Notes for Measure Theory 2nd edition, Donald Cohn

Lingfei Zhang, lfzhang@tamu.edu

July 2024

Introduction

For more details of Darboux's definition of the Riemann integral, see *Principles of Mathematical Analysis* by Walter Rudin, chapter 6.

For a slightly stronger version of the Fundamental Theorem of Calculus, see Rudin PMA, theorem 6.20.

For "Theorem 2 is valid for the Riemann integral...", see Rudin PMA, theorem 7.16.

After equation (2), "One can check this does not depend on the value of c...": Consider another number $s < c$ such that $f < s$ on $[a, b]$ (for $s > c$, switch the roles of s, c). Consider a partition $\mathcal{P}^c = \{a_i\}_{i=0}^k$ for $[0, c]$. Recall $a_0 = 0$. Let $j \in \{1, ..., k\}$ be the smallest index such that $s \leq a_j$, i.e., $j = \min\{i : a_i \geq s\}$. By defining $b_i = a_i, i = 0, ..., j-1$ and $b_j = s$. We see that all b_i thus defined form a partition $\mathcal{P}^s$ of $[0, s]$. Now $s(f, \mathcal{P}^s) = \sum_{i=1}^j b_{i-1}\text{meas}(B_i) = \sum_{i=1}^j a_{i-1}\text{meas}(B_i) = \sum_{i=1}^{j-1} a_{i-1}\text{meas}(A_i) + a_{j-1}\text{meas}(B_j)$ while $s(f, \mathcal{P}^c) = \sum_{i=1}^k a_{i-1}\text{meas}(A_i) = \sum_{i=1}^{j-1} a_{i-1}\text{meas}(A_i) + a_{j-1}\text{meas}(A_j) + \sum_{i=j+1}^k a_{i-1}\text{meas}(A_i)$. But notice $\sum_{i=j+1}^k a_{i-1}\text{meas}(A_i) = 0$ since no x is such that $f(x) \geq a_j$, so $s(f, \mathcal{P}^c) - s(f, \mathcal{P}^s) = a_{j-1}(\text{meas}(A_j) - \text{meas}(B_j))$. However, by definition (1), $A_j = B_j$. Therefore, $s(f, \mathcal{P}^c) = s(f, \mathcal{P}^s)$.

Now we've shown, for $s < c$, for any partition $\mathcal{P}^c$, we can find a partition $\mathcal{P}^s$ such that $s(f, \mathcal{P}^c) = s(f, \mathcal{P}^s)$, meaning the set of all possible values that $s(f, \mathcal{P}^c)$ can take is included by the set of all possible values that $s(f, \mathcal{P}^s)$ can take, implying $\sup_{\mathcal{P}^c} s(f, \mathcal{P}^c) \leq \sup_{\mathcal{P}^s} s(f, \mathcal{P}^s)$. Conversely, for any partition $\mathcal{P}^s = \{b_i\}_{i=1}^j$ of $[0, s]$, by simply including $b_{j+1} = c$ we have a partition $\mathcal{P}^c$ for $[0, c]$ with $s(f, \mathcal{P}^c) = s(f, \mathcal{P}^s)$, and by a similar argument we have $\sup_{\mathcal{P}^s} s(f, \mathcal{P}^s) \leq \sup_{\mathcal{P}^c} s(f, \mathcal{P}^c)$. Therefore, $\sup_{\mathcal{P}^s} s(f, \mathcal{P}^s) = \sup_{\mathcal{P}^c} s(f, \mathcal{P}^c)$ and hence "this does not depend on the value of c".

1 Measures

page 3

(d). We check for closure under finite unions. Suppose $A_1, ..., A_n \in \mathcal{A}$. Case 1: some A_i^c is finite, then $\cap_i A_i^c = (\cup_i A_i)^c$ is finite, so $\cup_i A_i \in \mathcal{A}$. Case 2: all A_i^c are infinite, which means all A_i must be finite, implying $\cup_i A_i$ is finite and so $\cup_i A_i \in \mathcal{A}$.

Note that $\mathcal{A}$ is not closed under countable unions. This is because we can divide X into two disjoint infinite sets X_1, X_2 (for countable X, we just put $x_1, x_3, ...$ into X_1 and $x_2, x_4, ...$ into X_2; for uncountable X we divide X' into such X_1', X_2' for a countable subset $X' \subset X$ and let one of X_1', X_2' be united with $X - X'$). Define $A_i = \{y_i\}, i \in \mathbb{N}$ for distinct $y_i \in X_1$. Then $\cup_i A_i$ is not finite, and neither is its complement as the complement includes X_2.

(f). To check for closure under countable union, recall the two cases in (d) above. Consider $A_1, ...$ instead of $A_1, ...A_n$; change "finite" to "countable".

page 4

For proposition 1.1.5, that $\mathcal{B}(\mathbb{R}^d) \subset \mathcal{B}_3$ is not proven. See theorem 1.23, *Probability Theory, Third Edition* by Achim Klenke for a more rigorous and comprehensive treatment regarding the Borel sigma algebra of $\mathbb{R}^d$ and its generators.

page 6

Fig 1.1. For example, we know from the proposition that $\mathcal{F} \subset \mathcal{G}_\delta$. So a countable intersection of sets in $\mathcal{F}$ is also a countable intersection of sets in $\mathcal{G}_\delta$, i.e., $\mathcal{F}_\sigma \subset \mathcal{G}_{\delta\sigma}$, etc.

page 8

(d). Let $A_1, ..., A_n \in \mathcal{A}$ be disjoint. We check for finite additivity. Case 1: some A_i are such that A_i^c is finite $\implies$ $A_i = \mathbb{N} - A_i^c$ is infinite. Note there can be only one such index i, otherwise, say A_j is also s.t. A_j^c is finite, then there exists $N_i, N_j \in \mathbb{N}$ such that all elements in A_i^c (respectively, A_j^c) are smaller than N_i (respectively, N_j). This means A_i^c (A_j^c) $\subset \{1, 2, ..., N_i\}$ ($\{1, 2, ..., N_j\}$) and that $\{N_i, N_i + 1, ...\} \subset A_i, \{N_j, N_j + 1, ...\} \subset A_j$, contradicting the fact that A_i, A_j are disjoint. So A_k^c is infinite for all indices $k \in \{1, ..., n\} - \{i\}$ meaning all those A_k are finite. With the fact that A_i is infinite, the countable additivity holds.

Case 2: all $A_i, i = 1, ..., n$ are such that A_i^c is infinite. Then by the definition of $\mathcal{A}$, all the A_i are finite, and countable additivity holds.

page 9

Line 11. On constructing B_i from A_i, note if we define $C_n = \cup_{i=1}^n A_i$, then $\cup_n C_n = \cup_i A_i = X$. It is obvious that we also have $C_n = \cup_{i=1}^n B_i$ and $X = \cup_n C_n = \cup_i B_i$.

page 14

Notice line 8, the first term. We need to justify that $\sum_j (b_j - a_j)$ makes sense. We consider an essentially the same example.

Let's say we have a nonnegative, doubly-indexed sequence (like a table extending to infinity in two directions) $(a_{ij})_{i,j\in\mathbb{N}}$, and assume $\sum_j a_{ij} = 1/2^i$.

By theorem 8.3, Rudin's PMA, we know $\sum_j \sum_i a_{ij} = \sum_i \sum_j a_{ij} = 1$. Rudin's theorem 3.55 also says, if series $\sum s_n$ converges absolutely, then any rearrangement of it converges to the same limit.

Now lets do the "diagonal trick" and convert the doubly-indexed sequence (a_{ij}) into just a normal sequence (a_n). Is there a way we can show that $\sum a_n = 1$?

Let $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ be a bijection. Denote the sequence $a_{f(n)}$ by b_n. We want to show $\sum_{n=1}^{\infty} b_n = 1$. Denote $f(n) = (x_n, y_n)$.

Let the partial sums be $p_m := \sum_{n=1}^{m} b_n$. It's easy to see each p_m is upper bounded by 1. (Otherwise, it would contradict the fact that $\sum_i \sum_j a_{ij} = 1$.)

Recall the value of a series whose terms are nonnegative is the supremum (also the limit, as the partial sums are increasing) of its partial sums. So we still have to show that 1 is the least upper bound of $\{p_m\}_{m=1}^{\infty}$.

Let $\mathbb{I}$ denote the indicator function. Pick $0 < \epsilon < 1$. There is an $N \in \mathbb{N}$ such that $\sum_{i=1}^{N} 1/2^i > 1 - \epsilon/2$. Also, there is an M such that if $m > M$, then $\{b_1, ..., b_m\}$ contains enough terms from $\{a_{i,j}\}_{j=1}^{\infty}$ such that $1/2^i \geq \sum_{n=1}^{m} b_n \mathbb{I}\{x_n = i\} > 1/2^i - \epsilon/2N$, for any $i \in \{1, 2, ..., N\}$. Therefore, for $m > M$, $\sum_{n=1}^{m} b_n \geq \sum_{n=1}^{m} b_n(\mathbb{I}\{x_n = 1\} + \mathbb{I}\{x_n = 2\} + ... + \mathbb{I}\{x_n = N\}) > \sum_{i=1}^{N}(1/2^i - \epsilon/2N) = \sum_{i=1}^{N}(1/2^i) - \epsilon/2 > 1 - \epsilon$.

This shows 1 is the limit of $\{p_m\}_{m=1}^{\infty}$ and hence the value of $\sum_{n=1}^{\infty} b_n$.

" It is easy to see that $\lambda^*([a, b]) \leq b - a$": Consider a sequence of open intervals (I_n) where $I_1 = (a - \epsilon/4, b + \epsilon/4)$, and the sum of lengths of $I_2, I_3, ...$ is $\epsilon/2$. This sequence covers $[a, b]$ and the sum of their lengths is $b - a + \epsilon$. So the infimum in the definition of Lebesgue outer measure is $\leq b - a + \epsilon$. Also, $\epsilon > 0$ is arbitrary. So, $\lambda^*([a, b]) \leq b - a$.

"It is easy to check that $b - a \leq \sum_{i=1}^{n}(b_i - a_i)$": I believe there is a way to argue rigorously, but to argue in a not-that-rigorous way, if we have freedom to define n open intervals with fixed lengths, then to cover "as more points as possible", we would want these n open intervals to be disjoint. And if $b - a > \sum_{i=1}^{n}(b_i - a_i)$, then even when all the $(a_i, b_i), i = 1, ..., n$ are disjoint, we cannot cover $[b - a]$.

page 15

"and are such that for each j the interior of K_j is included in some R_i": Don't know how to rigorously show this.

page 20

"It follows that $(t, x]$ is the union of a finite collection of ... in some $(a_n, b_n+\delta_n]$": We saw above that $(t, x] \subset [t, x] \subset \cup_{n=1}^N (a_n, b_n+\delta_n)$. Assume for each $n = 1, ..., N$, $(a_n, b_n+\delta_n)$ has nonempty intersection with $(t, x]$ otherwise we can remove $(a_n, b_n+\delta_n)$ from the cover. Now $(t, x] \subset \cup_{n=1}^N (a_n, b_n+\delta_n]$, so

$$\begin{aligned}(t, x] =&(t, x] \cap \Big(\cup_{n=1}^N (a_n, b_n + \delta_n]\Big) \\ =& \cup_{n=1}^N (c_n, d_n], \end{aligned} \tag{p20.1}$$

where $(c_n, d_n] = (t, x] \cap (a_n, b_n + \delta_n]$ is nonempty by our assumption above. Now we claim any union of nonempty half open half closed intervals in the form (p20.1) can be written as the union of **disjoint** nonempty half open half closed intervals in the same form. This clearly holds for $N = 2$ when we consider every possible scenario (disjoint, one included in the other, nonempty intersection and no inclusion relation) of $\cup_{n=1}^2 (c_n, d_n]$. Suppose the statement holds for $n = 1, ..., k$. For $\cup_{n=1}^{k+1} (c_n, d_n]$, write $\cup_{n=1}^{k} (c_n, d_n]$ as $\cup_{j=1}^J (c'_j, d'_j]$, a union of disjoint nonempty half open half closed intervals. So, $\cup_{n=1}^{k+1} (c_n, d_n]$ equals $\Big(\cup_{j=1}^J (c'_j, d'_j]\Big) \cup (c_{k+1}, d_{k+1}]$ which equals

$$\{\Big(\cup_{j=1}^J (c'_j, d'_j]\Big) \cap (c_{k+1}, d_{k+1}]\} \cup \{\Big(\cup_{j=1}^J (c'_j, d'_j]\Big)^c \cap (c_{k+1}, d_{k+1}]\} \cup \{\Big(\cup_{j=1}^J (c'_j, d'_j]\Big) \cap (c_{k+1}, d_{k+1}]^c\}.$$

We call the three terms inside the three pairs of curly brackets terms (A), (B), and (C), respectively. Clearly the three terms (A), (B), and (C) are disjoint, so it's enough to show each of them, if nonempty, is a union of disjoint nonempty half open half closed intervals. But to verify this for each of (A), (B), and (C) is straightforward, for example, by drawing the real number line and the corresponding intervals involved.

Line 19. We just showed for any cover of $(-\infty, x]$ in the form of $\{(a_n, b_n]\}$, we have $F(x) \leq \sum_{n=1}^\infty (F(b_n) - F(a_n)) + 2\epsilon, \forall \epsilon > 0$. This means $F(x) \leq \sum_{n=1}^\infty (F(b_n) - F(a_n))$ and hence the infimum of the sums in the form $\sum_{n=1}^\infty (F(b_n) - F(a_n))$.

Line 21. Let $B = (-\infty, b]$. We need an equation similar to (5) in the proof of proposition 1.3.7, i.e.,

$$F(b_n) - F(a_n) = \mu^*((a_n, b_n]) = \mu^*((a_n, b_n] \cap B) + \mu^*((a_n, b_n] \cap B^c).$$

For the first equality to hold, we need to show that μ^* assigns each half open half closed interval of the form $(a, c]$ the value $F(c) - F(a)$. Once we show this, since $(a, c] \cap B$ and $(a, c] \cap B^c$ are either (respectively) $(a, c]$ and empty, empty and $(a, c]$, or $(a, b]$ and $(b, c]$, the second equality will also hold.

I feel checking the first equality is doable but would be tedious so I'm skipping this step.

Notice theorem 1.3.10 basically says the mapping defined by $F(\cdot) := \mu((-\infty, \cdot])$ is a bijection between the set of measures $\{\mu : \mu \text{ is defined on } (\mathbb{R}, \mathcal{B}(\mathbb{R})), \mu(\mathbb{R}) < \infty\}$ and the set of all bounded, nondecreasing, right-continuous functions F on $\mathbb{R}$ with $\lim_{x \to -\infty} F(x) = 0$.

page 23

Why is $K \subset A$? $C - A \subset U \implies (C-A)^c = C^c \cup A \supset U^c \implies C \cap A = C \cap (C^c \cup A) \supset C \cap U^c = K$.

At the end of page 23, we have shown, for an arbitrary positive ϵ, there is a compact $K \subset A$ with $\lambda(A) - \epsilon < \lambda(K)$. Since $\lambda(A)$ is naturally an upper bound of $\{\lambda(K) : K \subset A, K \text{ compact}\}$, this means $\lambda(A)$ is the supremum of $\{\lambda(K) : K \subset A, K \text{ compact}\}$.

page 24

Line 31. For $i = 1, ..., k_0$, denote the cube in $\mathcal{C}_i$ that includes x by C_i. By the construction, $C_1, ..., C_{k_0 - 1}$ are not included in $\mathcal{D}$ but C_{k_0} is. This is because, C_{k_0} is included in every one of $C_1, ..., C_{k_0-1}$ so has empty intersection with any other cubes from $\mathcal{C}_1, ..., \mathcal{C}_{k_0 - 1}$ that are included in $\mathcal{D}$.

page 25

Line 13. We need that V is bounded in order to have $\lambda(V) < \infty$. If $\lambda(V) = \infty$, then as long as $\mu(V - A) = \lambda(V - A) = \infty$, we may have $\mu(A) < \lambda(A)$ while the equality of line 13 still holds.

Line 26. A sequence of open intervals $(R_i)_{i \in \mathbb{N}}$ covers A, i.e., $(R_i)_{i \in \mathbb{N}} \in \mathcal{C}_A$ if and only of $(R_i + x)_{i \in \mathbb{N}}$ covers $A + x$, i.e., $(R_i + x)_{i \in \mathbb{N}} \in \mathcal{C}_{A+x}$, and the volume of R_i equals that of $R_i + x$ for all $i \in \mathbb{N}$. So by the definition of Lebesgue outer measure, $\lambda^*(A) = \lambda^*(A + x)$

B is Lebesgue measurable if and only if for all (we can write any subset of A' of $\mathbb{R}$ in the form of $A - x, A \subset \mathbb{R}$ by taking $A = A' + x$) $A - x \subset \mathbb{R}$, in other words for all $A \subset \mathbb{R}$, $\lambda^*(A - x) = \lambda^*(A - x \cap B) + \lambda^*(A - x \cap B^c)$. This holds if and only if (by the translate invariant property, adding x to both $A - x$ and B) $\lambda^*(A) = \lambda^*(A \cap B + x) + \lambda^*(A \cap (B + x)^c)$.

page 27

"Each equivalent class under $\sim$ has the form $\mathbb{Q} + x$ for some x and so is dense in $\mathbb{R}$". We recall a basic fact form abstract algebra that if we have an equivalent relation $\sim$ on some set A, then $a \sim b \Leftrightarrow [a] = [b]$. Consider the equivalent class $[x]$, then $y \in [x] \Leftrightarrow y \sim x \Leftrightarrow y = x + r$ for some $r \in \mathbb{Q}$. Therefore, $[x] = \mathbb{Q} + x$.

Recall if we have some topology, then we say $A \subset X$ is dense in X if the closure of A, i.e., the union of A with all limit points of A, equals X. For any $y \in \mathbb{R}$, $y - x$ is either a rational number or the limit of some sequence of rational numbers, due to the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$. So, y is either in $\mathbb{Q}+x$ or the limit of some sequence of elements in $\mathbb{Q}+x$, so $\mathbb{Q}+x$ is also dense in $\mathbb{R}$.

"Each intersects the interval $(0,1)$". If $y \in \mathbb{Z}$, then obviously y can be written as $r + z, r \in \mathbb{Q}, z \in (0,1) \implies z \in [y]$. For any $y \in \mathbb{R} - \mathbb{Z}, y = \lfloor y \rfloor + (y - \lfloor y \rfloor)$, where $\lfloor y \rfloor \in \mathbb{Q}, (y - \lfloor y \rfloor) \in (0,1) \implies (y - \lfloor y \rfloor) \in [y]$.

"Let x be an arbitrary member of (0, 1), and let e be the member of E that satisfies $x \sim e$". Consider $[x]$. We know from the previous discussion that in forming E, we have taken one element $e \in [x] \cap (0,1)$.

page 28

Line 24. "strictly positive". Let A be a nonempty subset of $\mathbb{R}^n$. Then $d(x; A) = 0$ if and only if $x \in \bar{A}$. (easy to prove, or see Proposition 1.74, An Easy Path to Convex Analysis and Applications). Here since U^c is closed, $d(x; U^c) = 0$ iff $x \in U^c$.

page 29

"G_0 and G_1 are disjoint". Assume instead $x \in G_0 \cap G_1$, i.e., $x = r_0 + 2n_0\sqrt{2} = r_1 + (2n_1 + 1)\sqrt{2}$. Then, $\sqrt{2} = \frac{r_0 - r_1}{2(n_1 - n_0) + 1}$, a contradiction.

"$e_1 - e_2 + g_0 = g_1$ would imply that $e_1 = e_2$ and $g_0 = g_1$". $e_1 - e_2 + g_0 = g_1 \Leftrightarrow e_1 - e_2 = g_1 - g_0 \in G$, meaning $e_1 \sim e_2$. By the way E is constructed, $e_1 = e_2$.

"It is easy to check that $A^c = E + G_1$". Suppose $x \notin A = E + G_0$. By the way we construct E, there is $e \in E$ such that $[x] = [e], \Leftrightarrow x = e + g$ for some $g \in G$. If $g \in G_0$, this would contradict the fact that $x \notin A = E + G_0$. Therefore, $g \in G_1$, which means $A^c \subset E + G_1$. Conversely, if $x \in E + G_1$, then $x \notin A = E + G_0$, or, $x \in A^c$, otherwise $x = e_1 + g_1 = e_0 + g_0$, where $g_0 \in G_0, g_1 \in G_1, e_0, e_1 \in E$. Then $e_0 - e_1 = g_1 - g_0 \in G$ as G is a group. Therefore, $e_0 \sim e_1 \Rightarrow e_0 = e_1 \Rightarrow g_1 = g_0$, a contradiction as G_0 and G_1 are disjoint.

"A is not Lebesgue measurable". If it were, then let I be the unit cube. As $\mathcal{M}_{\lambda^*}$ is a sigma algebra, both $I \cap A, I \cap A^c$ would be measurable. We would then have $1 = \lambda(I) = \lambda(I \cap A) + \lambda(I \cap A^c)$. So at least one term of the right hand side is positive, and without loss of generality say it's $\lambda(I \cap A)$. Then, by monotonicity of measure, we know $\lambda(A) > 0$. As $A^c = A + \sqrt{2}$, we also have $\lambda(A^c) > 0$.

page 31

Line (15). Notice a symmetric argument. If $B \in \mathcal{A}, B \supset A$, then $\mu(B) \geq \mu(E) = \mu(F)$. This means $\mu(F) = \inf\{\mu(B) : B \in \mathcal{A}, B \supset A\}$.

Line (-9). Note that $F - E = E^c - F^c$.

Line (-5). Any $x \in (\cup_n F_n - \cup_n E_n)$, x belongs to some F_i but not any of E_j. So $x \notin E_i \Rightarrow x \in F_i - E_i \Rightarrow x \in \cup_n (F_n - E_n)$.

Line (-2). "It is an extension of μ". This means $\mu(A) = \bar{\mu}(A), \forall A \in \mathcal{A}$. This is because we saw before that $\bar{\mu}(A) := \mu(E)$ for any E that satisfies (1), and we can take $E = F = A$ in (1).

Check that $\bar{\mu}$ is complete: suppose $B \subset A \in \mathcal{A}_\mu, \bar{\mu}(A) = 0$. Then by (1), there exists $F \in \mathcal{A}, \mu(F) = 0$. Clearly $\emptyset \subset B \subset F$ with $\mu(F - \emptyset) = 0$, so $B \in \mathcal{A}_\mu$.

page 33

Check $\mu_*(A) \leq \mu^*(A)$. If RHS is infinity this holds. Suppose $c = \mu^*(A) < \infty$. Then for any positive ϵ, there exists $B \in \mathcal{A} : B \supset A, \mu(B) < c + \epsilon$. Now for any $C \in \mathcal{A}$ such that $C \subset A$, $\mu(C) \leq \mu(B) < c + \epsilon$. Taking the supremum over all such C, we know $\mu_*(A) \leq c + \epsilon$. Recall ϵ is arbitrary.

Line (-7). Notice if $\mu(F) = \infty$, then since $\mu(E) + \mu(F - E) = \mu(F)$, we must have $\mu(E) = \infty$ so line (-6) would imply $\mu^*(A) = \infty$, a contradiction. So $\mu(F) < \infty$.

page 35

Line 3. Let $A \in \mathcal{R}$, i.e., A satisfies (5) and (6). For $\epsilon > 0$, by (5), there exists U open such that $U \supset A, \mu(U) < \mu(A) + \epsilon/2$. Similarly by (6), there exists C closed, $C \subset A, \mu(C) > \mu(A) - \epsilon/2$. As μ is a finite measure, $\mu(U - C) = \mu(U) - \mu(C) = \mu(U) - \mu(A) + \mu(A) - \mu(C) < \epsilon$. Conversely, if A satisfies (7), then $\mu(A - C) < \epsilon, \mu(U - A) < \epsilon$, and (5) and (6) follow easily.

(8): Notice $U - C = \cup_k (U_k - C) \subset \cup_k (U_k - C_k)$

2 Functions and Integrals

page 42

Example 2.1.2 (b). Suppose the set $S = \{x \in I : f(x) < t\}$ is not empty. Case 1: $S = I$. Case 2: $S \neq I$. In this case, there exists $y \in I - S$. Since $S \neq \emptyset, \exists s \in S, f(s) < t$. As f is non-decreasing, we must have $y > s$. Therefore, the set $Y := I - S$ is bounded below by s. Let $i := \inf Y \geq s$. For any $x \in \mathbb{R}, x < i$, either $x \notin I$, or $x \in S \subset I$. Therefore, depending on whether $i \in S$, S is either $I \cap (-\infty, i]$, or $I \cap (-\infty, i)$, which means S may be a singleton or an interval.

page 45

Line 1. If $f(x) + g(x) < t$, then both $f(x), g(x)$ are upper-bounded, i.e., neither is infinity. There exists a positive number ϵ such that $f(x) + g(x) < t - \epsilon$. There exists a positive number $r < \epsilon$ such that $f(x) + r$ is rational. Then, $f(x) < f(x) + r, g(x) < t - \epsilon - f(x) < t - (f(x) + r)$.

page 46

Line 7. Note that although here the author only says "defined on a subset of X", by the definition of measurability, the domain, call it A, of such a measurable function also $\in \mathcal{A}$. Suppose f is measurable, then for $t \in \mathbb{R}, t > 0$, the preimage of $[t, \infty]$ under f^+ is $f^{-1}([t, \infty]) \in \mathcal{A}$; if $t \leq 0$, the preimage of $[t, \infty]$ under f^+ is A. Therefore, f^+ is measurable, and we can argue similarly for f^-. Conversely, suppose both f^+, f^- are measurable. For $t \in \mathbb{R}, t > 0, f^{-1}([t, \infty]) = f^{+-1}([t, \infty]) \in \mathcal{A}$. For $t = 0$, $f^{-1}([0, \infty]) = f^{+-1}([0, \infty]) \cap f^{--1}(\{0\}) \in \mathcal{A}$. For $t < 0$, $f^{-1}([t, \infty]) = f^{-1}([0, \infty]) \cup f^{-1}([t, 0))$. We already know $f^{-1}([0, \infty]) \in \mathcal{A}$, but note $f^{-1}([t, 0)) = f^{--1}((0, t]) \in \mathcal{A}$.

page 48

Line 17. Check continuity. First suppose $x \in (0, 1]$ and we prove left-continuity. Take a sequence $(x_n) \subset [0, 1], x_n \uparrow x, x_n < x$. Take $\epsilon > 0$. By definition of f and monotonicity, there exists $t \in [0, 1] - K, t < x$ with $f(x) - \epsilon < f(t) \leq f(x)$. There exists $N \in \mathbb{N}$ such that for all $n \geq N, t < x_n \leq x \Rightarrow f(x) - \epsilon < f(t) \leq f(x_n) \leq f(x)$.

Now suppose $x \in [0, 1)$, and suppose right-continuity doesn't hold, namely, there exists $\epsilon > 0, (x_n) \subset [0, 1], x_n \downarrow x, x_n > x$ with $f(x_n) > f(x) + \epsilon$. Recall the definition of f. As f is non-decreasing, we can write $f(x_n) = \sup\{f(t) : t \in [0, 1] - K, x \leq t < x_n\}$, i.e., all the t that are smaller than x doesn't affect the supremum. Now, we have that $f(x_n) - f(x) = \sup\{f(t) : t \in [0, 1] - K, x \leq t < x_n\} - \sup\{f(t) : t \in [0, 1] - K, t < x\} > \epsilon$, which is a contradiction if we recall the definition of f on $[0, 1] - K$ (some drawings may help understand), because as $(x_n) \downarrow x$, the lengths of the intervals $[x, x_n) \downarrow 0$.

Line 18. To see $f(1) = 1$, note that there exists a sequence $(x_n) \subset [0, 1] - K, x_n \uparrow x, f(x_n) \uparrow 1$. Then apply continuity.

Line 23. We check the values of g belong to K. If y has the form $m/2^n$, n a natural number, m an odd number less than 2^n, then there exists an open interval $(a, b) \subset [0, 1] - K$ such that $f = m/2^n = y$ on (a, b). Notice $a \in K$ and since f is continuous, $f(a) = f(b) = y$. For any $x \in [0, 1] - K - (a, b)$, obviously $f(x) \neq y$. Also, for $x \in K - [a, b]$, suppose n_0 is the smallest integer that x becomes an "end point" of K_{n_0}, e.g., for $x = 1/3, n_0 = 1$, and for $x = 2/9, n_0 = 2$. This means we have removed an open interval of the form (c, x) or (x, c) from K_{n_0-1} in constructing K_{n_0}. Then, by continuity of $f, f(x) = m'/2^{n'}$ where $m'/2^{n'}$ is the function value assigned to (c, x) or (x, c), and different than y. This means $f(x) \neq y$. This means the only points in $[0, 1]$ that take value y under f are $[a, b]$. Therefore, $g(y) = a \in K$.

Now suppose y does not have the form $m/2^n$. Then, there is no $x \in [0, 1] - K$ such that $f(x) = y$. Therefore, $g(y) = \inf\{x \in [0, 1] : f(x) = y\} = \inf\{x \in K : f(x) = y\} \in K$ because K is closed.

page 52

Line 9. For each n, $g_{0,n} = g_{1,n}$ holds except for at most a set S_n that is negligible. The set $S = \{x \in X : g_{0,n}(x) = g_{1,n}(x), \forall n\}$ includes $\cap_n(S_n^c)$, so its complement is included in $\cup_n S_n$, which is negligible. On S, we have $g_{0,n} = g_n = g_{1,n}, \forall n \Rightarrow f = \lim g_n = \lim g_{0,n}(= f_0) = \lim g_{1,n}(= f_1)$.

page 53

Line 14. If $f = 0$, then with either representation, its integral is zero. Suppose $f > 0$ on $E \neq \emptyset$ and $f = 0$ on E^c. If we eliminate all sets $A_i : a_i = 0$, and all sets $B_j : b_j = 0$, then the remaining A_i's and B_j's in the representation of f must be such that their corresponding a_i or b_j are > 0, and that A_i, $B_j \subset E$. Therefore, $\cup A_i, \cup B_j \subset E$. Conversely, it's easy to check that $E \subset \cup A_i, E \subset \cup B_j$. So $\cup A_i = \cup B_j = E$.

page 55

Line 16. Notice here $\lim_n \int f_n d\mu$ may be infinity. Recall when we say the limit of the sequence (a_n) exists, we mean $\lim_n a_n = c \in \mathbb{R}$ for some c. When we instead say $\lim_n a_n$ exists, we simply mean $\lim_n a_n$ is well defined and it can be a real number or infinity.

page 56

Line 16. We note that for an $\mathcal{A}$-measurable function f, it is integrable if and only if $|f|$ is integrable. Suppose f is so, which means $\int f^+, \int f^-$ are both finite. Then, $\int |f|^+ = \int |f| = \int f^+ + f^-$, which equals $\int f^+ + \int f^-$ by Proposition 2.3.4, which is finite. Also, $\int |f|^- = 0$ which is finite. Conversely, suppose $\int |f|^+, \int |f|^- = 0$ are both finite, then $\int |f|^+ = \int f^+ + f^- = \int f^+ + \int f^-$ is finite. This means both $\int f^+, \int f^-$ are finite.

page 58

Example 2.3.7, (a). Say f is bounded measurable on X: $|f| \leq M$ for some real M. Then by Proposition 2.3.4, $\int f^+ d\mu \leq \int M d\mu = \int M\chi_X d\mu = M\mu(X) < \infty$. Same argument for $\int f^- d\mu$.

(c). Consider f_n defined this way: $f_n(m) = f(m), m = 1, 2, ..., n$ and $f_n(m) = 0$ elsewhere.

For a not necessarily nonnegative f, we have that f^+, f^- integrable iff $\sum_{m=1}^{\infty} \max\{f(m), 0\} < \infty, \sum_{m=1}^{\infty} \max\{-f(m), 0\} < \infty$ iff $\sum_{m\in\mathbb{N}, f(m)\geq 0} f(m) < \infty, \sum_{m\in\mathbb{N}, f(m)<0} -f(m) < \infty$ iff $\sum_{m\in\mathbb{N}} |f(m)| < \infty$.

(d). Denote such a function by f, and say it doesn't vanish on a subset of $E, \mu(E) = 0$. Note that f may take ∞. Say $f = \infty$ on $F \subset E$. Note that $F = \cap_{n=1,2,\ldots}\{x : f(x) > n\}$ measurable. By defining f_n as one that takes n on F and equals f elsewhere, we have a sequence of real valued nonnegative simple functions (f_n). (Or we could simply apply Proposition 2.1.8.) Then Proposition 2.3.3 together with the definition of integral for simple functions implies $\int f d\mu = 0$.

page 60

It's easy to see $f\chi_A$ is integrable as f is so. If $\int f\chi_A > 0$, this means $\int (f\chi_A)^+ > 0$ because $\int (f\chi_A)^+ - \int (f\chi_A)^-$ has to be positive, by definition of integral. However, as $(f\chi_A)^+ = 0$, this is not possible.

Now, since $\int f\chi_A = \int (f\chi_A)^+ = 0$, we must have $\int (f\chi_A)^- = 0$ and therefore $\int |f|\chi_A = 0$, and then we can apply Proposition 2.3.12.

page 62

Line 12. Why is it that $f = \lim_n h_n$? It's easy to see that $\lim_n h_n \leq \lim_n f_n = f$. Suppose $\lim_n h_n \neq f$, i.e., there exists x such that $\lim_n h_n(x) < f(x)$. Then $\lim_n h_n(x)$ cannot be ∞ and there exists $r \in \mathbb{R} : \lim_n h_n(x) < r < f(x)$. As $f_n \to f$, there exists N such that $f_N(x) > r$. As $g_{N,k} \to f_N$, there exists $K > N$ such that $g_{N,K}(x) > r$. Therefore, $h_K(x) \geq g_{N,K}(x) > r \Rightarrow \lim_n h_n(x) \geq r$, contradicting that $\lim_n h_n(x) < r$.

page 69

Line 14. Say $g(x) = h(x)$ except possibly for points in the set E with measure 0. There are obviously only countably many points that are division points, so the set of division points D is of measure 0. Now if something holds on $E^c \cap D^c$, it holds almost everywhere. Say $x \in E^c \cap D^c$ but f is not continuous at x. Then there exists $\epsilon > 0$ and a sequence of points $(x_n) \to x$ with $|f(x_n) - f(x)| > \epsilon$, i.e., either $f(x_n) > f(x) + \epsilon$ or $f(x_n) < f(x) - \epsilon$ (at least one of the two must hold for infinitely many terms in (f_n)). Without loss of generality and up to considering a subsequence of (x_n), suppose $f(x_n) > f(x) + \epsilon, \forall n$. But this would mean that, if we recall the definition of h_n, $h_n(x) > f(x) + \epsilon, \forall n$, while from the definition of g_n, we know $g_n(x) \leq f(x)$. This would mean we cannot have $g(x) = h(x)$, a contradiction.

page 70

Line (-4). Recall "Each of these tagged partitions has mesh less than δ and so satisfies $|\mathcal{R}(f, \mathcal{P}) - L| < \epsilon$." Therefore, suppose $|\mathcal{R}(f, \mathcal{P}_1) - L| = (1-k)\epsilon, k \in (0, 1]$. Since in defining and calculating

$l(f, \mathcal{P}_0)$, for each subinterval we use the infimum of f over this subinterval, and there are only finitely many subintervals, say m, we can choose a $\mathcal{P}_1$ such that $f(x_i)(a_i - a_{i-1}) - \inf_{y \in [a_{i-1}, a_i]}\{f(y)\}(a_i - a_{i-1}) < k\epsilon/m, \forall i = 1, ..., m$. Then, we have that $|l(f, \mathcal{P}_0) - \mathcal{R}(f, \mathcal{P}_1)| < k\epsilon$, and so $|l(f, \mathcal{P}_0) - L| < \epsilon$.

page 71

Footnote. Notice for this function, the integrals of the positive part and negative part are both infinity, so the Lebesgue integral is not defined $(\infty - \infty)$. For the improper Riemann integral, recall (see for example, Rudin *Principles of Mathematical Analysis*, exercise 6.8) the definition says that if f is Riemann integrable on $[a, b]$, for all $b > a$ where a is fixed, then $\int_a^\infty f(x)dx := \lim_{b\to\infty} \int_a^b f(x)dx$, provided that this limit exists (and is finite). The improper integral exists because the alternating harmonic series is convergent.

page 74

Remark. After seeing Example 2.6.3, Proposition 2.6.4, we notice that, when we are referring to the definition of measurability in section 2.1, if f defined on X in the measure space $(X, \mathcal{A})$ takes values in $\overline{\mathbb{R}}$, then by saying f is $\mathcal{A}$-measurable, we actually mean f is a measurable mapping from $(X, \mathcal{A})$ to $\big(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}})\big)$. If such f is real-valued, then by saying f is $\mathcal{A}$-measurable, we mean f is a measurable mapping from $(X, \mathcal{A})$ to $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$.

page 75

Line 1. The product of two measurable complex-valued functions on X is measurable: suppose the two functions are $f = (f_1, f_2)$ and $g = (g_1, g_2)$. Recall that the product of two complex numbers (a, b) and (c, d), namely, of $a + bi$ and $c + di$, is $(ac - bd, ad + bc)$. So the product of the two functions is defined as $(f_1g_1 - f_2g_2, f_1g_2 + f_2g_1)$. Since both $f_1g_1 - f_2g_2$ and $f_1g_2 + f_2g_1$ are measurable, by Proposition 2.1.7, we have that $fg := (f_1g_1 - f_2g_2, f_1g_2 + f_2g_1)$ is measurable, by the discussion in Example 2.6.5.

Line (-7). We have that

$$w^{-1} \int f = R(w^{-1} \int f) = R(\int w^{-1} f) := \int R(w^{-1} f),$$

also note that w^{-1} is of unit length, and recall that left-multiplying a complex number by a unit complex number is equivalent to rotating in the $\mathbb{C}^2$ plane. So the norm of $w^{-1}f$ is the same as $|f|$, but since $w^{-1}f$ might have a nonzero projection on the imaginary axis, the norm of its real part is no more than $|f|$. So, we have $\int R(w^{-1}f) \leq \int |R(w^{-1}f)| \leq \int |f|$.

page 76

Line $(-8), (-3)$. We saw before in Proposition 1.4.4 that $\lambda(B + y) = \lambda(B), \forall B \in \mathcal{B}(\mathbb{R})$. The relation that $\lambda(B) = \lambda(-B)$ should be easy to verify according to the definition of Lebesgue outer measure.

Exercises

1.1.1

Intuition says the answer may be the sigma algebra $\mathcal{A}$ given by (f) in examples 1.1.1. Now let's try to prove it. Suppose $\mathcal{B}$ is a sigma algebra on $\mathbb{R}$ that contains all one-point sets. Then by closure under the formation of countable unions, $\mathcal{B}$ contains all countable subsets of $\mathbb{R}$. By closure of complementation, $\mathcal{B}$ contains all subsets of $\mathbb{R}$ whose complement is countable. Therefore, $\mathcal{B} \supset \mathcal{A}$.

1.1.2

Consider $(-\infty, b]$, where b is some rational number. By Proposition 1.1.4, $(-\infty, b] \in \mathcal{B}(\mathbb{R})$. Now suppose $\mathcal{S}$ is a sigma-algebra that contains all such $(-\infty, b]$. For any real number c, there is a sequence of rational numbers $(b_n) \downarrow c$, meaning $(-\infty, c] = \cap_n(-\infty, b_n]$. Therefore, $(-\infty, c] \in \mathcal{S}$. By Proposition 1.1.4, $\mathcal{S} \supset \mathcal{B}(\mathbb{R})$.

1.1.4

By Proposition 1.1.7, we only need to show that $\cup_n^\infty A_n \in \mathcal{A}$ if (A_n) is an increasing sequence of sets. Let (A_n) be such a sequence. Define $B_m := A_m - A_{m-1} \in \mathcal{A}, m \geq 2$ with $B_1 := A_1$, then $\cup_{n=1}^\infty A_n = \cup_{m=1}^\infty B_m$ (if not feeling comfortable, you can show this by arguing that LHS contains RHS, and RHS contains LHS). But each two different B_i and B_j are disjoint, so $\cup_{n=1}^\infty A_n = \cup_{m=1}^\infty B_m \in \mathcal{A}$.

1.1.5

Consider $X = \{1, 2, 3\}$. And let the two sigma algebras be $\{\{3\}, \{1, 2\}, \{1, 2, 3\}, \varnothing\}$ and $\{\{2\}, \{1, 3\}, \{1, 2, 3\}, \varnothing\}$. The union of these two is not a sigma algebra.

1.1.7

As it says that $\mathcal{S}$ is a "collection of subsets", we assume $\mathcal{S} \neq \varnothing$, i.e., there exists $Y \subset X, Y \in \mathcal{S}$. Let $\mathcal{A}$ be defined as in the hint. We first show it is a sigma algebra. Since $\{Y\}$ is a countable subfamily of $\mathcal{S}$, $\mathcal{A} \supset \sigma(\{Y\})$ which includes X and therefore $X \in \mathcal{A}$. If $A \in \mathcal{A}$, then by the definition of $\mathcal{A}$, there is some countable subfamily $\mathcal{C}$ of $\mathcal{S}$ such that $A \in \sigma(\mathcal{C})$, so $A^c \in \sigma(\mathcal{C}) \subset \mathcal{A}$. Now assume $A_1, A_2, ... \in \mathcal{A}$, i.e., $A_1 \in \sigma(\mathcal{C}_1), A_2 \in \sigma(\mathcal{C}_2), ...$ where $\mathcal{C}_i$ are countable subfamilies of $\mathcal{S}$, and $\mathcal{C}_i$ might be identical for different values of i. Now we define $\mathcal{C}_\infty := \cup_i \mathcal{C}_i$ which is also a countable subfamily. We notice for all $i \in \mathbb{N}$, $A_i \in \sigma(\mathcal{C}_i) \subset \sigma(\mathcal{C}_\infty)$, and so $\cup_i A_i \in \sigma(\mathcal{C}_\infty) \subset \mathcal{A}$. The above shows $\mathcal{A}$ is a sigma algebra.

Since $\mathcal{A} =$ the union of elements in $\{\sigma(\mathcal{C}) : \mathcal{C}$ a countable subfamily of $\mathcal{S}\}$, and $\sigma(\mathcal{C}) \subset \sigma(\mathcal{S})$ for all such $\mathcal{C}$, we know $\mathcal{A} \subset \sigma(\mathcal{S})$. For any element $Y \subset \mathcal{S}$, $\sigma(\{Y\}) \subset \mathcal{A}$. Therefore, the union of all such $\sigma(\{Y\})$ is a subset of $\mathcal{A}$. But the union of all such $\sigma(\{Y\})$ includes the union of all such $\{Y\}$ which is $\mathcal{S}$ itself. This means $\mathcal{S} \subset \mathcal{A}$. Now we have $\mathcal{S} \subset \mathcal{A} \subset \sigma(\mathcal{S}) \Rightarrow \mathcal{A} = \sigma(\mathcal{S})$. Therefore, for each A contained in $\sigma(\mathcal{S}) = \mathcal{A}$, A is contained in $\sigma(\mathcal{C}_0)$ for some $\mathcal{C}_0$ which is a countable subfamily of $\mathcal{S}$.

1.2.2

Let $(q_1, q_2, ...)$ be an enumeration of rational numbers, and let $\mathbb{I}$ denote the set of irrational numbers. Define $A_n := \{q_n\} \cup \mathbb{I}$. Then considering the sequence (A_n), we see μ is sigma-finite.

1.2.3

Take $A_n := \{n, n+1, n+2, ...\}$.

1.2.5

Assume $\delta_x = \delta_y$, i.e., for any $A \in \mathcal{A}, \delta_x(A) = \delta_y(A)$. Suppose it is not that x, y belong to exactly the same sets in $\mathcal{A}$. The case that neither of x, y belongs to any set of $\mathcal{A}$ counts as "belonging to exactly the same sets" so this case is not possible. Case 1: $x \in B \in \mathcal{A}$, and y belongs to no set in $\mathcal{A}$ (or, with roles of x and y switching). Then we have $\delta_x(B) \neq \delta_y(B)$, a contradiction. Case 2: both of x, y belong to some sets in $\mathcal{A}$, but the sets that either belongs to are not identical. Without loss of generality, suppose there is a $B \in \mathcal{A}$ s.t. $x \in B, y \notin B$. Then this again implies $\delta_x(B) \neq \delta_y(B)$.

Conversely, if x, y belong to exactly the same sets in $\mathcal{A}$, it is trivial to see that $\delta_x = \delta_y$.

1.2.6

(b). Once we have proven (a), consider the measures $\nu_k := \sum_{n=1}^k \mu_n$ and apply (a), then (b) follows.
(a). That $\mu(\emptyset) = 0$ is easy to see. For each $A \in \mathcal{A}, \mu(A)$ is nonnegative and well-defined as $\mu_n(A)$ is increasing in n. Now we check sigma additivity. Consider a disjoint sequence of sets $(A_k) \subset \mathcal{A}$. $\mu(\cup A_k) := \lim_{n\to\infty} \mu_n(\cup A_k) = \lim_{n\to\infty} \sum_{k=1}^\infty \mu_n(A_k)$. This actually equals (which we will prove shortly) $\sum_{k=1}^\infty \lim_{n\to\infty} \mu_n(A_k) = \sum_k^\infty \mu(A_k)$, which means countable additivity holds. Now we prove the claim used before.

Claim. Let $(a_{m,n})_{m,n\geq 0}$ be an array of nonnegative numbers such that, for all $n \geq 0, (a_{m,n})_{m\geq 0}$ is a nondecreasing sequence. Then, we have that $\lim_{m\to\infty} \sum_{n=0}^\infty a_{m,n} = \sum_{n=0}^\infty \lim_{m\to\infty} a_{m,n}$ in $[0, \infty]$.

Proof. We recall the fact that $\sum_i \sum_j a_{i,j} = \sum_j \sum_i a_{i,j}$ for a nonnegative doubly-indexed sequence $a_{i,j}$. Consider the array $A_{m,n} \geq 0$ given by $A_{0,n} = a_{0,n}$ and $A_{m,n} = a_{m,n} - a_{m-1,n}, m \geq 1$. Now, $\sum_{n=0}^\infty a_{m,n} = \sum_{n=0}^\infty \sum_{k=0}^m A_{k,n} = \sum_{k=0}^m \sum_{n=0}^\infty A_{k,n} \Rightarrow \lim_{m\to\infty} \sum_{n=0}^\infty a_{m,n} = \sum_{k=0}^\infty \sum_{n=0}^\infty A_{k,n}$. But $\sum_{k=0}^\infty \sum_{n=0}^\infty A_{k,n} = \sum_{n=0}^\infty \sum_{k=0}^\infty A_{k,n} = \sum_{n=0}^\infty \lim_{m\to\infty} \sum_{k=0}^m A_{k,n} = \sum_{n=0}^\infty \lim_{m\to\infty} a_{m,n}$.

Now back to the problem. If $\mu_n(A_k)$ is finite $\forall n, k$, then we have $\lim_{n\to\infty} \sum_{k=1}^\infty \mu_n(A_k)$ equals $\sum_{k=1}^\infty \lim_{n\to\infty} \mu_n(A_k)$ just as in the claim above. If some $\mu_n(A_k) = \infty$, we cannot resort to the claim as the difference of two infinities is not defined. But it is obvious that we still have the same

$\lim_{n\to\infty}\sum_{k=1}^{\infty}\mu_n(A_k)=\sum_{k=1}^{\infty}\lim_{n\to\infty}\mu_n(A_k)$ as both sides are infinity.

1.2.9

In fact, the limsup of such a sequence (A_n) is defined as $\limsup_n A_n := \cap_{n\geq 1}\cup_{k\geq n} A_k$. Notice that $\limsup_n A_n$ contains exactly points that appear in infinitely many A_k's. To see this, suppose x belongs to infinitely many A_k's, then x belongs to $\cup_{k\geq n}A_k$ for every n, and therefore belongs to $\cap_{n\geq 1}\cup_{k\geq n} A_k$. Conversely, suppose x belongs to only finitely many A_k's, then there exists $N\in\mathbb{N}$ such that for all $n>N, x\notin A_n$. So for all $n>N, x\notin \cup_{k\geq n}A_k \Rightarrow x\notin \cap_{n\geq 1}\cup_{k\geq n} A_k$.

Fix $\epsilon>0$, there exists $N\in\mathbb{N}, \sum_{n\geq N}\mu(A_n)<\epsilon$. Following the hint, we see that $\mu(\limsup A_n)\leq \mu(\cup_{n\geq N}A_n)\leq \sum_{n\geq N}\mu(A_n)<\epsilon$. As ϵ is arbitrary, $\mu(\limsup A_n)=0$.

This theorem is the Borel–Cantelli lemma. Here I give an intuitive explanation. Recall $\limsup A_n = \{x\in X : x\in A_n \text{ for infinitely many } n\}$. Suppose $\mu(\limsup A_n)>0$. For some $x\in\limsup A_n$, the measure mass that x occupies is counted (countably) infinitely many times during the calculation $\sum_k \mu(A_k)$. That is because, for those A_k that contains x, their measure $\geq \mu(\{x\})$. If we allow some non-rigor and let i denote "countably infinitely many", then we know $\sum_k \mu(A_k)\geq i\mu(\{x\})$. For another $y\in\limsup A_n$, its measure is also counted (countably) infinitely many times during the calculation $\sum_k\mu(A_k)$. Therefore, considering x, y together, we have $\sum_k\mu(A_k)\geq i\mu(\{x,y\})$. We make this argument for all elements in $\limsup A_n$ (not very rigorous, because if there are uncountably many such elements, how would we "argue" for each of them? In what order?). Then we would have $\sum_k\mu(A_k)\geq i\mu(\limsup A_n)$. Since i denotes countable infinity and $\mu(\limsup A_n)>0$, the RHS is ∞, meaning the LHS must be so as well.

1.3.3

If $\lambda^*(A)=\infty$, take $B=\mathbb{R}$. Now suppose $\lambda^*(A)<\infty$. For $n\in\mathbb{N}$, find a sequence of open intervals $\{(a_k^n,b_k^n)\}_{k=1}^{\infty}$ that covers A with $\sum_k b_k^n-a_k^n<\mu^*(A)+1/n$. Define $B^n=\cup_k(a_k^n,b_k^n)$. Note that B^n is Borel, contains A, and $\mu(B^n)\leq\sum_k b_k^n-a_k^n<\mu^*(A)+1/n$. Now $B:=\cap_n B^n$ is also Borel, contains A, and satisfies for any $n\in\mathbb{N}$, $\mu(B)\leq\mu(B^n)<\mu^*(A)+1/n$. Therefore, $\mu(B)\leq\mu^*(A)$ and so $\mu(B)=\mu^*(A)$

1.4.1

(a). We show $A :=$ the x-axis (and analogously y-axis) has zero measure. Consider any positive number c and the sequence of open rectangles $\{R_n=(-2^n,2^n)\times(-c2^{-2n},c2^{-2n})\}, n=1,2,\ldots$ Apparently $\cup_n R_n\supset A\Rightarrow\lambda^*(A)\leq\lambda^*(\cup_n R_n)\leq\sum_n\lambda^*(R_n)=\sum_n \mathrm{vol}(R_n)=4c$. c can be arbitrarily small, meaning $\lambda^*(A)=0$. By proposition 1.3.5, $\lambda(A)=0$. If we consider the rotation and translation of the x-axis and use exercise 1.4.3, we can see that any straight line of $\mathbb{R}^2$ has Lebesgue measure zero.

1.4.3

(a). Let T denote the rotation in $\mathbb{R}^2$, which is obviously a bijection of $\mathbb{R}^2$. Notice T^{-1} is also a rotation. We follow the hint. $\mathcal{A} := \{A \subset \mathbb{R}^2 : T(A) \in \mathcal{B}(\mathbb{R}^2)\}$ is easily seen to be a sigma algebra: $\mathbb{R}^2 \in \mathcal{A}$; if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ because $T(A^c) = T(A)^c$ as T is a bijection; if $(A_n)_n \subset \mathcal{A}$, then $\cup_n A_n \in \mathcal{A}$ because $T(\cup_n A_n) = \cup_n T(A_n)$ (easy to check inclusion of both directions).

Also, $\mathcal{A}$ includes all open sets: Let A be open. Since all open balls are a basis for the general topology on the Euclidean space, A is a union of some open balls. Now, $T(A)$ is the union of the rotation about origin of these open balls, which are still open balls. So, $T(A)$ is open and so $A \in \mathcal{A}$. Therefore, $\mathcal{A} \supset \mathcal{B}(\mathbb{R}^2)$. Similarly we can show $\mathcal{C} := \{A \subset \mathbb{R}^2 : T^{-1}(A) \in \mathcal{B}(\mathbb{R}^2)\} \supset \mathcal{B}(\mathbb{R}^2)$.

Now we show the converse inclusion. Let $A \in \mathcal{A} : T(A) \in \mathcal{B}(\mathbb{R}^2)$. Therefore, $T(A) \in \mathcal{C}$, which means $A = T^{-1} \circ T(A) \in \mathcal{B}(\mathbb{R}^2)$. This means $\mathcal{A} \subset \mathcal{B}(\mathbb{R}^2)$ and so the two sets are equal.

(b). Define μ by $\mu(A) = \lambda(T(A))$ on $\mathcal{B}(\mathbb{R}^2)$. It's easy to check μ is a nonzero measure. Then μ satisfies all conditions in proposition 1.4.5. For translation invariance, $\mu(A + x) = \lambda(T(A + x))$. But $T(A + x)$ is a translation of $T(A)$ so $\lambda(T(A + x)) = \lambda(T(A)) = \mu(A)$, by proposition 1.4.4. Therefore, μ is translation invariant. Now by proposition 1.4.5, we have $\mu(A) = c\lambda(A), \forall A \in \mathcal{B}(\mathbb{R}^2)$.

Now let A be the unit disc. A is closed so is measurable for the Borel sigma algebra. Also, A has positive Lebesgue measure as it includes a square, whose volume is positive. So $c = \lambda(A)/\mu(A) = \lambda(A)/\lambda(T(A)) = \lambda(A)/\lambda(A) = 1$.

1.4.4

Look up "Smith–Volterra–Cantor set" or "fat Cantor set". The following is from the Wikipedia.

In general, one can remove r_n from each remaining subinterval at the n-th step of the algorithm, and end up with a Cantor-like set. The resulting set will have positive measure if and only if the sum of the sequence is less than the measure of the initial interval. For instance, suppose the middle intervals of length a^n are removed from $[0, 1]$ for the n-th iteration, for some $0 \leq a \leq 1/3$. Then, the resulting set has Lebesgue measure

$$1 - \sum_{n=0}^{\infty} 2^n a^{n+1} = \frac{1 - 3a}{1 - 2a}.$$

which goes from 0 to 1 as a goes from 1/3 to 0.

1.5.1

By Proposition 1.5.1, we see $(\mathcal{A}_\mu)_{\overline{\mu}} \supset \mathcal{A}_\mu$. To show $(\mathcal{A}_\mu)_{\overline{\mu}} \subset \mathcal{A}_\mu$, take $A \in (\mathcal{A}_\mu)_{\overline{\mu}}$, and notice by definition there exist $E, F \in \mathcal{A}_\mu$ with $E \subset A \subset F, \overline{\mu}(F - E) = 0$. By the definition again, there exist $E_-, E_+, F_-, F_+ \in \mathcal{A}$ with $E_- \subset E \subset E_+, F_- \subset F \subset F_+, \mu(E_+ - E_-) = \mu(F_+ - F_-) = 0$. Since by Proposition 1.5.1, the restriction of $\overline{\mu}$ on $\mathcal{A}$ is μ, the last relation can also be written as $\overline{\mu}(E_+ - E_-) = \overline{\mu}(F_+ - F_-) = 0$. So, $\overline{\mu}(E - E_-) = \overline{\mu}(F_+ - F) = 0$

Recall the sets E_-, E, F, F_+ all belong to $\mathcal{A}_\mu$. We have

$$\begin{aligned}\mu(F_+ - E_-) &= \overline{\mu}(F_+ - E_-)\\ &= \overline{\mu}\big((F_+ - F) \cup (F - E) \cup (E - E_-)\big)\\ &\leq \overline{\mu}(F_+ - F) + \overline{\mu}(F - E) + \overline{\mu}(E - E_-) = 0,\end{aligned}$$

which means we have found E_-, F_+ that belong to $\mathcal{A}$ and "squeeze" A with $\mu(F_+ - E_-) = 0$. Therefore, $A \in \mathcal{A}_\mu$.

1.5.3
(a). Consider $X = \{0, 1, 2\}$ and $\mathcal{A} = \{\emptyset, \{0, 1, 2\}, \{0\}, \{1, 2\}\}$. Let μ be the counting measure, and let $\nu = 0$. Then $\mathcal{A}_\nu$ is the power set of X, which is different than $\mathcal{A}_\mu$.
(b). This is wrong. If $\mathcal{A}$ is already the power set of X, then no matter what μ and ν are, $\mathcal{A}_\mu = \mathcal{A}_\nu$.

1.5.4
Let X denote a (Lebesgue) non-measurable set. Consider $X \times \{0\}$. Notice this set is contained in $\mathbb{R} \times \{0\}$ which has Lebesgue measure zero (see exercise 1.4.1). Since Lebesgue measure is complete on the sigma algebra of all Lebesgue measurable sets, $X \times \{0\}$ is measurable and has measure zero.

1.5.5
By the definition of inner measure, there exists a sequence of sets $(B_n) \subset \mathcal{A}$ such that $B_n \subset A, \forall n$ and $\mu(B_n) > \mu_*(A) - 1/n$. Take $A_0 = \cup_n B_n$. The monotonicity of measure implies that $\mu(A_0) \geq \mu_*(A)$. Also since $A_0 \subset A$, the definition of inner measure implies $\mu(A_0) \leq \mu_*(A)$. Therefore, $\mu(A_0) = \mu_*(A)$. By a symmetric argument using the definition of outer measure, there is an $A_1 \supset A, \mu(A_1) = \mu^*(A)$.

1.5.8
Take $B = [0, 1] \cap A$ where A is the set constructed in Proposition 1.4.11. By the remark after Proposition 1.5.4, we know $\lambda_*(B) \leq \lambda_*(A) = 0 \Rightarrow \lambda_*(B) = 0$. Now because $B \subset [0, 1]$, by the monotonicity of outer measure we know $\lambda^*(B) \leq 1$. Suppose $\lambda^*(B) < 1$. Then by the definition of outer measure, there is a cover of B by countably many bounded open intervals, $\{(a_i, b_i)\}$, such that $\sum_{i=1}^{\infty} b_i - a_i < 1$. This means, even if those open intervals do not overlap with each other, they still cannot cover the whole interval $(0, 1)$ which has volume 1, and when this is the case, the set $[0, 1] - \cup_i(a_i, b_i)$ contains some small open interval I (some drawing might help understand). When some of the intervals of this open cover do overlap, the part of $[0, 1]$ that can be covered by $\cup_i(a_i, b_i)$ is even less, and again $[0, 1] - \cup_i(a_i, b_i)$ contains some small open interval I. Now in 1.4.11, as $A := E + G_0$ and G_0 is dense in $\mathbb{R}$, we know B is dense in $[0, 1]$. Therefore, these is some

element of B in I, which isn't covered by $\cup_i(a_i, b_i)$, a contradiction. So $\lambda^*(B) = 1$.

1.5.11

(a). Trivial.

(b). By definition of outer measure, since $A_1 \cap C_1 \supset A_1 \cap C$, we have $\mu^*(A_1 \cap C) \leq \mu(A_1 \cap C_1)$. We now show the inequality is actually an equality. Suppose it is instead a strict inequality, i.e., LHS < RHS, then there exists $B \supset A_1 \cap C, \mu(B) < \mu(A_1 \cap C_1)$. Note $C \subset B \cup (C - A_1) \subset B \cup (C_1 - A_1)$. But $\mu(B \cup (C_1 - A_1)) \leq \mu(B) + \mu(C_1 - A_1) < \mu(A_1 \cap C_1) + \mu(C_1 - A_1) = \mu(C_1)$, which contradicts that $\mu(C_1) = \mu^*(C)$. Therefore, $\mu^*(A_1 \cap C) = \mu(A_1 \cap C_1)$. The same argument for $A_2 \cap C_1$ shows that $\mu^*(A_2 \cap C) = \mu(A_2 \cap C_1)$.

(c). Say $B = A \cap C$ for some $A \in \mathcal{A}$. We want to show, for some C_1 as in part (b), $\mu(A \cap C_1) = \mu^*(A \cap C)$. But this is done in the beginning of the proof for part (b).

(d). The only nontrivial property to check is countable additivity. For a sequence $(B_n) \subset \mathcal{A}_C$ of disjoint sets, suppose $B_n = A_n \cap C$ for some $A_n \in \mathcal{A}, \forall n$. Then for some C_1 as in (b), (c), $\mu_C(\cup_n B_n) = \mu_C((\cup_n A_n) \cap C) := \mu((\cup_n A_n) \cap C_1) = \sum_n \mu(A_n \cap C_1) =: \sum_n \mu_C(A_n \cap C)$, but RHS is exactly $\sum_n \mu_C(B_n)$.

1.5.12

(a). Define $\mathcal{B} := \{(A_1 \cap C) \cup (A_2 \cap C^c) : A_1, A_2 \in \mathcal{A}\}$. We first check $\mathcal{B}$ is a sigma-algebra. Take $X = A_1 = A_2$, we see $X \in \mathcal{B}$. It is also easy to check $\mathcal{B}$ is closed under countable union. For $S = (A_1 \cap C) \cup (A_2 \cap C^c)$, notice

$$\begin{aligned}
S^c &= (A_1^c \cup C^c) \cap (A_2^c \cup C) \\
&= (A_1^c \cap C) \cup (A_2^c \cap C^c) \cup (A_1^c \cap A_2^c) \\
&= (A_1^c \cap C) \cup (A_2^c \cap C^c) \cup (A_1^c \cap A_2^c \cap C) \cup (A_1^c \cap A_2^c \cap C^c) \\
&= (A_1^c \cap C) \cup (A_2^c \cap C^c)
\end{aligned}$$

(b). Notice we have justified the existence of such C_0 and C_1 in exercise 1.5.5. We show only that μ_0 is a measure, as μ_1 is similar. Nonnegativity and that $\mu_0(\varnothing) = 0$ are obvious. For a sequence of disjoint sets $(A^n) \subset \sigma(\mathcal{A} \cup \{C\})$, we recall the fact shown in (a). Then we have that

$$\begin{aligned}
\mu_0(\cup_n A^n) &= \mu\big((\cup_n A_1^n) \cap C_0\big) + \mu_{C^c}\big((\cup_n A_2^n) \cap C^c\big) \\
&= \big(\sum_n \mu(A_1^n \cap C_0)\big) + \big(\sum_n \mu_{C^c}(A_2^n \cap C^c)\big) \\
&= \sum_n \mu(A_1^n \cap C_0) + \mu_{C^c}(A_2^n \cap C^c) \\
&= \sum_n \mu(A^n \cap C_0) + \mu_{C^c}(A^n \cap C^c) = \sum_n \mu_0(A^n).
\end{aligned}$$

This means countable additivity holds.

Now we want to show μ_0 agrees with μ on $\mathcal{A}$. For $A \in \mathcal{A}$,

$$\mu_0(A) = \mu(A \cap C_0) + \mu^*(A \cap C^c)$$

by exercise 1.5.11 (c). To show this equals $\mu(A)$, we can show $\mu(A \cap C_0^c) = \mu^*(A \cap C^c)$. We recall the definition of outer measure, which is the approximation "from above" by measure. Therefore, since $C_0^c \supset C^c$, we have $\mu(A \cap C_0^c) \geq \mu^*(A \cap C^c)$. Now we need to show $\mu(A \cap C_0^c) \leq \mu^*(A \cap C^c)$. How to show this hard direction?

The rest is easy to prove.

1.6.6

(a). Consider a collection $\{\mathcal{C}_i\}_{i \in I}$ where I is an index set, and $\mathcal{C}_i$ are monotone classes that include $\mathcal{A}$. Notice first that the power set of X is a monotone class of X that includes $\mathcal{A}$, therefore, the collection $\{\mathcal{C}_i\}_{i \in I}$ is not empty. It is then easy to argue that $\cap_{i \in I} \mathcal{C}_i$ is what we look for, i.e., $\cap_{i \in I} \mathcal{C}_i = m(\mathcal{A})$.
(b). See online resources.

2.1.3

Define functions $g_n(x) = \frac{f(x+1/n) - f(x)}{1/n}, n \in \mathbb{N}$. Each g_n is measurable as the difference of two continuous functions is continuous therefore measurable. $f' = \lim_n g_n$ is measurable by Proposition 2.1.5.

2.1.4

The set considered is $\{x \in X : \liminf f_n(x) = \limsup f_n(x)\} \cap \{x \in X : \limsup f_n(x) \in \mathbb{R}\}$. The former set appearing in this intersection is measurable because of Propositions 2.1.3, 2.1.5. The latter set equals $\cup_{m \in \mathbb{N}} \{x \in X : \limsup f_n(x) \in (-m, m)\} \in \mathcal{A}$.

2.1.7

(a). We use the notation from example 2.1.10. This is obvious by recalling the definition of f on the set $[0,1] - K$. Take the interval $I = [2/9, 3/9]$ for example. $\mu(I) = \mu((-\infty, 3/9]) - \mu((-\infty, 2/9))$, which by continuity is $F(3/9) - \lim_{y \uparrow 2/9} F(y) = \lim_{y \downarrow 3/9} F(y) - \lim_{y \uparrow 2/9} F(y) = 1/2 - 1/4 = 1/4$.
(b). Recall $K_n \supset K_{n+1}$ and that $\mu(K_n) = 2^n \times 1/2^n = 1$. $\mu(K) = \mu(\cap_n K_n) = \mu(\lim_n K_n) = 1$ by Proposition 1.2.5., the upper continuity of measure.
(c). Fix x. $\mu(\{x\}) = \mu((-\infty, x]) - \mu((-\infty, x)) = F(x) - \lim_{y \uparrow x} F(y) = 0$ because Cantor function is continuous.

2.1.9

Suppose f has the form $f_1\chi_C + f_2\chi_{C^c}$. Notice since $\sigma(\mathcal{A}\cup\{C\}) \supset \mathcal{A}$, both f_1 and f_2 are $\sigma(\mathcal{A}\cup\{C\})$-measurable. As $C \in \sigma(\mathcal{A}\cup\{C\}), \chi_C$ is $\sigma(\mathcal{A}\cup\{C\})$-measurable. Therefore, $f = f_1\chi_C + f_2\chi_{C^c}$ is so.

What about the hard direction?

2.2.1

Let f, g be constant and agree on $\mathbb{Q}$ but constant and disagree on $\mathbb{I}$. Note that singletons are closed sets so are all Borel, and so $\mathbb{Q}$ is Borel. By example 2.1.2. (d), f, g are Borel-measurable.

2.2.3

Suppose $f(x) < g(x)$. Then by continuity of $f - g$, there is a neighborhood (a, b) of x where $f - g < 0$. The interval (a, b) has positive measure, a contradiction.

2.2.5

There exist a negligible set E and a zero-measure $N \supset E$ such that $f_n \to f$ on N^c. For all n, define g_n as f_n on N^c and f on N. $g_n = f_n\chi_{N^c} + f\chi_N$ so are measurable.

2.3.1

Consider $\max(f, g) - g$, which equals $\max(f - g, 0) = (f - g)^+ \in \mathcal{L}^1$. Since $g \in \mathcal{L}^1$ we have that $\max(f, g) = g + (f - g)^+ \in \mathcal{L}^1$. Notice $\min(f, g) = -\max(-f, -g)$.

2.3.2

Take $f = g = \frac{1}{\sqrt{x}}\chi_{(0,1]}(x)$. Then $fg = \frac{1}{x}\chi_{(0,1]}(x)$. f is Riemann integrable on $(0, 1]$ with Riemann integral equal to 2, so its Lebesgue integral is also 2 (see section 2.5). To see fg is not Lebesgue integrable, we define $f_1 = 1\chi_{(0,1]}$, $f_2 = 2\chi_{(0,1/2]} + 1\chi_{(1/2,1]}$, $f_3 = 4\chi_{(0,1/4]} + 2\chi_{(1/4,1/2]} + 1\chi_{(1/2,1]}$, $f_4 = 8\chi_{(0,1/8]} + 4\chi_{(1/8,1/4]} + 2\chi_{(1/4,1/2]} + 1\chi_{(1/2,1]}$... We see that $fg \geq f_n, n = 1, 2, ...$, therefore by Proposition 2.3.4, $\int fg \geq \int f_n = \frac{1+n}{2} \Rightarrow \int fg = \infty$.

2.3.3

We follow the hint. Let A be the Lebesgue-nonmeasurable subset of $(0, 1)$, and let B be $(0, 1) - A$. To see that f is not Lebesgue integrable, note that $f^+ = \chi_A$ which is not event measurable (hence not integrable) by Examples 2.1.2 (d).

2.4.2

We notice from the relation $f_1 \leq f_2 \leq ...$ and $f = \lim f_n$ both holding a.e., we have that a.e.,

$0 \leq f_1^+ \leq f_2^+ \leq ... \leq f^+$, $f_1^- \geq f_2^- \geq ... \geq f^- \geq 0$, and $f^+ = \lim_n f_n^+$, $f^- = \lim_n f_n^-$. We have that

$$\int f^+ = \lim_n \int f_n^+,$$

by Theorem 2.4.1. We also have

$$\infty > \int f^- = \lim_n \int f_n^-$$

by Theorem 2.4.5, where f_1^- serves as the function g in 2.4.5. Also, by Theorem 2.4.5, each f_n^- is integrable, hence a finite integral. Now, we have that

$$\int f = \int f^+ - \int f^- = \left(\lim_n \int f_n^+\right) - \left(\lim_n \int f_n^-\right) = \lim_n \int f_n^+ - \int f_n^- := \lim_n \int f_n.$$

2.4.6

(a). Notice $f(x) = \sum_{n=1}^{\infty} \chi_{\{y:f(y)\geq n\}}(x)$. Define $f_m(x) = \sum_{n=1}^{m} \chi_{\{y:f(y)\geq n\}}(x)$, then the integral of each f_m is $\sum_{n=1}^{m} \mu(\{x : f(x) \geq n\})$. By Theorem 2.4.1 The Monotone Convergence Theorem, we have that $\int f d\mu = \sum_{n=1}^{\infty} \mu(\{x : f(x) \geq n\})$.

(b). Suppose f is integrable. By Corollary 2.3.14, $f < \infty$ a.e. We define f' as a function that agrees with f except that $f'(x)$ takes 0 if $f(x) = \infty$. It's easy to check that f' is measurable. By Proposition 2.3.9, $\int f' = \int f < \infty$. We define $g(x)$ as the floor function $\lfloor f'(x) \rfloor$. Notice that $g = \sum_{n=1}^{\infty} n\chi_{f'(x)\in[n,n+1)}$ and is easily seen to be measurable. We have that $0 \leq g \leq f$. By Proposition 2.3.4, we know $\int g$ is finite, and since g is integer-valued, by part (a) we know

$$\begin{aligned} \infty > \int g d\mu = \sum_{n=1}^{\infty} \mu(\{x : g(x) \geq n\}) &\geq \sum_{n=1}^{\infty} \mu(\{x : f(x) - 1 \geq n\}) \\ &= \sum_{n=2}^{\infty} \mu(\{x : f(x) \geq n\}) \Rightarrow \\ \sum_{n=1}^{\infty} \mu(\{x : f(x) \geq n\}) < \infty, \end{aligned}$$

where the last inequality is because μ is a finite measure so $\mu(\{x : f(x) \geq 1\}) < \infty$.

Now we turn to the other direction. Suppose the series $\sum_{n=1}^{\infty} \mu(\{x : f(x) \geq n\})$ is convergent. We define $h = \lceil f \rceil$ the ceiling function composed with f, and notice that $f \leq h \leq f + 1$. Then since μ is finite, we know $\int h = \sum_{n=1}^{\infty} \mu(\{x : h(x) \geq n\}) \leq \sum_{n=1}^{\infty} \mu(\{x : f(x) + 1 \geq n\}) = \sum_{n=0}^{\infty} \mu(\{x : f(x) \geq n\})$, which is convergent. Therefore, $\int f \leq \int h < \infty$.

2.4.7

(a) That this mapping is nonnegative and assigns zero to $\emptyset$ is easy to check. The countable additivity of this mapping can be checked applying the Monotone Convergence Theorem.
(b), (c). First suppose $f = \chi_B$ for some $B \in \mathcal{B}$. Then the mapping in (b) is simply $K(x, B)$ which is $\mathcal{A}$-measurable so (b) holds. It's also trivial to check that (c) holds. Now suppose f is a nonnegative simple function that is finite, i.e., $f = \sum_{n=1}^m \alpha_n \chi_{B_n}, \alpha_n > 0, \alpha_n \in \mathbb{R}, B_n \in \mathcal{B}$. By Proposition 2.3.4, Proposition 2.1.6, (b) holds. By Proposition 2.3.4, we can also check that (c) holds. Finally suppose $f \in [0, \infty]$. By Proposition 2.1.8, find a sequence (f_n) of nonnegative simple functions that are finite. By Proposition 2.4.1, Proposition 2.1.5, we know that (b) holds. By Proposition 2.4.1, it can also be seen that (c) holds.

2.4.8

(a), (b). First suppose $f = \chi_B$ for some $B \in \mathcal{B}$. By (ii), the measurability holds. By Proposition 2.3.4, $\int \chi_B K(x, dy) \leq \int \chi_Y K(x, dy) = K(x, Y) \geq 0$. Also by the assumption that $\sup_x \{K(x, Y)\} < \infty$, we know $\int \chi_B K(x, dy)$ is uniformly bounded for all x so (a) holds. It's easy to check (b) holds. Then suppose f is a real valued simple function, i.e., $f = \sum_{n=1}^m \alpha_n \chi_{B_n}, \alpha_n \neq 0, \alpha_n \in \mathbb{R}, B_n \in \mathcal{B}$. By Proposition 2.3.6, we know $\int f(y) K(x, dy) = \sum_{n=1}^m \alpha_n \int \chi_{B_n} K(x, dy)$, and by Proposition 2.1.7, the mapping from x into $\int f(y) K(x, dy)$ is measurable. To see boundedness, note that

$$\sum_{n=1}^m \alpha_n \int \chi_{B_n} K(x, dy) \leq \sum_{n=1}^m \alpha_n \int \chi_Y K(x, dy) = \sum_{n=1}^m \alpha_n K(x, Y),$$

which is bounded by the assumption $\sup_x \{K(x, Y)\} < \infty$. This means (a) holds. Now we check for (b). Recall that we have shown for any $B \in \mathcal{B}$,

$$\int \chi_B K(x, dy) = K(x, B)$$

is uniformly bounded for all x. Therefore,

$$\int \chi_B \nu(dy) = \nu(B) = \int K(x, B) \mu(dx) < \infty,$$

because of the uniform boundedness and that μ is finite. In other words, characteristic functions are integrable under the measure ν. Therefore, by Proposition 2.3.6,

$$\int f(y) \nu(dy) = \sum_{n=1}^m \alpha_n \int \chi_{B_n} \nu(dy) = \sum_{n=1}^m \alpha_n \int K(x, B_n) \mu(dx),$$

while recall $\int f(y)K(x,dy) = \sum_{n=1}^m \alpha_n \int \chi_{B_n} K(x,dy)$ so

$$\int \Big(\int f(y)K(x,dy) \Big) \mu(dx) = \int \sum_{n=1}^m \alpha_n \int \chi_{B_n} K(x,dy)\mu(dx) = \int \sum_{n=1}^m \alpha_n K(x,B_n)\mu(dx).$$

Now we only need to justify the exchange of summation and integration. But recall we already showed that $K(x,B_n) = \int \chi_B K(x,dy) \le \int \chi_Y K(x,dy) = K(x,Y) \le \sup_x\{K(x,Y)\} < \infty$, and so $K(x,B_n)$ is actually uniformly bounded for all x, B_n, and therefore integrable under the finite μ. Now the exchange of integration and summation is justified by Proposition 2.3.6.

Now suppose f is a measurable function bounded by $M > 0$. Notice since μ is finite, the constant function M is integrable. By the remark after Proposition 2.1.8, there is a sequence of real valued simple functions (f_n) whose pointwise limit is f. Without loss of generality, M can be chosen to be large so that $M \ge |f_n|, \forall n$. By Theorem 2.4.5, Proposition 2.1.5, (a) holds. We skip the details in checking (b), but it's worth noticing that by Theorem 2.4.5 (DCT), we have $\int f_n K(x,dy) \to \int f K(x,dy)$, and by DCT again,

$$\int \lim \Big(\int f_n K(x,dy) \Big) \mu(dx) = \lim \int \Big(\int f_n K(x,dy) \Big) \mu(dx).$$

2.4.9

Take any sequence $(t_n) \in [0,\infty), (t_n) \uparrow \infty$. By DCT, it can be seen $\int f = \lim_n f_{t_n}$. Therefore, $\int f = \lim_{t\to\infty} f_t$.

2.4.11

Since it's written as $|f_n - f|$, we have to assume that f_n, f never both take the value ∞ at the same x. By Corollary 2.3.14, and some simple arguments, it's easily seen that $f_n \to f, f_n < \infty, \forall n, f < \infty$ hold a.e. Assume they don't all hold on possible a null set E^c, so they all hold on E. Notice

$$|f - f_n|^+ = \max\{0, f - f_n\} \le f$$

holds on E, and notice $|f - f_n|^+ \to 0$ on E. By DCT,

$$\lim_n \int |f - f_n|^+ = \int \lim_n |f - f_n|^+ = \int 0 = 0.$$

It's given that $\int f = \lim_n \int f_n$, and since $f_n, f \in \mathcal{L}^1$, we can subtract $\int f$ from both sides and get

$$0 = \lim_n \int f_n - \int f = \lim_n \int f_n - f$$

so $0 = \lim_n \int f - f_n = \lim_n \int |f - f_n|^+ - |f - f_n|^-$. Now recall $\lim_n \int |f - f_n|^+ = 0$, and we know $\lim_n |f - f_n| = 0$.

2.5.3

We follow the hint and consider such B. In Proposition 2.1.11 we saw B is not a Borel set and so χ_B is not Borel measurable (Examples 2.1.2 (d)). Recall we denote the Cantor set by K, and χ_B is constant, i.e., zero, on the open set $\mathbb{R} - K$. In other words, for any $x \in \mathbb{R} - K$, find an open neighborhood of x that is contained in $\mathbb{R} - K$, then χ_B is constant in this neighborhood, so is continuous at x. Therefore, if we consider the interval $[0, 1]$, then χ_B is continuous almost everywhere there (except for possibly K, which has measure zero). By Theorem 2.5.4, it is Riemann integrable on $[0, 1]$.

2.5.5

Consider $f_n := f\chi_{[a_n,b_n]}$ where $(a_n) \searrow -\infty, (b_n) \nearrow \infty$. Then $f_n \to f, |f_n| \leq |f|$. By DCT, we have that $\int f d\lambda = \lim_n \int f_n d\lambda$.

2.5.8

$\frac{n}{n+j} = \frac{1}{1+j/n}$. The mean of these numbers as j ranges from 1 to n equals the Riemann sum of this integral, with subinterval lengths of $[a_{i-1}, a_i]$ being uniformly $1/n$ and tags x_i being i/n. By Proposition 2.5.7, we have the desired result.

2.6.1

By Example 2.6.5, we know the mapping $x \mapsto (f(x), g(x))$ is measurable. Consider the map from $\mathbb{R}^2$ into $\mathbb{R}$ defined by $h((x, y)) = x + y$, which is easily seen to be continuous, and hence the measurability by Example 2.1.2 (a). The composition map $x \mapsto h(f(x), g(x)) = f(x)+g(x)$ is therefore measurable. Similar arguments hold for fg.

2.6.2

Identify the mapping $x \mapsto f(x) = f_1(x) + if_2(x)$ with $x \mapsto (f_1(x), f_2(x))$, which is measurable by Example 2.6.5. Consider the map from $\mathbb{R}^2$ into $\mathbb{R}$ defined by $h((x, y)) = \sqrt{x^2 + y^2}$, which is continuous so is measurable. The composition map $x \mapsto |f(x)| = h(f(x)) = \sqrt{f_1^2(x) + f_2^2(x)}$ is therefore measurable.

2.6.3

Suppose $f(x) = f_1(x)+if_2(x), g(x) = g_1(x)+ig_2(x), f_1, f_2, g_1, g_2 \in \mathbb{R}$. Then $(f/g)(x) = \frac{f_1+if_2}{g_1+ig_2}(x) = \frac{(f_1+if_2)(g_1-ig_2)}{g_1^2+g_2^2}(x) = (\frac{f_1g_1+f_2g_2}{g_1^2+g_2^2}(x), \frac{f_2g_1-f_1g_2}{g_1^2+g_2^2}(x))$, which is measurable by Example 2.6.5, Propositions 2.1.6, 2.1.7.

2.6.4 (a)

We first check that $\mathcal{B}(\overline{\mathbb{R}})$ includes all sets in this form. In the following, $t \in \mathbb{R}$. $[-\infty, t] = (-\infty, t] \cup \{-\infty\}$, and so by the definition of $\mathcal{B}(\overline{\mathbb{R}})$ on page 74, is included in $\mathcal{B}(\overline{\mathbb{R}})$. Consider any sigma algebra $\mathcal{S}$ on $\overline{\mathbb{R}}$ that includes all sets of the form $[-\infty, t]$, we will check that $\mathcal{S} \supset \mathcal{B}(\overline{\mathbb{R}})$. Consider the sequence $[-\infty, -n], n = 1, 2, ...$ and their intersection, we can see that $\{-\infty\} \in \mathcal{S}$. Consider the union of such sequence, we see $[-\infty, \infty)$ and therefore $\{\infty\}$, is in $\mathcal{S}$. Then, $(-\infty, t] = [-\infty, t] \cap \{-\infty\}^c \in \mathcal{S}$. And, since $\mathcal{B}(\mathbb{R})$ is generated by sets of the form $(-\infty, t]$, we know $\mathcal{S} \supset \mathcal{B}(\mathbb{R})$. Now we have that $\mathcal{S} \supset \mathcal{B}(\mathbb{R})$ and $\mathcal{S} \supset \{-\infty, \infty\}$, so by definition on page 74 again, we know $\mathcal{S} \supset \mathcal{B}(\overline{\mathbb{R}})$.

2.6.5

Since f is a function and it doesn't make sense for a domain to be empty, we assume X is nonempty. Therefore, Y is also nonempty.

First, notice some identities that are easy to verify by logic (i.e., showing LHS includes RHS and RHS includes LHS): $f^{-1}(Y) = X$, $f^{-1}(\cap_{i\in I} B_i) = \cap_{i\in I} f^{-1}(B_i)$, $f^{-1}(\cup_{i\in I} B_i) = \cup_{i\in I} f^{-1}(B_i)$, and $f^{-1}(B^c) = \big(f^{-1}(B)\big)^c$, where I is an arbitrary index set and here $B_i, B \subset Y$.

Once we've established the above identities, (a) and (b) are trivial.

(c). See more discussions here. We here follow the top rated answer. We denote $\{f^{-1}(C) : C \in \mathcal{C}\}$ by $f^{-1}(\mathcal{C})$ and $\{f^{-1}(B) : B \in \sigma(\mathcal{C})\}$ by $f^{-1}(\sigma(\mathcal{C}))$. That the RHS $f^{-1}(\sigma(\mathcal{C}))$ includes $f^{-1}(\mathcal{C})$ is obvious. Therefore, $f^{-1}(\sigma(\mathcal{C})) \supset \sigma\big(f^{-1}(\mathcal{C})\big)$. Now we turn to the other direction of inclusion.

Let $\mathcal{D} := \{A \subset Y : f^{-1}(A) \in \sigma\big(f^{-1}(\mathcal{C})\big)\}$. By this definition, if a subset A of Y is in $\mathcal{D}$, then $f^{-1}(A) \in \sigma\big(f^{-1}(\mathcal{C})\big)$. Notice $\mathcal{D} \supset \mathcal{C}$. Recalling the identities we showed before, it is easy to check that $\mathcal{D}$ is a sigma-algebra. Therefore, $\mathcal{D} \supset \sigma(\mathcal{C})$. Taking inverse image, we see $f^{-1}(\mathcal{D}) \supset f^{-1}\big(\sigma(\mathcal{C})\big)$. But $f^{-1}(\mathcal{D}) = \{f^{-1}(A) : A \in \mathcal{D}\} \subset \sigma\big(f^{-1}(\mathcal{C})\big)$. This means $f^{-1}\big(\sigma(\mathcal{C})\big) \subset \sigma\big(f^{-1}(\mathcal{C})\big)$.

2.6.6

(a). Recall in chapter 1 we saw such F is bounded, nondecreasing and right-continuous, and that $\lim_{x\to-\infty} F(x) = 0$. That g is nondecreasing is easy to see. For any $x \in (0, \lim_{x\to\infty} F(x))$, considering the limit of F as t goes to infinity, we see there is $t \in \mathbb{R} : F(t) \geq x$. So the set $\{t \in \mathbb{R} : F(t) \geq x\}$ is not empty. We also recall the limit of F as t goes to minus infinity is 0, so there is a threshold $t_0 \in \mathbb{R}$ such that for all $t < t_0$, we have $F(t) < x$. Therefore, the set $\{t \in \mathbb{R} : F(t) \geq x\}$ is not empty and bounded below, so it has a finite infimum, which means $g(x)$ is finite.

Now we show g is Borel-Borel measurable. We claim that for x, the infimum of the set $\{t \in \mathbb{R} : F(t) \geq x\}$, denoted l_x, satisfies $F(l_x) \geq x$, i.e., the infimum itself belongs to this set. To see this, take a sequence (t_n) from this set that decreases to l_x, and recall $F(t_n) := \mu((-\infty, t_n]) \geq x$. Therefore, by the upper continuity of measure (Proposition 1.2.5), we have $F(l_x) \geq x$. So $g(x) = l_x$.

Now, for any real number $t, x \in g^{-1}((t, \infty))$ iff $l_x = g(x) > t$ iff $F(t) < x$. This means

$g^{-1}((t,\infty)) = (F(t),\infty)$ which is a Borel set. Therefore, g is (Borel-Borel) measurable.

(b). Consider a set $B = (-\infty, b], b \in \mathbb{R}$. Now $g^{-1}(B) = g^{-1}\big((b,\infty)^c\big) = \big(g^{-1}((b,\infty))\big)^c = (-\infty, F(b)]$. Now consider $A = (-\infty, a], a \in \mathbb{R}, a < b$. Then $B - A = (a, b]$, and $g^{-1}(B - A) = (F(a), F(b)]$ (see the discussion above about operations inside inverse mapping in solving exercise 2.6.5). Therefore, $\lambda g^{-1}(B-A) = F(b) - F(a)$. Also, it is easy to see that $\mu(B-A) = F(b) = F(a)$. By Corollary 1.6.3 and Proposition 1.1.4, we have the desired result.