an integrand on a null set does not change its integral)

$$0 = \int_{[a,b]} (F\varphi)'(x)\,dx = \int_{[a,b]} F'\varphi(x)\,dx + \int_{[a,b]} F\varphi'(x)\,dx.$$

Rearranging gives

$$\int_{[a,b]} F'\varphi(x)\,dx = -\int_{[a,b]} F\varphi'(x)\,dx,$$

as required. In the language of distributions, this says precisely that the weak (distributional) derivative of an absolutely continuous function is represented by its pointwise derivative F'; contrast the Cantor function F of Exercise 1.6.47, which by Exercise 1.6.48(v) is not absolutely continuous and for which $F' = 0$ almost everywhere. For that F the left-hand side above vanishes for every admissible φ, whereas the right-hand side does not. Indeed, for each $n \ge 3$ let $\varphi_n : [0,1] \to [0,1]$ be continuously differentiable, supported in the compact subset $[1/(2n), 1-1/(2n)]$ of $(0,1)$, equal to 1 on $[1/n, 1-1/n]$, non-decreasing on $[0,1/n]$ and non-increasing on $[1-1/n, 1]$. Since $0 \le F \le 1$ with F continuous, $F(0) = 0$ and $F(1) = 1$, and since $\int_{[0,1/n]} \varphi_n' = 1$ and $\int_{[1-1/n,1]} \varphi_n' = -1$ while $\varphi_n' = 0$ in between, we get

$$\left| \int_{[0,1]} F\varphi_n'(x)\,dx + 1 \right| \le \sup_{[0,1/n]} F + \sup_{[1-1/n,1]} (1-F) \longrightarrow 0$$

as $n \to \infty$, so $\int_{[0,1]} F\varphi_n' \to -1 \ne 0$. Hence the identity genuinely requires absolute continuity, not merely continuity plus almost-everywhere differentiability.

Problem 1.6.52. Let $F : [-1,1] \to \mathbb{R}$ be the function defined by setting $F(x) := x^2 \sin\left(\frac{1}{x^3}\right)$ when x is non-zero, and $F(0) := 0$. Show that F is everywhere differentiable, but the derivative F' is not absolutely integrable, and so the second fundamental theorem of calculus does not apply in this case (at least if we interpret $\int_{[a,b]} F'(x)\,dx$ using the absolutely convergent Lebesgue integral). See however the next exercise.

Solution. Step 1: F is everywhere differentiable.
At a point $x \neq 0$ the function $x \mapsto 1/x^3$ is differentiable with derivative $-3/x^4$, and sin is differentiable, so by the chain and product rules F is differentiable at x with

$$F'(x) = 2x \sin\left(\frac{1}{x^3}\right) + x^2 \cos\left(\frac{1}{x^3}\right) \cdot \left(-\frac{3}{x^4}\right) = 2x \sin\left(\frac{1}{x^3}\right) - \frac{3}{x^2} \cos\left(\frac{1}{x^3}\right).$$

At $x = 0$ we compute the difference quotient directly: for $h \neq 0$,

$$\left|\frac{F(h) - F(0)}{h}\right| = \left|\frac{h^2 \sin(1/h^3)}{h}\right| = |h| \left|\sin\left(\frac{1}{h^3}\right)\right| \leq |h| \to 0$$

as $h \to 0$. Hence F is differentiable at 0 with $F'(0) = 0$. So F is differentiable at every point of $[-1,1]$ (one-sidedly at the endpoints, which is the convention in force).
Note that F' is measurable: it is continuous on $[-1,1] \setminus \{0\}$ by the formula above, and $\{0\}$ is a null set. (Alternatively, Exercise 1.6.1 gives the measurability of the derivative of any everywhere differentiable function.)
Step 2: F' is not absolutely integrable.
Split $F' = G + H$ on $[-1,1]$, where

$$G(x) := 2x \sin\left(\frac{1}{x^3}\right), \qquad H(x) := -\frac{3}{x^2} \cos\left(\frac{1}{x^3}\right)$$

for $x \neq 0$, and $G(0) := H(0) := 0$. Both are measurable (continuous off the origin). Since $|G(x)| \leq 2|x| \leq 2$ on $[-1,1]$, G is bounded and hence absolutely integrable, with $\int_{[-1,1]} |G| \leq 4$. Therefore, if F' were absolutely integrable, then $H = F' - G$ would be absolutely integrable as well. So it suffices to prove

$$\int_{(0,1]} |H(x)|\,dx = \int_{(0,1]} \frac{3}{x^2} \left|\cos\left(\frac{1}{x^3}\right)\right| dx = +\infty.$$

For each integer $n \geq 1$ set

$$\alpha_n := \left(2\pi n + \frac{\pi}{3}\right)^{-1/3}, \qquad \beta_n := \left(2\pi n - \frac{\pi}{3}\right)^{-1/3}, \qquad I_n := [\alpha_n, \beta_n].$$

Since $t \mapsto t^{-1/3}$ is strictly decreasing on $(0,\infty)$ we have $0 < \alpha_n < \beta_n$, and $\beta_1 = (2\pi - \pi/3)^{-1/3} < 1$ because $2\pi - \pi/3 > 1$; as β_n is decreasing in n, all the I_n lie in $(0,1]$.
For $x \in I_n$ we have $1/x^3 \in [2\pi n - \pi/3,\ 2\pi n + \pi/3]$, an interval of radius $\pi/3$ about $2\pi n$, on which cos is at least $\cos(\pi/3) = 1/2$. Hence

$$\left|\cos\left(\frac{1}{x^3}\right)\right| \geq \frac{1}{2} \qquad (x \in I_n).$$

Moreover the intervals $[2\pi n - \pi/3, 2\pi n + \pi/3]$, $n \geq 1$, are pairwise disjoint (consecutive ones are separated by a gap of $2\pi - 2\pi/3 > 0$), and $x \mapsto 1/x^3$ is a bijection from $(0,\infty)$ to $(0,\infty)$, so the I_n are pairwise disjoint subsets of $(0,1]$.
We now bound the integrand and the length of I_n from below. Since $x \leq \beta_n$ on I_n,

$$\frac{3}{x^2} \geq \frac{3}{\beta_n^2} = 3\left(2\pi n - \frac{\pi}{3}\right)^{2/3} \geq 3(\pi n)^{2/3},$$

where we used $2\pi n - \pi/3 \geq \pi n$ for $n \geq 1$. For the length, apply the mean value theorem to $g(t) := t^{-1/3}$ on $[2\pi n - \pi/3, 2\pi n + \pi/3]$: there is ξ_n in that interval with

$$\beta_n - \alpha_n = \frac{2\pi}{3} \cdot \frac{1}{3}\xi_n^{-4/3} \geq \frac{2\pi}{9}\left(2\pi n + \frac{\pi}{3}\right)^{-4/3} \geq \frac{2\pi}{9}(3\pi n)^{-4/3},$$

using $2\pi n + \pi/3 \leq 3\pi n$ for $n \geq 1$.
Combining the three estimates, for each $n \geq 1$,

$$\int_{I_n} \frac{3}{x^2}\left|\cos\left(\frac{1}{x^3}\right)\right| dx \geq \frac{1}{2} \cdot 3(\pi n)^{2/3} \cdot \frac{2\pi}{9}(3\pi n)^{-4/3} = c\, n^{-2/3},$$

where $c := \frac{\pi}{3}\pi^{2/3}(3\pi)^{-4/3} > 0$ is a constant independent of n.
Since the I_n are disjoint measurable subsets of $(0,1]$ and $|H| \geq 0$, monotone convergence (or countable additivity of the integral of a non-negative function over a disjoint union) gives

$$\int_{(0,1]} |H(x)|\, dx \geq \sum_{n=1}^{\infty} \int_{I_n} |H(x)|\, dx \geq c \sum_{n=1}^{\infty} n^{-2/3} = +\infty,$$

because $\sum_n n^{-2/3}$ diverges ($2/3 \leq 1$).
Hence H is not absolutely integrable on $[-1,1]$, and therefore neither is $F' = G+H$: indeed $\int_{[-1,1]} |F'| \geq \int_{[-1,1]} |H| - \int_{[-1,1]} |G| = +\infty$.
Step 3: consequence.
Thus F is everywhere differentiable on $[-1,1]$, so the hypothesis of everywhere differentiability in Proposition 1.6.41 is satisfied, but the remaining hypothesis, absolute integrability of F', fails; the Lebesgue integral $\int_{[-1,1]} F'(x)\, dx$ simply does not exist as an absolutely convergent integral. Indeed both of its parts have infinite integral. On I_n we have, from the bounds of Step 2, $|H(x)| \geq \frac{3}{2\beta_n^2} = \frac{3}{2}\left(2\pi n - \frac{\pi}{3}\right)^{2/3} \geq \frac{3}{2}\pi^{2/3} > 2 \geq |G(x)|$, and $H < 0$ there because $\cos(1/x^3) \geq 1/2 > 0$; hence $F' = G+H$ is negative on I_n, with $(F')^- = |F'| \geq |H| - 2$ there. Summing over the disjoint intervals $I_n \subset (0,1]$,

$$\int_{[-1,1]} (F')^-(x)\, dx \geq \sum_{n=1}^{\infty} \left(\int_{I_n} |H(x)|\, dx - 2|I_n|\right) \geq \left(\sum_{n=1}^{\infty} \int_{I_n} |H(x)|\, dx\right) - 2 = +\infty,$$

since $\sum_n |I_n| \leq 1$ while $\sum_n \int_{I_n} |H| = +\infty$ by Step 2. Repeating Step 2 verbatim with $2\pi n$ replaced by $(2n+1)\pi$ produces disjoint intervals in $(0,1]$ on which $\cos(1/x^3) \leq -1/2$, so that $H > 0$ there, and the identical computation gives $\int_{[-1,1]} (F')^+(x)\, dx = +\infty$. So no meaning can be attached to the identity $\int_{[-1,1]} F'(x)\, dx = F(1) - F(-1)$ within the Lebesgue theory, even though the right-hand side is a perfectly well-defined real number. Incidentally, by Exercise 1.6.43 this also shows that F is not of bounded variation on $[-1,1]$, and hence (by Exercise 1.6.48(ii)) not absolutely continuous. The next exercise introduces the Henstock-Kurzweil integral, for which the identity is restored.

Problem 1.6.53. (Henstock-Kurzweil integral) Let $[a,b]$ be a compact interval of positive length. We say that a function $f : [a,b] \to \mathbb{R}$ is *Henstock-Kurzweil integrable* with integral $L \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists a gauge function $\delta : [a,b] \to (0,+\infty)$ such that one has

$$\left|\sum_{j=1}^{k} f(t_j^*)(t_j - t_{j-1}) - L\right| \leq \varepsilon$$

whenever $k \geq 1$ and $a = t_0 < t_1 < \cdots < t_k = b$ and $t_1^*, \ldots, t_k^*$ are such that $t_j^* \in [t_{j-1}, t_j]$ and $|t_j - t_{j-1}| \leq \delta(t_j^*)$ for every $1 \leq j \leq k$. When this occurs, we call L the Henstock-Kurzweil integral of f and write it as $\int_{[a,b]} f(x)\, dx$.
(i) Show that if a function is Henstock-Kurzweil integrable, it has a unique Henstock-Kurzweil

integral. (Hint: use Cousin's theorem.)
(ii) Show that if a function is Riemann integrable, then it is Henstock-Kurzweil integrable, and the Henstock-Kurzweil integral $\int_{[a,b]} f(x)\,dx$ is equal to the Riemann integral $\int_a^b f(x)\,dx$.
(iii) Show that if a function $f : [a,b] \to \mathbb{R}$ is everywhere defined, everywhere finite, and is absolutely integrable, then it is Henstock-Kurzweil integrable, and the Henstock-Kurzweil integral $\int_{[a,b]} f(x)\,dx$ is equal to the Lebesgue integral $\int_{[a,b]} f(x)\,dx$. (Hint: this is a variant of the proof of Theorem 1.6.40 or Proposition 1.6.41.)
(iv) Show that if $F : [a,b] \to \mathbb{R}$ is everywhere differentiable, then F' is Henstock-Kurzweil integrable, and the Henstock-Kurzweil integral $\int_{[a,b]} F'(x)\,dx$ is equal to $F(b) - F(a)$. (Hint: this is a variant of the proof of Theorem 1.6.40 or Proposition 1.6.41.)
(v) Explain why the above results give an alternate proof of Exercise 1.6.4 and of Proposition 1.6.41.

Solution. Throughout, call a datum $(k; t_0 < \cdots < t_k; t_1^*, \ldots, t_k^*)$ as in the definition a *tagged partition* of $[a,b]$, and say it is δ-*fine* if $t_j - t_{j-1} \le \delta(t_j^*)$ for all j. Write $\Delta_j := t_j - t_{j-1}$ and $\Sigma(f;\mathcal{P}) := \sum_{j=1}^k f(t_j^*)\Delta_j$ for the associated Riemann sum. The whole definition rests on Cousin's theorem (Exercise 1.6.23), which guarantees that for *every* gauge $\delta : [a,b] \to (0,+\infty)$ there exists at least one δ-fine tagged partition; without this the condition in the definition could be vacuously satisfied by every L.
(i) Uniqueness.
Suppose L and L' are both Henstock-Kurzweil integrals of f. Let $\varepsilon > 0$. Then there are gauges $\delta, \delta' : [a,b] \to (0,+\infty)$ such that every δ-fine tagged partition $\mathcal{P}$ satisfies $|\Sigma(f;\mathcal{P}) - L| \le \varepsilon$ and every δ'-fine tagged partition satisfies $|\Sigma(f;\mathcal{P}) - L'| \le \varepsilon$. Define $\delta''(x) := \min(\delta(x), \delta'(x))$; this is again a gauge, since the minimum of two strictly positive numbers is strictly positive. By Cousin's theorem (Exercise 1.6.23) there exists a δ''-fine tagged partition $\mathcal{P}$. Such a $\mathcal{P}$ is simultaneously δ-fine and δ'-fine, so by the triangle inequality

$$|L - L'| \le |L - \Sigma(f;\mathcal{P})| + |\Sigma(f;\mathcal{P}) - L'| \le 2\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $L = L'$.
(ii) The Henstock-Kurzweil integral extends the Riemann integral.
Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable, with Riemann integral $I := \int_a^b f(x)\,dx$. By definition of Riemann integrability, f is bounded; put $M := 1 + \sup_{x\in[a,b]} |f(x)| < \infty$. Recall that for a partition $P : a = s_0 < s_1 < \cdots < s_m = b$ the lower and upper Darboux sums are

$$L(f,P) := \sum_{i=1}^m \left(\inf_{[s_{i-1},s_i]} f \right)(s_i - s_{i-1}), \qquad U(f,P) := \sum_{i=1}^m \left(\sup_{[s_{i-1},s_i]} f \right)(s_i - s_{i-1}),$$

that refining a partition increases L and decreases U, that $L(f,P) \le I \le U(f,P)$ for every P, and that Riemann integrability means $\sup_P L(f,P) = \inf_P U(f,P) = I$.
Let $\varepsilon > 0$. Choose a partition $P : a = s_0 < \cdots < s_m = b$ with $U(f,P) - L(f,P) \le \varepsilon$; in particular

$$I - \varepsilon \le L(f,P) \le I \le U(f,P) \le I + \varepsilon.$$

Take the *constant* gauge $\delta(x) := \frac{\varepsilon}{4mM} > 0$ for all x, and let $\mathcal{P} : a = t_0 < \cdots < t_k = b$ with tags t_j^* be any δ-fine tagged partition. Split the indices $j \in \{1,\ldots,k\}$ into

$$A := \{j : (t_{j-1}, t_j) \text{ contains no } s_i\}, \qquad B := \{j : (t_{j-1}, t_j) \text{ contains some } s_i,\ 1 \le i \le m-1\}.$$

Since the open intervals (t_{j-1}, t_j) are disjoint, each s_i lies in at most one of them, so $\#B \le m-1 < m$, and the total length of the B-intervals is at most $m\delta = \frac{\varepsilon}{4M}$. Consequently

$$\left| \sum_{j\in B} f(t_j^*)\Delta_j \right| \le M \sum_{j\in B} \Delta_j \le \frac{\varepsilon}{4}. \tag{1}$$

Let $R := P \cup \{t_0, \dots, t_k\}$ be the common refinement. For $j \in A$ the interval $[t_{j-1}, t_j]$ contains no point of P in its interior, hence is exactly one of the intervals of R; and since $t_j^* \in [t_{j-1}, t_j]$,

$$\inf_{[t_{j-1},t_j]} f \;\le\; f(t_j^*) \;\le\; \sup_{[t_{j-1},t_j]} f.$$

Multiplying by Δ_j and summing over $j \in A$, and comparing with the full Darboux sums of R (whose remaining terms come from subdividing the B-intervals, and are bounded in absolute value by M times the total B-length, i.e. by $\varepsilon/4$), we obtain

$$L(f,R) - \frac{\varepsilon}{4} \;\le\; \sum_{j\in A} f(t_j^*)\Delta_j \;\le\; U(f,R) + \frac{\varepsilon}{4}.$$

Since R refines P we have $L(f,P) \le L(f,R)$ and $U(f,R) \le U(f,P)$, whence

$$I - \varepsilon - \frac{\varepsilon}{4} \;\le\; \sum_{j\in A} f(t_j^*)\Delta_j \;\le\; I + \varepsilon + \frac{\varepsilon}{4}. \tag{2}$$

Adding (1) and (2) gives

$$|\Sigma(f;\mathcal{P}) - I| \le \varepsilon + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} \le 2\varepsilon.$$

Since this holds for every δ-fine tagged partition, and since $\varepsilon > 0$ was arbitrary (replace ε by $\varepsilon/2$ throughout to get exactly ε), f is Henstock-Kurzweil integrable with Henstock-Kurzweil integral I; by (i) this is the only possible value.

(iii) The Henstock-Kurzweil integral extends the absolutely convergent Lebesgue integral.

Let $f : [a,b] \to \mathbb{R}$ be everywhere defined, everywhere finite, and absolutely integrable, and set $I := \int_{[a,b]} f(x)\,dx$ (Lebesgue). We follow the scheme of the proof of Proposition 1.6.41, with f in place of F'.

By Exercise 1.5.10(i) the constant sequence f is uniformly integrable, so by Exercise 1.5.13 we may find $\kappa > 0$, which we may also take to satisfy $\kappa \le \varepsilon$, such that

$$\int_U |f(x)|\,dx \le \varepsilon \quad \text{whenever } U \subset [a,b] \text{ is measurable with } m(U) \le \kappa.$$

Here $\varepsilon > 0$ is given, and we adopt the convention that f vanishes outside $[a,b]$.

Let $E \subset [a,b]$ consist of the two endpoints a, b together with all points of $[a,b]$ that are not Lebesgue points of f. By the Lebesgue differentiation theorem (Theorem 1.6.19), E is a null set. For each natural number $m = 1, 2, \dots$, use outer regularity (or simply the definition of Lebesgue outer measure) to find an open set $U_m \supset E$ with $m(U_m) \le \kappa/4^m$. Then $m\left(\bigcup_{m\ge1} U_m\right) \le \sum_m \kappa 4^{-m} \le \kappa$, and hence

$$\int_{\bigcup_m U_m} |f(x)|\,dx \le \varepsilon. \tag{3}$$

Define a gauge $\delta : [a,b] \to (0,+\infty)$ as follows.

- If $x \in E$: since $f(x)$ is finite, there is a first natural number $m = m(x) \ge 1$ with $|f(x)| \le 2^m$. As U_m is open and contains x, we may choose $\delta(x) > 0$ small enough that $(x-\delta(x), x+\delta(x)) \subset U_{m(x)}$.
- If $x \notin E$: then x is a Lebesgue point of f, i.e. $\frac{1}{2r}\int_{[x-r,x+r]} |f(y) - f(x)|\,dy \to 0$ as $r \to 0$. If Q is any interval containing x with $|Q| = h$, then $Q \subset [x-h, x+h]$, so
$$\frac{1}{|Q|}\int_Q |f(y)-f(x)|\,dy \le \frac{1}{h}\int_{[x-h,x+h]} |f(y)-f(x)|\,dy = 2\cdot\frac{1}{2h}\int_{[x-h,x+h]} |f(y)-f(x)|\,dy,$$
which tends to 0 as $h \to 0$. Hence we may choose $\delta(x) > 0$ small enough that

$$\int_Q f(y)\,dy = |Q|f(x) + O(\varepsilon|Q|) \tag{4}$$

for every interval Q containing x of length at most $\delta(x)$.

Now let $\mathcal{P} : a = t_0 < \cdots < t_k = b$ with tags t_j^* be an arbitrary δ-fine tagged partition (at least one exists, by Cousin's theorem). Split the indices into $J_1 := \{j : t_j^* \notin E\}$ and $J_2 := \{j : t_j^* \in E\}$.
For $j \in J_1$, the interval $[t_{j-1}, t_j]$ contains t_j^* and has length $\Delta_j \le \delta(t_j^*)$, so (4) gives

$$f(t_j^*)\Delta_j = \int_{[t_{j-1},t_j]} f(y)\,dy + O(\varepsilon\Delta_j).$$

Summing over $j \in J_1$ and writing $S := \bigcup_{j\in J_1}[t_{j-1}, t_j]$ (a finite union of intervals overlapping only in a null set of endpoints),

$$\sum_{j\in J_1} f(t_j^*)\Delta_j = \int_S f(y)\,dy + O(\varepsilon(b-a)).$$

Moreover, for $j \in J_2$ we have $[t_{j-1}, t_j] \subset (t_j^* - \delta(t_j^*), t_j^* + \delta(t_j^*)) \subset U_{m(t_j^*)}$, because $t_j^* \in [t_{j-1}, t_j]$ and $\Delta_j \le \delta(t_j^*)$. Hence $[a,b] \setminus S \subset \bigcup_{j\in J_2}[t_{j-1}, t_j] \subset \bigcup_m U_m$, and so by (3),

$$\int_S f(y)\,dy = I + O\left(\int_{[a,b]\setminus S} |f|\right) = I + O(\varepsilon).$$

Therefore

$$\sum_{j\in J_1} f(t_j^*)\Delta_j = I + O(\varepsilon(b-a)) + O(\varepsilon). \tag{5}$$

For $j \in J_2$ we use the size control instead. Group the indices of J_2 according to the value $m = m(t_j^*)$. For a fixed m, the corresponding intervals $[t_{j-1}, t_j]$ are non-overlapping subsets of U_m, so their total length is at most $m(U_m) \le \kappa/4^m$; and each such tag satisfies $|f(t_j^*)| \le 2^m$. Hence

$$\left|\sum_{j\in J_2} f(t_j^*)\Delta_j\right| \le \sum_{m=1}^{\infty} 2^m \cdot \frac{\kappa}{4^m} = \kappa \sum_{m=1}^{\infty} 2^{-m} = \kappa \le \varepsilon. \tag{6}$$

Adding (5) and (6),

$$|\Sigma(f;\mathcal{P}) - I| \le C\big(1 + (b-a)\big)\varepsilon$$

for an absolute constant C (coming from the implied constants in the $O(\cdot)$ notation). Given $\varepsilon' > 0$, running the above with $\varepsilon := \varepsilon'/(C(1+b-a))$ produces a gauge for which every δ-fine tagged partition has Riemann sum within ε' of I. Hence f is Henstock-Kurzweil integrable with Henstock-Kurzweil integral I, which by (i) is unique.
(iv) The fundamental theorem of calculus for the Henstock-Kurzweil integral.
Let $F : [a,b] \to \mathbb{R}$ be everywhere differentiable (one-sidedly at the endpoints). Here no integrability hypothesis is needed at all, and the argument is a stripped-down version of the proofs of Theorem 1.6.40 and Proposition 1.6.41: only the exceptional-set bookkeeping disappears, since there is no exceptional set.
Let $\varepsilon > 0$ and put $\varepsilon' := \frac{\varepsilon}{b-a} > 0$. For each $x \in [a,b]$, the differentiability of F at x means that

$$\lim_{y\to x,\, y\neq x} \frac{F(y) - F(x)}{y - x} = F'(x),$$

so we may choose $\delta(x) > 0$ small enough that

$$|F(y) - F(x) - (y-x)F'(x)| \le \varepsilon'|y-x| \qquad \text{whenever } y \in [a,b],\ |y-x| \le \delta(x). \tag{7}$$

(The inequality also holds trivially for $y = x$.) This defines a gauge $\delta : [a,b] \to (0, +\infty)$.
Let $\mathcal{P} : a = t_0 < \cdots < t_k = b$ with tags $t_j^* \in [t_{j-1}, t_j]$, $\Delta_j \le \delta(t_j^*)$, be any δ-fine tagged partition; at least one exists by Cousin's theorem (Exercise 1.6.23). Fix j. Since $t_j^* \in [t_{j-1}, t_j]$ we have $|t_j - t_j^*| \le \Delta_j \le \delta(t_j^*)$ and $|t_{j-1} - t_j^*| \le \Delta_j \le \delta(t_j^*)$, so (7) applies with $x = t_j^*$ and $y = t_j$, respectively $y = t_{j-1}$:

$$|F(t_j) - F(t_j^*) - (t_j - t_j^*)F'(t_j^*)| \le \varepsilon'(t_j - t_j^*),$$

$$|F(t_{j-1}) - F(t_j^*) - (t_{j-1} - t_j^*)F'(t_j^*)| \le \varepsilon'(t_j^* - t_{j-1}).$$

Subtracting the second from the first and using the triangle inequality, together with $(t_j - t_j^*) + (t_j^* - t_{j-1}) = \Delta_j$, gives

$$\left|F(t_j) - F(t_{j-1}) - F'(t_j^*)\Delta_j\right| \le \varepsilon'\Delta_j.$$

Summing over $j = 1, \dots, k$ and using that $\sum_j (F(t_j) - F(t_{j-1})) = F(b) - F(a)$ telescopes, and $\sum_j \Delta_j = b - a$,

$$\left|\sum_{j=1}^{k} F'(t_j^*)\Delta_j - (F(b) - F(a))\right| \le \varepsilon'(b-a) = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, F' is Henstock-Kurzweil integrable on $[a, b]$ with Henstock-Kurzweil integral $F(b) - F(a)$ (unique, by (i)).

(v) Alternate proofs of Exercise 1.6.4 and Proposition 1.6.41.

First, note that the zero function has Henstock-Kurzweil integral 0: every Riemann sum of the zero function vanishes, so any gauge (say $\delta \equiv 1$) works with $L = 0$, and by (i) no other value is possible.

Exercise 1.6.4. Let $F, G : [a, b] \to \mathbb{R}$ be differentiable. If $F = G + C$ for a constant C, then obviously $F' = G'$ everywhere. Conversely, suppose $F'(x) = G'(x)$ for every $x \in [a, b]$, and set $H := F - G$, which is everywhere differentiable with $H'(x) = 0$ for all x. Fix $x \in (a, b]$. Then H restricted to the compact interval $[a, x]$ of positive length is everywhere differentiable, so by part (iv) the function $H' \equiv 0$ is Henstock-Kurzweil integrable on $[a, x]$ with integral $H(x) - H(a)$. By the previous paragraph, that integral is 0, so $H(x) = H(a)$. As x was arbitrary in $(a, b]$ (and the case $x = a$ is trivial), H is the constant $C := H(a)$, i.e. $F(x) = G(x) + C$ for all $x \in [a, b]$.

Proposition 1.6.41. Let $F : [a, b] \to \mathbb{R}$ be everywhere differentiable with F' absolutely integrable, where $[a, b]$ has positive length. Since F is differentiable everywhere, F' is everywhere defined and everywhere finite, so part (iii) applies to $f := F'$ and shows that F' is Henstock-Kurzweil integrable with Henstock-Kurzweil integral equal to the Lebesgue integral $\int_{[a,b]} F'(x)\,dx$. On the other hand, part (iv) shows that F' is Henstock-Kurzweil integrable with Henstock-Kurzweil integral $F(b) - F(a)$. By the uniqueness in part (i), these two numbers coincide:

$$\int_{[a,b]} F'(x)\,dx = F(b) - F(a),$$

which is exactly Proposition 1.6.41. (The same combination of (ii) and (iv) likewise recovers Theorem 1.6.7: if F is differentiable with F' Riemann integrable, then by (ii) the Henstock-Kurzweil integral of F' is its Riemann integral, and by (iv) it is $F(b) - F(a)$. It does not, however, recover Theorem 1.6.40: an absolutely continuous function need only be differentiable almost everywhere, so part (iv) does not apply to it.) Note finally that Exercise 1.6.52 shows the Henstock-Kurzweil integral is a strict extension: there F' is Henstock-Kurzweil integrable by (iv), with integral $F(1) - F(-1)$, yet not absolutely integrable, so (iii) does not apply to it.

2.7 §1.7 — Outer Measures, Pre-measures, and Product Measures

Problem 1.7.1. (Null sets are Carathéodory measurable) Suppose that E is a null set for an outer measure μ^* (i.e. $\mu^*(E) = 0$). Show that E is Carathéodory measurable with respect to μ^*.

Solution. Let $A \subset X$ be arbitrary. By Definition 1.7.2 we must show

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E).$$

For the inequality $\le$, note that $A = (A \cap E) \cup (A \setminus E)$, so by the countable subadditivity axiom (iii) of Definition 1.7.1 (applied to the sequence $A \cap E, A \setminus E, \emptyset, \emptyset, \dots$, using the empty set axiom (i)) we get

$$\mu^*(A) \le \mu^*(A \cap E) + \mu^*(A \setminus E).$$

For the inequality $\ge$, observe that $A \cap E \subset E$, so by monotonicity (axiom (ii)) and the hypothesis $\mu^*(E) = 0$ we have $\mu^*(A \cap E) \le \mu^*(E) = 0$, hence $\mu^*(A \cap E) = 0$. Also $A \setminus E \subset A$, so monotonicity

gives $\mu^*(A \setminus E) \leq \mu^*(A)$. Adding,

$$\mu^*(A \cap E) + \mu^*(A \setminus E) = 0 + \mu^*(A \setminus E) \leq \mu^*(A).$$

The two inequalities combine to give the required identity for every $A \subset X$, so E is Carathéodory measurable with respect to μ^*.

Problem 1.7.2. (Compatibility with Lebesgue measurability) Show that a set $E \subset \mathbb{R}^d$ is Carathéodory measurable with respect to Lebesgue outer measure if and only if it is Lebesgue measurable. (Hint: one direction follows from Exercise 1.2.17. For the other direction, first verify simple cases, such as when E is a box, or when E or A are bounded.)

Solution. Throughout, m^* denotes Lebesgue outer measure and m elementary measure. Recall from Exercise 1.2.3 that m^* is an outer measure in the sense of Definition 1.7.1; in particular it is monotone and countably subadditive.

Suppose first that E is Carathéodory measurable with respect to m^*. Then in particular, for every elementary set $A \subset \mathbb{R}^d$ we have

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E).$$

By Lemma 1.2.6, $m^*(A) = m(A)$ for elementary A, so this says $m(A) = m^*(A \cap E) + m^*(A \setminus E)$ for every elementary set A. This is exactly condition (ii) of Exercise 1.2.17, and that exercise asserts that (ii) is equivalent to (i), i.e. to E being Lebesgue measurable. Hence E is Lebesgue measurable.

Conversely, suppose E is Lebesgue measurable. By Exercise 1.2.17 ((i) implies (iii)), for every box $B \subset \mathbb{R}^d$ we have

$$|B| = m^*(B \cap E) + m^*(B \setminus E).$$

We must upgrade this from boxes to arbitrary sets $A \subset \mathbb{R}^d$, i.e. show

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E).$$

The inequality $m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E)$ is immediate from finite subadditivity of m^*, since $A = (A \cap E) \cup (A \setminus E)$.

For the reverse inequality we may assume $m^*(A) < \infty$, since otherwise there is nothing to prove. Let $\varepsilon > 0$. By the definition of Lebesgue outer measure in Section 1.2, namely

$$m^*(A) = \inf\left\{ \sum_{n=1}^{\infty} |B_n| : A \subset \bigcup_{n=1}^{\infty} B_n,\ B_n \text{ boxes} \right\},$$

and the fact that this infimum is finite, there is a countable family of boxes $B_1, B_2, \ldots$ with

$$A \subset \bigcup_{n=1}^{\infty} B_n, \qquad \sum_{n=1}^{\infty} |B_n| \leq m^*(A) + \varepsilon.$$

Then $A \cap E \subset \bigcup_{n=1}^{\infty}(B_n \cap E)$ and $A \setminus E \subset \bigcup_{n=1}^{\infty}(B_n \setminus E)$, so by countable subadditivity of m^*,

$$m^*(A \cap E) \leq \sum_{n=1}^{\infty} m^*(B_n \cap E), \qquad m^*(A \setminus E) \leq \sum_{n=1}^{\infty} m^*(B_n \setminus E).$$

Adding these and using the box identity above term by term,

$$m^*(A \cap E) + m^*(A \setminus E) \leq \sum_{n=1}^{\infty} \big(m^*(B_n \cap E) + m^*(B_n \setminus E)\big) = \sum_{n=1}^{\infty} |B_n| \leq m^*(A) + \varepsilon.$$

(The rearrangement of the two sums into one is legitimate as all terms are non-negative.) Letting $\varepsilon \to 0$ gives $m^*(A \cap E) + m^*(A \setminus E) \leq m^*(A)$, as required.

Hence E is Carathéodory measurable with respect to m^*, and the two notions of measurability coincide.

Problem 1.7.3. Let $\mathcal{B}$ be a Boolean algebra on a set X. Show that $\mathcal{B}$ is a σ-algebra if and only if it is closed under countable disjoint unions, which means that $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$ whenever $E_1, E_2, E_3, \cdots \in \mathcal{B}$ are a countable sequence of disjoint sets in $\mathcal{B}$.

Solution. Recall (Definition 1.4.12) that a σ-algebra on X is a collection of subsets of X containing $\emptyset$, closed under complementation, and closed under countable unions; and that a Boolean algebra (Definition 1.4.1) satisfies the first two properties together with closure under finite unions.
If $\mathcal{B}$ is a σ-algebra, then it is closed under all countable unions, and in particular under countable disjoint unions. This is the trivial direction.
Conversely, suppose $\mathcal{B}$ is a Boolean algebra that is closed under countable disjoint unions. Since $\mathcal{B}$ is already a Boolean algebra, it contains $\emptyset$ and is closed under complements; so it suffices to show that $\mathcal{B}$ is closed under arbitrary countable unions.
Let $E_1, E_2, E_3, \cdots \in \mathcal{B}$ be arbitrary (not necessarily disjoint). Define

$$F_n := E_n \setminus \bigcup_{m=1}^{n-1} E_m \qquad (n = 1, 2, \dots),$$

with the convention that the empty union is $\emptyset$, so $F_1 = E_1$. Each F_n lies in $\mathcal{B}$: the union $\bigcup_{m=1}^{n-1} E_m$ is a finite union of elements of $\mathcal{B}$, hence lies in $\mathcal{B}$, and a Boolean algebra is closed under set difference (since $A \setminus B = (A^c \cup B)^c$, using closure under complements and finite unions).
The F_n are pairwise disjoint: if $n < n'$ then $F_{n'}$ is disjoint from E_n by construction, while $F_n \subset E_n$, so $F_n \cap F_{n'} = \emptyset$.
Finally, $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n$. Indeed, $F_n \subset E_n$ for each n gives the inclusion $\subset$. For the reverse inclusion, if $x \in \bigcup_{n=1}^{\infty} E_n$, let n be the least index with $x \in E_n$ (such an index exists by the well-ordering of the natural numbers); then $x \notin E_m$ for all $m < n$, so $x \in F_n$.
Since the F_n are disjoint elements of $\mathcal{B}$, the hypothesis gives $\bigcup_{n=1}^{\infty} F_n \in \mathcal{B}$, and hence $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$.
Thus $\mathcal{B}$ is closed under countable unions and is therefore a σ-algebra.

Problem 1.7.4. (i) Show that the requirement that μ_0 is finitely additive can be relaxed to the condition that $\mu_0(\emptyset) = 0$ without affecting the definition of a pre-measure.
(ii) Show that the condition $\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu_0(E_n)$ can be relaxed to $\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu_0(E_n)$ without affecting the definition of a pre-measure.
(iii) On the other hand, give an example to show that if one performs both of the above two relaxations at once, one starts admitting objects μ_0 that are not pre-measures.

Solution. Throughout, $\mathcal{B}_0$ is a Boolean algebra on a set X and $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$. Recall Definition 1.7.7: μ_0 is a pre-measure if it is a finitely additive measure (Definition 1.4.19: $\mu_0(\emptyset) = 0$ together with $\mu_0(E \cup F) = \mu_0(E) + \mu_0(F)$ for disjoint $E, F \in \mathcal{B}_0$) which in addition satisfies
(P) $\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu_0(E_n)$ whenever $E_1, E_2, \cdots \in \mathcal{B}_0$ are disjoint and $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_0$.
Part (i). We must show that a map μ_0 with $\mu_0(\emptyset) = 0$ satisfying (P) is automatically finitely additive; the converse implication is trivial since finite additivity is part of the original hypothesis and $\mu_0(\emptyset) = 0$ is part of Definition 1.4.19.
So suppose $\mu_0(\emptyset) = 0$ and (P) holds. Let $E, F \in \mathcal{B}_0$ be disjoint. Consider the sequence

$$E_1 := E, \quad E_2 := F, \quad E_n := \emptyset \ \text{ for } n \geq 3.$$

These sets all lie in $\mathcal{B}_0$ and are pairwise disjoint (the empty set is disjoint from every set, and $E \cap F = \emptyset$ by hypothesis). Their union is $E \cup F$, which lies in $\mathcal{B}_0$ because a Boolean algebra is closed under finite unions. Hence (P) applies and gives

$$\mu_0(E \cup F) = \sum_{n=1}^{\infty} \mu_0(E_n) = \mu_0(E) + \mu_0(F) + \sum_{n \geq 3} \mu_0(\emptyset) = \mu_0(E) + \mu_0(F),$$

using $\mu_0(\emptyset) = 0$. Thus μ_0 is finitely additive, and together with $\mu_0(\emptyset) = 0$ it is a finitely additive measure. So the two definitions of pre-measure coincide.

Part (ii). We must show that if μ_0 is a finitely additive measure obeying the inequality

$$\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu_0(E_n)$$

for all disjoint $E_1, E_2, \cdots \in \mathcal{B}_0$ whose union lies in $\mathcal{B}_0$, then the reverse inequality holds automatically, so that (P) holds. (Again the converse implication is trivial.)
Let such a disjoint sequence be given, and write $E := \bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_0$. Fix $N \geq 1$. The set $\bigcup_{n=1}^{N} E_n$ lies in $\mathcal{B}_0$ (finite unions) and is contained in E. Since μ_0 is a finitely additive measure, it is monotone and finitely additive on finite disjoint families by Exercise 1.4.20(i),(ii) (whose proofs use only the Boolean algebra axioms: monotonicity follows from $\mu_0(F) = \mu_0(E) + \mu_0(F \setminus E) \geq \mu_0(E)$ for $E \subset F$ in $\mathcal{B}_0$, and finite additivity for k sets follows by induction). Hence

$$\sum_{n=1}^{N} \mu_0(E_n) = \mu_0\left(\bigcup_{n=1}^{N} E_n\right) \leq \mu_0(E).$$

This holds for every N, so letting $N \to \infty$ (the partial sums of a series of non-negative terms increase to the sum) gives

$$\sum_{n=1}^{\infty} \mu_0(E_n) \leq \mu_0(E).$$

Combined with the assumed inequality in the other direction we obtain $\mu_0(E) = \sum_{n=1}^{\infty} \mu_0(E_n)$, which is (P).
Part (iii). Now relax both conditions at once: we ask only that $\mu_0(\emptyset) = 0$ and that

$$\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu_0(E_n)$$

for all disjoint sequences in $\mathcal{B}_0$ with union in $\mathcal{B}_0$. We exhibit a μ_0 satisfying these but which is not a pre-measure.
Take $X := \{1, 2\}$ and $\mathcal{B}_0 := 2^X = \{\emptyset, \{1\}, \{2\}, X\}$, which is a Boolean algebra (indeed the discrete algebra of Example 1.4.3). Define

$$\mu_0(\emptyset) := 0, \qquad \mu_0(\{1\}) := \mu_0(\{2\}) := 1, \qquad \mu_0(X) := 1.$$

We check the relaxed hypotheses. Certainly $\mu_0(\emptyset) = 0$. Now let $E_1, E_2, \cdots \in \mathcal{B}_0$ be disjoint with union E (automatically in $\mathcal{B}_0$, as $\mathcal{B}_0 = 2^X$). Discarding the terms equal to $\emptyset$ (which contribute 0 to the sum and nothing to the union), the non-empty terms are pairwise disjoint non-empty subsets of the two-point set X, so there are only three possibilities:

- no non-empty terms: then $E = \emptyset$ and $\mu_0(E) = 0 \leq 0$;
- exactly one non-empty term F: then $E = F$ and $\mu_0(E) = \mu_0(F)$, so the inequality holds with equality;
- exactly two non-empty terms, necessarily $\{1\}$ and $\{2\}$: then $E = X$ and $\mu_0(E) = 1 \leq 1 + 1 = 2$.

In every case the countable subadditivity inequality holds. So μ_0 satisfies both relaxed conditions.
However μ_0 is not a pre-measure: it is not even finitely additive, since $\{1\}$ and $\{2\}$ are disjoint elements of $\mathcal{B}_0$ with

$$\mu_0(\{1\} \cup \{2\}) = \mu_0(X) = 1 \neq 2 = \mu_0(\{1\}) + \mu_0(\{2\}).$$

Equivalently, taking $E_1 = \{1\}$, $E_2 = \{2\}$ and $E_n = \emptyset$ for $n \geq 3$ shows that (P) fails. Hence performing both relaxations simultaneously genuinely enlarges the class of admissible μ_0 beyond the pre-measures.

Problem 1.7.5. Without using the theory of Lebesgue measure, show that elementary measure (on the elementary Boolean algebra) is a pre-measure. (Hint: use Lemma 1.2.6. Note that one has to

also deal with co-elementary sets as well as elementary sets in the elementary Boolean algebra.)

Solution. Let $\mathcal{E}[\mathbb{R}^d]$ denote the elementary Boolean algebra of Exercise 1.4.1, consisting of all sets $E \subset \mathbb{R}^d$ that are either elementary or co-elementary (that is, E^c is elementary). Following the convention of Example 1.4.21 we extend elementary measure to this algebra by

$$\mu_0(E) := m(E) \text{ if } E \text{ is elementary}, \qquad \mu_0(E) := +\infty \text{ if } E \text{ is co-elementary}.$$

This is well defined: no set is both elementary and co-elementary. Indeed an elementary set is a finite union of boxes and hence bounded, whereas if E is co-elementary then E^c is bounded, say $E^c \subset [-R,R]^d$, so $E \supset \mathbb{R}^d \setminus [-R,R]^d$ is unbounded.
We record two consequences of boundedness that will be used repeatedly.
(A) Two co-elementary sets are never disjoint. For if E, F are co-elementary and $E \cap F = \emptyset$, then $E^c \cup F^c = \mathbb{R}^d$; but E^c, F^c are elementary hence bounded, so their union is bounded, contradicting the unboundedness of $\mathbb{R}^d$ (here $d \geq 1$).
(B) If $E \in \mathcal{E}[\mathbb{R}^d]$ contains a co-elementary set F, then E is co-elementary. Indeed F is unbounded, hence so is E, hence E is not elementary; being in the algebra it must be co-elementary.
We now verify the two requirements of Definition 1.7.7.
Finite additivity. Clearly $\mu_0(\emptyset) = m(\emptyset) = 0$. Let $E, F \in \mathcal{E}[\mathbb{R}^d]$ be disjoint. By (A) at most one of them is co-elementary. If both are elementary then $E \cup F$ is elementary and $\mu_0(E \cup F) = m(E \cup F) = m(E) + m(F) = \mu_0(E) + \mu_0(F)$ by the finite additivity of elementary measure recorded after Lemma 1.1.2. If exactly one is co-elementary, say E, then $E \subset E \cup F$, so by (B) the set $E \cup F$ is co-elementary and

$$\mu_0(E \cup F) = +\infty = (+\infty) + m(F) = \mu_0(E) + \mu_0(F).$$

Thus μ_0 is a finitely additive measure on $\mathcal{E}[\mathbb{R}^d]$.
The pre-measure property. Let $E_1, E_2, \cdots \in \mathcal{E}[\mathbb{R}^d]$ be disjoint with $E := \bigcup_{n=1}^{\infty} E_n \in \mathcal{E}[\mathbb{R}^d]$. We must show $\mu_0(E) = \sum_{n=1}^{\infty} \mu_0(E_n)$. (All terms are non-negative, so the sum is unambiguous.)
Case 1: some E_{n_0} is co-elementary. Then $\sum_n \mu_0(E_n) = +\infty$ since one of the terms is $+\infty$. Also $E_{n_0} \subset E$, so by (B) the set E is co-elementary and $\mu_0(E) = +\infty$. The identity holds.
In the remaining cases every E_n is elementary, so $\mu_0(E_n) = m(E_n)$ for all n. Here we use Lemma 1.2.6, which says $m^*(F) = m(F)$ for every elementary set F, together with the fact (Exercise 1.2.3) that Lebesgue outer measure m^* is monotone and countably subadditive. Note that only these facts about m^* are used; no property of Lebesgue measure itself is invoked.
Case 2: every E_n is elementary and E is elementary. For the lower bound, fix N; then $E_1, \ldots, E_N$ are disjoint elementary sets whose union is elementary and contained in E, so by finite additivity and monotonicity of elementary measure,

$$\sum_{n=1}^{N} m(E_n) = m\left(\bigcup_{n=1}^{N} E_n\right) \leq m(E).$$

Letting $N \to \infty$ gives $\sum_{n=1}^{\infty} m(E_n) \leq m(E)$. For the upper bound, Lemma 1.2.6 and countable subadditivity of m^* give

$$m(E) = m^*(E) = m^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m^*(E_n) = \sum_{n=1}^{\infty} m(E_n).$$

Hence $\mu_0(E) = m(E) = \sum_{n=1}^{\infty} \mu_0(E_n)$.
Case 3: every E_n is elementary but E is co-elementary. Then $\mu_0(E) = +\infty$, and we must show $\sum_{n=1}^{\infty} m(E_n) = +\infty$. Suppose for contradiction that $S := \sum_{n=1}^{\infty} m(E_n) < \infty$. Since E is co-elementary, E^c is elementary and hence bounded, say $E^c \subset [-R,R]^d$; consequently

$$\mathbb{R}^d \setminus [-R,R]^d \subset E.$$

For each integer $k > R$ consider the elementary set $Q_k := [-k,k]^d \setminus [-R,R]^d$, which satisfies $Q_k \subset E$ (as Q_k is disjoint from $[-R,R]^d \supset E^c$) and, by finite additivity of elementary measure applied to the

partition $[-k,k]^d = [-R,R]^d \cup Q_k$, $m(Q_k) = (2k)^d - (2R)^d$ (the subtraction being legitimate as both quantities are finite). On the other hand $Q_k \subset E = \bigcup_{n=1}^{\infty} E_n$, so by Lemma 1.2.6, monotonicity and countable subadditivity of m^*,

$$(2k)^d - (2R)^d = m(Q_k) = m^*(Q_k) \leq \sum_{n=1}^{\infty} m^*(E_n) = \sum_{n=1}^{\infty} m(E_n) = S.$$

Letting $k \to \infty$ the left-hand side tends to $+\infty$, contradicting $S < \infty$. Hence $S = +\infty = \mu_0(E)$, as required.
Case 4: every E_n is elementary and E is neither elementary nor co-elementary. This cannot occur, since $E \in \mathcal{E}[\mathbb{R}^d]$ by hypothesis.
In all cases the pre-measure identity holds, so elementary measure on the elementary Boolean algebra is a pre-measure.

Problem 1.7.6. Construct a finitely additive measure $\mu_0 : \mathcal{B}_0 \to [0,+\infty]$ that is not a pre-measure. (Hint: take X to be the natural numbers, take $\mathcal{B}_0 = 2^{\mathbb{N}}$ to be the discrete algebra, and define μ_0 separately for finite and infinite sets.)

Solution. Take $X := \mathbb{N} = \{1,2,3,\dots\}$ and $\mathcal{B}_0 := 2^{\mathbb{N}}$, the discrete algebra of Example 1.4.3, which is certainly a Boolean algebra (indeed a σ-algebra). Define $\mu_0 : \mathcal{B}_0 \to [0,+\infty]$ by

$$\mu_0(E) := 0 \text{ if } E \text{ is finite}, \qquad \mu_0(E) := +\infty \text{ if } E \text{ is infinite}.$$

μ_0 is a finitely additive measure. The empty set is finite, so $\mu_0(\emptyset) = 0$, giving axiom (i) of Definition 1.4.19. For axiom (ii), let $E, F \subset \mathbb{N}$ be disjoint. If both E and F are finite then $E \cup F$ is finite and

$$\mu_0(E \cup F) = 0 = 0 + 0 = \mu_0(E) + \mu_0(F).$$

If at least one of E, F is infinite, then $E \cup F$ is infinite (it contains an infinite set), so $\mu_0(E \cup F) = +\infty$, while the right-hand side $\mu_0(E) + \mu_0(F)$ also equals $+\infty$ since one of its two non-negative terms is $+\infty$. In both cases finite additivity holds. (Note that disjointness was not even needed here.)
μ_0 is not a pre-measure. Take $E_n := \{n\}$ for $n = 1,2,3,\dots$. These are pairwise disjoint elements of $\mathcal{B}_0$, and their union is

$$\bigcup_{n=1}^{\infty} E_n = \mathbb{N} \in \mathcal{B}_0.$$

Each E_n is a singleton, hence finite, so $\mu_0(E_n) = 0$ and

$$\sum_{n=1}^{\infty} \mu_0(E_n) = 0.$$

On the other hand $\mathbb{N}$ is infinite, so

$$\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) = \mu_0(\mathbb{N}) = +\infty \neq 0 = \sum_{n=1}^{\infty} \mu_0(E_n).$$

Thus the defining property of a pre-measure in Definition 1.7.7 fails, and μ_0 is a finitely additive measure that is not a pre-measure.
By the discussion preceding Definition 1.7.7, it follows in particular that this μ_0 admits no extension to a countably additive measure on any σ-algebra refining $\mathcal{B}_0$: any such extension μ would satisfy $\mu(\mathbb{N}) = \sum_{n=1}^{\infty} \mu(\{n\}) = 0$, contradicting $\mu(\mathbb{N}) = \mu_0(\mathbb{N}) = +\infty$.

Problem 1.7.7. Let $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ be a pre-measure, let $\mu : \mathcal{B} \to [0, +\infty]$ be the Hahn-Kolmogorov extension of μ_0, and let $\mu' : \mathcal{B}' \to [0, +\infty]$ be another countably additive extension of μ_0. Suppose also that μ_0 is σ-finite, which means that one can express the whole space X as the countable union of sets $E_1, E_2, \cdots \in \mathcal{B}_0$ for which $\mu_0(E_n) < \infty$ for all n. Show that μ and μ' agree on their common domain of definition. In other words, show that $\mu(E) = \mu'(E)$ for all $E \in \mathcal{B} \cap \mathcal{B}'$. (Hint: first show that $\mu'(E) \leq \mu^*(E)$ for all $E \in \mathcal{B}'$.)

Solution. Recall the construction in the proof of Theorem 1.7.8. One defines the outer measure

$$\mu^*(E) := \inf \left\{ \sum_{n=1}^{\infty} \mu_0(F_n) : E \subset \bigcup_{n=1}^{\infty} F_n,\ F_n \in \mathcal{B}_0 \text{ for all } n \right\}$$

for every $E \subset X$; then $\mathcal{B}$ is the σ-algebra of sets Carathéodory measurable with respect to μ^*, and μ is the restriction of μ^* to $\mathcal{B}$. It was shown there that $\mathcal{B}_0 \subset \mathcal{B}$ and that μ extends μ_0. By hypothesis $\mathcal{B}'$ is also a σ-algebra containing $\mathcal{B}_0$ and $\mu' : \mathcal{B}' \to [0, +\infty]$ is countably additive with $\mu'|_{\mathcal{B}_0} = \mu_0$.
Step 1: we show that $\mu'(E) \leq \mu^*(E)$ for every $E \in \mathcal{B}'$.
Let $E \in \mathcal{B}'$ and let $F_1, F_2, \cdots \in \mathcal{B}_0$ be any countable family with $E \subset \bigcup_{n=1}^{\infty} F_n$. Since $\mathcal{B}_0 \subset \mathcal{B}'$ and $\mathcal{B}'$ is a σ-algebra, the union $\bigcup_n F_n$ lies in $\mathcal{B}'$. As μ' is a countably additive measure, it is monotone (Exercise 1.4.20(i), noting that countably additive measures are finitely additive) and countably subadditive (Exercise 1.4.23(i)). Hence

$$\mu'(E) \leq \mu' \left(\bigcup_{n=1}^{\infty} F_n \right) \leq \sum_{n=1}^{\infty} \mu'(F_n) = \sum_{n=1}^{\infty} \mu_0(F_n),$$

the last equality because μ' extends μ_0 and $F_n \in \mathcal{B}_0$. Taking the infimum over all such covers gives $\mu'(E) \leq \mu^*(E)$. (If E admits no countable cover by sets of $\mathcal{B}_0$ at all, then $\mu^*(E) = \inf \emptyset = +\infty$ and the inequality is trivial; in fact the σ-finiteness hypothesis guarantees that covers always exist, since X itself is such a union.)
In particular, if $E \in \mathcal{B} \cap \mathcal{B}'$ then $\mu(E) = \mu^*(E)$, so

$$\mu'(E) \leq \mu(E) \qquad \text{for all } E \in \mathcal{B} \cap \mathcal{B}'.$$

Step 2: reduction to sets of finite measure.
By σ-finiteness, write $X = \bigcup_{n=1}^{\infty} E_n$ with $E_n \in \mathcal{B}_0$ and $\mu_0(E_n) < \infty$. Disjointify these as in Exercise 1.7.3 by setting

$$X_n := E_n \setminus \bigcup_{m=1}^{n-1} E_m.$$

Each X_n lies in $\mathcal{B}_0$ (a Boolean algebra is closed under finite unions and set differences), the X_n are pairwise disjoint, $\bigcup_n X_n = \bigcup_n E_n = X$, and $X_n \subset E_n$ gives $\mu_0(X_n) \leq \mu_0(E_n) < \infty$ by monotonicity of the finitely additive measure μ_0 (Exercise 1.4.20(i)).
Now let $E \in \mathcal{B} \cap \mathcal{B}'$. Since $X_n \in \mathcal{B}_0 \subset \mathcal{B} \cap \mathcal{B}'$ and both $\mathcal{B}$ and $\mathcal{B}'$ are σ-algebras, the sets $E \cap X_n$ lie in $\mathcal{B} \cap \mathcal{B}'$; they are pairwise disjoint with union E. By countable additivity of μ and of μ',

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E \cap X_n), \qquad \mu'(E) = \sum_{n=1}^{\infty} \mu'(E \cap X_n).$$

So it suffices to prove $\mu(F) = \mu'(F)$ for every $F \in \mathcal{B} \cap \mathcal{B}'$ contained in some X_n.
Step 3: the finite-measure case.
Fix n and write $Y := X_n$, so $Y \in \mathcal{B}_0$ and $\mu_0(Y) < \infty$. Let $F \in \mathcal{B} \cap \mathcal{B}'$ with $F \subset Y$, and put $G := Y \setminus F$, which also lies in $\mathcal{B} \cap \mathcal{B}'$ (both are σ-algebras containing Y and F). Then F and G are disjoint with $F \cup G = Y$, so finite additivity of μ and μ' gives

$$\mu(F) + \mu(G) = \mu(Y) = \mu_0(Y), \qquad \mu'(F) + \mu'(G) = \mu'(Y) = \mu_0(Y),$$

using that both μ and μ' extend μ_0 and that $Y \in \mathcal{B}_0$. Since $\mu_0(Y) < \infty$, all four of $\mu(F), \mu(G), \mu'(F), \mu'(G)$ are finite, so subtraction is legitimate.
By the inequality established in Step 1 we have $\mu'(F) \le \mu(F)$ and $\mu'(G) \le \mu(G)$. If the first inequality were strict, then adding would give

$$\mu_0(Y) = \mu'(F) + \mu'(G) < \mu(F) + \mu(G) = \mu_0(Y),$$

a contradiction (again legitimate because all quantities are finite). Hence $\mu'(F) = \mu(F)$.
Conclusion. Combining Steps 2 and 3: for any $E \in \mathcal{B} \cap \mathcal{B}'$, each $E \cap X_n$ is a set of $\mathcal{B} \cap \mathcal{B}'$ contained in X_n, so $\mu(E \cap X_n) = \mu'(E \cap X_n)$ for every n, and therefore

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E \cap X_n) = \sum_{n=1}^{\infty} \mu'(E \cap X_n) = \mu'(E).$$

Thus μ and μ' agree on their common domain $\mathcal{B} \cap \mathcal{B}'$, so the Hahn-Kolmogorov extension of a σ-finite pre-measure is unique within its own σ-algebra.

Problem 1.7.8. The purpose of this exercise is to show that the σ-finite hypothesis in Exercise 1.7.7 cannot be removed. Let $\mathcal{A}$ be the collection of all subsets in $\mathbb{R}$ that can be expressed as finite unions of half-open intervals $[a, b)$. Let $\mu_0 : \mathcal{A} \to [0, +\infty]$ be the function such that $\mu_0(E) = +\infty$ for non-empty E and $\mu_0(\emptyset) = 0$.

- (i) Show that μ_0 is a pre-measure.
- (ii) Show that $\langle \mathcal{A} \rangle$ is the Borel σ-algebra $\mathcal{B}[\mathbb{R}]$.
- (iii) Show that the Hahn-Kolmogorov extension $\mu : \mathcal{B}[\mathbb{R}] \to [0, +\infty]$ of μ_0 assigns an infinite measure to any non-empty Borel set.
- (iv) Show that counting measure $\#$ (or more generally, $c\#$ for any $c \in (0, +\infty]$) is another extension of μ_0 on $\mathcal{B}[\mathbb{R}]$.

Solution. As in the proof of Theorem 1.7.9 we allow the endpoints to be infinite, so that $a \in [-\infty, +\infty)$ and $b \in (-\infty, +\infty]$ are permitted in $[a, b)$ (with $[-\infty, b) := (-\infty, b)$). This convention is forced on us: without it $\mathcal{A}$ would not be closed under complements (the complement of $[0, 1)$ is $(-\infty, 0) \cup [1, +\infty)$) and so would not be a Boolean algebra. With the convention, $\mathcal{A}$ is a Boolean algebra: it contains $\emptyset = [0, 0)$ and $\mathbb{R} = (-\infty, +\infty)$, it is closed under finite unions by construction, it is closed under finite intersections since $[a, b) \cap [c, d) = [\max(a, c), \min(b, d))$ is again a half-open interval (and intersection distributes over finite unions), and it is closed under complements since the complement of $[a, b)$ is $(-\infty, a) \cup [b, +\infty) \in \mathcal{A}$, and the complement of a finite union is the finite intersection of the complements.
The single structural fact we use repeatedly is this: a set $E \in \mathcal{A}$ is non-empty if and only if it contains a half-open interval $[a, b)$ with $a < b$, and any such interval is uncountable. Indeed, if $E = [a_1, b_1) \cup \cdots \cup [a_k, b_k)$ is non-empty then at least one of the constituent intervals is non-empty, i.e. has $a_i < b_i$.
(i) By Exercise 1.7.4(i) it suffices to check that $\mu_0(\emptyset) = 0$ (which holds by definition) and that μ_0 is countably additive on $\mathcal{A}$ in the sense of Definition 1.7.7; but let us simply verify both the finite and the countable additivity directly, as both are immediate.
Suppose $E_1, E_2, \cdots \in \mathcal{A}$ are disjoint and $E := \bigcup_{n=1}^{\infty} E_n$ lies in $\mathcal{A}$ (the finitely additive case is the special case in which all but finitely many E_n are empty). If every E_n is empty, then E is empty and both $\mu_0(E)$ and $\sum_n \mu_0(E_n)$ equal 0. Otherwise some E_{n_0} is non-empty; then $E \supset E_{n_0}$ is non-empty, so $\mu_0(E) = +\infty$, while $\sum_n \mu_0(E_n) \ge \mu_0(E_{n_0}) = +\infty$. In either case

$$\mu_0\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} \mu_0(E_n).$$

Thus μ_0 is a finitely additive measure obeying the countable additivity requirement of Definition 1.7.7, i.e. a pre-measure.

(ii) Every $[a,b)$ is Borel, being the intersection of the closed set $[a,+\infty)$ (or of $\mathbb{R}$, if $a=-\infty$) with the open set $(-\infty,b)$; hence $\mathcal{A}\subset\mathcal{B}[\mathbb{R}]$, and since $\mathcal{B}[\mathbb{R}]$ is a σ-algebra, $\langle\mathcal{A}\rangle\subset\mathcal{B}[\mathbb{R}]$.
Conversely, every open interval (a,b) with $-\infty\leq a<b\leq+\infty$ can be written as the countable union

$$(a,b)=\bigcup_{n=1}^{\infty}[a_n,b)$$

where a_n is any sequence of reals in (a,b) decreasing to a (take $a_n\to-\infty$ if $a=-\infty$); each $[a_n,b)\in\mathcal{A}$, so $(a,b)\in\langle\mathcal{A}\rangle$. Every open subset U of $\mathbb{R}$ is a countable union of open intervals (e.g. the union of all intervals with rational endpoints contained in U), hence lies in $\langle\mathcal{A}\rangle$. Since the open sets generate $\mathcal{B}[\mathbb{R}]$ by definition, $\mathcal{B}[\mathbb{R}]\subset\langle\mathcal{A}\rangle$. Combining, $\langle\mathcal{A}\rangle=\mathcal{B}[\mathbb{R}]$.
(iii) Recall from the proof of the Hahn-Kolmogorov theorem (Theorem 1.7.8) that the extension μ is the restriction to the Carathéodory measurable sets of the outer measure

$$\mu^*(E):=\inf\Big\{\sum_{n=1}^{\infty}\mu_0(E_n):E\subset\bigcup_{n=1}^{\infty}E_n,\ E_n\in\mathcal{A}\text{ for all }n\Big\}.$$

Let $E\subset\mathbb{R}$ be non-empty and let $E_1,E_2,\cdots\in\mathcal{A}$ cover E. Since $E\neq\emptyset$, at least one E_n must be non-empty, whence $\sum_n\mu_0(E_n)=+\infty$. Taking the infimum over all covers gives $\mu^*(E)=+\infty$. Of course $\mu^*(\emptyset)=0$ (cover $\emptyset$ by $\emptyset$). So

$$\mu^*(E)=\begin{cases}0, & E=\emptyset\\ +\infty, & E\neq\emptyset.\end{cases}$$

In particular $\mu(E)=\mu^*(E)=+\infty$ for every non-empty Borel set E, as claimed. (For completeness: every subset of $\mathbb{R}$ is Carathéodory measurable for this μ^*. Indeed, given $E,A\subset\mathbb{R}$, if $A=\emptyset$ then $\mu^*(A)=\mu^*(A\cap E)=\mu^*(A\setminus E)=0$; and if $A\neq\emptyset$ then at least one of $A\cap E$, $A\setminus E$ is non-empty, so both sides of the Carathéodory identity are $+\infty$. Thus $\mathcal{B}=2^{\mathbb{R}}$, and restricting to $\mathcal{B}[\mathbb{R}]$ is legitimate.)
(iv) Counting measure $\#$ on $\mathcal{B}[\mathbb{R}]$ is a countably additive measure, and for $c\in(0,+\infty]$ so is $c\#$, by Exercise 1.4.22(i). It remains to check that $c\#$ agrees with μ_0 on $\mathcal{A}$. We have $c\#(\emptyset)=c\cdot0=0=\mu_0(\emptyset)$ (using the convention $\infty\cdot0=0$ when $c=+\infty$). If $E\in\mathcal{A}$ is non-empty then, as noted at the outset, E contains an interval $[a,b)$ with $a<b$, which is an infinite (indeed uncountable) set, so $\#(E)=+\infty$ and hence $c\#(E)=+\infty=\mu_0(E)$ since $c>0$. Thus $c\#$ is a countably additive extension of μ_0 to $\langle\mathcal{A}\rangle=\mathcal{B}[\mathbb{R}]$.
Finally, this settles the point of the exercise. The measures μ and $\#$ are both countably additive extensions of μ_0 defined on the same σ-algebra $\mathcal{B}[\mathbb{R}]$, yet they disagree: $\#(\{0\})=1$ while $\mu(\{0\})=+\infty$. The hypothesis of Exercise 1.7.7 that fails is σ-finiteness: the only element of $\mathcal{A}$ of finite μ_0-measure is $\emptyset$, so $\mathbb{R}$ is certainly not a countable union of sets in $\mathcal{A}$ of finite μ_0-measure. Hence the σ-finite hypothesis in Exercise 1.7.7 cannot be removed.

Problem 1.7.9. Let $\mu_0:\mathcal{B}_0\to[0,+\infty]$ be a pre-measure which is σ-finite (thus X is the countable union of sets in $\mathcal{B}_0$ of finite μ_0-measure), and let $\mu:\mathcal{B}\to[0,+\infty]$ be the Hahn-Kolmogorov extension of μ_0.

- (i) Show that if $E\in\mathcal{B}$, then there exists $F\in\langle\mathcal{B}_0\rangle$ containing E such that $\mu(F\setminus E)=0$ (thus F consists of the union of E and a null set). Furthermore, show that F can be chosen to be a countable intersection $F=\bigcap_{n=1}^{\infty}F_n$ of sets F_n, each of which is a countable union $F_n=\bigcup_{m=1}^{\infty}F_{n,m}$ of sets $F_{n,m}$ in $\mathcal{B}_0$.
- (ii) If $E\in\mathcal{B}$ has finite measure (i.e. $\mu(E)<\infty$), and $\varepsilon>0$, show that there exists $F\in\mathcal{B}_0$ such that $\mu(E\Delta F)\leq\varepsilon$.
- (iii) Conversely, if E is a set such that for every $\varepsilon>0$ there exists $F\in\mathcal{B}_0$ such that $\mu^*(E\Delta F)\leq\varepsilon$, show that $E\in\mathcal{B}$.

Solution. Throughout, μ^* denotes the outer measure

$$\mu^*(E) := \inf\Big\{ \sum_{n=1}^{\infty} \mu_0(E_n) : E \subset \bigcup_{n=1}^{\infty} E_n,\ E_n \in \mathcal{B}_0 \Big\}$$

built in the proof of the Hahn-Kolmogorov theorem (Theorem 1.7.8), $\mathcal{B}$ is the σ-algebra of μ^*-Carathéodory measurable sets, and $\mu = \mu^*|_{\mathcal{B}}$. Recall from that proof that $\mathcal{B}_0 \subset \mathcal{B}$ and $\mu^* = \mu_0$ on $\mathcal{B}_0$; since $\mathcal{B}$ is a σ-algebra containing $\mathcal{B}_0$ we also have $\langle \mathcal{B}_0 \rangle \subset \mathcal{B}$.
Two conventions will save repetition. First, by σ-finiteness we may write $X = \bigcup_{k=1}^{\infty} X_k$ with $X_k \in \mathcal{B}_0$ and $\mu_0(X_k) < \infty$; replacing X_k by $X_k \setminus (X_1 \cup \cdots \cup X_{k-1})$ (which lies in $\mathcal{B}_0$, as $\mathcal{B}_0$ is a Boolean algebra, and has smaller μ_0-measure by monotonicity) we may and do assume the X_k are disjoint. Second, call a set a $\mathcal{B}_0$-union if it is a countable union of sets in $\mathcal{B}_0$. The intersection of two $\mathcal{B}_0$-unions is again a $\mathcal{B}_0$-union, since

$$\Big(\bigcup_m A_m\Big) \cap \Big(\bigcup_l B_l\Big) = \bigcup_{m,l} (A_m \cap B_l)$$

is a countable union of sets $A_m \cap B_l \in \mathcal{B}_0$.
(i) Step 1: the case $\mu(E) < \infty$. Fix $n \geq 1$. By the definition of μ^* there are $F_{n,m} \in \mathcal{B}_0$ $(m = 1, 2, \ldots)$ with $E \subset F_n := \bigcup_m F_{n,m}$ and

$$\sum_{m=1}^{\infty} \mu_0(F_{n,m}) \leq \mu^*(E) + \tfrac{1}{n}.$$

By countable subadditivity of the outer measure μ^* and $\mu^*(F_{n,m}) = \mu_0(F_{n,m})$,

$$\mu^*(F_n) \leq \sum_{m=1}^{\infty} \mu_0(F_{n,m}) \leq \mu(E) + \tfrac{1}{n}.$$

Put $F := \bigcap_{n=1}^{\infty} F_n$. Then $F \in \langle \mathcal{B}_0 \rangle \subset \mathcal{B}$, $E \subset F$, and by monotonicity $\mu(F) \leq \mu^*(F_n) \leq \mu(E) + \frac{1}{n}$ for every n, so $\mu(F) \leq \mu(E)$; the reverse inequality is monotonicity again, so $\mu(F) = \mu(E) < \infty$. Since $E, F \in \mathcal{B}$ with $E \subset F$ and $\mu(E)$ finite, additivity of μ on $\mathcal{B}$ gives

$$\mu(F \setminus E) = \mu(F) - \mu(E) = 0.$$

By construction F is a countable intersection of $\mathcal{B}_0$-unions. Moreover we may replace each F_n by $F_1 \cap \cdots \cap F_n$, which is still a $\mathcal{B}_0$-union by the remark above and does not change F; so we may assume the F_n are decreasing. Finally, if $Y \in \mathcal{B}_0$ is any set with $E \subset Y$, we may replace each F_n by $F_n \cap Y = \bigcup_m (F_{n,m} \cap Y)$, again a $\mathcal{B}_0$-union; this shrinks F to $F \cap Y$, which still contains E and still has $\mu(F \cap Y \setminus E) = 0$. So in the finite-measure case we may additionally demand $F \subset Y$ for any prescribed $\mathcal{B}_0$-set $Y \supset E$.
Step 2: the general case. Let $E \in \mathcal{B}$ be arbitrary and set $E^{(k)} := E \cap X_k \in \mathcal{B}$. Then $\mu(E^{(k)}) \leq \mu(X_k) = \mu_0(X_k) < \infty$, so Step 1 applies: there are $G_n^{(k)}$, each a $\mathcal{B}_0$-union, decreasing in n, with

$$F^{(k)} := \bigcap_{n=1}^{\infty} G_n^{(k)} \supset E^{(k)}, \qquad \mu(F^{(k)} \setminus E^{(k)}) = 0,$$

and (taking $Y = X_k$ in the last remark of Step 1) with $F^{(k)} \subset X_k$; indeed we may take $G_n^{(k)} \subset X_k$ for every n.
Set $F := \bigcup_{k=1}^{\infty} F^{(k)}$. Then $E = \bigcup_k E^{(k)} \subset F$, and

$$F \setminus E = \bigcup_k \big(F^{(k)} \setminus E\big) \subset \bigcup_k \big(F^{(k)} \setminus E^{(k)}\big),$$

so by countable subadditivity $\mu(F \setminus E) = 0$.
It remains to exhibit F in the required form. I claim

$$F = \bigcap_{n=1}^{\infty} F_n, \qquad F_n := \bigcup_{k=1}^{\infty} G_n^{(k)}.$$

Each F_n is a countable union of $\mathcal{B}_0$-unions, hence itself a $\mathcal{B}_0$-union, i.e. $F_n = \bigcup_m F_{n,m}$ with $F_{n,m} \in \mathcal{B}_0$ after re-indexing the countable family by a single index m. For the claim: if $x \in F$ then $x \in F^{(k)}$ for some k, so $x \in G_n^{(k)} \subset F_n$ for every n, giving $x \in \bigcap_n F_n$. Conversely suppose $x \in \bigcap_n F_n$. For each n pick k_n with $x \in G_n^{(k_n)} \subset X_{k_n}$. Since the X_k are disjoint, x lies in at most one of them, so all the k_n are equal to a common value k. Hence $x \in G_n^{(k)}$ for every n, i.e. $x \in F^{(k)} \subset F$. This proves the claim, and hence part (i).
(ii) Let $E \in \mathcal{B}$ with $\mu(E) < \infty$ and let $\varepsilon > 0$. By the definition of $\mu^*(E) = \mu(E)$ there are $E_1, E_2, \cdots \in \mathcal{B}_0$ with $E \subset \bigcup_n E_n$ and

$$\sum_{n=1}^{\infty} \mu_0(E_n) \leq \mu(E) + \tfrac{\varepsilon}{2} < \infty.$$

Since the series converges, there is N with $\sum_{n>N} \mu_0(E_n) \leq \varepsilon/2$. Put $F := E_1 \cup \cdots \cup E_N \in \mathcal{B}_0$ (a Boolean algebra is closed under finite unions).
On one hand $E \setminus F \subset \bigcup_{n>N} E_n$, so by countable subadditivity

$$\mu(E \setminus F) \leq \sum_{n>N} \mu_0(E_n) \leq \tfrac{\varepsilon}{2}.$$

On the other hand, writing $G := \bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$, we have $F \setminus E \subset G \setminus E$ and $\mu(G) \leq \sum_n \mu_0(E_n) \leq \mu(E) + \varepsilon/2 < \infty$; since $E \subset G$ are both in $\mathcal{B}$ and $\mu(E) < \infty$,

$$\mu(F \setminus E) \leq \mu(G \setminus E) = \mu(G) - \mu(E) \leq \tfrac{\varepsilon}{2}.$$

Adding, $\mu(E \Delta F) = \mu(E \setminus F) + \mu(F \setminus E) \leq \varepsilon$, as required.
(iii) Let $E \subset X$ be such that for every $\varepsilon > 0$ there is $F \in \mathcal{B}_0$ with $\mu^*(E \Delta F) \leq \varepsilon$. We verify the Carathéodory criterion for E: for every $A \subset X$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E).$$

The inequality $\leq$ is subadditivity of μ^*, so only $\geq$ needs proof, and we may assume $\mu^*(A) < \infty$, the claim being trivial otherwise.
Fix $\varepsilon > 0$ and choose $F \in \mathcal{B}_0$ with $\mu^*(E \Delta F) \leq \varepsilon$. Since $E \setminus F$ and $F \setminus E$ are both subsets of $E \Delta F$, monotonicity gives $\mu^*(E \setminus F) \leq \varepsilon$ and $\mu^*(F \setminus E) \leq \varepsilon$. Now

$$A \cap E \subset (A \cap F) \cup (E \setminus F), \qquad A \setminus E \subset (A \setminus F) \cup (F \setminus E),$$

because a point of $A \cap E$ either lies in F, and then in $A \cap F$, or does not, and then lies in $E \setminus F$; similarly for the second inclusion. Hence by monotonicity and subadditivity of μ^*,

$$\mu^*(A \cap E) + \mu^*(A \setminus E) \leq \mu^*(A \cap F) + \mu^*(A \setminus F) + \mu^*(E \setminus F) + \mu^*(F \setminus E).$$

Since $F \in \mathcal{B}_0 \subset \mathcal{B}$, F is Carathéodory measurable, so $\mu^*(A \cap F) + \mu^*(A \setminus F) = \mu^*(A)$. Therefore

$$\mu^*(A \cap E) + \mu^*(A \setminus E) \leq \mu^*(A) + 2\varepsilon.$$

As $\varepsilon > 0$ was arbitrary and $\mu^*(A) < \infty$, we conclude $\mu^*(A \cap E) + \mu^*(A \setminus E) \leq \mu^*(A)$. Thus E is Carathéodory measurable, i.e. $E \in \mathcal{B}$.
(Note that parts (ii) and (iii) did not need the σ-finiteness hypothesis; it was used only in Step 2 of part (i), where it is genuinely needed, as Exercise 1.7.8 shows.)

Problem 1.7.10. Verify the claims marked 'Exercise!' in the above proof.
[The proof in question is that of Theorem 1.7.9 (existence of Lebesgue-Stieltjes measure), and the four claims it marks 'Exercise!' are, in order:

- (a) the additivity property $|I \cup J|_F = |I|_F + |J|_F$ whenever I, J are disjoint intervals that share

a common endpoint;
- (b) that the measure μ_0 on the Boolean algebra $\mathcal{B}_0$ generated by the (possibly infinite) intervals, defined by $\mu_0(E) = |I_1|_F + \cdots + |I_k|_F$ whenever $E = I_1 \cup \cdots \cup I_k$ is the disjoint union of finitely many intervals, is well-defined and finitely additive;
- (c) the identity (1.35), namely $\mu_0(E) = \sup_{K \subset E} \mu_0(K)$, where K ranges over all compact intervals contained in the interval E;
- (d) the identity $\mu_0(E) = \inf_{U \supset E} \mu_0(U)$, where U ranges over all open intervals containing the interval E.]

The measure μ_F given by the above theorem is known as the Lebesgue-Stieltjes measure μ_F of F. (In some texts, this measure is only defined when F is right-continuous, or equivalently if $F = F_+$.)

Solution. Throughout, $F : \mathbb{R} \to \mathbb{R}$ is monotone non-decreasing, $F_-(x) = \sup_{y<x} F(y)$, $F_+(x) = \inf_{y>x} F(y)$, with the conventions $F_-(+\infty) = \sup_{y\in\mathbb{R}} F(y)$ and $F_+(-\infty) = \inf_{y\in\mathbb{R}} F(y)$ (these are exactly the same formulas, read at $x = \pm\infty$). Intervals are allowed to be unbounded.

It is convenient to record the F-volume in one formula. Given a non-empty interval I with left endpoint $a \in [-\infty, +\infty)$ and right endpoint $b \in (-\infty, +\infty]$, define its left and right heights

$$\ell(I) := \begin{cases} F_-(a) & \text{if } a \in I \\ F_+(a) & \text{if } a \notin I \end{cases} \qquad r(I) := \begin{cases} F_+(b) & \text{if } b \in I \\ F_-(b) & \text{if } b \notin I. \end{cases}$$

Comparing with the four definitions of $|I|_F$ (and with $|\{a\}|_F = F_+(a) - F_-(a)$, the case $a = b$), we see in every case

$$|I|_F = r(I) - \ell(I),$$

and we set $|\emptyset|_F = 0$. Note $\ell(I) \le r(I)$: if $a < b$ pick $a < y < b$, then $\ell(I) \le F_+(a) \le F(y) \le F_-(b) \le r(I)$; while if $a = b$ the only non-empty interval is $\{a\}$, for which $\ell = F_-(a) \le F_+(a) = r$. Hence $|I|_F \in [0, +\infty]$, as claimed in the text. Also, for finite x the numbers $F_-(x), F(x), F_+(x)$ are all real (they are squeezed between $F(x-1)$ and $F(x+1)$), so infinities can only enter through an infinite endpoint.

We shall use two elementary one-sided limit facts.

Lemma A. For any $a \in [-\infty, +\infty)$, $\inf_{c>a} F_-(c) = F_+(a)$; consequently $F_-(c_n) \to F_+(a)$ whenever $c_n \downarrow a$ with $c_n > a$ (the sequence $F_-(c_n)$ being monotone).

Indeed, for $c > a$ we have $F_-(c) = \sup_{y<c} F(y) \ge \inf_{y>a} F(y) = F_+(a)$, because $F_-(c) \ge F(y)$ for any y with $a < y < c$. Conversely, given any $y > a$, choose c with $a < c < y$; every $z < c$ satisfies $z < y$, so $F_-(c) \le F(y)$, whence $\inf_{c>a} F_-(c) \le \inf_{y>a} F(y) = F_+(a)$.

Lemma B. For any $b \in (-\infty, +\infty]$, $\sup_{d<b} F_+(d) = F_-(b)$; consequently $F_+(d_n) \to F_-(b)$ whenever $d_n \uparrow b$ with $d_n < b$.

Indeed, for $d < b$, choosing $d < y < b$ gives $F_+(d) \le F(y) \le F_-(b)$; and for any $y < b$, choosing $y < d < b$ gives $F_+(d) \ge F(d) \ge F(y)$, so $\sup_{d<b} F_+(d) \ge \sup_{y<b} F(y) = F_-(b)$.

(a) Additivity for adjacent intervals. Let I, J be disjoint intervals sharing a common endpoint and with $I \cup J$ again an interval (this is implicit in the assertion; note that if c lay in neither I nor J then $I \cup J$ would not be an interval unless one of them is empty). If one of them is empty the claim is trivial, so assume both are non-empty; then, relabelling, I lies to the left of J, I has right endpoint c, J has left endpoint c, and c is finite. Let a be the left endpoint of I and b the right endpoint of J, so that $K := I \cup J$ is the interval with endpoints a, b, containing a if and only if I does and containing b if and only if J does; thus

$$\ell(K) = \ell(I), \qquad r(K) = r(J).$$

Since I, J are disjoint and $I \cup J$ is an interval, c lies in exactly one of them. If $c \in I$ then $r(I) = F_+(c)$ and, as $c \notin J$, $\ell(J) = F_+(c)$. If instead $c \in J$ then $r(I) = F_-(c)$ and $\ell(J) = F_-(c)$. Either way $r(I) = \ell(J) =: t$, and t is a real number since c is finite. Therefore

$$|I|_F + |J|_F = (t - \ell(I)) + (r(J) - t) = r(J) - \ell(I) = |K|_F,$$

the cancellation of the finite number t being legitimate even if $\ell(I) = -\infty$ or $r(J) = +\infty$.

By induction we get the stated corollary: if an interval I is partitioned into finitely many disjoint sub-intervals $I_1, \dots, I_k$, then $|I|_F = |I_1|_F + \cdots + |I_k|_F$. Indeed, discard the empty I_j and order the rest from left to right (they are disjoint intervals, so they are totally ordered by position). Then for each j, $I_1 \cup \cdots \cup I_j$ is an interval: it is the set of points of I lying to the left of I_{j+1}, hence an interval, and it is adjacent to I_{j+1}, sharing the endpoint separating them. Applying (a) repeatedly gives the claim.
(b) μ_0 is well-defined and finitely additive. First, every $E \in \mathcal{B}_0$ is a finite union of disjoint intervals. Indeed, write $E = I_1 \cup \cdots \cup I_k$ as a finite union of intervals and let $P \subset \mathbb{R}$ be the finite set of all finite endpoints of the I_j. The set P cuts $\mathbb{R}$ into finitely many atoms: the points of P, and the finitely many open intervals into which $\mathbb{R} \setminus P$ decomposes. Each atom is either contained in a given I_j or disjoint from it (an atom cannot straddle an endpoint of I_j), hence each atom is either contained in E or disjoint from E; so E is the union of the atoms it contains, a finite union of disjoint intervals.
Now suppose $E = I_1 \sqcup \cdots \sqcup I_k = J_1 \sqcup \cdots \sqcup J_{k'}$ are two such decompositions. Each I_j is partitioned by the disjoint intervals $I_j \cap J_l$ ($l = 1, \dots, k'$), which are intervals since an intersection of two intervals is an interval; so by the corollary in (a), $|I_j|_F = \sum_l |I_j \cap J_l|_F$. Symmetrically $|J_l|_F = \sum_j |I_j \cap J_l|_F$. Summing over j and over l respectively (all terms being non-negative, so the finite sums may be rearranged freely even when infinite),

$$\sum_{j=1}^{k} |I_j|_F = \sum_{j,l} |I_j \cap J_l|_F = \sum_{l=1}^{k'} |J_l|_F.$$

So $\mu_0(E)$ does not depend on the chosen decomposition, i.e. μ_0 is well-defined on $\mathcal{B}_0$.
Finite additivity is now immediate: if $E = E_1 \sqcup \cdots \sqcup E_n$ with all sets in $\mathcal{B}_0$, decompose each E_i into disjoint intervals $I_{i,1}, \dots, I_{i,k_i}$; the totality of these intervals is a decomposition of E into finitely many disjoint intervals, so by well-definedness

$$\mu_0(E) = \sum_{i=1}^{n} \sum_{j=1}^{k_i} |I_{i,j}|_F = \sum_{i=1}^{n} \mu_0(E_i).$$

In particular $\mu_0(\emptyset) = 0$ and μ_0 is monotone ($A \subset B$ in $\mathcal{B}_0$ gives $\mu_0(B) = \mu_0(A) + \mu_0(B \setminus A) \geq \mu_0(A)$) and finitely subadditive.
(c) Inner regularity by compact intervals, i.e. (1.35). Let E be an interval with endpoints $a \leq b$; we must show $|E|_F = \sup_K |K|_F$ over compact intervals $K \subset E$. The inequality $\geq$ is the monotonicity of μ_0 just noted. For $\leq$, we may assume $E \neq \emptyset$. Choose sequences c_n, d_n as follows: if $a \in E$ set $c_n := a$ for all n, otherwise let $c_n \downarrow a$ with $c_n > a$ and $c_n \in E$; if $b \in E$ set $d_n := b$, otherwise let $d_n \uparrow b$ with $d_n < b$ and $d_n \in E$. (If E is a single point this reads $c_n = d_n = a$.) For n large enough $c_n \leq d_n$, and $K_n := [c_n, d_n]$ is a compact interval contained in E, with

$$|K_n|_F = F_+(d_n) - F_-(c_n).$$

If $a \in E$ then $F_-(c_n) = F_-(a) = \ell(E)$; if $a \notin E$ then $F_-(c_n) \to F_+(a) = \ell(E)$ by Lemma A (this covers $a = -\infty$, where necessarily $a \notin E$). Likewise, if $b \in E$ then $F_+(d_n) = F_+(b) = r(E)$, and if $b \notin E$ then $F_+(d_n) \to F_-(b) = r(E)$ by Lemma B. Hence

$$|K_n|_F \to r(E) - \ell(E) = |E|_F$$

in $[0, +\infty]$, which gives $\sup_K |K|_F \geq |E|_F$ and so (1.35).
(d) Outer regularity by open intervals. Again let E be an interval with endpoints $a \leq b$; the inequality $\inf_U |U|_F \geq |E|_F$ is monotonicity, and for $E = \emptyset$ we may take $U = \emptyset$, so assume $E \neq \emptyset$ and prove $\inf_U |U|_F \leq |E|_F$. Choose c_n and d_n this time from outside E: if $a \notin E$ set $c_n := a$, otherwise let $c_n \uparrow a$ with $c_n < a$; if $b \notin E$ set $d_n := b$, otherwise let $d_n \downarrow b$ with $d_n > b$. Then $U_n := (c_n, d_n)$ is an open interval containing E (its endpoints lie weakly outside those of E, and any endpoint of E belonging to E has been strictly passed), and

$$|U_n|_F = F_-(d_n) - F_+(c_n).$$

If $a \notin E$ then $F_+(c_n) = F_+(a) = \ell(E)$; if $a \in E$ then $c_n \uparrow a$ with $c_n < a$, so $F_+(c_n) \to F_-(a) = \ell(E)$ by Lemma B. Symmetrically, if $b \notin E$ then $F_-(d_n) = F_-(b) = r(E)$, and if $b \in E$ then $F_-(d_n) \to F_+(b) = r(E)$ by Lemma A. Hence $|U_n|_F \to r(E) - \ell(E) = |E|_F$, so $\inf_U |U|_F \leq |E|_F$, and the identity follows.

Problem 1.7.11. Define a Radon measure on $\mathbb{R}$ to be a Borel measure μ obeying the following additional properties:

- (i) (Local finiteness) $\mu(K) < \infty$ for every compact K.
- (ii) (Inner regularity) One has $\mu(E) = \sup_{K \subset E,\, K \text{ compact}} \mu(K)$ for every Borel set E.
- (iii) (Outer regularity) One has $\mu(E) = \inf_{U \supset E,\, U \text{ open}} \mu(U)$ for every Borel set E.

Show that for every monotone function $F : \mathbb{R} \to \mathbb{R}$, the Lebesgue-Stieltjes measure μ_F is a Radon measure on $\mathbb{R}$; conversely, if μ is a Radon measure on $\mathbb{R}$, show that there exists a monotone function $F : \mathbb{R} \to \mathbb{R}$ such that $\mu = \mu_F$.

Radon measures are studied in more detail in Section 1.10 of An epsilon of room, Vol. I.

Solution. (Here and below, monotone means monotone non-decreasing, since that is the setting in which Theorem 1.7.9 defines μ_F.)

Part 1: μ_F is Radon. Write $\mathcal{B}_0$ for the Boolean algebra generated by the intervals, $\mu_0(E) = |I_1|_F + \cdots + |I_k|_F$ for the pre-measure of Theorem 1.7.9, and recall that μ_F is the restriction to $\mathcal{B}[\mathbb{R}]$ of the Hahn-Kolmogorov extension of μ_0; in particular $\mu_F = \mu_0$ on $\mathcal{B}_0$, and $\mu_F(E) = \mu^*(E)$ for Borel E, where

$$\mu^*(E) = \inf\Big\{ \sum_{n=1}^{\infty} \mu_0(E_n) : E \subset \bigcup_{n=1}^{\infty} E_n,\ E_n \in \mathcal{B}_0 \Big\}.$$

Splitting each E_n into its finitely many component intervals (which only refines the cover and, by finite additivity of μ_0, does not change the sum), we may restrict the infimum to covers by intervals I_n:

$$\mu_F(E) = \inf\Big\{ \sum_{n=1}^{\infty} |I_n|_F : E \subset \bigcup_{n=1}^{\infty} I_n,\ I_n \text{ intervals} \Big\}.$$

(i) Local finiteness. A compact $K \subset \mathbb{R}$ is bounded, so $K \subset [-N, N]$ for some N, and by monotonicity and (1.33),

$$\mu_F(K) \le \mu_F([-N, N]) = F_+(N) - F_-(-N) < \infty,$$

since $F_-(-N)$ and $F_+(N)$ are real numbers (F takes real values and is monotone, so $F(-N-1) \le F_-(-N) \le F_+(N) \le F(N+1)$).

(iii) Outer regularity. The inequality $\mu_F(E) \le \inf_{U \supset E} \mu_F(U)$ is monotonicity. For the converse, we may assume $\mu_F(E) < \infty$ (otherwise take $U = \mathbb{R}$). Let $\varepsilon > 0$. By the displayed formula for $\mu_F(E)$ there are intervals I_n covering E with $\sum_n |I_n|_F \le \mu_F(E) + \varepsilon/2$. In particular each $|I_n|_F$ is finite, so by claim (d) of Exercise 1.7.10 (outer regularity of $|\cdot|_F$ on intervals by open intervals) we may choose, for each n, an open interval $U_n \supset I_n$ with

$$|U_n|_F \le |I_n|_F + \varepsilon 2^{-n-1}.$$

Then $U := \bigcup_n U_n$ is open and contains E, and by countable subadditivity of μ_F,

$$\mu_F(U) \le \sum_{n=1}^{\infty} \mu_F(U_n) = \sum_{n=1}^{\infty} |U_n|_F \le \mu_F(E) + \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \mu_F(E) + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $\inf_{U \supset E} \mu_F(U) \le \mu_F(E)$.

(ii) Inner regularity. Again $\sup_{K \subset E} \mu_F(K) \le \mu_F(E)$ is monotonicity, so only the reverse is at issue.

Step 1: suppose first $E \subset [-N, N]$ for some N. Put $E' := [-N, N] \setminus E$, a Borel set; both E and E' have finite measure by local finiteness. Let $\varepsilon > 0$. By outer regularity, choose an open $U \supset E'$ with $\mu_F(U) \le \mu_F(E') + \varepsilon$. Set $K := [-N, N] \setminus U$. Then K is closed and bounded, hence compact, and $K \subset [-N, N] \setminus E' = E$. Moreover $U \cap [-N, N]$ is the disjoint union of E' (which is contained in U) and $E \cap U$, so

$$\mu_F(E \cap U) = \mu_F(U \cap [-N, N]) - \mu_F(E') \le \mu_F(U) - \mu_F(E') \le \varepsilon,$$

all quantities being finite. Since $E \setminus K = E \cap U$, we get $\mu_F(K) = \mu_F(E) - \mu_F(E \cap U) \ge \mu_F(E) - \varepsilon$.

Step 2: general Borel E. Let $E_N := E \cap [-N, N]$. Then E_N increases to E, so by upward monotone convergence for measures (Exercise 1.4.23), $\mu_F(E_N) \to \mu_F(E)$. Given any real $t < \mu_F(E)$, pick N with $\mu_F(E_N) > t$, and then by Step 1 (applied with $\varepsilon := \mu_F(E_N) - t$) a compact $K \subset E_N \subset E$ with $\mu_F(K) > t$. Hence $\sup_{K \subset E} \mu_F(K) \geq \mu_F(E)$, completing the proof that μ_F is Radon.
Part 2: every Radon measure is a Lebesgue-Stieltjes measure. Let μ be a Radon measure on $\mathbb{R}$; in fact only local finiteness will be used. Define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(x) := \begin{cases} \mu([0, x)) & \text{if } x \geq 0 \\ -\mu([x, 0)) & \text{if } x < 0. \end{cases}$$

These are finite numbers, since $[0, x)$ and $[x, 0)$ are contained in compact intervals. F is monotone non-decreasing: for $0 \leq x \leq y$ we have $[0, x) \subset [0, y)$; for $x \leq y < 0$ we have $[y, 0) \subset [x, 0)$ so $-\mu([y, 0)) \geq -\mu([x, 0))$; and for $x < 0 \leq y$ we have $F(x) \leq 0 \leq F(y)$.
The basic identity, which we label $(*)$, is

$$\mu([a, b)) = F(b) - F(a) \qquad (-\infty < a < b < \infty).$$

Indeed, if $0 \leq a < b$ then $[0, b) = [0, a) \sqcup [a, b)$ and all measures are finite, so $\mu([a, b)) = \mu([0, b)) - \mu([0, a)) = F(b) - F(a)$. If $a < b \leq 0$ then $[a, 0) = [a, b) \sqcup [b, 0)$, so $\mu([a, b)) = \mu([a, 0)) - \mu([b, 0)) = -F(a) + F(b)$. If $a < 0 \leq b$ then $[a, b) = [a, 0) \sqcup [0, b)$, so $\mu([a, b)) = -F(a) + F(b)$.
Next we identify F_- and F_+. For $x \in \mathbb{R}$ and $y < x$, $(*)$ gives $F(x) - F(y) = \mu([y, x))$, and $[y, x) \downarrow \emptyset$ as $y \uparrow x$ along any sequence, with $\mu([y_0, x)) < \infty$; by downward monotone convergence (Exercise 1.4.23) $\mu([y, x)) \to 0$. Hence $F(y) \to F(x)$ as $y \uparrow x$, i.e.

$$F_-(x) = \sup_{y<x} F(y) = F(x).$$

Similarly, for $y > x$, $(*)$ gives $F(y) - F(x) = \mu([x, y))$, and $[x, y) \downarrow \{x\}$ as $y \downarrow x$, with finite measure, so $\mu([x, y)) \to \mu(\{x\})$ and

$$F_+(x) = \inf_{y>x} F(y) = F(x) + \mu(\{x\}) = F_-(x) + \mu(\{x\}).$$

We can now verify that μ satisfies (1.33) and (1.34) for this F, on all bounded intervals. From the last display,

$$\mu(\{a\}) = F_+(a) - F_-(a),$$

which is (1.34). From $(*)$ and $F = F_-$,

$$\mu([a, b)) = F_-(b) - F_-(a).$$

Since all the sets involved have finite measure, additivity then gives

$$\mu([a, b]) = \mu([a, b)) + \mu(\{b\}) = F_-(b) - F_-(a) + F_+(b) - F_-(b) = F_+(b) - F_-(a),$$

$$\mu((a, b]) = \mu([a, b]) - \mu(\{a\}) = F_+(b) - F_-(a) - (F_+(a) - F_-(a)) = F_+(b) - F_+(a),$$

$$\mu((a, b)) = \mu([a, b)) - \mu(\{a\}) = F_-(b) - F_-(a) - (F_+(a) - F_-(a)) = F_-(b) - F_+(a).$$

Thus μ is a Borel measure obeying (1.33) and (1.34) for the monotone non-decreasing function F. By the uniqueness assertion of Theorem 1.7.9, $\mu = \mu_F$.
(Combining the two halves, we see that on $\mathbb{R}$ a Borel measure is Radon if and only if it is locally finite: local finiteness alone forced $\mu = \mu_F$, and every μ_F is Radon by Part 1.)

Problem 1.7.12. (Near uniqueness) If $F, F' : \mathbb{R} \to \mathbb{R}$ are monotone non-decreasing functions, show that $\mu_F = \mu_{F'}$ if and only if there exists a constant $C \in \mathbb{R}$ such that $F_+(x) = F'_+(x) + C$ and $F_-(x) = F'_-(x) + C$ for all $x \in \mathbb{R}$. Note that this implies that the value of F at its points of

discontinuity are irrelevant for the purposes of determining the Lebesgue-Stieltjes measure μ_F; in particular, $\mu_F = \mu_{F_+} = \mu_{F_-}$.
In the special case when $F_+(-\infty) = 0$ and $F_-(+\infty) = 1$, then μ_F is a probability measure, and $F_+(x) = \mu_F((-\infty, x])$ is known as the cumulative distribution function of μ_F.

Solution. Recall from (1.33) and (1.34) that for all real $a < b$ and all real a,

$$\mu_F([a,b]) = F_+(b) - F_-(a), \qquad \mu_F(\{a\}) = F_+(a) - F_-(a),$$

and note that all the numbers $F_\pm(x)$ at finite x are real (being squeezed between $F(x-1)$ and $F(x+1)$), so the identities above may be manipulated by ordinary arithmetic.
Sufficiency. Suppose there is $C \in \mathbb{R}$ with $F_+ = F'_+ + C$ and $F_- = F'_- + C$ pointwise. Then for all $-\infty < a < b < \infty$,

$$F_+(b) - F_-(a) = (F'_+(b) + C) - (F'_-(a) + C) = F'_+(b) - F'_-(a),$$

and likewise $F_-(b) - F_-(a) = F'_-(b) - F'_-(a)$, $F_+(b) - F_+(a) = F'_+(b) - F'_+(a)$, and $F_-(b) - F_+(a) = F'_-(b) - F'_+(a)$; also $F_+(a) - F_-(a) = F'_+(a) - F'_-(a)$. In other words, the Borel measure $\mu_{F'}$ satisfies the defining identities (1.33) and (1.34) with F' replaced by F. By the uniqueness assertion of Theorem 1.7.9 (which says there is exactly one Borel measure obeying (1.33) and (1.34) for the given F), $\mu_{F'} = \mu_F$.
Necessity. Suppose conversely $\mu_F = \mu_{F'}$. Applying both measures to $[a,b]$ with $a < b$ gives

$$F_+(b) - F_-(a) = F'_+(b) - F'_-(a),$$

i.e., writing $G_+ := F_+ - F'_+$ and $G_- := F_- - F'_-$ (real-valued functions on $\mathbb{R}$),

$$G_+(b) = G_-(a) \qquad \text{whenever } a < b.$$

Fix any $b_0 \in \mathbb{R}$ and set $C := G_+(b_0)$. For an arbitrary $x \in \mathbb{R}$: choosing $a := \min(x, b_0) - 1$, which is smaller than both x and b_0, we get $G_+(x) = G_-(a) = G_+(b_0) = C$; and choosing $b := \max(x, b_0) + 1$, which exceeds both x and b_0, we get $G_-(x) = G_+(b) = G_-(b_0 - 1) = G_+(b_0) = C$ (the last step using $b_0 - 1 < b_0$). Hence $G_+ \equiv G_- \equiv C$, i.e. $F_+ = F'_+ + C$ and $F_- = F'_- + C$, as required.
The stated consequences. First, $\mu_F = \mu_{F_+} = \mu_{F_-}$. By the criterion just proved (with $C = 0$) it suffices to check

$$(F_+)_+ = F_+, \quad (F_+)_- = F_-, \quad (F_-)_+ = F_+, \quad (F_-)_- = F_-,$$

noting that F_+ and F_- are themselves monotone non-decreasing real-valued functions. Now $(F_-)_+(x) = \inf_{y>x} F_-(y) = F_+(x)$ is exactly Lemma A of Exercise 1.7.10, and $(F_+)_-(x) = \sup_{y<x} F_+(y) = F_-(x)$ is Lemma B. For $(F_+)_+(x) = \inf_{y>x} F_+(y)$: monotonicity gives $F_+(y) \geq F_+(x)$ for $y > x$, so the infimum is $\geq F_+(x)$; conversely for $z > x$ pick $x < y < z$, then $F_+(y) \leq F(z)$, so the infimum is $\leq \inf_{z>x} F(z) = F_+(x)$. Symmetrically $(F_-)_-(x) = \sup_{y<x} F_-(y) = F_-(x)$: it is $\leq F_-(x)$ by monotonicity, and for $z < x$ picking $z < y < x$ gives $F_-(y) \geq F(z)$, so it is $\geq \sup_{z<x} F(z) = F_-(x)$.
Second, the values of F at its discontinuity points are irrelevant. Let S be the set of points at which F is continuous; since F is monotone, its discontinuity set is at most countable, so S is dense in $\mathbb{R}$. Suppose $G : \mathbb{R} \to \mathbb{R}$ is another monotone non-decreasing function with $G = F$ on S. For any x and any $y < x$, density of S gives $y' \in S$ with $y < y' < x$, and $F(y) \leq F(y')$; hence

$$F_-(x) = \sup_{y<x} F(y) = \sup_{y<x,\ y\in S} F(y) = \sup_{y<x,\ y\in S} G(y) = G_-(x),$$

the last equality by the same density argument applied to the monotone function G. Symmetrically $F_+ = G_+$. So by the criterion (with $C = 0$), $\mu_F = \mu_G$: modifying F at its discontinuity points (in any way that keeps it monotone) does not change μ_F.
Finally, in the special case $F_+(-\infty) = 0$ and $F_-(+\infty) = 1$, the interval $(-\infty, +\infty)$ has $\mu_F(\mathbb{R}) = F_-(+\infty) - F_+(-\infty) = 1$, so μ_F is a probability measure; and $\mu_F((-\infty, x]) = F_+(x) - F_+(-\infty) = F_+(x)$, the cumulative distribution function.

Problem 1.7.13. (Lebesgue-Stieltjes measure, absolutely continuous case)

- (i) If $F : \mathbb{R} \to \mathbb{R}$ is the identity function $F(x) = x$, show that μ_F is equal to Lebesgue measure m.
- (ii) If $F : \mathbb{R} \to \mathbb{R}$ is monotone non-decreasing and absolutely continuous (which in particular implies that F' exists and is absolutely integrable), show that $\mu_F = m_{F'}$ in the sense of Exercise 1.4.49, thus

$$\mu_F(E) = \int_E F'(x)\,dx$$

for any Borel measurable E, and

$$\int_{\mathbb{R}} f(x)\,d\mu_F(x) = \int_{\mathbb{R}} f(x)F'(x)\,dx$$

for any unsigned Borel measurable $f : \mathbb{R} \to [0, +\infty]$.
In view of the above exercise, the integral $\int_{\mathbb{R}} f\,d\mu_F$ is often abbreviated $\int_{\mathbb{R}} f\,dF$, and referred to as the Lebesgue-Stieltjes integral of f with respect to F. In particular, observe the identity

$$\int_{[a,b]} dF = F_+(b) - F_-(a)$$

for any monotone non-decreasing $F : \mathbb{R} \to \mathbb{R}$ and any $-\infty < a < b < +\infty$, which can be viewed as yet another formulation of the fundamental theorem of calculus.

Solution. (i) If $F(x) = x$ then F is continuous, so $F_-(x) = F_+(x) = x$ for every x. Thus the identities (1.33), (1.34) characterising μ_F read

$$\mu_F([a,b]) = \mu_F([a,b)) = \mu_F((a,b]) = \mu_F((a,b)) = b - a, \qquad \mu_F(\{a\}) = 0.$$

Let m denote the restriction of Lebesgue measure to $\mathcal{B}[\mathbb{R}]$, which is a Borel measure. Lebesgue measure of an interval is its length, and singletons are Lebesgue null; so m satisfies precisely the identities above. By the uniqueness assertion in Theorem 1.7.9, $\mu_F = m$ on $\mathcal{B}[\mathbb{R}]$.
(ii) Let F be monotone non-decreasing and absolutely continuous. By Exercise 1.6.48(ii), F is of bounded variation on every compact interval and hence, by Exercise 1.6.43, differentiable almost everywhere with locally absolutely integrable derivative.
Before proceeding we must be careful about which σ-algebra the derivative is measurable for: Exercise 1.4.49 will be applied on the Borel measure space $(\mathbb{R}, \mathcal{B}[\mathbb{R}], m)$, so we need a genuinely Borel measurable representative, not merely a Lebesgue measurable one. This is available here because F is continuous (Exercise 1.6.48(i)). Indeed, for $h \neq 0$ put $Q_h(x) := (F(x+h) - F(x))/h$; each Q_h is continuous in x, and for fixed x the map $h \mapsto Q_h(x)$ is continuous on $h \neq 0$. Hence for each $k \geq 1$ the supremum of $Q_h(x)$ over $0 < |h| < 1/k$ may be taken over rational h alone, so it is a countable supremum of continuous functions and therefore Borel measurable; letting $k \to \infty$ shows that $\overline{D}(x) := \limsup_{h\to 0} Q_h(x)$ is Borel measurable, and symmetrically so is $\underline{D}(x) := \liminf_{h\to 0} Q_h(x)$. Consequently the set $D := \{x : \overline{D}(x) = \underline{D}(x) \in \mathbb{R}\}$ on which F is differentiable is Borel, and we may define $F' : \mathbb{R} \to [0, +\infty)$ to be $\overline{D}$ on D and 0 off D; this is a Borel measurable, everywhere finite function, non-negative because all difference quotients of a monotone non-decreasing function are non-negative, and it agrees with the derivative of F outside the Lebesgue-null set $\mathbb{R} \setminus D$.
By Exercise 1.4.49 applied to the measure space $(\mathbb{R}, \mathcal{B}[\mathbb{R}], m)$ and $g = F'$, the formula

$$m_{F'}(E) := \int_{\mathbb{R}} 1_E F'\,dm = \int_E F'(x)\,dx$$

defines a Borel measure $m_{F'}$ on $\mathbb{R}$.
We check that $m_{F'}$ satisfies (1.33) and (1.34) for F. Absolute continuity implies continuity, so $F_- = F = F_+$. Since a singleton is Lebesgue-null, $m_{F'}(\{a\}) = 0 = F_+(a) - F_-(a)$, giving (1.34). For a compact interval $[a,b]$, the second fundamental theorem of calculus for absolutely continuous functions (Theorem 1.6.40), applied to the restriction of F to $[a,b]$ and combined with the fact that F' as defined

above agrees with the derivative of F outside a Lebesgue-null set (so that the two integrals agree), gives

$$m_{F'}([a,b]) = \int_{[a,b]} F'(x)\,dx = F(b) - F(a) = F_+(b) - F_-(a).$$

Because the endpoints carry no $m_{F'}$-measure, the same value is obtained for $[a,b)$, $(a,b]$ and (a,b), which matches (1.33) since $F_\pm = F$. By the uniqueness assertion of Theorem 1.7.9, $\mu_F = m_{F'}$; that is,

$$\mu_F(E) = \int_E F'(x)\,dx$$

for every Borel set E.

It remains to prove the integral identity for unsigned Borel measurable $f : \mathbb{R} \to [0,+\infty]$. We argue by the usual three-step ladder.

Indicators. If $f = 1_E$ with E Borel, then by Exercise 1.4.34(ii) and the measure identity just proved,

$$\int_{\mathbb{R}} 1_E\,d\mu_F = \mu_F(E) = \int_{\mathbb{R}} 1_E F'\,dm.$$

Simple functions. If $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ is an unsigned simple function with E_j Borel and $c_j \in [0,+\infty]$, then by finite additivity and homogeneity of the simple integral (Exercise 1.4.34(iii),(iv)) applied on the left, and by the corresponding linearity of the unsigned Lebesgue integral applied on the right,

$$\int_{\mathbb{R}} f\,d\mu_F = \sum_{j=1}^k c_j \mu_F(E_j) = \sum_{j=1}^k c_j \int_{\mathbb{R}} 1_{E_j} F'\,dm = \int_{\mathbb{R}} f F'\,dm.$$

General unsigned f. Set

$$f_n := \min\big(2^{-n}\lfloor 2^n f\rfloor,\ n\big),$$

with the convention $f_n := n$ where $f = +\infty$. Each f_n is Borel measurable and takes only the finitely many values $\{j2^{-n} : 0 \le j \le n2^n\}$, hence is an unsigned simple function; and $0 \le f_1 \le f_2 \le \dots$ with $f_n(x) \to f(x)$ for every x (if $f(x) < \infty$ then $0 \le f(x) - f_n(x) \le 2^{-n}$ once $n \ge f(x)$, and if $f(x) = \infty$ then $f_n(x) = n \to \infty$). By the monotone convergence theorem (Theorem 1.4.44) applied in the measure space $(\mathbb{R}, \mathcal{B}[\mathbb{R}], \mu_F)$,

$$\int_{\mathbb{R}} f\,d\mu_F = \lim_{n\to\infty} \int_{\mathbb{R}} f_n\,d\mu_F.$$

On the other hand $f_n F' \uparrow f F'$ pointwise (multiplication by the fixed non-negative $F'(x)$ preserves monotone convergence, with the convention $0 \cdot \infty = 0$), so by the monotone convergence theorem in $(\mathbb{R}, \mathcal{B}[\mathbb{R}], m)$,

$$\int_{\mathbb{R}} f F'\,dm = \lim_{n\to\infty} \int_{\mathbb{R}} f_n F'\,dm.$$

By the simple-function case the two sequences of integrals coincide term by term, so their limits agree:

$$\int_{\mathbb{R}} f(x)\,d\mu_F(x) = \int_{\mathbb{R}} f(x) F'(x)\,dx,$$

as required.

Finally, the closing remark of the exercise is immediate from (1.33): for monotone non-decreasing F and $-\infty < a < b < +\infty$,

$$\int_{[a,b]} dF = \int_{\mathbb{R}} 1_{[a,b]}\,d\mu_F = \mu_F([a,b]) = F_+(b) - F_-(a).$$

Problem 1.7.14. (Lebesgue-Stieltjes measure, pure point case)

- (i) If $H : \mathbb{R} \to \mathbb{R}$ is the Heaviside function $H := 1_{[0,+\infty)}$, show that μ_H is equal to the Dirac

measure δ_0 at the origin (defined in Example 1.4.22).

- (ii) If $F = \sum_n c_n J_n$ is a jump function (as defined in Definition 1.6.30), show that μ_F is equal to the linear combination $\sum_n c_n \delta_{x_n}$ of delta functions (as defined in Exercise 1.4.22), where x_n is the point of discontinuity for the basic jump function J_n.

Solution. Both parts are verifications of the identities (1.33) and (1.34), after which the uniqueness assertion in Theorem 1.7.9 does the rest. Recall from Example 1.4.22 that $\delta_x(E) = 1_E(x)$, and from Example 1.4.29 that Dirac measures are countably additive.
(i) The Heaviside function $H = 1_{[0,+\infty)}$ is monotone non-decreasing. Its one-sided limits are

$$H_-(x) = \sup_{y<x} H(y) = \begin{cases} 0 & x \le 0 \\ 1 & x > 0 \end{cases} \qquad H_+(x) = \inf_{y>x} H(y) = \begin{cases} 0 & x < 0 \\ 1 & x \ge 0 \end{cases}$$

so that $H_- = 1_{(0,+\infty)}$ and $H_+ = 1_{[0,+\infty)} = H$.
Now compare with δ_0, which is a Borel measure. For $a \in \mathbb{R}$,

$$H_+(a) - H_-(a) = \begin{cases} 1 - 0 = 1 & a = 0 \\ 0 - 0 = 0 & a < 0 \\ 1 - 1 = 0 & a > 0 \end{cases}$$

which is exactly $1_{\{a\}}(0) = \delta_0(\{a\})$, giving (1.34). For $-\infty < a < b < \infty$: $H_+(b) - H_-(a)$ equals 1 precisely when $H_+(b) = 1$ and $H_-(a) = 0$, i.e. when $b \ge 0$ and $a \le 0$, and equals 0 otherwise (if $b < 0$ both terms vanish, and if $a > 0$ both terms equal 1; note $H_-(a) \le H_+(b)$ always, so the difference is never negative). But $\delta_0([a,b]) = 1$ precisely when $0 \in [a,b]$, i.e. when $a \le 0 \le b$. So $\delta_0([a,b]) = H_+(b) - H_-(a)$. The same bookkeeping handles the remaining three cases:

$$H_-(b) - H_-(a) = 1 \iff (b > 0 \text{ and } a \le 0) \iff 0 \in [a,b),$$

$$H_+(b) - H_+(a) = 1 \iff (b \ge 0 \text{ and } a < 0) \iff 0 \in (a,b],$$

$$H_-(b) - H_+(a) = 1 \iff (b > 0 \text{ and } a < 0) \iff 0 \in (a,b),$$

and in each case the difference is 0 when the membership fails. Hence δ_0 is a Borel measure satisfying (1.33) and (1.34) for H, and by the uniqueness in Theorem 1.7.9, $\mu_H = \delta_0$.
(ii) Let $F = \sum_n c_n J_n$ be a jump function as in Definition 1.6.30: n ranges over an at most countable index set, $c_n > 0$ with $\sum_n c_n < \infty$, and each J_n is a basic jump function with point of discontinuity x_n and fraction $\theta_n \in [0,1]$, i.e. $J_n(x) = 0$ for $x < x_n$, $J_n(x_n) = \theta_n$, $J_n(x) = 1$ for $x > x_n$. Since $0 \le J_n \le 1$ and $\sum_n c_n < \infty$, the series defining F converges absolutely and uniformly, and F is monotone non-decreasing and real-valued. Exactly as in part (i),

$$(J_n)_- = 1_{(x_n,+\infty)}, \qquad (J_n)_+ = 1_{[x_n,+\infty)}.$$

Step 1: the one-sided limits of F are obtained termwise, namely

$$F_-(x) = \sum_n c_n (J_n)_-(x) = \sum_{n: x_n < x} c_n, \qquad F_+(x) = \sum_n c_n (J_n)_+(x) = \sum_{n: x_n \le x} c_n.$$

For the first identity: if $y < x$ then $J_n(y) \le \sup_{z<x} J_n(z) = (J_n)_-(x)$ for every n, so $F(y) \le \sum_n c_n (J_n)_-(x)$, whence $F_-(x) \le \sum_n c_n (J_n)_-(x)$. Conversely, let $\varepsilon > 0$ and choose a finite subset N of the index set with $\sum_{n \notin N} c_n \le \varepsilon$. Since N is finite, we may choose $y < x$ with $y > x_n$ for every $n \in N$ having $x_n < x$; for such y we have $J_n(y) = (J_n)_-(x)$ for every $n \in N$ (both sides are 1 if $x_n < x$, and both are 0 if $x_n \ge x$, since then $x_n > y$). Hence

$$F_-(x) \ge F(y) \ge \sum_{n \in N} c_n J_n(y) = \sum_{n \in N} c_n (J_n)_-(x) \ge \sum_n c_n (J_n)_-(x) - \varepsilon,$$

and letting $\varepsilon \to 0$ gives the first identity. The second is proved in the same way with $y > x$ chosen so that $y < x_n$ for all $n \in N$ with $x_n > x$, giving $J_n(y) = (J_n)_+(x)$ for $n \in N$.
Step 2: comparison with $\nu := \sum_n c_n \delta_{x_n}$. By Exercise 1.4.22(i),(ii), ν is a countably additive Borel measure, and $\nu(\mathbb{R}) = \sum_n c_n < \infty$, so ν is a finite measure. Directly from the definition of ν,

$$\nu((-\infty, x)) = \sum_{n: x_n < x} c_n = F_-(x), \qquad \nu((-\infty, x]) = \sum_{n: x_n \le x} c_n = F_+(x),$$

by Step 1. Since ν is finite, all subtractions below are legitimate. For $a \in \mathbb{R}$,

$$\nu(\{a\}) = \nu((-\infty, a]) - \nu((-\infty, a)) = F_+(a) - F_-(a),$$

which is (1.34). For $-\infty < a < b < \infty$,

$$\nu([a, b]) = \nu((-\infty, b]) - \nu((-\infty, a)) = F_+(b) - F_-(a),$$

$$\nu([a, b)) = \nu((-\infty, b)) - \nu((-\infty, a)) = F_-(b) - F_-(a),$$

$$\nu((a, b]) = \nu((-\infty, b]) - \nu((-\infty, a]) = F_+(b) - F_+(a),$$

$$\nu((a, b)) = \nu((-\infty, b)) - \nu((-\infty, a]) = F_-(b) - F_+(a),$$

which is (1.33). Thus ν is a Borel measure obeying (1.33) and (1.34) for F, and by the uniqueness assertion of Theorem 1.7.9,

$$\mu_F = \sum_n c_n \delta_{x_n}.$$

(Part (i) is the special case of a single basic jump function with $c = 1$, $x_0 = 0$ and fraction $\theta = 1$.)

Problem 1.7.15. (Lebesgue-Stieltjes measure, singular continuous case).
(i) If $F : \mathbb{R} \to \mathbb{R}$ is a monotone non-decreasing function, show that F is continuous if and only if $\mu_F(\{x\}) = 0$ for all $x \in \mathbb{R}$.
(ii) If F is the Cantor function (defined in Exercise 1.6.47), show that μ_F is a probability measure supported on the middle-thirds Cantor set (see Exercise 1.2.9) in the sense that $\mu_F(\mathbb{R} \setminus C) = 0$. The measure μ_F is known as Cantor measure.
(iii) If μ_F is Cantor measure, establish the self-similarity properties $\mu(\frac{1}{3} \cdot E) = \frac{1}{2}\mu(E)$ and $\mu(\frac{1}{3} \cdot E + \frac{2}{3}) = \frac{1}{2}\mu(E)$ for every Borel-measurable $E \subset [0, 1]$, where $\frac{1}{3} \cdot E := \{\frac{1}{3}x : x \in E\}$.

Solution. (i) Recall from Theorem 1.7.9 the definitions

$$F_-(x) = \sup_{y<x} F(y), \qquad F_+(x) = \inf_{y>x} F(y),$$

and the identity (1.34), which states that

$$\mu_F(\{x\}) = F_+(x) - F_-(x)$$

for every $x \in \mathbb{R}$. Since F is monotone non-decreasing, $F_-(x)$ is exactly the left limit $\lim_{y \to x^-} F(y)$ and $F_+(x)$ is exactly the right limit $\lim_{y \to x^+} F(y)$ (a monotone function has one-sided limits everywhere, and for a non-decreasing function the left limit is the supremum of the earlier values and the right limit is the infimum of the later values). Moreover monotonicity gives $F_-(x) \le F(x) \le F_+(x)$.
Now F is continuous at x precisely when both one-sided limits exist and equal $F(x)$, i.e. precisely when $F_-(x) = F(x) = F_+(x)$. Given the sandwich $F_-(x) \le F(x) \le F_+(x)$, this happens if and only if $F_+(x) - F_-(x) = 0$, i.e. if and only if $\mu_F(\{x\}) = 0$. Taking the conjunction over all $x \in \mathbb{R}$: F is continuous (everywhere) if and only if $\mu_F(\{x\}) = 0$ for every $x \in \mathbb{R}$.
(ii) Let $F : [0, 1] \to [0, 1]$ be the Cantor function of Exercise 1.6.47, extended to all of $\mathbb{R}$ by setting $F(x) := 0$ for $x < 0$ and $F(x) := 1$ for $x > 1$. By Exercise 1.6.47(iv), F is continuous and monotone non-decreasing on $[0, 1]$ with $F(0) = 0$ and $F(1) = 1$, so the extension is monotone non-decreasing and continuous on all of $\mathbb{R}$.

Since F is continuous, $F_- = F = F_+$ everywhere by part (i); and with the conventions of Theorem 1.7.9 we have $F_+(-\infty) = \inf_y F(y) = 0$ and $F_-(+\infty) = \sup_y F(y) = 1$. Hence by (1.33) applied to the interval $(-\infty, +\infty)$,

$$\mu_F(\mathbb{R}) = F_-(+\infty) - F_+(-\infty) = 1 - 0 = 1,$$

so μ_F is a probability measure.
Now we show $\mu_F(\mathbb{R} \setminus C) = 0$, where $C \subset [0,1]$ is the middle-thirds Cantor set of Exercise 1.2.9. Write

$$\mathbb{R} \setminus C = (-\infty, 0) \cup \big([0,1] \setminus C\big) \cup (1, +\infty).$$

For the two unbounded pieces, (1.33) and the continuity of F give

$$\mu_F((-\infty, 0)) = F_-(0) - F_+(-\infty) = 0 - 0 = 0, \qquad \mu_F((1, +\infty)) = F_-(+\infty) - F_+(1) = 1 - 1 = 0.$$

For the middle piece: by construction (Exercise 1.2.9), $[0,1] \setminus C$ is the union of the countably many open middle-third intervals (a_j, b_j), $j = 1, 2, \ldots$, that are deleted at the successive stages of the construction. By Exercise 1.6.47(v), every point $x \in [0,1] \setminus C$ has a neighbourhood on which F is constant; thus F is locally constant on the open set $[0,1] \setminus C$, and since each (a_j, b_j) is connected, F is constant on each (a_j, b_j). Consequently, using (1.33) and continuity of F again,

$$\mu_F((a_j, b_j)) = F_-(b_j) - F_+(a_j) = F(b_j) - F(a_j) = 0,$$

where the last equality holds because F is constant on (a_j, b_j) and continuous at the endpoints, so $F(a_j) = F(b_j)$. By countable additivity (the intervals are disjoint),

$$\mu_F([0,1] \setminus C) = \sum_{j=1}^{\infty} \mu_F((a_j, b_j)) = 0.$$

Adding the three pieces gives $\mu_F(\mathbb{R} \setminus C) = 0$, as claimed. In particular $\mu_F(C) = 1$, so Cantor measure is supported on C; and by part (i) it has no atoms, since F is continuous.
(iii) Write $\mu := \mu_F$ for Cantor measure. We first record the two functional equations satisfied by the Cantor function. From the recursion in Exercise 1.6.47,

$$F_n(x) = \tfrac{1}{2} F_{n-1}(3x) \ \ (x \in [0, \tfrac{1}{3}]), \qquad F_n(x) = \tfrac{1}{2} \ \ (x \in (\tfrac{1}{3}, \tfrac{2}{3})), \qquad F_n(x) = \tfrac{1}{2} + \tfrac{1}{2} F_{n-1}(3x-2) \ \ (x \in [\tfrac{2}{3}, 1]).$$

Since $F_n \to F$ uniformly on $[0,1]$ (Exercise 1.6.47(iii)), we may pass to the limit in the first and third relations to obtain, for all $x \in [0,1]$,

$$F\big(\tfrac{x}{3}\big) = \tfrac{1}{2} F(x), \qquad F\big(\tfrac{x+2}{3}\big) = \tfrac{1}{2} + \tfrac{1}{2} F(x).$$

Call this pair of identities $(\star)$.
Note that if $E \subset [0,1]$ is Borel then so are $\frac{1}{3} \cdot E = \varphi^{-1}(E)$ and $\frac{1}{3} \cdot E + \frac{2}{3} = \psi^{-1}(E)$, where $\varphi(x) := 3x$ and $\psi(x) := 3x - 2$ are continuous, hence Borel measurable. Define, for Borel $E \subset [0,1]$,

$$\nu(E) := 2\,\mu\big(\tfrac{1}{3} \cdot E\big), \qquad \tau(E) := 2\,\mu\big(\tfrac{1}{3} \cdot E + \tfrac{2}{3}\big).$$

Both ν and τ are countably additive on the Borel σ-algebra of $[0,1]$: they assign 0 to $\emptyset$, and if $E_1, E_2, \ldots$ are disjoint Borel subsets of $[0,1]$ then the sets $\frac{1}{3} \cdot E_n$ (respectively $\frac{1}{3} \cdot E_n + \frac{2}{3}$) are disjoint with union $\frac{1}{3} \cdot \bigcup_n E_n$ (respectively $\frac{1}{3} \cdot \bigcup_n E_n + \frac{2}{3}$), since $x \mapsto x/3$ and $x \mapsto x/3 + 2/3$ are bijections. So ν, τ are finite Borel measures on $[0,1]$; and $\nu([0,1]) = 2\mu([0, \frac{1}{3}]) = 2(F(\frac{1}{3}) - F(0)) = 2 \cdot \frac{1}{2} = 1$ by $(\star)$ with $x = 1$, and likewise $\tau([0,1]) = 2(F(1) - F(\frac{2}{3})) = 2(1 - \frac{1}{2}) = 1$.
We claim $\nu = \tau = \mu$ on the Borel subsets of $[0,1]$. First we check they agree on intervals. Let $0 \le c \le d \le 1$. Since F is continuous, part (i) shows μ has no atoms, and the same is then true of ν and τ (e.g. $\nu(\{t\}) = 2\mu(\{t/3\}) = 0$); so all four interval types $[c,d], [c,d), (c,d], (c,d)$ receive the same measure under each of μ, ν, τ, and it suffices to treat closed intervals. By (1.33) and continuity,

$$\mu([c,d]) = F(d) - F(c).$$

On the other hand $\frac{1}{3}\cdot[c,d]=[\frac{c}{3},\frac{d}{3}]$, so by $(\star)$,

$$\nu([c,d])=2\big(F(\tfrac{d}{3})-F(\tfrac{c}{3})\big)=2\big(\tfrac{1}{2}F(d)-\tfrac{1}{2}F(c)\big)=F(d)-F(c)=\mu([c,d]),$$

and $\frac{1}{3}\cdot[c,d]+\frac{2}{3}=[\frac{c+2}{3},\frac{d+2}{3}]$, so

$$\tau([c,d])=2\big(F(\tfrac{d+2}{3})-F(\tfrac{c+2}{3})\big)=2\big((\tfrac{1}{2}+\tfrac{1}{2}F(d))-(\tfrac{1}{2}+\tfrac{1}{2}F(c))\big)=F(d)-F(c)=\mu([c,d]).$$

Now let $\mathcal{A}$ be the collection of finite unions of subintervals of $[0,1]$ (allowing degenerate and empty intervals). This is a Boolean algebra on $[0,1]$, every element of which can be written as a finite disjoint union of intervals, and $\langle\mathcal{A}\rangle$ is the Borel σ-algebra of $[0,1]$. By finite additivity and the interval computation above, μ,ν,τ all agree on $\mathcal{A}$; call the common restriction $\mu_0:\mathcal{A}\to[0,+\infty]$. Being the restriction of a countably additive measure to a Boolean subalgebra, μ_0 is a pre-measure, and it is finite (hence σ-finite) since $\mu_0([0,1])=1$. Let λ be the Hahn-Kolmogorov extension of μ_0, defined on a σ-algebra $\mathcal{B}\supset\langle\mathcal{A}\rangle$. Each of μ,ν,τ is a countably additive extension of μ_0 defined on $\langle\mathcal{A}\rangle$, so by Exercise 1.7.7 (applicable because μ_0 is σ-finite) each agrees with λ on $\mathcal{B}\cap\langle\mathcal{A}\rangle=\langle\mathcal{A}\rangle$. Hence $\mu=\nu=\tau$ on all Borel subsets of $[0,1]$, which is exactly the pair of self-similarity identities

$$\mu\big(\tfrac{1}{3}\cdot E\big)=\tfrac{1}{2}\mu(E),\qquad \mu\big(\tfrac{1}{3}\cdot E+\tfrac{2}{3}\big)=\tfrac{1}{2}\mu(E)$$

for every Borel $E\subset[0,1]$.

Problem 1.7.16. (Connection with Riemann-Stieltjes integral). Let $F:\mathbb{R}\to\mathbb{R}$ be monotone non-decreasing, let $[a,b]$ be a compact interval, and let $f:[a,b]\to\mathbb{R}$ be continuous. Suppose that F is continuous at the endpoints a,b of the interval. Show that for every $\varepsilon>0$ there exists $\delta>0$ such that

$$\Big|\sum_{i=1}^{n}f(t_i^*)\big(F(t_i)-F(t_{i-1})\big)-\int_{[a,b]}f\,dF\Big|\le\varepsilon$$

whenever $a=t_0<t_1<\cdots<t_n=b$ and $t_i^*\in[t_{i-1},t_i]$ for $1\le i\le n$ are such that $\sup_{1\le i\le n}|t_i-t_{i-1}|\le\delta$. In the language of the Riemann-Stieltjes integral, this result asserts that the Lebesgue-Stieltjes integral extends the Riemann-Stieltjes integral.

Solution. Throughout, $\int_{[a,b]}f\,dF$ denotes $\int_{[a,b]}f\,d\mu_F$, and we abbreviate $G:=F_+$, so that G is monotone non-decreasing and $F(x)\le G(x)$ for all x. Since F is continuous at a and at b, part (i) of Exercise 1.7.15 gives $F_-(a)=F(a)=F_+(a)$ and $F_-(b)=F(b)=F_+(b)$; in particular $G(a)=F(a)$ and $G(b)=F(b)$.
Preliminaries. By (1.33),

$$L:=\mu_F([a,b])=F_+(b)-F_-(a)=F(b)-F(a)<\infty.$$

Since f is continuous on the compact interval $[a,b]$, it is bounded, say $|f|\le M$, and uniformly continuous there. As f is continuous it is Borel measurable, and $\int_{[a,b]}|f|\,d\mu_F\le ML<\infty$, so f is absolutely integrable on $[a,b]$ with respect to μ_F and the integral $\int_{[a,b]}f\,dF$ is well defined.
Choice of δ. Let $\varepsilon>0$ be given and set $\varepsilon':=\varepsilon/(2L+1)$. By uniform continuity choose $\eta>0$ so that $|f(s)-f(t)|\le\varepsilon'$ whenever $s,t\in[a,b]$ with $|s-t|\le\eta$, and put $\delta:=\eta/2$.
Now fix a partition $a=t_0<t_1<\cdots<t_n=b$ of mesh $\le\delta$ and tags $t_i^*\in[t_{i-1},t_i]$. Write $S:=\sum_{i=1}^{n}f(t_i^*)(F(t_i)-F(t_{i-1}))$ for the Riemann-Stieltjes sum.
Step 1: replace F by G in the sum. Consider

$$S':=\sum_{i=1}^{n}f(t_i^*)\big(G(t_i)-G(t_{i-1})\big).$$

Let $h:=F-G$, so $h\le0$ and, as noted above, $h(t_0)=h(a)=0$ and $h(t_n)=h(b)=0$. Then, writing

$f_i := f(t_i^*)$ and $h_i := h(t_i)$, summation by parts gives

$$S - S' = \sum_{i=1}^{n} f_i(h_i - h_{i-1}) = f_n h_n - f_1 h_0 + \sum_{i=1}^{n-1} h_i\big(f_i - f_{i+1}\big) = \sum_{i=1}^{n-1} h_i\big(f(t_i^*) - f(t_{i+1}^*)\big).$$

For $1 \le i \le n-1$ we have $t_i^*, t_{i+1}^* \in [t_{i-1}, t_{i+1}]$, an interval of length at most $2\delta = \eta$, so $|f(t_i^*) - f(t_{i+1}^*)| \le \varepsilon'$. Also, by (1.34),

$$|h_i| = G(t_i) - F(t_i) = F_+(t_i) - F(t_i) \le F_+(t_i) - F_-(t_i) = \mu_F(\{t_i\}),$$

and since the points $t_1, \dots, t_{n-1}$ are distinct elements of $[a, b]$, additivity and monotonicity of μ_F give $\sum_{i=1}^{n-1} \mu_F(\{t_i\}) = \mu_F(\{t_1, \dots, t_{n-1}\}) \le \mu_F([a, b]) = L$. Therefore

$$|S - S'| \le \varepsilon' \sum_{i=1}^{n-1} |h_i| \le \varepsilon' L.$$

Step 2: S' is the integral of a step function. Set $I_1 := [t_0, t_1]$ and $I_i := (t_{i-1}, t_i]$ for $2 \le i \le n$. These are disjoint Borel sets whose union is $[a, b]$. By (1.33),

$$\mu_F(I_1) = F_+(t_1) - F_-(t_0) = G(t_1) - G(t_0), \qquad \mu_F(I_i) = F_+(t_i) - F_+(t_{i-1}) = G(t_i) - G(t_{i-1}) \ \ (i \ge 2),$$

where for I_1 we used $F_-(a) = F(a) = G(a) = G(t_0)$. Hence, letting $g := \sum_{i=1}^{n} f(t_i^*) 1_{I_i}$ (a simple function supported in $[a, b]$), linearity of the integral gives

$$\int_{[a,b]} g \, d\mu_F = \sum_{i=1}^{n} f(t_i^*)\, \mu_F(I_i) = S'.$$

Step 3: g approximates f. If $x \in I_i$ then x and t_i^* both lie in $[t_{i-1}, t_i]$, an interval of length at most $\delta \le \eta$, so $|f(x) - g(x)| = |f(x) - f(t_i^*)| \le \varepsilon'$. Thus $|f - g| \le \varepsilon'$ pointwise on $[a, b]$, and by the triangle inequality for integrals and monotonicity,

$$\Big| \int_{[a,b]} f \, dF - S' \Big| = \Big| \int_{[a,b]} (f - g) \, d\mu_F \Big| \le \int_{[a,b]} |f - g| \, d\mu_F \le \varepsilon' \mu_F([a, b]) = \varepsilon' L.$$

Conclusion. Combining Steps 1 and 3,

$$\Big| S - \int_{[a,b]} f \, dF \Big| \le |S - S'| + \Big| S' - \int_{[a,b]} f \, dF \Big| \le 2\varepsilon' L = \frac{2L}{2L+1} \varepsilon \le \varepsilon,$$

which is the required bound. Since δ was chosen depending only on ε, f and F (and not on the partition or the tags), the claim is proved.

Note that the hypothesis that F be continuous at the endpoints is genuinely needed: it is what makes h vanish at t_0 and t_n in Step 1, and it is what pins the total mass $\mu_F([a, b])$ to $F(b) - F(a)$. At interior points F may jump freely, because a jump at an interior partition point t_i is multiplied by the small quantity $f(t_i^*) - f(t_{i+1}^*)$.

Problem 1.7.17. (Integration by parts formula). Let $F, G : \mathbb{R} \to \mathbb{R}$ be monotone non-decreasing and continuous. Show that

$$\int_{[a,b]} F \, dG = - \int_{[a,b]} G \, dF + F(b)G(b) - F(a)G(a)$$

for any compact interval $[a, b]$. (Hint: use Exercise 1.7.16.) This formula can be partially extended to the case when one or both of F, G have discontinuities, but care must be taken when F and G are simultaneously discontinuous at the same location.

Solution. If $a = b$ the claim is trivial: F and G are continuous, so $\mu_F(\{a\}) = \mu_G(\{a\}) = 0$ by Exercise 1.7.15(i), and both sides of the identity vanish. So assume $a < b$.
Both integrals are well defined and finite: F and G are continuous, hence Borel measurable and bounded on the compact interval $[a, b]$, while $\mu_G([a, b]) = G(b) - G(a)$ and $\mu_F([a, b]) = F(b) - F(a)$ are finite by (1.33) and continuity. Moreover the hypotheses of Exercise 1.7.16 are met in both of the following two ways:

- with integrand $f := F|_{[a,b]}$ (continuous) and integrator G (monotone non-decreasing and continuous at a and b);
- with integrand $f := G|_{[a,b]}$ (continuous) and integrator F (monotone non-decreasing and continuous at a and b).

For each $n \geq 1$ take the uniform partition $t_i := a + \frac{i}{n}(b - a)$, $0 \leq i \leq n$, whose mesh is $(b - a)/n$. Consider the two Riemann-Stieltjes sums

$$S_n := \sum_{i=1}^{n} F(t_i)\big(G(t_i) - G(t_{i-1})\big), \qquad T_n := \sum_{i=1}^{n} G(t_{i-1})\big(F(t_i) - F(t_{i-1})\big).$$

These are legitimate tagged Riemann-Stieltjes sums for the two situations above: in S_n the tag for the i-th subinterval is $t_i^* := t_i \in [t_{i-1}, t_i]$, and in T_n it is $t_i^* := t_{i-1} \in [t_{i-1}, t_i]$.
The algebraic identity. Summing the two expressions, the cross terms cancel:

$$\begin{aligned} S_n + T_n &= \sum_{i=1}^{n} \Big(F(t_i)G(t_i) - F(t_i)G(t_{i-1}) + G(t_{i-1})F(t_i) - G(t_{i-1})F(t_{i-1}) \Big) \\ &= \sum_{i=1}^{n} \Big(F(t_i)G(t_i) - F(t_{i-1})G(t_{i-1}) \Big) \\ &= F(t_n)G(t_n) - F(t_0)G(t_0) = F(b)G(b) - F(a)G(a), \end{aligned}$$

the last line by telescoping. Note this identity holds exactly, for every n.
Passing to the limit. Let $\varepsilon > 0$. Applying Exercise 1.7.16 to the first situation gives a $\delta_1 > 0$ such that any tagged partition of $[a, b]$ of mesh at most δ_1 has its Riemann-Stieltjes sum within ε of $\int_{[a,b]} F\, dG$; applying it to the second situation gives a corresponding $\delta_2 > 0$ for $\int_{[a,b]} G\, dF$. Choose N so large that $(b - a)/N \leq \min(\delta_1, \delta_2)$. Then for all $n \geq N$,

$$\Big| S_n - \int_{[a,b]} F\, dG \Big| \leq \varepsilon \quad \text{and} \quad \Big| T_n - \int_{[a,b]} G\, dF \Big| \leq \varepsilon.$$

Combining with the algebraic identity and the triangle inequality, for such n,

$$\Big| \int_{[a,b]} F\, dG + \int_{[a,b]} G\, dF - \big(F(b)G(b) - F(a)G(a)\big) \Big| = \Big| \Big(\int_{[a,b]} F\, dG - S_n \Big) + \Big(\int_{[a,b]} G\, dF - T_n \Big) \Big| \leq 2\varepsilon.$$

The left-hand side does not depend on n or on ε, and $\varepsilon > 0$ was arbitrary, so it must be zero. That is,

$$\int_{[a,b]} F\, dG + \int_{[a,b]} G\, dF = F(b)G(b) - F(a)G(a),$$

which rearranges to the asserted formula

$$\int_{[a,b]} F\, dG = - \int_{[a,b]} G\, dF + F(b)G(b) - F(a)G(a).$$

Remark on the role of continuity. The asymmetry in the choice of tags (t_i on the left, t_{i-1} on the right) is what makes the sum telescope. When F and G are both discontinuous at a common point c, the two sums no longer both converge to the corresponding Lebesgue-Stieltjes integrals with the same conventions, and the identity acquires a correction term supported on the common discontinuities; this is the phenomenon alluded to in the statement.

Problem 1.7.18. Let $(X, \mathcal{B}_X)$ and $(Y, \mathcal{B}_Y)$ be measurable spaces.
(i) Show that $\mathcal{B}_X \times \mathcal{B}_Y$ is the σ-algebra generated by the sets $E \times F$ with $E \in \mathcal{B}_X$, $F \in \mathcal{B}_Y$. In other words, $\mathcal{B}_X \times \mathcal{B}_Y$ is the coarsest σ-algebra on $X \times Y$ with the property that the product of a $\mathcal{B}_X$-measurable set and a $\mathcal{B}_Y$-measurable set is always $\mathcal{B}_X \times \mathcal{B}_Y$ measurable.
(ii) Show that $\mathcal{B}_X \times \mathcal{B}_Y$ is the coarsest σ-algebra on $X \times Y$ that makes the projection maps π_X, π_Y both measurable morphisms (see Remark 1.4.33).
(iii) If $E \in \mathcal{B}_X \times \mathcal{B}_Y$, show that the sets $E_x := \{y \in Y : (x, y) \in E\}$ lie in $\mathcal{B}_Y$ for every $x \in X$, and similarly that the sets $E^y := \{x \in X : (x, y) \in E\}$ lie in $\mathcal{B}_X$ for every $y \in Y$.
(iv) If $f : X \times Y \to [0, +\infty]$ is measurable (with respect to $\mathcal{B}_X \times \mathcal{B}_Y$), show that the function $f_x : y \mapsto f(x, y)$ is $\mathcal{B}_Y$-measurable for every $x \in X$, and similarly that the function $f^y : x \mapsto f(x, y)$ is $\mathcal{B}_X$-measurable for every $y \in Y$.
(v) If $E \in \mathcal{B}_X \times \mathcal{B}_Y$, show that the slices $E_x := \{y \in Y : (x, y) \in E\}$ lie in a countably generated σ-algebra. In other words, show that there exists an at most countable collection $\mathcal{A} = \mathcal{A}_E$ of sets (which can depend on E) such that $\{E_x : x \in X\} \subset \langle \mathcal{A} \rangle$. Conclude in particular that the number of distinct slices E_x is at most $\mathfrak{c}$, the cardinality of the continuum. (The last part of this exercise is only suitable for students who are comfortable with cardinal arithmetic.)

Solution. Recall the definition in force: $\mathcal{B}_X \times \mathcal{B}_Y := \langle \pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y) \rangle$, where $\pi_X^*(\mathcal{B}_X) = \{E \times Y : E \in \mathcal{B}_X\}$ and $\pi_Y^*(\mathcal{B}_Y) = \{X \times F : F \in \mathcal{B}_Y\}$. Write $\mathcal{P} := \{E \times F : E \in \mathcal{B}_X,\ F \in \mathcal{B}_Y\}$ for the collection of measurable rectangles.
(i) We show $\langle \mathcal{P} \rangle = \mathcal{B}_X \times \mathcal{B}_Y$, by two inclusions.
Since $Y \in \mathcal{B}_Y$ and $X \in \mathcal{B}_X$, every set $E \times Y$ with $E \in \mathcal{B}_X$ and every set $X \times F$ with $F \in \mathcal{B}_Y$ belongs to $\mathcal{P}$. Hence $\pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y) \subset \mathcal{P} \subset \langle \mathcal{P} \rangle$, and since $\langle \mathcal{P} \rangle$ is a σ-algebra we conclude $\mathcal{B}_X \times \mathcal{B}_Y \subset \langle \mathcal{P} \rangle$. Conversely, for $E \in \mathcal{B}_X$ and $F \in \mathcal{B}_Y$ we have the identity

$$E \times F = (E \times Y) \cap (X \times F),$$

and both sets on the right lie in $\mathcal{B}_X \times \mathcal{B}_Y$; since a σ-algebra is closed under finite intersections, $E \times F \in \mathcal{B}_X \times \mathcal{B}_Y$. Thus $\mathcal{P} \subset \mathcal{B}_X \times \mathcal{B}_Y$ and hence $\langle \mathcal{P} \rangle \subset \mathcal{B}_X \times \mathcal{B}_Y$. This proves equality.
The "coarsest" reformulation is immediate: $\mathcal{B}_X \times \mathcal{B}_Y$ contains $\mathcal{P}$, and if $\mathcal{B}$ is any σ-algebra on $X \times Y$ containing $\mathcal{P}$, then $\mathcal{B} \supset \langle \mathcal{P} \rangle = \mathcal{B}_X \times \mathcal{B}_Y$.
(ii) Let $\mathcal{B}$ be a σ-algebra on $X \times Y$. By Remark 1.4.33, $\pi_X : (X \times Y, \mathcal{B}) \to (X, \mathcal{B}_X)$ is a measurable morphism precisely when $\pi_X^{-1}(E) \in \mathcal{B}$ for every $E \in \mathcal{B}_X$, i.e. precisely when $E \times Y \in \mathcal{B}$ for every $E \in \mathcal{B}_X$, i.e. precisely when $\pi_X^*(\mathcal{B}_X) \subset \mathcal{B}$. Likewise π_Y is a measurable morphism precisely when $\pi_Y^*(\mathcal{B}_Y) \subset \mathcal{B}$. So $\mathcal{B}$ makes both projections measurable morphisms if and only if

$$\pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y) \subset \mathcal{B},$$

which, since $\mathcal{B}$ is a σ-algebra, holds if and only if $\mathcal{B} \supset \langle \pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y) \rangle = \mathcal{B}_X \times \mathcal{B}_Y$. As $\mathcal{B}_X \times \mathcal{B}_Y$ itself has this property (take $\mathcal{B} = \mathcal{B}_X \times \mathcal{B}_Y$), it is the coarsest such σ-algebra.
(iii) Fix $x \in X$ and define the inclusion map $\iota_x : Y \to X \times Y$ by $\iota_x(y) := (x, y)$. Then for any $E \subset X \times Y$ we have $\iota_x^{-1}(E) = E_x$. We claim ι_x is a measurable morphism from $(Y, \mathcal{B}_Y)$ to $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$; this gives the assertion.
To see this, let

$$\mathcal{S} := \{E \subset X \times Y : E_x \in \mathcal{B}_Y\}.$$

Since preimages commute with all Boolean operations, $\mathcal{S}$ is a σ-algebra: $(X \times Y)_x = Y \in \mathcal{B}_Y$; $((X \times Y) \setminus E)_x = Y \setminus E_x$; and $\left(\bigcup_{n=1}^{\infty} E_n\right)_x = \bigcup_{n=1}^{\infty} (E_n)_x$. Moreover $\mathcal{S}$ contains every measurable rectangle: for $E \in \mathcal{B}_X$, $F \in \mathcal{B}_Y$,

$$(E \times F)_x = \begin{cases} F & \text{if } x \in E, \\ \emptyset & \text{if } x \notin E, \end{cases}$$

and both F and $\emptyset$ lie in $\mathcal{B}_Y$. Hence $\mathcal{P} \subset \mathcal{S}$, so by (i) $\mathcal{B}_X \times \mathcal{B}_Y = \langle \mathcal{P} \rangle \subset \mathcal{S}$. That is, $E_x \in \mathcal{B}_Y$ for every $E \in \mathcal{B}_X \times \mathcal{B}_Y$.
The statement for E^y is proved identically, using the map $\jmath_y : X \to X \times Y$, $\jmath_y(x) := (x, y)$, and the fact that $(E \times F)^y$ equals E if $y \in F$ and $\emptyset$ otherwise.

(iv) Fix $x \in X$. As observed in (iii), $\iota_x : (Y, \mathcal{B}_Y) \to (X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$ is a measurable morphism, and $f_x = f \circ \iota_x$. By Remark 1.4.33 an unsigned measurable function $f : X \times Y \to [0, +\infty]$ is the same thing as a measurable morphism into $[0, +\infty]$ equipped with its Borel σ-algebra, and compositions of measurable morphisms are measurable morphisms; hence f_x is a measurable morphism from $(Y, \mathcal{B}_Y)$ to $[0, +\infty]$, i.e. f_x is $\mathcal{B}_Y$-measurable.
Concretely, for every Borel set $U \subset [0, +\infty]$ we have

$$f_x^{-1}(U) = \iota_x^{-1}\big(f^{-1}(U)\big) = \big(f^{-1}(U)\big)_x,$$

which lies in $\mathcal{B}_Y$ by (iii), since $f^{-1}(U) \in \mathcal{B}_X \times \mathcal{B}_Y$ by measurability of f. The argument for $f^y = f \circ \jmath_y$ is symmetric.
(v) Let

$$\mathcal{S} := \{E \subset X \times Y : \text{there is an at most countable } \mathcal{A}_E \subset 2^Y \text{ with } \{E_x : x \in X\} \subset \langle \mathcal{A}_E \rangle\}.$$

We check $\mathcal{S}$ is a σ-algebra containing $\mathcal{P}$.

- $X \times Y \in \mathcal{S}$: all slices equal Y, so take $\mathcal{A} := \emptyset$, whose generated σ-algebra is $\{\emptyset, Y\}$.
- Rectangles: if $E \in \mathcal{B}_X$, $F \in \mathcal{B}_Y$, then every slice of $E \times F$ is F or $\emptyset$, so $\mathcal{A}_{E\times F} := \{F\}$ works, since $\langle \{F\} \rangle = \{\emptyset, F, Y \setminus F, Y\}$.
- Complements: $\big((X \times Y) \setminus E\big)_x = Y \setminus E_x$, and $\langle \mathcal{A}_E \rangle$ is closed under complements, so $\mathcal{A}_{E^c} := \mathcal{A}_E$ works.
- Countable unions: if $E = \bigcup_{n=1}^\infty E_n$ with each $E_n \in \mathcal{S}$, put $\mathcal{A}_E := \bigcup_{n=1}^\infty \mathcal{A}_{E_n}$, an at most countable collection. Then for each x, $E_x = \bigcup_n (E_n)_x$ with each $(E_n)_x \in \langle \mathcal{A}_{E_n} \rangle \subset \langle \mathcal{A}_E \rangle$, so $E_x \in \langle \mathcal{A}_E \rangle$.

Hence $\mathcal{P} \subset \mathcal{S}$ and $\mathcal{S}$ is a σ-algebra, so by (i) $\mathcal{B}_X \times \mathcal{B}_Y \subset \mathcal{S}$. This is the assertion.
It remains to bound the cardinality. Fix $E \in \mathcal{B}_X \times \mathcal{B}_Y$ and let $\mathcal{A}_E$ be as produced above, so that $\{E_x : x \in X\} \subset \langle \mathcal{A}_E \rangle$ and hence

$$\big|\{E_x : x \in X\}\big| \le \big|\langle \mathcal{A}_E \rangle\big|.$$

If $\mathcal{A}_E$ is finite then $\langle \mathcal{A}_E \rangle$ is finite (a Boolean algebra generated by k sets has at most 2^{2^k} elements, and for finitely many generators the generated Boolean algebra is already a σ-algebra), so certainly $|\langle \mathcal{A}_E \rangle| \le \mathfrak{c}$. If $\mathcal{A}_E$ is countably infinite, then it is an infinite family of cardinality $\kappa = \aleph_0$, and Exercise 1.4.16 gives

$$\big|\langle \mathcal{A}_E \rangle\big| \le \kappa^{\aleph_0} = \aleph_0^{\aleph_0} = 2^{\aleph_0} = \mathfrak{c}.$$

(Here $\aleph_0^{\aleph_0} = 2^{\aleph_0}$ because $2^{\aleph_0} \le \aleph_0^{\aleph_0} \le (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$.) In either case the number of distinct slices E_x is at most $\mathfrak{c}$, as claimed.
For the reader who prefers to see the mechanism, Exercise 1.4.16 is proved from the recursive description of Exercise 1.4.15: one builds $\langle \mathcal{A} \rangle$ as $\bigcup_{\alpha<\omega_1} \mathcal{A}_\alpha$, where $\mathcal{A}_{\alpha+1}$ consists of the at most countable unions of members of $\mathcal{A}_\alpha$ together with the complements of such unions, and $\mathcal{A}_\alpha := \bigcup_{\beta<\alpha} \mathcal{A}_\beta$ at limit ordinals. (This union is already a σ-algebra because ω_1 is regular: a countable set of countable ordinals has a countable supremum, so any countable family drawn from the union lies in a single $\mathcal{A}_\alpha$.) Each stage multiplies the cardinality by at most $\mathfrak{c}^{\aleph_0} = \mathfrak{c}$, and there are $\aleph_1 \le \mathfrak{c}$ stages, giving the bound $\mathfrak{c}$.

Problem 1.7.19. (i) Show that the product of two trivial σ-algebras (on two different spaces X, Y) is again trivial.
(ii) Show that the product of two atomic σ-algebras is again atomic.
(iii) Show that the product of two finite σ-algebras is again finite.
(iv) Show that the product of two Borel σ-algebras (on two Euclidean spaces $\mathbb{R}^d, \mathbb{R}^{d'}$ with $d, d' \ge 1$) is again the Borel σ-algebra (on $\mathbb{R}^d \times \mathbb{R}^{d'} \equiv \mathbb{R}^{d+d'}$).
(v) Show that the product of two Lebesgue σ-algebras (on two Euclidean spaces $\mathbb{R}^d, \mathbb{R}^{d'}$ with $d, d' \ge 1$) is not the Lebesgue σ-algebra. (Hint: argue by contradiction and use Exercise 1.7.18(iii).)
(vi) However, show that the Lebesgue σ-algebra on $\mathbb{R}^{d+d'}$ is the completion (see Exercise 1.4.26) of the product of the Lebesgue σ-algebras of $\mathbb{R}^d$ and $\mathbb{R}^{d'}$ with respect to $d + d'$-dimensional Lebesgue measure.

(vii) This part of the exercise is only for students who are comfortable with cardinal arithmetic. Give an example to show that the product of two discrete σ-algebras is not necessarily discrete.
(viii) On the other hand, show that the product of two discrete σ-algebras $2^X, 2^Y$ is again a discrete σ-algebra if at least one of the domains X, Y is at most countably infinite.

Solution. Throughout we use Exercise 1.7.18(i): $\mathcal{B}_X \times \mathcal{B}_Y$ is the σ-algebra generated by the measurable rectangles $E \times F$, $E \in \mathcal{B}_X$, $F \in \mathcal{B}_Y$.
(i) Suppose $\mathcal{B}_X = \{\emptyset, X\}$ and $\mathcal{B}_Y = \{\emptyset, Y\}$. The measurable rectangles are then $\emptyset \times \emptyset = \emptyset$, $\emptyset \times Y = \emptyset$, $X \times \emptyset = \emptyset$ and $X \times Y$. So the generating family is $\{\emptyset, X \times Y\}$, which is already a σ-algebra on $X \times Y$; hence $\mathcal{B}_X \times \mathcal{B}_Y = \{\emptyset, X \times Y\}$ is the trivial σ-algebra on $X \times Y$. (If X or Y is empty then $X \times Y = \emptyset$ and the conclusion is vacuous but still correct.)
(ii) As printed this claim is false in general: it is contradicted by part (vii) below. Let us first see why, and then prove the corrected statement.
Recall from Example 1.4.7 that a σ-algebra is atomic exactly when it is of the form $\mathcal{A}((C_\gamma)_{\gamma \in K})$ for some partition of the underlying space into atoms; deleting the empty atoms changes nothing, so we may take all atoms non-empty. A discrete σ-algebra 2^X is atomic, with the singletons as atoms. Now take X, Y as in part (vii), so that $2^X \times 2^Y \neq 2^{X \times Y}$, and suppose for contradiction that the product $2^X \times 2^Y$ were atomic, with non-empty atoms $(C_\gamma)_{\gamma \in K}$ partitioning $X \times Y$. Every singleton $\{(x, y)\} = \{x\} \times \{y\}$ is a measurable rectangle, hence lies in $2^X \times 2^Y$, hence is a union of some subfamily of the C_γ; as the C_γ are non-empty and disjoint, that subfamily has exactly one member and $\{(x, y)\} = C_\gamma$ for some γ. Thus every singleton is an atom, and since the atoms partition $X \times Y$ the atoms are precisely the singletons; therefore $2^X \times 2^Y = \mathcal{A}((\{p\})_{p \in X \times Y}) = 2^{X \times Y}$, contradicting the choice of X, Y. So the product of two atomic σ-algebras need not be atomic.
What is true is the following: if at least one of the two atom-partitions is at most countable, then $\mathcal{B}_X \times \mathcal{B}_Y$ is atomic, with atoms $(A_\alpha \times B_\beta)_{(\alpha,\beta) \in I \times J}$. Let $\mathcal{B}_X = \mathcal{A}((A_\alpha)_{\alpha \in I})$ and $\mathcal{B}_Y = \mathcal{A}((B_\beta)_{\beta \in J})$ with all atoms non-empty, and say I is at most countable. Write $\mathcal{C} := \mathcal{A}((A_\alpha \times B_\beta)_{(\alpha,\beta) \in I \times J})$; the sets $A_\alpha \times B_\beta$ do indeed partition $X \times Y$, and $\mathcal{C}$ is a σ-algebra by Exercise 1.4.10.
$\mathcal{B}_X \times \mathcal{B}_Y \subset \mathcal{C}$: if $E \in \mathcal{B}_X$ and $F \in \mathcal{B}_Y$ then $E = \bigcup_{\alpha \in I'} A_\alpha$ and $F = \bigcup_{\beta \in J'} B_\beta$ for some $I' \subset I$, $J' \subset J$, whence $E \times F = \bigcup_{(\alpha,\beta) \in I' \times J'} A_\alpha \times B_\beta \in \mathcal{C}$. So $\mathcal{C}$ contains all measurable rectangles, and being a σ-algebra it contains the σ-algebra they generate.
$\mathcal{C} \subset \mathcal{B}_X \times \mathcal{B}_Y$: let $S = \bigcup_{(\alpha,\beta) \in K} A_\alpha \times B_\beta$ for some $K \subset I \times J$. For $\alpha \in I$ set $K_\alpha := \{\beta \in J : (\alpha, \beta) \in K\}$ and $F_\alpha := \bigcup_{\beta \in K_\alpha} B_\beta$, which lies in $\mathcal{B}_Y$ because an atomic algebra is closed under arbitrary unions of atoms. Since the A_α partition X, a point (x, y) with $x \in A_{\alpha_0}$ lies in S if and only if $y \in F_{\alpha_0}$; hence

$$S = \bigcup_{\alpha \in I} (A_\alpha \times F_\alpha).$$

As I is at most countable, this is an at most countable union of measurable rectangles, so $S \in \mathcal{B}_X \times \mathcal{B}_Y$. This proves the corrected claim.
(iii) Let $\mathcal{B}_X, \mathcal{B}_Y$ be finite σ-algebras. By Exercise 1.4.4, a finite Boolean algebra is atomic; so $\mathcal{B}_X = \mathcal{A}((A_i)_{i=1}^n)$ and $\mathcal{B}_Y = \mathcal{A}((B_j)_{j=1}^m)$ for finite partitions into non-empty atoms. Both index sets are finite, hence at most countable, so the corrected form of (ii) proved above applies: $\mathcal{B}_X \times \mathcal{B}_Y$ is the atomic algebra with the nm atoms $A_i \times B_j$. An atomic algebra with nm atoms has at most 2^{nm} elements, so $\mathcal{B}_X \times \mathcal{B}_Y$ is finite. (In fact it has exactly 2^{nm} elements, since the atoms $A_i \times B_j$ are non-empty and hence distinct unions of them are distinct sets.)
(iv) We show $\mathcal{B}[\mathbb{R}^d] \times \mathcal{B}[\mathbb{R}^{d'}] = \mathcal{B}[\mathbb{R}^{d+d'}]$.
$\subset$: the coordinate projections $\pi_1 : \mathbb{R}^{d+d'} \to \mathbb{R}^d$ and $\pi_2 : \mathbb{R}^{d+d'} \to \mathbb{R}^{d'}$ are continuous, hence Borel measurable morphisms: $\{S \subset \mathbb{R}^d : \pi_1^{-1}(S) \in \mathcal{B}[\mathbb{R}^{d+d'}]\}$ is a σ-algebra containing all open sets, hence contains $\mathcal{B}[\mathbb{R}^d]$ by Exercise 1.4.14(i). Therefore, for Borel $E \subset \mathbb{R}^d$ and Borel $F \subset \mathbb{R}^{d'}$,

$$E \times F = \pi_1^{-1}(E) \cap \pi_2^{-1}(F) \in \mathcal{B}[\mathbb{R}^{d+d'}],$$

and since $\mathcal{B}[\mathbb{R}^{d+d'}]$ is a σ-algebra containing all such rectangles, Exercise 1.7.18(i) gives $\mathcal{B}[\mathbb{R}^d] \times \mathcal{B}[\mathbb{R}^{d'}] \subset \mathcal{B}[\mathbb{R}^{d+d'}]$.

$\supset$: let $U \subset \mathbb{R}^{d+d'}$ be open. For each point $p \in U$ there is an open box $B \times B' \subset U$ containing p, with $B \subset \mathbb{R}^d$ and $B' \subset \mathbb{R}^{d'}$ open boxes with rational vertices; the collection of such rational boxes is countable, so U is an at most countable union of sets $B \times B'$. Each such $B \times B'$ is a measurable rectangle (open boxes are Borel), hence $U \in \mathcal{B}[\mathbb{R}^d] \times \mathcal{B}[\mathbb{R}^{d'}]$. Since the product σ-algebra contains every open set, Exercise 1.4.14(i) gives $\mathcal{B}[\mathbb{R}^{d+d'}] \subset \mathcal{B}[\mathbb{R}^d] \times \mathcal{B}[\mathbb{R}^{d'}]$.

(v) First note that for every $n \geq 1$ there is a non-Lebesgue-measurable subset of $\mathbb{R}^n$. The proof of Proposition 1.2.18 goes through with $\mathbb{Q}$ replaced by the countable dense additive subgroup $\mathbb{Q}^n$ of $\mathbb{R}^n$: each coset of $\mathbb{Q}^n$ in $\mathbb{R}^n$ meets $[0,1]^n$, so by the axiom of choice we may pick one representative from each coset, forming a set $A \subset [0,1]^n$ containing exactly one point of each coset. The translates $A + q$, $q \in \mathbb{Q}^n \cap [-1,1]^n$, are then pairwise disjoint, their union contains $[0,1]^n$ and is contained in $[-1,2]^n$. Were A Lebesgue measurable, countable additivity and translation invariance would give $1 \leq \sum_{q \in \mathbb{Q}^n \cap [-1,1]^n} m^n(A) \leq 3^n$, whereas that sum is 0 if $m^n(A) = 0$ and $+\infty$ if $m^n(A) > 0$; contradiction. Apply this with $n := d$ to obtain a non-measurable $A \subset [0,1]^d$.

Next, the "horizontal hyperplane" $H := \mathbb{R}^d \times \{0\} \subset \mathbb{R}^{d+d'}$ is a null set. Indeed for each n and each $\varepsilon > 0$, the set $[-n,n]^d \times \{0\}$ is contained in the box $[-n,n]^d \times [-\varepsilon,\varepsilon]^{d'}$ of volume $(2n)^d (2\varepsilon)^{d'}$, so it has Lebesgue outer measure 0; and $H = \bigcup_n [-n,n]^d \times \{0\}$ has outer measure 0 by countable subadditivity. By Lemma 1.2.13(iii), H and, by monotonicity of outer measure, every subset of H is Lebesgue measurable (with measure zero).

In particular $E := A \times \{0\} \subset H$ lies in $\mathcal{L}[\mathbb{R}^{d+d'}]$. Suppose for contradiction that $\mathcal{L}[\mathbb{R}^d] \times \mathcal{L}[\mathbb{R}^{d'}] = \mathcal{L}[\mathbb{R}^{d+d'}]$. Then E would belong to the product σ-algebra, and Exercise 1.7.18(iii) applied with $y := 0 \in \mathbb{R}^{d'}$ would give

$$E^0 = \{x \in \mathbb{R}^d : (x,0) \in E\} = A \in \mathcal{L}[\mathbb{R}^d],$$

contradicting the non-measurability of A. Hence the product of the two Lebesgue σ-algebras is strictly coarser than $\mathcal{L}[\mathbb{R}^{d+d'}]$.

(vi) Write $\mathcal{P} := \mathcal{L}[\mathbb{R}^d] \times \mathcal{L}[\mathbb{R}^{d'}]$. By Exercise 1.2.22(ii), $E \times F$ is Lebesgue measurable in $\mathbb{R}^{d+d'}$ with $m^{d+d'}(E \times F) = m^d(E) m^{d'}(F)$ whenever E, F are Lebesgue measurable; since $\mathcal{L}[\mathbb{R}^{d+d'}]$ is a σ-algebra, Exercise 1.7.18(i) gives $\mathcal{P} \subset \mathcal{L}[\mathbb{R}^{d+d'}]$. Moreover the restriction of $m^{d+d'}$ to $\mathcal{P}$ is a measure obeying (1.36), and $(\mathbb{R}^d, m^d)$, $(\mathbb{R}^{d'}, m^{d'})$ are σ-finite, so by the uniqueness half of Proposition 1.7.11,

$$m^d \times m^{d'} = m^{d+d'}\big|_{\mathcal{P}},$$

call this identity (†); it is Example 1.7.13. Let $\overline{\mathcal{P}}$ denote the completion of $(\mathbb{R}^{d+d'}, \mathcal{P}, m^d \times m^{d'})$.

$\overline{\mathcal{P}} \subset \mathcal{L}[\mathbb{R}^{d+d'}]$: by Exercise 1.4.26, an element of $\overline{\mathcal{P}}$ has the form $S \,\Delta\, N$ where $S \in \mathcal{P}$ and $N \subset N'$ for some $N' \in \mathcal{P}$ with $(m^d \times m^{d'})(N') = 0$. By (†), $m^{d+d'}(N') = 0$, so N has Lebesgue outer measure zero and is Lebesgue measurable by Lemma 1.2.13(iii); also $S \in \mathcal{P} \subset \mathcal{L}[\mathbb{R}^{d+d'}]$. Hence $S \,\Delta\, N \in \mathcal{L}[\mathbb{R}^{d+d'}]$.

$\mathcal{L}[\mathbb{R}^{d+d'}] \subset \overline{\mathcal{P}}$: let E be Lebesgue measurable in $\mathbb{R}^{d+d'}$. By the definition of Lebesgue measurability, for each $n \geq 1$ there is an open $U_n \supset E$ with $(m^{d+d'})^*(U_n \setminus E) \leq 1/n$; put $G := \bigcap_n U_n$, a Borel set with $E \subset G$ and $m^{d+d'}(G \setminus E) = 0$. Applying the same to the measurable set $\mathbb{R}^{d+d'} \setminus E$ gives a Borel $G' \supset \mathbb{R}^{d+d'} \setminus E$ with $m^{d+d'}(G' \setminus (\mathbb{R}^{d+d'} \setminus E)) = 0$; set $F := \mathbb{R}^{d+d'} \setminus G'$, a Borel set with $F \subset E$ and $m^{d+d'}(E \setminus F) = 0$. Then $F \subset E \subset G$ with $m^{d+d'}(G \setminus F) = 0$. By part (iv), F and G are in $\mathcal{B}[\mathbb{R}^{d+d'}] = \mathcal{B}[\mathbb{R}^d] \times \mathcal{B}[\mathbb{R}^{d'}] \subset \mathcal{P}$, and by (†), $(m^d \times m^{d'})(G \setminus F) = 0$. Since $E = F \cup (E \setminus F)$ with $F \in \mathcal{P}$ and $E \setminus F \subset G \setminus F$ a $\mathcal{P}$-subnull set, Exercise 1.4.26 gives $E \in \overline{\mathcal{P}}$.

Thus $\mathcal{L}[\mathbb{R}^{d+d'}] = \overline{\mathcal{P}}$. Finally the measures agree: with $F \subset E \subset G$ as above, the completed measure of E equals $(m^d \times m^{d'})(F) = m^{d+d'}(F) = m^{d+d'}(E)$. So $(\mathbb{R}^{d+d'}, \mathcal{L}[\mathbb{R}^{d+d'}], m^{d+d'})$ is exactly the completion of $(\mathbb{R}^{d+d'}, \mathcal{P}, m^d \times m^{d'})$.

(vii) Take $X = Y := 2^{\mathbb{R}}$, so that $|X| = 2^{\mathfrak{c}} > \mathfrak{c}$, and equip both with the discrete σ-algebra. Let $E := \{(x,x) : x \in X\} \subset X \times Y$ be the diagonal, which certainly lies in $2^{X \times Y}$. Its slices are $E_x = \{x\}$, and these are pairwise distinct as x ranges over X; so E has exactly $|X| = 2^{\mathfrak{c}}$ distinct slices. By Exercise 1.7.18(v), any set in $2^X \times 2^Y$ has at most $\mathfrak{c}$ distinct slices. Since $2^{\mathfrak{c}} > \mathfrak{c}$ by Cantor's theorem, $E \notin 2^X \times 2^Y$. Hence $2^X \times 2^Y \neq 2^{X \times Y}$: the product of two discrete σ-algebras need not be discrete.

(viii) Suppose X is at most countable (the case of Y at most countable is symmetric). Let $E \subset X \times Y$ be arbitrary. Then

$$E = \bigcup_{x \in X} \{x\} \times E_x,$$

where $E_x = \{y \in Y : (x,y) \in E\}$: indeed $(x',y) \in E$ iff $y \in E_{x'}$ iff $(x',y) \in \{x'\} \times E_{x'}$. Each $\{x\} \times E_x$ is a measurable rectangle for the pair $(2^X, 2^Y)$, and the union is over the at most countable set X. Hence $E \in 2^X \times 2^Y$. Since E was arbitrary, $2^X \times 2^Y = 2^{X \times Y}$, which is the discrete σ-algebra on $X \times Y$.

Problem 1.7.20. Let $(X, \mathcal{B}_X)$, $(Y, \mathcal{B}_Y)$ be measurable spaces.
(i) Show that the product of two Dirac measures on $(X, \mathcal{B}_X)$, $(Y, \mathcal{B}_Y)$ is a Dirac measure on $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$.
(ii) If X, Y are at most countable, show that the product of the two counting measures on $(X, \mathcal{B}_X)$, $(Y, \mathcal{B}_Y)$ is the counting measure on $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$.

Solution. In both parts the strategy is the same: exhibit a measure on $\mathcal{B}_X \times \mathcal{B}_Y$ that obeys the base-times-height law (1.36), and invoke the uniqueness half of Proposition 1.7.11.
(i) Let $x_0 \in X$ and $y_0 \in Y$, and let $\delta_{x_0} : \mathcal{B}_X \to [0, +\infty]$ and $\delta_{y_0} : \mathcal{B}_Y \to [0, +\infty]$ be the corresponding Dirac measures of Example 1.4.22, so $\delta_{x_0}(E) = 1$ if $x_0 \in E$ and 0 otherwise, and likewise for δ_{y_0}. Both are probability measures, so both spaces are σ-finite and Proposition 1.7.11 applies: the product measure $\delta_{x_0} \times \delta_{y_0}$ exists and is the unique measure on $\mathcal{B}_X \times \mathcal{B}_Y$ satisfying (1.36).
Let $\delta_{(x_0,y_0)}$ be the Dirac measure at the point (x_0, y_0) on the measurable space $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$; this is a measure on any σ-algebra on $X \times Y$, no measurability of the singleton being required. For $E \in \mathcal{B}_X$ and $F \in \mathcal{B}_Y$ we have $(x_0, y_0) \in E \times F$ if and only if $x_0 \in E$ and $y_0 \in F$, so

$$\delta_{(x_0,y_0)}(E \times F) = 1_{x_0 \in E}\, 1_{y_0 \in F} = \delta_{x_0}(E)\, \delta_{y_0}(F).$$

Thus $\delta_{(x_0,y_0)}$ obeys (1.36), and by the uniqueness assertion of Proposition 1.7.11,

$$\delta_{x_0} \times \delta_{y_0} = \delta_{(x_0,y_0)},$$

which is a Dirac measure on $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$.
(ii) Let $\#_X : \mathcal{B}_X \to [0, +\infty]$, $\#_X(E) := |E|$, be counting measure on $(X, \mathcal{B}_X)$, and similarly $\#_Y$ and $\#_{X \times Y}$.
A preliminary remark on hypotheses. Proposition 1.7.11 defines the product measure only for σ-finite factors, and countability of X alone does not give this: for instance if $X = \mathbb{N}$ with the trivial σ-algebra $\mathcal{B}_X = \{\emptyset, X\}$, the only sets available have measure 0 and $+\infty$, so $(X, \mathcal{B}_X, \#_X)$ is not σ-finite and $\#_X \times \#_Y$ is not defined. So we read the exercise with the natural convention that the singletons are measurable (as they are for the discrete σ-algebra, which is the natural home of counting measure); then $X = \bigcup_{x \in X} \{x\}$ is an at most countable union of sets of measure 1, so $(X, \mathcal{B}_X, \#_X)$ is σ-finite, and similarly for Y. Under this hypothesis Proposition 1.7.11 applies.
Now let $E \in \mathcal{B}_X$ and $F \in \mathcal{B}_Y$. We claim

$$\#_{X \times Y}(E \times F) = |E \times F| = |E|\,|F| = \#_X(E)\, \#_Y(F),$$

with the convention $0 \cdot \infty = \infty \cdot 0 = 0$ in force throughout the text. Indeed: if $E = \emptyset$ or $F = \emptyset$ then $E \times F = \emptyset$ and both sides are 0. If E and F are both non-empty and finite, then $|E \times F| = |E||F|$ by elementary counting. If both are non-empty and at least one is infinite, then (since $E \subset X$ and $F \subset Y$ are at most countable) that one is countably infinite, and $E \times F$ contains a copy of it, hence is infinite; so both sides equal $+\infty$.
Therefore the restriction of $\#_{X \times Y}$ to $\mathcal{B}_X \times \mathcal{B}_Y$ is a measure on $\mathcal{B}_X \times \mathcal{B}_Y$ obeying (1.36). By the uniqueness assertion of Proposition 1.7.11,

$$\#_X \times \#_Y = \#_{X \times Y}\big|_{\mathcal{B}_X \times \mathcal{B}_Y},$$

i.e. the product of the two counting measures is counting measure on $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$.
(Note that the countability hypothesis is used twice: to obtain σ-finiteness, and to rule out the case of two uncountable factors, where one would need $|E \times F| = |E||F|$ for infinite cardinals (which is in fact still true under the axiom of choice), but where σ-finiteness genuinely fails.)

Problem 1.7.21. (Associativity of product). Let $(X, \mathcal{B}_X, \mu_X)$, $(Y, \mathcal{B}_Y, \mu_Y)$, $(Z, \mathcal{B}_Z, \mu_Z)$ be σ-finite sets. We may identify the Cartesian products $(X \times Y) \times Z$ and $X \times (Y \times Z)$ with each other in the obvious manner. If we do so, show that $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z)$ and $(\mu_X \times \mu_Y) \times \mu_Z = \mu_X \times (\mu_Y \times \mu_Z)$.

Solution. We identify $(X \times Y) \times Z$, $X \times (Y \times Z)$ and $X \times Y \times Z$ throughout, so that the set $(E \times F) \times G$ is the same as $E \times (F \times G)$ and we may write it as the box $E \times F \times G$.
Part 1: the σ-algebras. Let

$$\mathcal{R} := \{E \times F \times G : E \in \mathcal{B}_X,\ F \in \mathcal{B}_Y,\ G \in \mathcal{B}_Z\}$$

and $\mathcal{D} := \langle \mathcal{R} \rangle$. We show $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{D}$; by symmetry the same argument gives $\mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z) = \mathcal{D}$, whence the two agree.
Write $\mathcal{C} := (\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z$. By Exercise 1.7.18(i), $\mathcal{C}$ is generated by the sets $S \times G$ with $S \in \mathcal{B}_X \times \mathcal{B}_Y$ and $G \in \mathcal{B}_Z$.
$\mathcal{D} \subset \mathcal{C}$: a box $E \times F \times G$ is $(E \times F) \times G$ with $E \times F \in \mathcal{B}_X \times \mathcal{B}_Y$ (Exercise 1.7.18(i)) and $G \in \mathcal{B}_Z$, so it is one of the generators of $\mathcal{C}$. Hence $\mathcal{R} \subset \mathcal{C}$ and therefore $\mathcal{D} = \langle \mathcal{R} \rangle \subset \mathcal{C}$.
$\mathcal{C} \subset \mathcal{D}$: fix $G \in \mathcal{B}_Z$ and let

$$\mathcal{S}_G := \{S \subset X \times Y : S \times G \in \mathcal{D}\}.$$

We check $\mathcal{S}_G$ is a σ-algebra on $X \times Y$. First, $(X \times Y) \times G = X \times Y \times G \in \mathcal{R} \subset \mathcal{D}$, so $X \times Y \in \mathcal{S}_G$. Next, if $S \in \mathcal{S}_G$ then

$$\big((X \times Y) \setminus S\big) \times G = \big(X \times Y \times G\big) \setminus \big(S \times G\big) \in \mathcal{D},$$

since $\mathcal{D}$ is a σ-algebra and hence closed under set differences; so $\mathcal{S}_G$ is closed under complements. Finally, if $S_1, S_2, \cdots \in \mathcal{S}_G$ then $\big(\bigcup_n S_n\big) \times G = \bigcup_n (S_n \times G) \in \mathcal{D}$. Moreover $\mathcal{S}_G$ contains every measurable rectangle $E \times F$ ($E \in \mathcal{B}_X$, $F \in \mathcal{B}_Y$), since $(E \times F) \times G = E \times F \times G \in \mathcal{R}$. By Exercise 1.7.18(i) we conclude $\mathcal{B}_X \times \mathcal{B}_Y \subset \mathcal{S}_G$; that is, $S \times G \in \mathcal{D}$ for every $S \in \mathcal{B}_X \times \mathcal{B}_Y$ and every $G \in \mathcal{B}_Z$. So all generators of $\mathcal{C}$ lie in $\mathcal{D}$, giving $\mathcal{C} \subset \mathcal{D}$.
Hence $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{D} = \mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z)$.
Part 2: the measures. First, both iterated products are defined. Write $X = \bigcup_n X_n$, $Y = \bigcup_m Y_m$, $Z = \bigcup_k Z_k$ with $\mu_X(X_n), \mu_Y(Y_m), \mu_Z(Z_k) < \infty$. Then $X \times Y = \bigcup_{n,m} X_n \times Y_m$ is a countable union of sets of finite $\mu_X \times \mu_Y$-measure, by (1.36); so $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y, \mu_X \times \mu_Y)$ is σ-finite and Proposition 1.7.11 may be applied a second time to form $(\mu_X \times \mu_Y) \times \mu_Z$. Symmetrically for $\mu_X \times (\mu_Y \times \mu_Z)$.
Set $\mu' := (\mu_X \times \mu_Y) \times \mu_Z$ and $\mu'' := \mu_X \times (\mu_Y \times \mu_Z)$; by Part 1 both are measures on $\mathcal{D}$. For a box, two applications of (1.36) give

$$\mu'(E \times F \times G) = (\mu_X \times \mu_Y)(E \times F) \cdot \mu_Z(G) = \mu_X(E)\mu_Y(F)\mu_Z(G),$$

and likewise

$$\mu''(E \times F \times G) = \mu_X(E) \cdot (\mu_Y \times \mu_Z)(F \times G) = \mu_X(E)\mu_Y(F)\mu_Z(G).$$

So μ' and μ'' agree on $\mathcal{R}$.
Let $\mathcal{B}_0$ be the collection of finite unions of boxes. This is a Boolean algebra on $X \times Y \times Z$: it is closed under finite unions by definition; the intersection of two boxes is a box, so it is closed under finite intersections; and the complement of a box is the disjoint union of three boxes,

$$(X \times Y \times Z) \setminus (E \times F \times G) = \big((X \setminus E) \times Y \times Z\big) \cup \big(E \times (Y \setminus F) \times Z\big) \cup \big(E \times F \times (Z \setminus G)\big),$$

so by de Morgan it is closed under complements. Clearly $\langle \mathcal{B}_0 \rangle = \langle \mathcal{R} \rangle = \mathcal{D}$.
Every $S \in \mathcal{B}_0$ is a finite disjoint union of boxes: if $S = \bigcup_{i=1}^k E_i \times F_i \times G_i$, let $a_1, \ldots, a_p$ be the atoms of the finite Boolean algebra generated by $E_1, \ldots, E_k$ in $\mathcal{B}_X$, and similarly $b_1, \ldots, b_q$ and $c_1, \ldots, c_r$ for the F_i in $\mathcal{B}_Y$ and the G_i in $\mathcal{B}_Z$ (Exercise 1.4.4). The boxes $a_u \times b_v \times c_w$ are pairwise disjoint, each $E_i \times F_i \times G_i$ is a union of some of them, and hence so is S.
Since μ' and μ'' are finitely additive and agree on boxes, they therefore agree on all of $\mathcal{B}_0$. Let $\mu_0 := \mu'|_{\mathcal{B}_0} = \mu''|_{\mathcal{B}_0}$. Being the restriction of a countably additive measure to a Boolean subalgebra,

μ_0 is a pre-measure. It is σ-finite: $X \times Y \times Z = \bigcup_{n,m,k} X_n \times Y_m \times Z_k$ is a countable union of boxes with $\mu_0(X_n \times Y_m \times Z_k) = \mu_X(X_n)\mu_Y(Y_m)\mu_Z(Z_k) < \infty$.
Let λ be the Hahn-Kolmogorov extension of μ_0, defined on a σ-algebra $\mathcal{B} \supset \langle \mathcal{B}_0 \rangle = \mathcal{D}$. Both μ' and μ'' are countably additive extensions of μ_0 defined on $\mathcal{D}$, so by Exercise 1.7.7 (applicable since μ_0 is σ-finite) each agrees with λ on $\mathcal{B} \cap \mathcal{D} = \mathcal{D}$. Hence

$$(\mu_X \times \mu_Y) \times \mu_Z = \mu' = \lambda|_{\mathcal{D}} = \mu'' = \mu_X \times (\mu_Y \times \mu_Z)$$

on $\mathcal{D} = (\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z)$, as required.

Problem 1.7.22. The purpose of this exercise is to demonstrate that Tonelli's theorem can fail if the σ-finite hypothesis is removed, and also that product measure need not be unique. Let X be the unit interval $[0,1]$ with Lebesgue measure m (and the Lebesgue σ-algebra $\mathcal{L}([0,1])$) and Y be the unit interval $[0,1]$ with counting measure $\#$ (and the discrete σ-algebra $2^{[0,1]}$). Let $f := 1_E$ be the indicator function of the diagonal $E := \{(x,x) : x \in [0,1]\}$.

- (i) Show that f is measurable in the product σ-algebra.
- (ii) Show that $\int_X \left(\int_Y f(x,y)\, d\#(y)\right) dm(x) = 1$.
- (iii) Show that $\int_Y \left(\int_X f(x,y)\, dm(x)\right) d\#(y) = 0$.
- (iv) Show that there is more than one measure μ on $\mathcal{L}([0,1]) \times 2^{[0,1]}$ with the property that $\mu(E \times F) = m(E)\#(F)$ for all $E \in \mathcal{L}([0,1])$ and $F \in 2^{[0,1]}$. (Hint: use the two different ways to perform a double integral to create two different measures.)

Solution. Throughout, $\mathcal{B} := \mathcal{L}([0,1]) \times 2^{[0,1]}$ denotes the product σ-algebra on $[0,1] \times [0,1]$. Note that $(Y, 2^{[0,1]}, \#)$ is **not** σ-finite: a set of finite counting measure is finite, and $[0,1]$ is not a countable union of finite sets.
(i) The diagonal lies in the product σ-algebra.
For each $n \geq 1$ and each $1 \leq j \leq n$ let

$$I_{n,j} := \left[\tfrac{j-1}{n}, \tfrac{j}{n}\right) \cap [0,1] \quad (1 \leq j < n), \qquad I_{n,n} := \left[\tfrac{n-1}{n}, 1\right],$$

so that $I_{n,1}, \ldots, I_{n,n}$ partition $[0,1]$ into intervals of length at most $1/n$. Each $I_{n,j}$ is Lebesgue measurable as a subset of X, and is (trivially) an element of $2^{[0,1]}$ as a subset of Y, so each product $I_{n,j} \times I_{n,j}$ lies in $\mathcal{B}$, and hence so does the finite union

$$E_n := \bigcup_{j=1}^{n} I_{n,j} \times I_{n,j}.$$

I claim $E = \bigcap_{n=1}^{\infty} E_n$. If $x \in [0,1]$ then for each n the point x lies in exactly one $I_{n,j}$, whence $(x,x) \in I_{n,j} \times I_{n,j} \subseteq E_n$; so $E \subseteq \bigcap_n E_n$. Conversely, if $(x,y) \in E_n$ then x and y lie in a common $I_{n,j}$, so $|x-y| \leq 1/n$. If this holds for every n then $x = y$, i.e. $(x,y) \in E$. Thus $E = \bigcap_n E_n$ is a countable intersection of elements of $\mathcal{B}$, hence lies in $\mathcal{B}$, and $f = 1_E$ is $\mathcal{B}$-measurable.
(ii) The first iterated integral is 1.
Fix $x \in [0,1]$. The section $f_x : y \mapsto 1_E(x,y)$ equals $1_{\{x\}}$, which is $2^{[0,1]}$-measurable (every function on Y is). Its integral against counting measure is

$$\int_Y 1_{\{x\}}(y)\, d\#(y) = \#(\{x\}) = 1.$$

So the inner integral is the constant function 1 on X, which is certainly Lebesgue measurable, and

$$\int_X \left(\int_Y f(x,y)\, d\#(y)\right) dm(x) = \int_X 1\, dm = m([0,1]) = 1.$$

(iii) The second iterated integral is 0.

Fix $y \in [0,1]$. The section $f^y : x \mapsto 1_E(x,y)$ equals $1_{\{y\}}$, which is Lebesgue measurable (it is the indicator of a single point), and

$$\int_X 1_{\{y\}}(x)\, dm(x) = m(\{y\}) = 0.$$

So the inner integral is the constant function 0 on Y, and

$$\int_Y \left(\int_X f(x,y)\, dm(x) \right) d\#(y) = \int_Y 0\, d\# = 0.$$

Thus the two iterated integrals of the unsigned $\mathcal{B}$-measurable function f are 1 and 0. This does not contradict Tonelli's theorem (Theorem 1.7.15), because Y is not σ-finite; it shows that the σ-finiteness hypothesis there cannot be dropped.

(iv) Non-uniqueness of product measure.

We exhibit two distinct measures on $\mathcal{B}$ obeying the base-times-height formula $\mu(A \times B) = m(A)\#(B)$ (with the usual convention $0 \cdot \infty = 0$).

The measure μ_2 (integrate in x first, then in y). For $S \in \mathcal{B}$ define

$$\mu_2(S) := \int_Y \left(\int_X 1_S(x,y)\, dm(x) \right) d\#(y) = \sum_{y \in [0,1]} m(S^y),$$

where $S^y := \{x \in X : (x,y) \in S\}$ and the second expression is the (unsigned) sum over the arbitrary index set $[0,1]$, which agrees with the integral against counting measure by Exercise 1.4.40. This is well defined: by Exercise 1.7.18(iii) each section S^y lies in $\mathcal{L}([0,1])$, so $m(S^y)$ makes sense; and the resulting function $y \mapsto m(S^y)$ is automatically $2^{[0,1]}$-measurable since the σ-algebra on Y is discrete, so the outer integral exists. Note that no appeal to Tonelli's theorem is being made here: the definition of μ_2 is simply a formula, and we verify by hand that it is a measure.

μ_2 is a measure. Clearly $\mu_2(\emptyset) = 0$. If $S_1, S_2, \cdots \in \mathcal{B}$ are disjoint with union S, then for each fixed y the sections S_n^y are disjoint with union S^y, so by countable additivity of m we have $m(S^y) = \sum_n m(S_n^y)$. Summing over $y \in [0,1]$ and interchanging the two unsigned sums (Tonelli's theorem for series over arbitrary sets, Exercise 0.0.2) gives

$$\mu_2(S) = \sum_{y \in [0,1]} \sum_{n=1}^{\infty} m(S_n^y) = \sum_{n=1}^{\infty} \sum_{y \in [0,1]} m(S_n^y) = \sum_{n=1}^{\infty} \mu_2(S_n).$$

μ_2 obeys (1.36): if $S = A \times B$ then $S^y = A$ for $y \in B$ and $S^y = \emptyset$ otherwise, so $\mu_2(A \times B) = \sum_{y \in B} m(A) = m(A)\#(B)$ (this is correct in all cases, including $m(A) = 0$ with B infinite, where both sides vanish, and $m(A) > 0$ with B infinite, where both sides are $+\infty$).

Finally, by part (iii), $\mu_2(E) = 0$.

The measure μ_1 (Hahn-Kolmogorov). Let $\mathcal{B}_0$ be the Boolean algebra of finite unions of products $A \times B$ with $A \in \mathcal{L}([0,1])$, $B \in 2^{[0,1]}$, and let $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ be the elementary product pre-measure of Proposition 1.7.11, namely $\mu_0(\bigcup_{j=1}^k A_j \times B_j) := \sum_{j=1}^k m(A_j)\#(B_j)$ for disjoint decompositions. Inspecting the proof of Proposition 1.7.11, the verification that μ_0 is well defined and finitely additive uses only the finite additivity of m and of $\#$, and the verification that μ_0 is in fact a pre-measure uses only the pointwise identity $1_{A \times B} = \sum_n 1_{A_n \times B_n}$ together with two applications of the monotone convergence theorem (Theorem 1.4.44), one on Y and one on X. At no point in that argument is σ-finiteness invoked; it enters Proposition 1.7.11 only at the very last step, to deduce **uniqueness** from Exercise 1.7.7. Hence μ_0 is a pre-measure in our (non-σ-finite) setting as well, and the Hahn-Kolmogorov theorem (Theorem 1.7.8) extends it to a countably additive measure defined on the σ-algebra of Caratheodory-measurable sets; that σ-algebra contains $\mathcal{B}_0$, hence contains $\langle \mathcal{B}_0 \rangle = \mathcal{B}$ (the identification $\langle \mathcal{B}_0 \rangle = \mathcal{B}$ is Exercise 1.7.18(i)). Let μ_1 be the restriction of that extension to $\mathcal{B}$. By construction $\mu_1(A \times B) = \mu_0(A \times B) = m(A)\#(B)$, so μ_1 also obeys (1.36).

Now we compute $\mu_1(E)$. In the Hahn-Kolmogorov construction, μ_1 is the restriction to $\mathcal{B}$ of the outer measure

$$\mu_0^*(S) = \inf \Big\{ \sum_{n=1}^{\infty} \mu_0(S_n) : S \subseteq \bigcup_{n=1}^{\infty} S_n,\ S_n \in \mathcal{B}_0 \Big\}.$$

Since every $S_n \in \mathcal{B}_0$ is a finite disjoint union of products with μ_0 equal to the corresponding sum, we may refine any such cover into a countable cover by products without increasing the total; hence it suffices to consider covers $E \subseteq \bigcup_n A_n \times B_n$. For such a cover and any $x \in [0,1]$, the point (x,x) lies in some $A_n \times B_n$, i.e. $x \in A_n \cap B_n$. Thus

$$[0,1] = \bigcup_{n=1}^{\infty} (A_n \cap B_n),$$

and so, writing m^* for Lebesgue outer measure (the sets $A_n \cap B_n$ need not be measurable, as B_n is arbitrary), countable subadditivity of outer measure gives $1 = m^*([0,1]) \leq \sum_n m^*(A_n \cap B_n)$. In particular there is some n with $m^*(A_n \cap B_n) > 0$. For that n, monotonicity of outer measure gives $m(A_n) = m^*(A_n) \geq m^*(A_n \cap B_n) > 0$. Also $A_n \cap B_n$ cannot be countable, since a countable set has Lebesgue outer measure zero; hence $B_n \supseteq A_n \cap B_n$ is an infinite set, so $\#(B_n) = +\infty$. Hence $\mu_0(A_n \times B_n) = m(A_n)\#(B_n) = +\infty$, and the cover has total mass $+\infty$. As every cover behaves this way,

$$\mu_1(E) = \mu_0^*(E) = +\infty.$$

Since $\mu_1(E) = +\infty \neq 0 = \mu_2(E)$, the measures μ_1 and μ_2 are distinct, yet both satisfy $\mu(A \times B) = m(A)\#(B)$ for all $A \in \mathcal{L}([0,1])$ and $B \in 2^{[0,1]}$. So product measure is not unique once σ-finiteness is dropped, as claimed in Remark 1.7.12.

Problem 1.7.23. Give an example of a Borel measurable function $f : [0,1]^2 \to \mathbb{R}$ such that the integrals $\int_{[0,1]} f(x,y)\,dy$ and $\int_{[0,1]} f(x,y)\,dx$ exist and are absolutely integrable for all $x \in [0,1]$ and $y \in [0,1]$ respectively, and that $\int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dy \right) dx$ and $\int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dx \right) dy$ exist and are absolutely integrable, but such that

$$\int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dy \right) dx \neq \int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dx \right) dy.$$

(Hint: adapt the example from Remark 0.0.3.) Thus we see that Fubini's theorem fails when one drops the hypothesis that f is absolutely integrable with respect to the product space.

Solution. We transplant the counterexample of Remark 0.0.3 (the doubly indexed array $x_{n,m}$ equal to $+1$ when $n = m$, -1 when $n = m+1$, and 0 otherwise) from $\mathbb{N}^2$ to $[0,1]^2$, replacing each lattice point by a normalised bump on a small interval.
The construction. For $n = 1, 2, 3, \ldots$ set

$$I_n := \left(\tfrac{1}{n+1}, \tfrac{1}{n} \right], \qquad |I_n| = \tfrac{1}{n} - \tfrac{1}{n+1} = \tfrac{1}{n(n+1)},$$

so that the I_n are disjoint intervals partitioning $(0,1]$. Let

$$g_n := \tfrac{1}{|I_n|} 1_{I_n}, \qquad \text{so that} \qquad \int_{[0,1]} g_n(t)\,dt = 1 \quad \text{for every } n,$$

and set $g_0 := 0$. Define $f : [0,1]^2 \to \mathbb{R}$ by

$$f(x,y) := \sum_{n=1}^{\infty} \big(g_n(x) - g_{n+1}(x) \big) g_n(y).$$

f is well defined, real valued and Borel measurable. If $x = 0$ or $y = 0$ then every term vanishes and $f(x,y) = 0$. If $x, y \in (0,1]$, then y lies in exactly one interval I_M, so $g_n(y) = 0$ for all $n \neq M$ and the series has exactly one possibly nonzero term; in particular it converges (indeed it is a finite sum at each point) and f is finite everywhere. Each partial sum $\sum_{n=1}^{N} (g_n - g_{n+1})(x) g_n(y)$ is a finite linear combination of indicator functions of Borel rectangles $I_j \times I_n$, hence Borel measurable on $[0,1]^2$;

and f is the pointwise limit of these partial sums, hence Borel measurable, since a pointwise limit of measurable functions is measurable (Exercise 1.4.29(vi)).
Explicitly, for $x \in I_N$ and $y \in (0,1]$,

$$f(x,y) = \tfrac{1}{|I_N|}\big(g_N(y) - g_{N-1}(y)\big),$$

since $g_n(x) \neq 0$ only for $n = N$ and $g_{n+1}(x) \neq 0$ only for $n = N-1$. Dually, for $y \in I_M$ and $x \in (0,1]$,

$$f(x,y) = \tfrac{1}{|I_M|}\big(g_M(x) - g_{M+1}(x)\big).$$

Rows: integrating in y first. Fix $x \in [0,1]$. If $x = 0$ then $f(x,\cdot) \equiv 0$. Otherwise $x \in I_N$ for a unique N, and by the first displayed formula $f(x,\cdot) = \frac{1}{|I_N|}(g_N - g_{N-1})$ is a bounded Borel function, so it is absolutely integrable on $[0,1]$, with

$$\int_{[0,1]} f(x,y)\,dy = \tfrac{1}{|I_N|}\left(\int_{[0,1]} g_N - \int_{[0,1]} g_{N-1}\right) = \begin{cases} \frac{1}{|I_1|} = 2, & N = 1, \\ \frac{1}{|I_N|}(1-1) = 0, & N \geq 2, \end{cases}$$

using $g_0 = 0$. Hence the function $F(x) := \int_{[0,1]} f(x,y)\,dy$ equals $2 \cdot 1_{I_1}(x)$, which is bounded and Borel measurable, therefore absolutely integrable, and

$$\int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dy\right) dx = 2\,|I_1| = 2 \cdot \tfrac{1}{2} = 1.$$

Columns: integrating in x first. Fix $y \in [0,1]$. If $y = 0$ then $f(\cdot,y) \equiv 0$. Otherwise $y \in I_M$ for a unique M, and $f(\cdot,y) = \frac{1}{|I_M|}(g_M - g_{M+1})$ is bounded and Borel, hence absolutely integrable, with

$$\int_{[0,1]} f(x,y)\,dx = \tfrac{1}{|I_M|}\left(\int_{[0,1]} g_M - \int_{[0,1]} g_{M+1}\right) = \tfrac{1}{|I_M|}(1-1) = 0.$$

So $G(y) := \int_{[0,1]} f(x,y)\,dx$ is identically 0, which is certainly absolutely integrable, and

$$\int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dx\right) dy = 0.$$

Conclusion. Both iterated integrals exist, with absolutely integrable inner integrals and absolutely integrable outer integrands, yet

$$1 = \int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dy\right) dx \neq \int_{[0,1]} \left(\int_{[0,1]} f(x,y)\,dx\right) dy = 0.$$

Consistency with Fubini's theorem. By Fubini's theorem (Theorem 1.7.21) this forces f to fail to be absolutely integrable on the product space, and indeed it is: for $x \in I_N$ with $N \geq 2$ the functions g_N and g_{N-1} have disjoint supports, so $\int_{[0,1]} |f(x,y)|\,dy = \frac{1}{|I_N|}(1+1) = \frac{2}{|I_N|}$, and Tonelli's theorem (Theorem 1.7.15) applied to $|f|$ gives

$$\int_{[0,1]^2} |f(x,y)|\,d(x,y) \geq \sum_{N=2}^{\infty} \int_{I_N} \tfrac{2}{|I_N|}\,dx = \sum_{N=2}^{\infty} 2 = +\infty.$$

Thus the absolute integrability hypothesis in Fubini's theorem cannot be replaced by the mere existence of the two iterated integrals.

Problem 1.7.24. (Area interpretation of integral) Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space, and let $\mathbb{R}$ be equipped with Lebesgue measure m and the Borel σ-algebra $\mathcal{B}[\mathbb{R}]$. Show that $f : X \to [0, +\infty]$ is measurable if and only if the set $\{(x,t) \in X \times \mathbb{R} : 0 \le t \le f(x)\}$ is measurable in $\mathcal{B} \times \mathcal{B}[\mathbb{R}]$, in which case we have

$$(\mu \times m)\big(\{(x,t) \in X \times \mathbb{R} : 0 \le t \le f(x)\}\big) = \int_X f(x)\, d\mu(x).$$

Similarly if we replace $\{(x,t) \in X \times \mathbb{R} : 0 \le t \le f(x)\}$ by $\{(x,t) \in X \times \mathbb{R} : 0 \le t < f(x)\}$.

Solution. Write

$$\Sigma_{\le} := \{(x,t) \in X \times \mathbb{R} : 0 \le t \le f(x)\}, \qquad \Sigma_{<} := \{(x,t) \in X \times \mathbb{R} : 0 \le t < f(x)\}.$$

Both $(X, \mathcal{B}, \mu)$ and $(\mathbb{R}, \mathcal{B}[\mathbb{R}], m)$ are σ-finite: the former by hypothesis, and $\mathbb{R} = \bigcup_{n=1}^{\infty}[-n, n]$ with $m([-n, n]) = 2n < \infty$. So the product measure $\mu \times m$ on $\mathcal{B} \times \mathcal{B}[\mathbb{R}]$ exists and is unique (Proposition 1.7.11), and Tonelli's theorem (Theorem 1.7.15) is available.

If f is measurable, then $\Sigma_{\le}$ and $\Sigma_{<}$ are product measurable. The point of the argument is to express the relevant regions as countable unions of measurable rectangles, using the density of the rationals. I claim

$$\{(x,t) \in X \times \mathbb{R} : t > f(x)\} = \bigcup_{q \in \mathbb{Q}} \{x \in X : f(x) < q\} \times (q, +\infty).$$

Indeed, if $t > f(x)$ then (since $f(x) < t < +\infty$) we may choose a rational q with $f(x) < q < t$, and then $x \in \{f < q\}$ and $t \in (q, +\infty)$. Conversely if $f(x) < q$ and $t > q$ then $t > f(x)$. Each set $\{f < q\}$ lies in $\mathcal{B}$ because f is measurable (Exercise 1.4.29(i) and (iii)), and each $(q, +\infty)$ is Borel, so the right-hand side is a countable union of measurable rectangles, hence lies in $\mathcal{B} \times \mathcal{B}[\mathbb{R}]$. Taking complements, $\{(x,t) : t \le f(x)\}$ is product measurable, and therefore so is

$$\Sigma_{\le} = \big(X \times [0, +\infty)\big) \cap \{(x,t) : t \le f(x)\}.$$

Likewise

$$\{(x,t) \in X \times \mathbb{R} : t < f(x)\} = \bigcup_{q \in \mathbb{Q}} \{x \in X : f(x) > q\} \times (-\infty, q)$$

(if $t < f(x)$ pick a rational q with $t < q < f(x)$, which is possible whether or not $f(x) = +\infty$; the converse inclusion is immediate), so this set is product measurable and hence so is $\Sigma_{<} = (X \times [0, +\infty)) \cap \{(x,t) : t < f(x)\}$.

Conversely, if $\Sigma_{\le}$ (or $\Sigma_{<}$) is product measurable, then f is measurable. Suppose $\Sigma_{\le} \in \mathcal{B} \times \mathcal{B}[\mathbb{R}]$. By Exercise 1.7.18(iii), every horizontal slice $\Sigma_{\le}^t := \{x \in X : (x,t) \in \Sigma_{\le}\}$ lies in $\mathcal{B}$. For $\lambda \ge 0$ this slice is exactly $\{x \in X : f(x) \ge \lambda\}$. Hence $\{f \ge \lambda\} \in \mathcal{B}$ for every $\lambda \ge 0$, and trivially $\{f \ge \lambda\} = X \in \mathcal{B}$ for $\lambda < 0$. Consequently, for any real λ,

$$\{x \in X : f(x) > \lambda\} = \bigcup_{n=1}^{\infty} \{x \in X : f(x) \ge \lambda + \tfrac{1}{n}\} \in \mathcal{B},$$

so f is measurable by Exercise 1.4.29(i). The same argument works for $\Sigma_{<}$: its slice at height $\lambda \ge 0$ is $\{x : f(x) > \lambda\}$, which together with the trivial case $\lambda < 0$ gives measurability of all level sets $\{f > \lambda\}$ directly. Note this direction did not use σ-finiteness.

The area formula. Assume now f is measurable, so that $\Sigma_{\le}$ is product measurable. Since the integral of an indicator function is the measure of the underlying set, Tonelli's theorem (Theorem 1.7.15) applied to the unsigned measurable function $1_{\Sigma_{\le}}$ gives

$$(\mu \times m)(\Sigma_{\le}) = \int_{X \times \mathbb{R}} 1_{\Sigma_{\le}}(x,t)\, d\mu \times m(x,t) = \int_X \left(\int_{\mathbb{R}} 1_{\Sigma_{\le}}(x,t)\, dm(t) \right) d\mu(x).$$

For each fixed $x \in X$ the vertical slice is

$$(\Sigma_{\leq})_x = \{t \in \mathbb{R} : 0 \leq t \leq f(x)\} = \begin{cases} [0, f(x)], & f(x) < +\infty, \\ [0, +\infty), & f(x) = +\infty, \end{cases}$$

whose Lebesgue measure is $f(x)$ in either case (an interval of length $f(x)$ when $f(x)$ is finite, and $+\infty = f(x)$ when $f(x) = +\infty$). Hence the inner integral equals $f(x)$, and

$$(\mu \times m)(\Sigma_{\leq}) = \int_X f(x)\, d\mu(x),$$

as required.
For $\Sigma_<$ the argument is identical: the vertical slice at x is $[0, f(x))$ when $f(x) < +\infty$ (which is the empty set when $f(x) = 0$) and $[0, +\infty)$ when $f(x) = +\infty$, and in every case its Lebesgue measure is again $f(x)$. Tonelli's theorem therefore gives

$$(\mu \times m)(\Sigma_<) = \int_X \left(\int_{\mathbb{R}} 1_{\Sigma_<}(x,t)\, dm(t) \right) d\mu(x) = \int_X f(x)\, d\mu(x).$$

(In particular the graph $\Sigma_{\leq} \setminus \Sigma_< = \{(x,t) : t = f(x) < +\infty\}$ is a $\mu \times m$-null set, as one expects.)

Problem 1.7.25. (Distribution formula) Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space, and let $f : X \to [0, +\infty]$ be measurable. Show that

$$\int_X f(x)\, d\mu(x) = \int_{[0,+\infty]} \mu(\{x \in X : f(x) \geq \lambda\})\, d\lambda.$$

(Note that the integrand on the right-hand side is monotone and thus Lebesgue measurable.) Similarly if we replace $\{x \in X : f(x) \geq \lambda\}$ by $\{x \in X : f(x) > \lambda\}$.

Solution. As in Exercise 1.7.24, equip $\mathbb{R}$ with Lebesgue measure m and the Borel σ-algebra $\mathcal{B}[\mathbb{R}]$; this space is σ-finite, as is $(X, \mathcal{B}, \mu)$ by hypothesis, so the product measure $\mu \times m$ exists (Proposition 1.7.11) and Tonelli's theorem (Theorem 1.7.15) applies on $X \times \mathbb{R}$.
First a remark on the domain of integration on the right. The single point $\{+\infty\}$ is a null set for Lebesgue measure on $[0, +\infty]$, so

$$\int_{[0,+\infty]} \mu(\{f \geq \lambda\})\, d\lambda = \int_{[0,+\infty)} \mu(\{f \geq \lambda\})\, d\lambda,$$

and it suffices to work over $[0, +\infty)$. Also, as the problem notes, the map $\lambda \mapsto \mu(\{f \geq \lambda\})$ is non-increasing (if $\lambda \leq \lambda'$ then $\{f \geq \lambda'\} \subseteq \{f \geq \lambda\}$, so monotonicity of μ applies), and a monotone function $[0, +\infty) \to [0, +\infty]$ has level sets $\{\lambda : \mu(\{f \geq \lambda\}) > c\}$ that are intervals, hence Lebesgue measurable; so the right-hand side is well defined. (The measurability also drops out of Tonelli's theorem below.)
The formula with $\geq$. Let

$$\Sigma_{\leq} := \{(x,t) \in X \times \mathbb{R} : 0 \leq t \leq f(x)\}.$$

By Exercise 1.7.24 this set lies in $\mathcal{B} \times \mathcal{B}[\mathbb{R}]$ and satisfies

$$(\mu \times m)(\Sigma_{\leq}) = \int_X f(x)\, d\mu(x).$$

Now evaluate the same product measure by slicing in the other direction. Applying Tonelli's theorem (Theorem 1.7.15) to the unsigned measurable function $1_{\Sigma_{\leq}}$, and integrating in x before t, we get

$$(\mu \times m)(\Sigma_{\leq}) = \int_{\mathbb{R}} \left(\int_X 1_{\Sigma_{\leq}}(x,t)\, d\mu(x) \right) dm(t) = \int_{\mathbb{R}} \mu\big((\Sigma_{\leq})^t\big)\, dm(t),$$

where the horizontal slice is

$$(\Sigma_{\leq})^t = \{x \in X : 0 \leq t \leq f(x)\} = \begin{cases} \{x \in X : f(x) \geq t\}, & t \geq 0, \\ \emptyset, & t < 0. \end{cases}$$

(Tonelli's theorem also asserts that $t \mapsto \mu((\Sigma_{\leq})^t)$ is Borel measurable.) Hence

$$\int_X f(x)\, d\mu(x) = (\mu \times m)(\Sigma_{\leq}) = \int_{[0,+\infty)} \mu(\{x \in X : f(x) \geq \lambda\})\, d\lambda = \int_{[0,+\infty]} \mu(\{x \in X : f(x) \geq \lambda\})\, d\lambda,$$

which is the claim.
The formula with $>$. Repeat the argument with

$$\Sigma_{<} := \{(x,t) \in X \times \mathbb{R} : 0 \leq t < f(x)\},$$

which by Exercise 1.7.24 is also product measurable with $(\mu \times m)(\Sigma_{<}) = \int_X f\, d\mu$. Its horizontal slice at height t is

$$(\Sigma_{<})^t = \{x \in X : 0 \leq t < f(x)\} = \begin{cases} \{x \in X : f(x) > t\}, & t \geq 0, \\ \emptyset, & t < 0, \end{cases}$$

so Tonelli's theorem gives

$$\int_X f(x)\, d\mu(x) = \int_{[0,+\infty)} \mu(\{x \in X : f(x) > \lambda\})\, d\lambda = \int_{[0,+\infty]} \mu(\{x \in X : f(x) > \lambda\})\, d\lambda.$$

Remark. The two integrands agree except on $J := \{\lambda \geq 0 : \mu(\{f = \lambda\}) > 0\}$, since $\mu(\{f \geq \lambda\}) = \mu(\{f > \lambda\}) + \mu(\{f = \lambda\})$ whenever the right-hand side is not of the form $\infty + \infty$. The set J is at most countable: writing $X = \bigcup_n X_n$ with $\mu(X_n) < \infty$, for each n and each $k \geq 1$ there are at most $k\mu(X_n)$ values of λ with $\mu(X_n \cap \{f = \lambda\}) > 1/k$ (the sets $\{f = \lambda\}$ being disjoint for distinct λ), so $\{\lambda : \mu(X_n \cap \{f = \lambda\}) > 0\}$ is a countable union of finite sets, and J is the union over n of these. Thus J is Lebesgue null and the two right-hand sides in the exercise agree integrand-by-integrand almost everywhere, which is an alternative route to the second identity once the first is known.

Problem 1.7.26. (Approximations to the identity) Let $P : \mathbb{R}^d \to \mathbb{R}^+$ be a good kernel (see Exercise 1.6.27), and let $P_t(x) := \frac{1}{t^d} P\left(\frac{x}{t}\right)$ be the associated rescaled functions. Show that if $f : \mathbb{R}^d \to \mathbb{C}$ is absolutely integrable, then $f * P_t$ converges in L^1 norm to f as $t \to 0$. (Hint: use the density argument. You will need an upper bound on $\|f * P_t\|_{L^1(\mathbb{R}^d)}$ which can be obtained using Tonelli's theorem.)

Solution. Recall from Exercise 1.6.27 that a good kernel is a measurable $P : \mathbb{R}^d \to \mathbb{R}^+$ which is non-negative, radial (so $P(x) = \tilde{P}(|x|)$ for some $\tilde{P} : [0,+\infty) \to \mathbb{R}^+$), radially non-increasing (so $\tilde{P}$ is non-increasing), and has total mass $\int_{\mathbb{R}^d} P(x)\, dx = 1$.
Step 0: preliminaries on P and P_t.
First, P is in fact Borel measurable. Indeed, for any $\lambda \in \mathbb{R}$ the set $\{r \geq 0 : \tilde{P}(r) > \lambda\}$ is, by monotonicity of $\tilde{P}$, an interval of the form $[0, r_\lambda)$ or $[0, r_\lambda]$ (possibly empty, possibly all of $[0,+\infty)$), so $\{x \in \mathbb{R}^d : P(x) > \lambda\}$ is an open ball, a closed ball, the empty set, or $\mathbb{R}^d$, and in every case Borel. Hence each P_t is Borel, and therefore the function

$$(x,y) \mapsto P_t(x-y)$$

is Borel measurable on $\mathbb{R}^d \times \mathbb{R}^d$, being the composition of the continuous map $(x,y) \mapsto x - y$ with a Borel function. Since $\mathcal{B}[\mathbb{R}^{2d}] = \mathcal{B}[\mathbb{R}^d] \times \mathcal{B}[\mathbb{R}^d]$ (Exercise 1.7.19(iv)), this function is measurable with respect to $\mathcal{L}[\mathbb{R}^d] \times \mathcal{L}[\mathbb{R}^d]$. Together with the fact that $(x,y) \mapsto |h(y)|$ is $\mathcal{L}[\mathbb{R}^d] \times \mathcal{L}[\mathbb{R}^d]$-measurable for any Lebesgue measurable h (compose $|h|$ with the measurable projection π_Y, Exercise 1.7.18(ii)), we conclude that

$$(x,y) \mapsto |h(y)| P_t(x-y)$$

is an unsigned $\mathcal{L}[\mathbb{R}^d] \times \mathcal{L}[\mathbb{R}^d]$-measurable function, so Tonelli's theorem (Theorem 1.7.15) applies to it; note both factors are σ-finite.
Next, each P_t has total mass one: applying the linear change of variables $x \mapsto tx$ (Exercise 1.3.16, with $|\det| = t^d$),

$$\int_{\mathbb{R}^d} P_t(x)\,dx = \tfrac{1}{t^d}\int_{\mathbb{R}^d} P(x/t)\,dx = \tfrac{1}{t^d}\cdot t^d \int_{\mathbb{R}^d} P(x)\,dx = 1.$$

The same change of variables gives, for any $\delta > 0$,

$$\int_{|y|>\delta} P_t(y)\,dy = \int_{|s|>\delta/t} P(s)\,ds =: \varepsilon(\delta/t),$$

where $\varepsilon(R) := \int_{|s|>R} P(s)\,ds$. The functions $P1_{\{|s|\le n\}}$ increase pointwise to P as $n \to \infty$ through the integers, so the monotone convergence theorem (Theorem 1.4.44) gives $\int_{|s|\le n} P \to \int_{\mathbb{R}^d} P = 1$; since $\int P = 1$ is finite we may subtract to get $\varepsilon(n) = 1 - \int_{|s|\le n} P \to 0$. As ε is non-increasing in R, it follows that $\varepsilon(R) \to 0$ as $R \to \infty$ through the reals as well. Hence for each fixed $\delta > 0$,

$$\int_{|y|>\delta} P_t(y)\,dy \longrightarrow 0 \qquad \text{as } t \to 0. \qquad (*)$$

Step 1: the quantitative bound $\|h * P_t\|_{L^1} \le \|h\|_{L^1}$.
Let $h : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable. For fixed y, the change of variables $x \mapsto x + y$ followed by the reflection $x \mapsto -x$ (Exercise 1.3.15 and Exercise 1.3.16 with $T = -\mathrm{id}$, which has $|\det T| = 1$) gives $\int_{\mathbb{R}^d} P_t(x-y)\,dx = \int_{\mathbb{R}^d} P_t(x)\,dx = 1$. Hence by Tonelli's theorem (Theorem 1.7.15; both factors are σ-finite),

$$\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |h(y)|P_t(x-y)\,dy\right)dx = \int_{\mathbb{R}^d} |h(y)|\left(\int_{\mathbb{R}^d} P_t(x-y)\,dx\right)dy = \int_{\mathbb{R}^d} |h(y)|\,dy = \|h\|_{L^1(\mathbb{R}^d)},$$

which is finite. Consequently the inner integral is finite for almost every x, so $h*P_t(x) = \int_{\mathbb{R}^d} h(y)P_t(x-y)\,dy$ is well defined for almost every x; moreover $(x,y) \mapsto h(y)P_t(x-y)$ is absolutely integrable on the product space, so Fubini's theorem (Theorem 1.7.21) applies – both factor spaces $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d], m)$ are complete and σ-finite – and shows that $x \mapsto h * P_t(x)$ is (almost everywhere defined and) absolutely integrable, in particular measurable. By the triangle inequality for integrals,

$$\|h * P_t\|_{L^1(\mathbb{R}^d)} \le \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |h(y)|P_t(x-y)\,dy\right)dx = \|h\|_{L^1(\mathbb{R}^d)}. \qquad (1)$$

This is the required upper bound. Note also that convolution against P_t is linear, so $(f-g)*P_t = f*P_t - g*P_t$ almost everywhere.
Step 2: the dense subclass.
By Littlewood's second principle in the form of Theorem 1.3.20(iii), the continuous, compactly supported functions are dense in $L^1(\mathbb{R}^d)$.
Step 3: convergence for continuous compactly supported g.
Let $g : \mathbb{R}^d \to \mathbb{C}$ be continuous and supported in the closed ball $\overline{B(0,R)}$ for some $R > 0$. Since $\|g\|_{L^\infty} < \infty$ and $\int P_t = 1$, we have $\int |g(y)|P_t(x-y)\,dy \le \|g\|_{L^\infty} < \infty$, so the integral defining $g*P_t(x)$ converges absolutely for every x, and $g*P_t$ is defined everywhere (and agrees with the almost-everywhere defined object of Step 1). Making the change of variables $y = x - z$ (the composition of a reflection and a translation, Exercises 1.3.15 and 1.3.16; these are stated for unsigned integrands, and extend to absolutely integrable complex ones by splitting into real and imaginary, then positive and negative, parts) and using $\int P_t = 1$,

$$g * P_t(x) - g(x) = \int_{\mathbb{R}^d} \big(g(x-z) - g(x)\big)P_t(z)\,dz.$$

A continuous compactly supported function is uniformly continuous, so its modulus of continuity

$$\omega(\delta) := \sup_{|z|\le\delta}\ \sup_{x\in\mathbb{R}^d} |g(x-z) - g(x)|$$

satisfies $\omega(\delta) \to 0$ as $\delta \to 0$. Splitting the integral at $|z| = \delta$ and using $P_t \geq 0$,

$$|g * P_t(x) - g(x)| \leq \omega(\delta) \int_{|z|\leq\delta} P_t(z)\, dz + 2\|g\|_{L^\infty} \int_{|z|>\delta} P_t(z)\, dz \leq \omega(\delta) + 2\|g\|_{L^\infty} \int_{|z|>\delta} P_t(z)\, dz,$$

uniformly in x. Given $\eta > 0$, first choose $\delta > 0$ with $\omega(\delta) \leq \eta$, and then, by $(*)$, choose t small enough that $2\|g\|_{L^\infty} \int_{|z|>\delta} P_t \leq \eta$. Hence

$$\sup_{x\in\mathbb{R}^d} |g * P_t(x) - g(x)| \longrightarrow 0 \qquad \text{as } t \to 0. \qquad (2)$$

Uniform convergence alone does not give L^1 convergence on the unbounded domain $\mathbb{R}^d$, so we also control the tail. For $|x| > 2R$ we have $g(x) = 0$, and by Tonelli's theorem

$$\int_{|x|>2R} |g * P_t(x)|\, dx \leq \int_{|y|\leq R} |g(y)| \left(\int_{|x|>2R} P_t(x - y)\, dx \right) dy.$$

If $|y| \leq R$ and $|x| > 2R$ then $|x - y| \geq |x| - |y| > R$, so the inner integral is at most $\int_{|z|>R} P_t(z)\, dz$. Therefore

$$\int_{|x|>2R} |g * P_t(x) - g(x)|\, dx \leq \|g\|_{L^1} \int_{|z|>R} P_t(z)\, dz \longrightarrow 0 \qquad \text{as } t \to 0, \qquad (3)$$

again by $(*)$ (with $\delta = R$). Combining (2) and (3),

$$\|g * P_t - g\|_{L^1(\mathbb{R}^d)} \leq m(B(0, 2R)) \sup_x |g * P_t(x) - g(x)| + \|g\|_{L^1} \int_{|z|>R} P_t(z)\, dz \longrightarrow 0$$

as $t \to 0$, since $m(B(0, 2R)) < \infty$. Thus $g * P_t \to g$ in L^1 norm for every continuous compactly supported g.

Step 4: the density argument.

Let $f : \mathbb{R}^d \to \mathbb{C}$ be absolutely integrable and let $\eta > 0$. By Step 2 there is a continuous compactly supported g with $\|f - g\|_{L^1(\mathbb{R}^d)} \leq \eta$. For any $t > 0$, the triangle inequality and the quantitative bound (1) applied to $h := f - g$ give

$$\|f * P_t - f\|_{L^1} \leq \|(f - g) * P_t\|_{L^1} + \|g * P_t - g\|_{L^1} + \|g - f\|_{L^1} \leq \|f - g\|_{L^1} + \|g * P_t - g\|_{L^1} + \|f - g\|_{L^1},$$

so

$$\|f * P_t - f\|_{L^1} \leq 2\eta + \|g * P_t - g\|_{L^1}.$$

By Step 3, $\|g * P_t - g\|_{L^1} \leq \eta$ for all sufficiently small $t > 0$, whence $\|f * P_t - f\|_{L^1} \leq 3\eta$ for all sufficiently small t. As $\eta > 0$ was arbitrary,

$$\lim_{t\to 0} \|f * P_t - f\|_{L^1(\mathbb{R}^d)} = 0,$$

i.e. $f * P_t$ converges to f in L^1 norm, as claimed.

3 Related Articles

Contents

3.1 §2.2 — The Rademacher Differentiation Theorem 299
3.2 §2.3 — Probability Spaces 305
3.3 §2.4 — Infinite Product Spaces and the Kolmogorov Extension Theorem 307

3.1 §2.2 — The Rademacher Differentiation Theorem

Problem 2.2.1. Show that Lipschitz continuous functions are uniformly continuous, and hence continuous. Then give an example of a uniformly continuous function $f : [0,1] \to [0,1]$ that is not Lipschitz continuous.

Solution. Lipschitz implies uniformly continuous. Let $f : X \to Y$ be Lipschitz continuous between metric spaces (X, d_X) and (Y, d_Y), so by Definition 2.2.1 there is a constant $C > 0$ with

$$d_Y(f(x), f(x')) \leq C\, d_X(x, x') \quad \text{for all } x, x' \in X.$$

Let $\varepsilon > 0$ and set $\delta := \varepsilon / C > 0$. If $x, x' \in X$ satisfy $d_X(x, x') < \delta$, then

$$d_Y(f(x), f(x')) \leq C\, d_X(x, x') < C\delta = \varepsilon.$$

The choice of δ depended only on ε and not on the points x, x', so f is uniformly continuous. Uniform continuity trivially implies continuity: given $x \in X$ and $\varepsilon > 0$, the same δ witnesses continuity of f at x.
Now for an example of a uniformly continuous function on $[0,1]$ that is not Lipschitz. Take

$$f : [0,1] \to [0,1], \qquad f(x) := \sqrt{x}.$$

This does map $[0,1]$ into $[0,1]$, since $0 \leq x \leq 1$ forces $0 \leq \sqrt{x} \leq 1$.
To see that f is uniformly continuous, we first check the elementary inequality

$$|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|} \qquad (x, y \geq 0).$$

Indeed, by symmetry we may assume $x \geq y \geq 0$, and then

$$(\sqrt{x} - \sqrt{y})^2 \leq (\sqrt{x} - \sqrt{y})(\sqrt{x} + \sqrt{y}) = x - y = |x - y|,$$

where the inequality used $\sqrt{x} - \sqrt{y} \geq 0$ and $\sqrt{x} + \sqrt{y} \geq \sqrt{x} - \sqrt{y}$. Taking square roots gives the claim. Consequently, given $\varepsilon > 0$ we may take $\delta := \varepsilon^2$: whenever $|x - y| < \delta$ we get $|f(x) - f(y)| \leq \sqrt{|x - y|} < \varepsilon$. So f is uniformly continuous.
On the other hand f is not Lipschitz continuous. Suppose it were, with constant $C > 0$. Testing the Lipschitz bound at the pair $x = 1/n$, $x' = 0$ for a positive integer n gives

$$\frac{1}{\sqrt{n}} = \left| f\left(\tfrac{1}{n}\right) - f(0) \right| \leq C \left| \tfrac{1}{n} - 0 \right| = \frac{C}{n},$$

i.e. $\sqrt{n} \leq C$ for every $n \geq 1$, which is absurd. Hence no Lipschitz constant exists, and $f(x) = \sqrt{x}$ is uniformly continuous but not Lipschitz continuous.

Problem 2.2.2. (Total differentiability implies directional and partial differentiability) Show that if $f : \mathbb{R}^d \to \mathbb{R}$ is totally differentiable at x_0, then it is directionally differentiable at x_0 in each direction $v \in \mathbb{R}^d$, and one has the formula

$$D_v f(x_0) = v \cdot \nabla f(x_0). \tag{2.2}$$

In particular, the partial derivatives $\frac{\partial f}{\partial x_i}(x_0)$ exist for $i = 1, \ldots, d$ and

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_d}(x_0) \right). \tag{2.3}$$

Solution. By Definition 2.2.2, total differentiability of f at x_0 means that there is a vector $\nabla f(x_0) \in \mathbb{R}^d$ such that

$$\lim_{k \to 0;\, k \in \mathbb{R}^d \setminus \{0\}} \frac{f(x_0 + k) - f(x_0) - k \cdot \nabla f(x_0)}{|k|} = 0.$$

It is convenient to name the error quotient: for $k \in \mathbb{R}^d \setminus \{0\}$ put

$$\varepsilon(k) := \frac{f(x_0 + k) - f(x_0) - k \cdot \nabla f(x_0)}{|k|},$$

so that total differentiability says precisely that $\varepsilon(k) \to 0$ as $k \to 0$ with $k \neq 0$, and that

$$f(x_0 + k) - f(x_0) = k \cdot \nabla f(x_0) + |k|\, \varepsilon(k) \qquad (k \neq 0).$$

Fix a direction $v \in \mathbb{R}^d$. We must show that the limit

$$D_v f(x_0) = \lim_{h \to 0;\, h \in \mathbb{R} \setminus \{0\}} \frac{f(x_0 + hv) - f(x_0)}{h}$$

exists and equals $v \cdot \nabla f(x_0)$.
Consider first the degenerate case $v = 0$. Then $f(x_0 + hv) - f(x_0) = 0$ for every h, so the difference quotient is identically 0 and the limit exists and equals $0 = 0 \cdot \nabla f(x_0)$, which is (2.2) in this case.
Now suppose $v \neq 0$. For $h \in \mathbb{R} \setminus \{0\}$ the vector $k := hv$ is a nonzero element of $\mathbb{R}^d$, and $|k| = |h|\,|v|$. Substituting into the displayed identity above,

$$\frac{f(x_0 + hv) - f(x_0)}{h} = \frac{(hv) \cdot \nabla f(x_0) + |h|\,|v|\,\varepsilon(hv)}{h} = v \cdot \nabla f(x_0) + \frac{|h|}{h} |v|\, \varepsilon(hv).$$

Since $\left|\frac{|h|}{h}\right| = 1$, we obtain the bound

$$\left| \frac{f(x_0 + hv) - f(x_0)}{h} - v \cdot \nabla f(x_0) \right| = |v|\, |\varepsilon(hv)|.$$

As $h \to 0$ with $h \neq 0$ we have $hv \to 0$ in $\mathbb{R}^d$ with $hv \neq 0$ (here we use $v \neq 0$), so $\varepsilon(hv) \to 0$ and hence the right-hand side tends to 0. Explicitly: given $\eta > 0$, pick $\rho > 0$ such that $|\varepsilon(k)| < \eta/|v|$ whenever $0 < |k| < \rho$; then for $0 < |h| < \rho/|v|$ the above quantity is less than η. Therefore the limit defining $D_v f(x_0)$ exists and

$$D_v f(x_0) = v \cdot \nabla f(x_0),$$

which is (2.2).
Finally, specialise to $v = e_i$, the i-th standard basis vector, for $i = 1, \ldots, d$. By the above, $D_{e_i} f(x_0)$ exists, i.e. the partial derivative $\frac{\partial f}{\partial x_i}(x_0)$ exists in the sense of Definition 2.2.2, and

$$\frac{\partial f}{\partial x_i}(x_0) = D_{e_i} f(x_0) = e_i \cdot \nabla f(x_0),$$

which is exactly the i-th coordinate of the vector $\nabla f(x_0)$. Since this holds for every i, the vector $\nabla f(x_0)$ has coordinates $\left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_d}(x_0)\right)$, which is (2.3).
As a byproduct, (2.3) shows that the vector $\nabla f(x_0)$ appearing in Definition 2.2.2 is uniquely determined by f and x_0, so the notation $\nabla f(x_0)$ for the gradient is unambiguous.

Problem 2.2.3. (Continuous partial differentiability implies total differentiability) Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that the partial derivatives $\frac{\partial f}{\partial x_i} : \mathbb{R}^d \to \mathbb{R}$ exist everywhere and are continuous. Then show that f is totally differentiable everywhere, which in particular implies that the gradient is given by the formula (2.3) and the directional derivatives are given by (2.2).

Solution. Fix $x_0 \in \mathbb{R}^d$ and set

$$g := \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_d}(x_0) \right) \in \mathbb{R}^d.$$

We claim that f is totally differentiable at x_0 with $\nabla f(x_0) = g$; since x_0 was arbitrary this gives total differentiability everywhere, and then (2.3) and (2.2) follow from Exercise 2.2.2.
Let $\varepsilon > 0$. By hypothesis each $\frac{\partial f}{\partial x_i}$ is continuous at x_0, so we may pick $\delta > 0$ such that

$$\left| \frac{\partial f}{\partial x_i}(y) - \frac{\partial f}{\partial x_i}(x_0) \right| \le \varepsilon \qquad \text{for all } i = 1, \dots, d \text{ and all } y \text{ with } |y - x_0| \le \delta$$

(take the minimum of the d radii supplied by continuity of the individual partial derivatives).
Now let $h = (h_1, \dots, h_d) \in \mathbb{R}^d$ with $0 < |h| \le \delta$. We walk from x_0 to $x_0 + h$ one coordinate at a time. Define the intermediate points

$$x^{(0)} := x_0, \qquad x^{(i)} := x_0 + \sum_{j=1}^{i} h_j e_j \quad (i = 1, \dots, d),$$

so $x^{(d)} = x_0 + h$, and telescope:

$$f(x_0 + h) - f(x_0) = \sum_{i=1}^{d} \left[f(x^{(i)}) - f(x^{(i-1)}) \right].$$

Fix i and consider the segment joining $x^{(i-1)}$ to $x^{(i)} = x^{(i-1)} + h_i e_i$. Every point of this segment has the form

$$x_0 + \sum_{j<i} h_j e_j + t h_i e_i \qquad (t \in [0,1]),$$

whose distance from x_0 has square $\sum_{j<i} h_j^2 + t^2 h_i^2 \le \sum_{j=1}^{d} h_j^2 = |h|^2$; so the whole segment lies in the closed ball of radius $|h| \le \delta$ about x_0, where the above estimate on the partial derivatives is available.
If $h_i = 0$ then $f(x^{(i)}) - f(x^{(i-1)}) = 0 = h_i \frac{\partial f}{\partial x_i}(y_i)$ for any choice of point y_i on the segment, so set $y_i := x^{(i-1)}$. If $h_i \neq 0$, consider the one-variable function

$$\varphi_i(t) := f\big(x^{(i-1)} + t h_i e_i\big), \qquad t \in [0,1].$$

Since the partial derivative $\frac{\partial f}{\partial x_i}$ exists at every point of $\mathbb{R}^d$, the function $s \mapsto f(x^{(i-1)} + s e_i)$ is differentiable in s everywhere, and by the chain rule for the linear reparametrisation $s = t h_i$ we get that φ_i is differentiable on $[0,1]$ with

$$\varphi_i'(t) = h_i \frac{\partial f}{\partial x_i}\big(x^{(i-1)} + t h_i e_i\big).$$

In particular φ_i is continuous on $[0,1]$ and differentiable on $(0,1)$, so the mean value theorem supplies a $t_i \in (0,1)$ with $\varphi_i(1) - \varphi_i(0) = \varphi_i'(t_i)$, that is,

$$f(x^{(i)}) - f(x^{(i-1)}) = h_i \frac{\partial f}{\partial x_i}(y_i), \qquad y_i := x^{(i-1)} + t_i h_i e_i,$$

and y_i lies on the segment above, hence $|y_i - x_0| \le |h| \le \delta$.
Combining the telescoping identity with these d mean value identities and subtracting the linear approximation $h \cdot g = \sum_{i=1}^{d} h_i \frac{\partial f}{\partial x_i}(x_0)$, we obtain

$$f(x_0 + h) - f(x_0) - h \cdot g = \sum_{i=1}^{d} h_i \left[\frac{\partial f}{\partial x_i}(y_i) - \frac{\partial f}{\partial x_i}(x_0) \right].$$

Estimating each term using $|y_i - x_0| \le \delta$ and $|h_i| \le |h|$,

$$\big|f(x_0+h) - f(x_0) - h\cdot g\big| \le \sum_{i=1}^{d} |h_i|\,\varepsilon \le d\,\varepsilon\,|h|.$$

Dividing by $|h| > 0$,

$$\frac{\big|f(x_0+h) - f(x_0) - h\cdot g\big|}{|h|} \le d\,\varepsilon \qquad \text{whenever } 0 < |h| \le \delta.$$

Since $\varepsilon > 0$ was arbitrary (and d is a fixed constant), this says exactly that

$$\lim_{h\to 0;\, h\in\mathbb{R}^d\setminus\{0\}} \frac{f(x_0+h) - f(x_0) - h\cdot g}{|h|} = 0,$$

so f is totally differentiable at x_0 with $\nabla f(x_0) = g$. As $x_0 \in \mathbb{R}^d$ was arbitrary, f is totally differentiable everywhere. By Exercise 2.2.2 the gradient is then given by (2.3) (which is consistent with our choice of g) and all directional derivatives exist and satisfy $D_v f(x_0) = v \cdot \nabla f(x_0)$, which is (2.2).

Problem 2.2.4. (Directional differentiability does not imply total differentiability) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by setting $f(0,0) := 0$ and

$$f(x_1,x_2) := \frac{x_1 x_2^2}{x_1^2 + x_2^2} \qquad \text{for } (x_1,x_2) \in \mathbb{R}^2 \setminus \{(0,0)\}.$$

Show that the directional derivatives $D_v f(x)$ exist for all $x, v \in \mathbb{R}^2$ (so in particular, the partial derivatives exist), but that f is not totally differentiable at the origin $(0,0)$.

Solution. Existence of the directional derivatives away from the origin. Let $x \in \mathbb{R}^2 \setminus \{(0,0)\}$ and $v \in \mathbb{R}^2$. On the punctured plane f is the quotient of the polynomial $P(y) := y_1 y_2^2$ by the polynomial $Q(y) := y_1^2 + y_2^2$, and Q does not vanish there. Consider the one-variable function

$$t \mapsto f(x+tv) = \frac{P(x+tv)}{Q(x+tv)},$$

defined for all t in a neighbourhood of $t = 0$: indeed $Q(x+tv) = |x+tv|^2$ is a polynomial in t which equals $|x|^2 > 0$ at $t = 0$, hence is nonzero for all t in some interval $(-\rho, \rho)$. Numerator and denominator are polynomials in t, so the quotient is differentiable at $t = 0$ by the quotient rule for functions of one real variable. Its derivative at $t = 0$ is precisely

$$D_v f(x) = \lim_{h\to 0;\, h\in\mathbb{R}\setminus\{0\}} \frac{f(x+hv) - f(x)}{h},$$

so this directional derivative exists.
Existence of the directional derivatives at the origin. Note first that f is homogeneous of degree 1: for any $\lambda \in \mathbb{R}$ and $y \in \mathbb{R}^2 \setminus \{0\}$ with $\lambda \neq 0$,

$$f(\lambda y) = \frac{(\lambda y_1)(\lambda y_2)^2}{(\lambda y_1)^2 + (\lambda y_2)^2} = \frac{\lambda^3 y_1 y_2^2}{\lambda^2(y_1^2 + y_2^2)} = \lambda f(y),$$

and $f(\lambda y) = 0 = \lambda f(y)$ trivially when $\lambda = 0$ or $y = 0$. Hence for $v \in \mathbb{R}^2$ and $h \in \mathbb{R} \setminus \{0\}$,

$$\frac{f((0,0)+hv) - f(0,0)}{h} = \frac{f(hv)}{h} = \frac{h f(v)}{h} = f(v),$$

a quantity independent of h. So the limit as $h \to 0$ exists trivially, and

$$D_v f(0,0) = f(v) = \frac{v_1 v_2^2}{v_1^2 + v_2^2} \quad (v \neq 0), \qquad D_0 f(0,0) = 0.$$

Thus $D_v f(x)$ exists for all $x, v \in \mathbb{R}^2$; taking $v = e_1, e_2$ shows in particular that both partial derivatives exist everywhere.
Failure of total differentiability at the origin. Taking $v = e_1 = (1,0)$ and $v = e_2 = (0,1)$ in the formula just derived gives

$$\frac{\partial f}{\partial x_1}(0,0) = f(1,0) = \frac{1 \cdot 0}{1} = 0, \qquad \frac{\partial f}{\partial x_2}(0,0) = f(0,1) = \frac{0 \cdot 1}{1} = 0.$$

Suppose, for contradiction, that f were totally differentiable at the origin, with gradient $\nabla f(0,0)$. By Exercise 2.2.2, formula (2.3) would force

$$\nabla f(0,0) = \left(\frac{\partial f}{\partial x_1}(0,0), \frac{\partial f}{\partial x_2}(0,0) \right) = (0,0),$$

and formula (2.2) would then force $D_v f(0,0) = v \cdot (0,0) = 0$ for every $v \in \mathbb{R}^2$. But taking $v = (1,1)$ we computed

$$D_{(1,1)} f(0,0) = f(1,1) = \frac{1 \cdot 1}{1+1} = \frac{1}{2} \neq 0,$$

a contradiction. Hence f is not totally differentiable at $(0,0)$.
The obstruction can also be seen directly from Definition 2.2.2 without invoking Exercise 2.2.2: whatever candidate gradient $w \in \mathbb{R}^2$ one proposes, the difference quotient

$$\frac{f(h) - f(0,0) - h \cdot w}{|h|}$$

does not tend to 0. Indeed, along the ray $h = (t,t)$ with $t \to 0^+$ this quotient equals

$$\frac{\frac{t}{2} - t(w_1 + w_2)}{t\sqrt{2}} = \frac{\frac{1}{2} - (w_1 + w_2)}{\sqrt{2}},$$

while along the ray $h = (t,0)$ with $t \to 0^+$ it equals $-w_1$, and along $h = (0,t)$, $t \to 0^+$, it equals $-w_2$. For all three limits to vanish we would need $w_1 = w_2 = 0$ and simultaneously $w_1 + w_2 = 1/2$, which is impossible. This confirms directly that no gradient vector exists at the origin, even though every directional derivative there does; the failure is precisely that $v \mapsto D_v f(0,0) = f(v)$ is not a linear function of v.

Problem 2.2.5. Let $F : \mathbb{R}^d \to \mathbb{R}$ be a locally integrable function with the property that $\int_{\mathbb{R}^d} F(x) g(x)\, dx = 0$ whenever g is a compactly supported, continuously differentiable function. Show that F is zero almost everywhere. (Hint: if not, use the Lebesgue differentiation theorem to find a Lebesgue point x_0 of F for which $F(x_0) \neq 0$, then pick a g which is supported in a sufficiently small neighbourhood of x_0.)

Solution. We first record that the pairing in the hypothesis makes sense. If g is continuously differentiable and supported in a compact set, choose $R > 0$ with $\operatorname{supp}(g) \subset B(0,R)$. Then g is bounded (a continuous function on a compact set), say $|g| \leq M$, and $|Fg| \leq M|F| 1_{B(0,R)}$, which is absolutely integrable because F is locally integrable: by Exercise 1.6.14(i), local integrability is equivalent to $\int_{B(0,R)} |F(x)|\, dx < \infty$ for every $R > 0$. So $\int_{\mathbb{R}^d} Fg$ is an absolutely convergent integral.
Note also that F is Lebesgue measurable: local integrability asserts in particular that F is absolutely integrable on a neighbourhood of each point, and absolute integrability includes measurability, so F is measurable on a neighbourhood of each point and hence (covering $\mathbb{R}^d$ by countably many balls $B(0,n)$) measurable on all of $\mathbb{R}^d$. Consequently the set $A := \{x \in \mathbb{R}^d : F(x) \neq 0\}$ is Lebesgue measurable.
Suppose for contradiction that F is not zero almost everywhere, i.e. that $m(A) > 0$. By the Lebesgue differentiation theorem (Theorem 1.6.19), extended to locally integrable functions by Exercise 1.6.14(ii),

the set N of points of $\mathbb{R}^d$ which are not Lebesgue points of F is a null set, i.e.

$$\lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)}|F(y)-F(x)|\,dy=0$$

for every $x\notin N$. Since $m(A\setminus N)=m(A)>0$, the set $A\setminus N$ is nonempty; pick $x_0\in A\setminus N$. Thus x_0 is a Lebesgue point of F and $F(x_0)\neq 0$.
Replacing F by $-F$ if necessary (which changes neither the hypothesis, since $\int(-F)g=-\int Fg=0$, nor the conclusion), we may assume $c:=F(x_0)>0$.
Because x_0 is a Lebesgue point, we may choose $r>0$ so small that the mean oscillation bound

$$\frac{1}{m(B(x_0,r))}\int_{B(x_0,r)}|F(y)-F(x_0)|\,dy<\frac{c}{2^{d+1}}$$

holds; call this the oscillation bound, and fix such an r for the rest of the argument.
Next we build a suitable test function. Define $\varphi:\mathbb{R}\to[0,1]$ by

$$\varphi(s):=\begin{cases}1, & s\le\frac14,\\ \cos^2\!\big(\frac{2\pi}{3}\,(s-\frac14)\big), & \frac14\le s\le 1,\\ 0, & s\ge 1.\end{cases}$$

This is well defined and continuous: at $s=1/4$ the middle formula gives $\cos^2 0=1$, and at $s=1$ it gives $\cos^2(\pi/2)=0$. It is moreover continuously differentiable: on the middle interval

$$\varphi'(s)=-\frac{2\pi}{3}\sin\!\big(\tfrac{4\pi}{3}\,(s-\tfrac14)\big),$$

which vanishes at both $s=1/4$ (as $\sin 0=0$) and $s=1$ (as $\sin\pi=0$), matching the derivative 0 of the constant pieces on either side; so φ' exists everywhere and is continuous. Also $0\le\varphi\le 1$.
Now set

$$g(x):=\varphi\!\left(\frac{|x-x_0|^2}{r^2}\right),\qquad x\in\mathbb{R}^d.$$

The map $x\mapsto|x-x_0|^2/r^2$ is a polynomial, hence continuously differentiable, so by the chain rule g is continuously differentiable on $\mathbb{R}^d$. Moreover:

$$0\le g\le 1;\qquad g(x)=1 \text{ when } |x-x_0|\le r/2;\qquad g(x)=0 \text{ when } |x-x_0|\ge r.$$

In particular g vanishes identically outside the open ball $B(x_0,r)$ (note $\varphi(s)=0$ already at $s=1$, so g vanishes on the sphere $|x-x_0|=r$ as well), so the support of g is contained in the closed ball of radius r about x_0, which is compact. Thus g is an admissible test function for the hypothesis, and therefore

$$\int_{\mathbb{R}^d}F(x)g(x)\,dx=0.$$

We now estimate this integral and derive a contradiction. Write

$$0=\int_{\mathbb{R}^d}Fg=F(x_0)\int_{\mathbb{R}^d}g(x)\,dx+\int_{\mathbb{R}^d}\big(F(x)-F(x_0)\big)g(x)\,dx,$$

which is legitimate since both integrals on the right are absolutely convergent (g is bounded with compact support and F is locally integrable; the constant $F(x_0)$ times g is certainly integrable).
For the first term, since $g\ge 0$ everywhere and $g=1$ on $B(x_0,r/2)$, monotonicity of the integral gives

$$\int_{\mathbb{R}^d}g(x)\,dx\ge m\big(B(x_0,r/2)\big)=2^{-d}m\big(B(x_0,r)\big),$$

using the scaling property of Lebesgue measure. As $c=F(x_0)>0$, the first term is at least $c\,2^{-d}m(B(x_0,r))$.

For the second term, since $|g| \le 1$ and g vanishes outside $B(x_0, r)$, the triangle inequality and the oscillation bound give

$$\left| \int_{\mathbb{R}^d} \big(F(x) - F(x_0)\big) g(x)\, dx \right| \le \int_{B(x_0,r)} |F(x) - F(x_0)|\, dx < \frac{c}{2^{d+1}}\, m\big(B(x_0, r)\big).$$

Combining, and writing $V := m(B(x_0, r))$, which satisfies $0 < V < \infty$,

$$0 = \int_{\mathbb{R}^d} Fg = F(x_0) \int_{\mathbb{R}^d} g + \int_{\mathbb{R}^d} \big(F - F(x_0)\big) g \ge c\, 2^{-d} V - \left| \int_{\mathbb{R}^d} \big(F - F(x_0)\big) g \right| > c\, 2^{-d} V - c\, 2^{-d-1} V = c\, 2^{-d-1} V > 0,$$

which is absurd. (Here g vanishing outside $B(x_0, r)$ was used to confine both estimates to that ball, and the factor 2^{-d} from the inner ball is exactly what forced the choice of threshold $c/2^{d+1}$ in the oscillation bound.)
This contradiction shows that $m(A) = 0$, i.e. $F(x) = 0$ for almost every $x \in \mathbb{R}^d$, as required.

3.2 §2.3 — Probability Spaces

Problem 2.3.1. (No translation-invariant random integer). Show that there is no probability measure $\mathbf{P}$ on the integers $\mathbb{Z}$ with the discrete σ-algebra $2^{\mathbb{Z}}$ with the translation-invariance property $\mathbf{P}(E + n) = \mathbf{P}(E)$ for every event $E \in 2^{\mathbb{Z}}$ and every integer n.

Solution. Suppose for contradiction that $\mathbf{P}$ is a probability measure on the measurable space $(\mathbb{Z}, 2^{\mathbb{Z}})$ obeying $\mathbf{P}(E + n) = \mathbf{P}(E)$ for all $E \subset \mathbb{Z}$ and all $n \in \mathbb{Z}$. Since the σ-algebra is the discrete one, every subset of $\mathbb{Z}$ is an event, so all the quantities below are defined.
Set

$$p := \mathbf{P}(\{0\}) \in [0, 1].$$

For any integer n we have the set identity $\{n\} = \{0\} + n$, and hence by translation-invariance

$$\mathbf{P}(\{n\}) = \mathbf{P}(\{0\} + n) = \mathbf{P}(\{0\}) = p.$$

Thus every singleton has the same probability p.
Now enumerate the integers as $\mathbb{Z} = \{n_1, n_2, n_3, \dots\}$ with the n_k distinct (possible since $\mathbb{Z}$ is countably infinite; for instance $0, 1, -1, 2, -2, \dots$). The singletons $\{n_k\}$ are pairwise disjoint and their union is all of $\mathbb{Z}$, so countable additivity of the measure $\mathbf{P}$ gives

$$1 = \mathbf{P}(\mathbb{Z}) = \sum_{k=1}^{\infty} \mathbf{P}(\{n_k\}) = \sum_{k=1}^{\infty} p.$$

(The series on the right is a series of unsigned terms, so it converges in $[0, +\infty]$ and its value does not depend on the enumeration chosen.)
We now examine the two possible cases for p, and find both to be absurd.
If $p = 0$, then every partial sum $\sum_{k=1}^{K} p$ equals 0, hence $\sum_{k=1}^{\infty} p = 0 \ne 1$.
If $p > 0$, then the partial sums are $\sum_{k=1}^{K} p = Kp$, which tend to $+\infty$ as $K \to \infty$, hence $\sum_{k=1}^{\infty} p = +\infty \ne 1$.
In either case we contradict the identity $\sum_{k=1}^{\infty} p = 1$ derived above. Since p must be either zero or positive, no such $\mathbf{P}$ can exist, which is the claim.
Two remarks are worth making about where the hypotheses were used. First, countable additivity is essential and not merely convenient: a finitely additive translation-invariant probability measure on $2^{\mathbb{Z}}$ does exist (a Banach mean, whose construction requires a choice principle; compare the finitely additive translation-invariant extensions of Lebesgue measure mentioned in Section 1.2), so the obstruction found above is genuinely a countable phenomenon. Second, the argument used only that $\mathbb{Z}$ is a countably infinite set on which the translations act transitively; the same proof shows that no countably infinite set admits a probability measure on its discrete σ-algebra which is invariant under a transitive group of bijections, and indeed that no probability measure on a countably infinite discrete space can

give all points equal probability. This is exactly the statement that there is no way to choose an integer uniformly at random.

Problem 2.3.2. (No translation-invariant random real). Show that there is no probability measure $\mathbf{P}$ on the reals $\mathbb{R}$ with the Lebesgue σ-algebra $\mathcal{L}[\mathbb{R}]$ with the translation-invariance property $\mathbf{P}(E+x) = \mathbf{P}(E)$ for every event $E \in \mathcal{L}[\mathbb{R}]$ and every real x.

Solution. Suppose for contradiction that $\mathbf{P}$ is a probability measure on $(\mathbb{R}, \mathcal{L}[\mathbb{R}])$ with $\mathbf{P}(E+x) = \mathbf{P}(E)$ for every Lebesgue measurable E and every real x. (Implicit in the hypothesis is the fact, established in Exercise 1.2.20, that $E + x$ is Lebesgue measurable whenever E is; so the left-hand side makes sense.) Rather than work with singletons, which now carry no useful information, we chop the line into a countable family of congruent pieces. For each integer n put

$$I_n := [n, n+1).$$

Each I_n is an interval, hence Borel, hence lies in $\mathcal{L}[\mathbb{R}]$ and is a legitimate event. These sets are pairwise disjoint: if $m \neq n$, say $m < n$, then every element of I_m is less than $m + 1 \leq n$ while every element of I_n is at least n, so $I_m \cap I_n = \emptyset$. They also cover the line: given $x \in \mathbb{R}$, the integer $n := \lfloor x \rfloor$ satisfies $n \leq x < n + 1$, so $x \in I_n$. Thus

$$\mathbb{R} = \biguplus_{n \in \mathbb{Z}} I_n$$

is a partition of $\mathbb{R}$ into countably many events.
Moreover $I_n = I_0 + n$ for every integer n, since $[0, 1) + n = [n, n + 1)$. Hence, by the assumed translation-invariance applied with the real number $x = n$,

$$\mathbf{P}(I_n) = \mathbf{P}(I_0 + n) = \mathbf{P}(I_0) =: c$$

for every $n \in \mathbb{Z}$, where $c \in [0, 1]$ is a single fixed number.
Enumerate $\mathbb{Z}$ as $n_1, n_2, n_3, \ldots$ with the n_k distinct. By countable additivity applied to the disjoint events $I_{n_1}, I_{n_2}, \ldots$ whose union is $\mathbb{R}$,

$$1 = \mathbf{P}(\mathbb{R}) = \sum_{k=1}^{\infty} \mathbf{P}(I_{n_k}) = \sum_{k=1}^{\infty} c,$$

the series being one of unsigned terms and so having a well-defined value in $[0, +\infty]$ independent of the enumeration.
If $c = 0$ then every partial sum vanishes and the series equals 0; if $c > 0$ then the K-th partial sum is $Kc \to +\infty$ and the series equals $+\infty$. Neither value is 1, so in both cases we have a contradiction. Since c is either zero or positive, no such $\mathbf{P}$ exists, as required.
Some remarks. First, note that the argument never used any property of Lebesgue measure beyond the fact that intervals are measurable; the same proof shows there is no translation-invariant probability measure on the Borel σ-algebra $\mathcal{B}[\mathbb{R}]$, nor on any translation-invariant σ-algebra containing the intervals $[n, n+1)$. In particular no appeal to a non-measurable (Vitali) set is needed here: the obstruction is the pure fact that the line is a disjoint union of countably many translates of a single set, and a probability measure cannot assign a common value to countably infinitely many disjoint events summing to 1.
Second, the failure is one of normalisation rather than of existence: Lebesgue measure m itself is a translation-invariant measure on $(\mathbb{R}, \mathcal{L}[\mathbb{R}])$, but it has $m(\mathbb{R}) = +\infty$, so the recipe of Example 2.3.2 for normalising a measure to a probability measure is unavailable. This exercise says that no other normalisation can succeed either: there is no uniform distribution on the whole real line. The same argument applies verbatim in $\mathbb{R}^d$, chopping $\mathbb{R}^d$ into the countably many unit cubes $[0, 1)^d + n$, $n \in \mathbb{Z}^d$.

3.3 §2.4 — Infinite Product Spaces and the Kolmogorov Extension Theorem

Problem 2.4.1. Let $((X_\alpha, \mathcal{B}_\alpha))_{\alpha\in A}$ be a family of measurable spaces. For any $B \subset A$, write $\mathcal{B}_B := \prod_{\beta\in B} \mathcal{B}_\beta$.

- (1) Show that $\mathcal{B}_A$ is the coarsest σ-algebra on X_A that makes the projection maps π_β measurable morphisms for all $\beta \in A$.
- (2) Show that for each $B \subset A$, that π_B is a measurable morphism from $(X_A, \mathcal{B}_A)$ to $(X_B, \mathcal{B}_B)$.
- (3) If $E_A \in \mathcal{B}_A$, show that there exists an at most countable set $B \subset A$ and a set $E_B \in \mathcal{B}_B$ such that $E_A = \pi_B^{-1}(E_B)$. Informally, this asserts that a measurable event can only depend on at most countably many of the coefficients.
- (4) If $f : X_A \to [0, +\infty]$ is $\mathcal{B}_A$-measurable, show that there exists an at most countable set $B \subset A$ and a $\mathcal{B}_B$-measurable function $f_B : X_B \to [0, +\infty]$ such that $f = f_B \circ \pi_B$.
- (5) If A is at most countable, show that $\mathcal{B}_A$ is the σ-algebra generated by the sets $\prod_{\beta\in A} E_\beta$ with $E_\beta \in \mathcal{B}_\beta$ for all $\beta \in A$.
- (6) On the other hand, show that if A is uncountable and the $\mathcal{B}_\alpha$ are all non-trivial, then $\mathcal{B}_A$ is not the σ-algebra generated by the sets $\prod_{\beta\in A} E_\beta$ with $E_\beta \in \mathcal{B}_\beta$ for all $\beta \in A$.
- (7) If $B \subset A$, $E \in \mathcal{B}_A$, and $x_{A\setminus B} \in X_{A\setminus B}$, show that the set $E_{x_{A\setminus B},B} := \{x_B \in X_B : (x_B, x_{A\setminus B}) \in E\}$ lies in $\mathcal{B}_B$, where we identify $X_B \times X_{A\setminus B}$ with X_A in the obvious manner.
- (8) If $B \subset A$, $f : X_A \to [0, +\infty]$ is $\mathcal{B}_A$-measurable, and $x_{A\setminus B} \in X_{A\setminus B}$, show that the function $f_{x_{A\setminus B},B} : x_B \mapsto f(x_B, x_{A\setminus B})$ is $\mathcal{B}_B$-measurable.

Solution. Notation and two preliminary lemmas.

For $C \subset B \subset A$ we write $\pi_{C\leftarrow B} : X_B \to X_C$ for the partial subprojection $(x_\alpha)_{\alpha\in B} \mapsto (x_\alpha)_{\alpha\in C}$, so that $\pi_B = \pi_{B\leftarrow A}$ and $\pi_\beta = \pi_{\{\beta\}\leftarrow A}$ (identifying $X_{\{\beta\}}$ with X_β), and we recall the composition law $\pi_{D\leftarrow C} \circ \pi_{C\leftarrow B} = \pi_{D\leftarrow B}$ for $D \subset C \subset B \subset A$. By definition,

$$\mathcal{B}_B \;=\; \Big\langle \bigcup_{\beta\in B} \pi^*_{\beta\leftarrow B}(\mathcal{B}_\beta) \Big\rangle, \qquad \pi^*_{\beta\leftarrow B}(\mathcal{B}_\beta) := \{\pi^{-1}_{\beta\leftarrow B}(E_\beta) : E_\beta \in \mathcal{B}_\beta\},$$

where $\langle\mathcal{F}\rangle$ denotes the σ-algebra generated by $\mathcal{F}$, i.e. the coarsest σ-algebra containing $\mathcal{F}$ (Exercise 1.4.13 and the discussion following it). Note that the definition of $\mathcal{B}_B$ refers only to the subfamily $((X_\beta, \mathcal{B}_\beta))_{\beta\in B}$, so every statement proved below for the index set A applies verbatim with A replaced by any subset B.

Lemma A (pullback of a σ-algebra). If $\varphi : X \to Y$ is any map and $\mathcal{C}$ is a σ-algebra on Y, then $\varphi^*(\mathcal{C}) := \{\varphi^{-1}(F) : F \in \mathcal{C}\}$ is a σ-algebra on X. Indeed $\varphi^{-1}(\emptyset) = \emptyset$, $\varphi^{-1}(Y \setminus F) = X \setminus \varphi^{-1}(F)$, and $\varphi^{-1}(\bigcup_n F_n) = \bigcup_n \varphi^{-1}(F_n)$, so the three axioms for $\mathcal{C}$ transfer to $\varphi^*(\mathcal{C})$.

Lemma B (generating criterion for measurability). Let $\varphi : X \to Y$ be a map, let $\mathcal{B}$ be a σ-algebra on X, and let $\mathcal{F}$ be a collection of subsets of Y with $\varphi^{-1}(F) \in \mathcal{B}$ for every $F \in \mathcal{F}$. Then φ is a measurable morphism from $(X, \mathcal{B})$ to $(Y, \langle\mathcal{F}\rangle)$. Indeed, put $\mathcal{G} := \{F \subset Y : \varphi^{-1}(F) \in \mathcal{B}\}$. By the same three identities as in Lemma A, $\mathcal{G}$ is a σ-algebra on Y; by hypothesis it contains $\mathcal{F}$, hence it contains $\langle\mathcal{F}\rangle$, which is precisely the assertion that $\varphi^{-1}(F) \in \mathcal{B}$ for all $F \in \langle\mathcal{F}\rangle$.

Part (1).

First, $\mathcal{B}_A$ does make every π_β a measurable morphism from $(X_A, \mathcal{B}_A)$ to $(X_\beta, \mathcal{B}_\beta)$: for $E_\beta \in \mathcal{B}_\beta$ we have $\pi_\beta^{-1}(E_\beta) \in \pi^*_\beta(\mathcal{B}_\beta) \subset \bigcup_{\gamma\in A} \pi^*_\gamma(\mathcal{B}_\gamma) \subset \mathcal{B}_A$.

Conversely, suppose $\mathcal{C}$ is any σ-algebra on X_A making every π_β a measurable morphism. Then for every $\beta \in A$ and every $E_\beta \in \mathcal{B}_\beta$ we have $\pi_\beta^{-1}(E_\beta) \in \mathcal{C}$, i.e. $\pi^*_\beta(\mathcal{B}_\beta) \subset \mathcal{C}$. Hence $\bigcup_{\beta\in A} \pi^*_\beta(\mathcal{B}_\beta) \subset \mathcal{C}$, and since $\mathcal{B}_A$ is the coarsest σ-algebra containing that union, $\mathcal{B}_A \subset \mathcal{C}$. Thus $\mathcal{B}_A$ is contained in every σ-algebra with the stated property and has that property itself, i.e. it is the coarsest such σ-algebra. (Recall from Remark 1.4.33 that $\mathcal{B}_A \subset \mathcal{C}$ is exactly the statement that $\mathcal{B}_A$ is coarser than $\mathcal{C}$.)

Part (2).

By definition $\mathcal{B}_B = \langle\mathcal{F}\rangle$ with $\mathcal{F} := \bigcup_{\beta\in B} \pi^*_{\beta\leftarrow B}(\mathcal{B}_\beta)$, so by Lemma B it suffices to show $\pi_B^{-1}(F) \in \mathcal{B}_A$ for every $F \in \mathcal{F}$. Such an F has the form $F = \pi^{-1}_{\beta\leftarrow B}(E_\beta)$ with $\beta \in B$ and $E_\beta \in \mathcal{B}_\beta$, and by the composition law $\pi_{\beta\leftarrow B} \circ \pi_{B\leftarrow A} = \pi_{\beta\leftarrow A} = \pi_\beta$ we get

$$\pi_B^{-1}(F) = \pi_B^{-1}\big(\pi^{-1}_{\beta\leftarrow B}(E_\beta)\big) = (\pi_{\beta\leftarrow B} \circ \pi_B)^{-1}(E_\beta) = \pi_\beta^{-1}(E_\beta) \in \mathcal{B}_A$$

by Part (1). Hence π_B is a measurable morphism from $(X_A, \mathcal{B}_A)$ to $(X_B, \mathcal{B}_B)$. Applying this statement to the index set B in place of A, we also record the more general fact, referred to below as the transfer property: whenever $C \subset B \subset A$, the map $\pi_{C \leftarrow B}$ is a measurable morphism from $(X_B, \mathcal{B}_B)$ to $(X_C, \mathcal{B}_C)$.
Part (3).
Let

$$\mathcal{C} := \{E \subset X_A : E = \pi_B^{-1}(E_B) \text{ for some at most countable } B \subset A \text{ and some } E_B \in \mathcal{B}_B\}.$$

We claim $\mathcal{C}$ is a σ-algebra on X_A containing $\bigcup_{\beta \in A} \pi_\beta^*(\mathcal{B}_\beta)$.
It contains the generators: given $\beta \in A$ and $E_\beta \in \mathcal{B}_\beta$, take $B := \{\beta\}$, which is countable; under the identification of $X_{\{\beta\}}$ with X_β the map $\pi_{\beta \leftarrow \{\beta\}}$ is the identity, so $\mathcal{B}_{\{\beta\}} = \langle \mathcal{B}_\beta \rangle = \mathcal{B}_\beta$, and $\pi_\beta^{-1}(E_\beta) = \pi_{\{\beta\}}^{-1}(E_\beta)$ with $E_\beta \in \mathcal{B}_{\{\beta\}}$.
It contains $\emptyset$ and is closed under complements: if $E = \pi_B^{-1}(E_B)$ with B countable and $E_B \in \mathcal{B}_B$, then $X_A \setminus E = \pi_B^{-1}(X_B \setminus E_B)$ and $X_B \setminus E_B \in \mathcal{B}_B$; and $\emptyset = \pi_B^{-1}(\emptyset)$ for any such B (e.g. $B = \emptyset$, for which $X_\emptyset$ is the one-point set consisting of the empty tuple and $\mathcal{B}_\emptyset = \{\emptyset, X_\emptyset\}$).
It is closed under countable unions: suppose $E_n = \pi_{B_n}^{-1}(F_n)$ with $B_n \subset A$ at most countable and $F_n \in \mathcal{B}_{B_n}$, for $n = 1, 2, \ldots$. Put $B := \bigcup_n B_n$, an at most countable subset of A. By the transfer property from Part (2), $\pi_{B_n \leftarrow B}$ is a measurable morphism from $(X_B, \mathcal{B}_B)$ to $(X_{B_n}, \mathcal{B}_{B_n})$, so $G_n := \pi_{B_n \leftarrow B}^{-1}(F_n) \in \mathcal{B}_B$; and by the composition law $\pi_{B_n \leftarrow B} \circ \pi_B = \pi_{B_n}$ we get $\pi_B^{-1}(G_n) = \pi_{B_n}^{-1}(F_n) = E_n$. Hence

$$\bigcup_n E_n = \bigcup_n \pi_B^{-1}(G_n) = \pi_B^{-1}\Big(\bigcup_n G_n\Big),$$

and $\bigcup_n G_n \in \mathcal{B}_B$, so $\bigcup_n E_n \in \mathcal{C}$.
Thus $\mathcal{C}$ is a σ-algebra containing $\bigcup_{\beta \in A} \pi_\beta^*(\mathcal{B}_\beta)$, whence $\mathcal{B}_A \subset \mathcal{C}$ by minimality of the generated σ-algebra. This is exactly the assertion: every $E_A \in \mathcal{B}_A$ is of the form $\pi_B^{-1}(E_B)$ with B at most countable and $E_B \in \mathcal{B}_B$. (Conversely $\mathcal{C} \subset \mathcal{B}_A$ by Part (2), so in fact $\mathcal{C} = \mathcal{B}_A$.)
Part (4).
If $X_A = \emptyset$ the claim is vacuous: take $B := \emptyset$ and $f_B := 0$ on the one-point space $X_\emptyset$; then f and $f_B \circ \pi_B$ are both the empty function. So we may assume $X_A \neq \emptyset$, which forces every X_α to be non-empty, and hence (by the axiom of choice) $X_{A'} \neq \emptyset$ for every $A' \subset A$; in particular every $\pi_B : X_A \to X_B$ is surjective, since any $x_B \in X_B$ may be completed by an arbitrary point of $X_{A \setminus B}$.
By Exercise 1.4.29(i), f is $\mathcal{B}_A$-measurable if and only if $\{f > \lambda\} := \{x_A \in X_A : f(x_A) > \lambda\} \in \mathcal{B}_A$ for every $\lambda \in \mathbb{R}$ (equivalently, for every $\lambda \in [0, +\infty)$, since $\{f > \lambda\} = X_A$ for $\lambda < 0$). For each rational $q \geq 0$, Part (3) supplies an at most countable $B_q \subset A$ and a set $F_q \in \mathcal{B}_{B_q}$ with $\{f > q\} = \pi_{B_q}^{-1}(F_q)$. Set

$$B := \bigcup_{q \in \mathbb{Q},\, q \geq 0} B_q,$$

an at most countable subset of A. Exactly as in Part (3), $G_q := \pi_{B_q \leftarrow B}^{-1}(F_q) \in \mathcal{B}_B$ and $\{f > q\} = \pi_B^{-1}(G_q)$. Thus every set $\{f > q\}$, $q \in \mathbb{Q}$, $q \geq 0$, lies in the σ-algebra $\mathcal{D} := \pi_B^*(\mathcal{B}_B)$ on X_A (a σ-algebra by Lemma A). Since for arbitrary real $\lambda \geq 0$

$$\{f > \lambda\} = \bigcup_{q \in \mathbb{Q},\, q > \lambda} \{f > q\},$$

a countable union, and $\{f > \lambda\} = X_A \in \mathcal{D}$ for $\lambda < 0$, we conclude that f is $\mathcal{D}$-measurable, again by Exercise 1.4.29(i).
We now show that a $\mathcal{D}$-measurable f factors through π_B. First, f is constant on the fibres of π_B: suppose $x_A, x'_A \in X_A$ satisfy $\pi_B(x_A) = \pi_B(x'_A)$. For each λ write $\{f > \lambda\} = \pi_B^{-1}(H_\lambda)$ with $H_\lambda \in \mathcal{B}_B$; then $x_A \in \{f > \lambda\}$ if and only if $\pi_B(x_A) \in H_\lambda$, if and only if $\pi_B(x'_A) \in H_\lambda$, if and only if $x'_A \in \{f > \lambda\}$. Thus $f(x_A) > \lambda \iff f(x'_A) > \lambda$ for every real λ, which forces $f(x_A) = f(x'_A)$.
Since π_B is surjective, we may therefore define $f_B : X_B \to [0, +\infty]$ unambiguously by $f_B(x_B) := f(x_A)$ for any $x_A \in \pi_B^{-1}(\{x_B\})$; by construction $f = f_B \circ \pi_B$. Finally f_B is $\mathcal{B}_B$-measurable: for each real λ,

$$\pi_B^{-1}(\{f_B > \lambda\}) = \{x_A : f(x_A) > \lambda\} = \pi_B^{-1}(H_\lambda),$$

and since π_B is surjective the operation $S \mapsto \pi_B^{-1}(S)$ is injective on subsets of X_B (indeed $\pi_B(\pi_B^{-1}(S)) = S$ for surjective π_B), so $\{f_B > \lambda\} = H_\lambda \in \mathcal{B}_B$. By Exercise 1.4.29(i), f_B is measurable, as required.
Part (5).
Let $\mathcal{P} := \{\prod_{\beta\in A} E_\beta : E_\beta \in \mathcal{B}_\beta \text{ for all } \beta \in A\}$ and let $\mathcal{G} := \langle\mathcal{P}\rangle$. We show $\mathcal{G} = \mathcal{B}_A$ when A is at most countable.
$\mathcal{B}_A \subset \mathcal{G}$: for $\beta \in A$ and $E_\beta \in \mathcal{B}_\beta$, the set $\pi_\beta^{-1}(E_\beta)$ equals $\prod_{\gamma\in A} E_\gamma$ where $E_\gamma := X_\gamma$ for $\gamma \neq \beta$; since $X_\gamma \in \mathcal{B}_\gamma$, this lies in $\mathcal{P} \subset \mathcal{G}$. Hence $\bigcup_{\beta\in A} \pi_\beta^*(\mathcal{B}_\beta) \subset \mathcal{G}$ and so $\mathcal{B}_A \subset \mathcal{G}$. (This direction does not use countability of A.)
$\mathcal{G} \subset \mathcal{B}_A$: for $E_\beta \in \mathcal{B}_\beta$ ($\beta \in A$) we have the identity

$$\prod_{\beta\in A} E_\beta = \bigcap_{\beta\in A} \pi_\beta^{-1}(E_\beta),$$

since a tuple $(x_\alpha)_{\alpha\in A}$ lies in the left-hand side precisely when $x_\beta \in E_\beta$ for every β. As A is at most countable, this is an at most countable intersection of sets $\pi_\beta^{-1}(E_\beta) \in \mathcal{B}_A$, hence lies in the σ-algebra $\mathcal{B}_A$ (a σ-algebra is closed under countable intersections, by De Morgan). Thus $\mathcal{P} \subset \mathcal{B}_A$ and so $\mathcal{G} \subset \mathcal{B}_A$.
Part (6).
Now let A be uncountable, and suppose each $\mathcal{B}_\alpha$ is non-trivial, i.e. $\mathcal{B}_\alpha \neq \{\emptyset, X_\alpha\}$, so that for each $\alpha \in A$ we may fix $E_\alpha \in \mathcal{B}_\alpha$ with $E_\alpha \neq \emptyset$ and $E_\alpha \neq X_\alpha$. (In particular every X_α is non-empty, since a σ-algebra on the empty set is trivial.) With $\mathcal{P}, \mathcal{G}$ as in Part (5), we still have $\mathcal{B}_A \subset \mathcal{G}$ by the first half of Part (5); we show the inclusion is strict, by exhibiting an element of $\mathcal{P}$ that is not in $\mathcal{B}_A$.
Consider $E := \prod_{\alpha\in A} E_\alpha \in \mathcal{P} \subset \mathcal{G}$. By the axiom of choice we may pick $x_\alpha \in E_\alpha$ for every α, so $x := (x_\alpha)_{\alpha\in A} \in E$ and in particular $E \neq \emptyset$.
Suppose for contradiction that $E \in \mathcal{B}_A$. By Part (3) there is an at most countable $B \subset A$ and a set $E_B \in \mathcal{B}_B$ with $E = \pi_B^{-1}(E_B)$. Since A is uncountable and B is at most countable, we may pick $\gamma \in A \setminus B$. Since $E_\gamma \neq X_\gamma$, we may pick $y_\gamma \in X_\gamma \setminus E_\gamma$. Define $x' = (x'_\alpha)_{\alpha\in A}$ by $x'_\alpha := x_\alpha$ for $\alpha \neq \gamma$ and $x'_\gamma := y_\gamma$. As $\gamma \notin B$ we have $\pi_B(x') = \pi_B(x) \in E_B$, so $x' \in \pi_B^{-1}(E_B) = E$; but $x'_\gamma = y_\gamma \notin E_\gamma$, so $x' \notin \prod_{\alpha\in A} E_\alpha = E$, a contradiction.
Hence $E \in \mathcal{G} \setminus \mathcal{B}_A$, so $\mathcal{B}_A \subsetneq \mathcal{G}$ and in particular $\mathcal{B}_A$ is not the σ-algebra generated by the product sets $\prod_{\beta\in A} E_\beta$.
Part (7).
Fix $B \subset A$ and $x_{A\setminus B} \in X_{A\setminus B}$, and define the insertion map

$$\iota : X_B \to X_A, \qquad \iota(x_B) := (x_B, x_{A\setminus B}),$$

i.e. $\iota((x_\beta)_{\beta\in B})$ is the tuple in X_A whose β-coordinate is x_β for $\beta \in B$ and whose α-coordinate is the given x_α for $\alpha \in A \setminus B$. Then by definition

$$E_{x_{A\setminus B},B} = \iota^{-1}(E),$$

so it suffices to show that ι is a measurable morphism from $(X_B, \mathcal{B}_B)$ to $(X_A, \mathcal{B}_A)$. By Lemma B applied to the generating family $\bigcup_{\beta\in A} \pi_\beta^*(\mathcal{B}_\beta)$ of $\mathcal{B}_A$, it is enough to check that $\iota^{-1}(\pi_\beta^{-1}(E_\beta)) \in \mathcal{B}_B$ for every $\beta \in A$ and $E_\beta \in \mathcal{B}_\beta$. Now $\iota^{-1}(\pi_\beta^{-1}(E_\beta)) = (\pi_\beta \circ \iota)^{-1}(E_\beta)$, and there are two cases:
If $\beta \in B$, then $\pi_\beta \circ \iota = \pi_{\beta\leftarrow B}$ (both send x_B to its β-coordinate), so $(\pi_\beta \circ \iota)^{-1}(E_\beta) = \pi_{\beta\leftarrow B}^{-1}(E_\beta) \in \pi_{\beta\leftarrow B}^*(\mathcal{B}_\beta) \subset \mathcal{B}_B$.
If $\beta \in A \setminus B$, then $\pi_\beta \circ \iota$ is the constant map with value x_β, so $(\pi_\beta \circ \iota)^{-1}(E_\beta)$ is X_B if $x_\beta \in E_\beta$ and $\emptyset$ otherwise; either way it lies in $\mathcal{B}_B$.
Hence ι is a measurable morphism, and $E_{x_{A\setminus B},B} = \iota^{-1}(E) \in \mathcal{B}_B$ for every $E \in \mathcal{B}_A$.
Part (8).
With ι as in Part (7), we have $f_{x_{A\setminus B},B} = f \circ \iota$. For any real λ,

$$\{x_B \in X_B : f_{x_{A\setminus B},B}(x_B) > \lambda\} = \iota^{-1}\big(\{f > \lambda\}\big),$$

and $\{f > \lambda\} \in \mathcal{B}_A$ by the $\mathcal{B}_A$-measurability of f (Exercise 1.4.29(i)); hence this level set lies in $\mathcal{B}_B$ by Part (7) (i.e. by the measurability of the morphism ι). Since all level sets of $f_{x_{A\setminus B},B}$ are $\mathcal{B}_B$-measurable, Exercise 1.4.29(i) shows that $f_{x_{A\setminus B},B}$ is $\mathcal{B}_B$-measurable. (Equivalently: f is a measurable morphism into $[0, +\infty]$ with its Borel σ-algebra by Exercise 1.4.29(iii) and Remark 1.4.33, and a composition of measurable morphisms is a measurable morphism.)