Math 245A Note 6

qshuyu

February 2024

1 Selected Exercises in Note 6

Exercise 3

Proof. By countable subadditivity, for every set $A \subset X$, $\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \backslash E)$. By monotonicity, $\mu^*(A \cap E) = 0$ and thus $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \backslash E)$. Hence $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \backslash E)$ and E is Carathéodory measurable with respect to μ^*. $\square$

Exercise 4

Proof. If $E \subset \mathbf{R}^d$ is Carathéodory measurable with respect to Lebesgue outer measure, then E is Lebesgue measurable by Exercise 17 of note 1.

Conversely, let E be Lebesgue measurable. By outer regularity, there exists an open set $U \supset A$ with $m^*(U) \leq m^*(A) + \varepsilon$. Then by additivity of Lebesgue measure,

$$m^*(A \cap E) + m^*(A \backslash E) \leq m^*(U \cap E) + m^*(U \backslash E) = m^*(U) \leq m^*(A) + \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, we conclude that

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

by subadditivity.

Alternatively, we assume first that E is a box. For every $A \subset \mathbf{R}^d$, let $\bigcup_{n=1}^{\infty} B_n \supset A$ be a countable union of disjoint boxes. Since the intersection of boxes are boxes, we can divide those boxes B_n with $B_n \cap (A \cap E) \neq \emptyset$ into two disjoint sub-boxes $B_n \cap E$ and $B_n \backslash E$, thus obtain two disjoint families of boxes that cover $A \cap E$ and $A \backslash E$ respectively. Denote these two families by F_1 and F_2, respectively. By the definition of Lebesgue outer measure, we get

$$m^*(A) = \inf_{\bigcup_{n=1}^{\infty} B_n \supset A} \sum_{n=1}^{\infty} |B_n| = \inf_{\bigcup_{n=1}^{\infty} B_n \supset A} \sum_{B_i' \in F_1} |B_i'| + \inf_{\bigcup_{n=1}^{\infty} B_n \supset A} \sum_{B_j' \in F_2} |B_j'|,$$

where the infimum runs over all countable disjoint union of boxes $\bigcup_{n=1}^{\infty} B_n \supset A$. By definition, the first term on the right hand side is no less than $m^*(A \cap E)$, while the second term on the right hand side is no less than $m^*(A \backslash E)$, and thus we obtain that

$$m^*(A) \geq m^*(A \cap E) + m^*(A \backslash E).$$

In view of subadditivity, we conclude that E is Carathéodory measurable. Now, by Carathéodory lemma and Exercise 14 of note 3, all Borel sets are Carathéodory measurable. Clearly, all sets in the null σ-algebra are Carathéodory measurable. Since the Lebesgue σ-algebra on $\mathbf{R}^d$ is generated by the union of the Borel σ-algebra and the null σ-algebra by Exercise 19 of note 3, we see that Lebesgue measurable sets are Carathéodory measurable. $\square$

Exercise 6

Proof. If $\mathcal{B}$ is a σ-algebra, it clearly is closed under countable disjoint unions. Conversely, since every countable union $\bigcup_{n=1}^{\infty} E_n$ of $\mathcal{B}$-measurable sets can be written as a countable disjoint union $\bigcup_{n=1}^{\infty}(E_n \backslash \bigcup_{m \neq n} E_m)$, we see that $\mathcal{B}$ is a σ-algebra if it is closed under countable disjoint unions. $\square$

Exercise 11

Proof. (1). If $\mu_0(\emptyset) := 0$, then for any disjoint finite sequence $E_1, \ldots, E_n \in \mathcal{B}_0$, viewed as a disjoint infinite sequence $E_1, \ldots, E_n, \emptyset, \ldots \in \mathcal{B}_0$, we have by countable additivity of μ_0 that

$$\mu_0(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu_0(E_i) = \sum_{i=1}^{n} \mu_0(E_i),$$

i.e. μ_0 is finitely additive as well.

(2). By monotonicity and finite additivity of μ_0, we have

$$\mu_0(\bigcup_{n=1}^{\infty} E_n) \geq \mu_0(\bigcup_{n=1}^{N} E_n) = \sum_{n=1}^{N} \mu_0(E_n)$$

for every $N \geq 1$. Hence $\mu_0(\bigcup_{n=1}^{\infty} E_n) \geq \sum_{n=1}^{\infty} \mu_0(E_n)$. By condition,

$$\mu_0(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu_0(E_n).$$

i.e. μ_0 is countably additive within $\mathcal{B}_0$.

(3). If one performs both of the above two relaxations at once, it is possible to have $\mu_0(E)$, $\mu_0(F) > 0$ for a pair of disjoint sets $E, F \in \mathcal{B}_0$ such that $\mu_0(E \cup F) < \mu_0(E) + \mu_0(F)$. □

Exercise 12

Proof. We extend the elementary measure m to the elementary algebra $\mathcal{E}\overline{[\mathbf{R}^d]}$ by setting $m(E) := +\infty$ for every co-elementary set E. Let $E_1, E_2, \ldots \in \mathcal{E}\overline{[\mathbf{R}^d]}$ be a sequence of disjoint sets, such that $\bigcup_{n=1}^{\infty} E_n \in \mathcal{E}\overline{[\mathbf{R}^d]}$. If the E_n are elementary, by lemma 6 of note 1, and countable subadditivity of the Lebesgue outer measure m^*,

$$m(\bigcup_{n=1}^{\infty} E_n) = m^*(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} m^*(E_n) = \sum_{n=1}^{\infty} m(E_n).$$

Note that the first equality holds even if $\bigcup_{n=1}^{\infty} E_n$ is a co-elementary set, since the Lebesgue outer measure of any co-elementary set can be easily shown to be infinite. On the other hand, if some of the E_n (say E_m for some $m \geq 1$) are co-elementary, then by monotonicity of m we have

$$m(\bigcup_{n=1}^{\infty} E_n) \geq m(E_m) = +\infty, \ \sum_{n=1}^{\infty} m(E_n) = +\infty,$$

so in this case $m(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n) = +\infty$. Since in both cases we have

$$m(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} m(E_n),$$

the elementary measure (on the elementary Boolean algebra) is a pre-measure by Exercise 11. □

Exercise 13

Proof. Take $X := \mathbf{N}$ to be the natural number, and $\mathcal{B}_0 := 2^{\mathbf{N}}$ to be the discrete algebra. Let $\mu_0(E) = \mu_0(\emptyset) := 0$ for any finite set $E \subset \mathbf{N}$, and let $\mu_0(E) := 1$ (or any other positive number) for any infinite set $E \subset \mathbf{N}$. Then $\mu_0 : \mathcal{B}_0 :\to [0, +\infty]$ is a finitely additive measure that is not a pre-measure. $\square$

Exercise 15

Proof. Let $E \in \mathcal{B}'$. Since μ' is an extension of μ_0, by monotonicity and countable subadditivity,

$$\mu'(E) \leq \sum_{n=1}^{\infty} \mu'(E_n) = \sum_{n=1}^{\infty} \mu_0(E_n)$$

whenever $E_1, E_2, \ldots \in \mathcal{B}_0$ cover E. Hence $\mu'(E) \leq \mu^*(E)$ for all $E \in \mathcal{B}'$.

Now let $E \in \mathcal{B} \cap \mathcal{B}'$, which is a σ-algebra by Exercise 13 of note 3. $E \in \mathcal{B}'$, thus we have

$$\mu'(E) \leq \mu(E).$$

To show the converse inequality, it suffices to show that

$$\sum_{n=1}^{\infty} \mu_0(E_n) \geq \mu'(E)$$

whenever $E_1, E_2, \ldots \in \mathcal{B}_0$ cover E. By replacing each E_n with the smaller set $E_n \backslash \bigcup_{m=1}^{n-1} E_m$ (which still lies in $\mathcal{B}_0$, and still covers E), we may assume without loss of generality (thanks to the monotonicity of μ_0) that the E_n are disjoint. Similarly, by replacing each E_n with the smaller set $E_n \cap E$ we may assume without loss of generality that the union of the E_n is exactly equal to E. But then the claim follows from the hypothesis that μ' is a countably additive extension of μ_0. $\square$

Exercise 16

Proof. (1). Clearly μ_0 is a finitely additive measure. Let $E_1, E_2, E_3, \ldots \in \mathcal{A}$ be disjoint sets such that $\bigcup_{n=1}^{\infty} E_n$ is in $\mathcal{A}$. If any of the E_n is non-empty, we have $\mu_0(\bigcup_{n=1}^{\infty} E_n) = +\infty = \sum_{n=1}^{\infty} \mu_0(E_n)$. Otherwise the E_n are empty and $\mu_0(\bigcup_{n=1}^{\infty} E_n) = 0 = \sum_{n=1}^{\infty} \mu_0(E_n)$. By definition, μ_0 is thus a pre-measure.

(2). Let $a < b$ be real numbers. $(a, b) = \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b)$, so $(a, b) \in \langle\mathcal{A}\rangle$. Picking $c > b$, then $(a, c)\backslash(b, c) = (a, b] \in \langle\mathcal{A}\rangle$. Picking $a < c < d < b$, then $[a, d) \cup (c, b] = [a, b] \in \langle\mathcal{A}\rangle$. Hence we see that $\langle\mathcal{A}\rangle$ contains the intervals. By Exercise 14 of note 3, the Borel σ-algebra $\mathcal{B}[\mathbf{R}]$ is generated by the intervals, so $\mathcal{B}[\mathbf{R}] \subset \langle\mathcal{A}\rangle$. Clearly, $\mathcal{A} \subset \mathcal{B}[\mathbf{R}]$, so $\langle\mathcal{A}\rangle \subset \mathcal{B}[\mathbf{R}]$ and the two algebras are equal.

(3). Let $E \in \mathcal{B}[\mathbf{R}]$ be a non-empty Borel set. By the construction of the Hahn-Kolmogorov extension, it suffices to show that if $E_1, E_2, \ldots \in \mathcal{A}$ is a sequence covering E, some of the E_n must be non-empty. This clearly holds since otherwise $\bigcup_{n=1}^{\infty} E_n = \emptyset$.

(4). By definition, the general counting measure $c\#$ is countably additive. $\forall E \in \mathcal{A},\ E \neq \emptyset$, E contains at least one non-degenerate interval, and is thus an infinite set, so $c\#(E) = +\infty$. By definition, $c\#(\emptyset) = 0$. Hence we see that $c\#$ is another extension of μ_0 on $\mathcal{B}[\mathbf{R}]$. □

Exercise 17

Proof. (1). Suppose that $\mu(E) < +\infty$. By definition, $\forall n \geq 1$, there exist $F_{n,1}, F_{n,2}, \ldots \in \mathcal{B}_0$ covering E such that $\sum_{m=1}^{\infty} \mu(F_{n,m}) \leq \mu(E) + \frac{1}{n}$. By countable additivity, we have

$$\mu(\bigcup_{m=1}^{\infty} F_{n,m}\backslash E) = \mu(\bigcup_{m=1}^{\infty} F_{n,m}) - \mu(E) \leq \frac{1}{n}.$$

Take $F_n := \bigcup_{m=1}^{\infty} F_{n,m}$, and $F := \bigcap_{n=1}^{\infty} F_n \in \langle\mathcal{B}_0\rangle$. By construction, $F \supset E$, and by monotonicity we have

$$\mu(F\backslash E) \leq \mu(F_n\backslash E) \leq \frac{1}{n}$$

for all $n \geq 1$. Hence $\mu(F\backslash E) = 0$.

Now suppose that $\mu(E) = +\infty$. As μ_0 is σ-finite, there exist $A_1, A_2, \ldots \in \mathcal{B}_0$ such that $X = \bigcup_{n=1}^{\infty} A_n$ with $\mu(A_n)$ finite. Note that

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

where the $E_n \in \mathcal{B}$ have finite measure. Thus, there exists $F_n \in \langle\mathcal{B}_0\rangle$ containing E_n with $\mu(F_n\backslash E_n) = 0$. Take $F := \bigcup_{n=1}^{\infty} F_n$, we then have

$$\mu(F\backslash E) \leq \mu(\bigcup_{n=1}^{\infty} E_n\backslash F_n) \leq \sum_{n=1}^{\infty} \mu(E_n\backslash F_n) = 0,$$

as desired.

(2). By definition there exist $F_1, F_2 \ldots \in \mathcal{B}_0$ covering E such that

$$\sum_{n=1}^{\infty} \mu_0(F_n) \leq \mu(E) + \frac{\varepsilon}{2} < \infty.$$

By replacing each F_n with the smaller set $F_n \backslash \bigcup_{m=1}^{n-1} F_m$ (which still lies in $\mathcal{B}_0$, and still covers E), we may assume without loss of generality (thanks to the monotonicity of μ_0) that the F_n are disjoint. Hence the above inequality implies

$$\mu(\bigcup_{n=1}^{\infty} F_n \backslash E) \leq \frac{\varepsilon}{2}.$$

The series $\sum_{n=1}^{\infty} \mu_0(F_n)$ being convergent, there exists $N \geq 1$ such that

$$\mu(\bigcup_{n=1}^{\infty} F_n \backslash \bigcup_{m=1}^{N} F_m) = \sum_{n=1}^{\infty} \mu_0(F_n) - \sum_{m=1}^{N} \mu_0(F_m) \leq \frac{\varepsilon}{2}.$$

Take $F := \bigcup_{m=1}^{N} F_m$. By the triangle inequality and monotonicity of μ, we get

$$\mu(E \Delta F) = \mu(E \backslash F) + \mu(F \backslash E) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and the claim follows.

(3). Suppose for contradiction that $E \notin \mathcal{B}$. By construction of the Hahn-Kolmogorov theorem, there exists $A \subset X$ such that

$$\mu^*(A) < \mu^*(A \cap E) + \mu^*(A \backslash E).$$

Let $F \in \mathcal{B}_0$ be such that $\mu^*(E \Delta F) \leq \varepsilon$. By the Hahn-Kolmogorov theorem again, we have

$$\mu^*(A) = \mu^*(A \cap F) + \mu^*(A \backslash F).$$

Without loss of generality, assume that $\mu^*(A \cap E) \geq \mu^*(A \cap F)$ (and thus $\mu^*(A \backslash E) \leq \mu^*(A \backslash F)$). Then

$$\mu^*(A \cap E) - \mu^*(A \cap F) \leq \mu^*(A \cap E \backslash A \cap F) \leq \mu^*(E \backslash F) \leq \frac{\varepsilon}{2},$$

and similarly

$$\mu^*(A \backslash F) - \mu^*(A \backslash E) \leq \mu^*(A \cap E \backslash F) \leq \mu^*(E \backslash F) \leq \frac{\varepsilon}{2}.$$

Combine these two inequalities, we get

$$|\mu^*(A \cap E) + \mu^*(A \backslash E) - \mu^*(A)| \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \backslash E)$, a contradiction. We conclude that $E \in \mathcal{B}$. $\square$

Exercise 19

Proof. (1). Let I, J be disjoint intervals that share a common endpoint, such that $I \cup J$ is an interval. Thus we have

$$(a,b)\cup[b,c),\ [a,b)\cup[b,c),\ (a,b)\cup[b,c],\ [a,b)\cup[b,c],\ (-\infty,b)\cup[b,c),$$
$$(-\infty,b)\cup[b,c],\ (a,b)\cup[b,+\infty),\ [a,b)\cup[b,+\infty),\ (-\infty,b)\cup[b,+\infty)$$

as well as

$$(a,b]\cup(b,c),\ [a,b]\cup(b,c),\ (a,b]\cup(b,c],\ [a,b]\cup(b,c],\ (-\infty,b]\cup(b,c),$$
$$(-\infty,b]\cup(b,c],\ (a,b]\cup(b,+\infty),\ [a,b]\cup(b,+\infty),\ (-\infty,b]\cup(b,+\infty),$$

for a total of eighteen cases. By definition, we check that

1.$|(a,b)\cup[b,c)|_F = |(a,c)|_F = F_-(c) - F_+(a) = F_-(b) - F_+(a) + F_-(c) - F_-(b)$
$= |(a,b)|_F + |[b,c)|_F$.
2.$|[a,b)\cup[b,c)|_F = |[a,c)|_F = F_-(c) - F_-(a) = F_-(b) - F_-(a) + F_-(c) - F_-(b)$
$= |[a,b)|_F + |[b,c)|_F$.
3.$|(a,b)\cup[b,c]|_F = |(a,c]|_F = F_+(c) - F_+(a) = F_-(b) - F_+(a) + F_+(c) - F_-(b)$
$= |(a,b)|_F + |[b,c]|_F$.
4.$|[a,b)\cup[b,c]|_F = |[a,c]|_F = F_+(c) - F_-(a) = F_-(b) - F_-(a) + F_+(c) - F_-(b)$
$= |(a,b)|_F + |[b,c]|_F$.
5.$|(-\infty,b)\cup[b,c)|_F = |(-\infty,c)|_F = F_-(c) - F_+(-\infty)$
$= F_-(b) - F_+(-\infty) + F_-(c) - F_-(b)$
$= |(-\infty,b)|_F + |[b,c)|_F$.
6.$|(-\infty,b)\cup[b,c]|_F = |(-\infty,c]|_F = F_+(c) - F_+(-\infty)$
$= F_-(b) - F_+(-\infty) + F_+(c) - F_-(b)$
$= |(-\infty,b)|_F + |[b,c]|_F$.
7.$|(a,b)\cup[b,+\infty)|_F = |(a,+\infty)|_F = F_-(+\infty) - F_+(a)$
$= F_-(b) - F_+(a) + F_-(+\infty) - F_-(b)$
$= |(a,b)|_F + |[b,+\infty)|_F$.
8.$|[a,b)\cup[b,+\infty)|_F = |[a,+\infty)|_F$
$= F_-(+\infty) - F_-(a)$
$= F_-(b) - F_-(a) + F_-(+\infty) - F_-(b)$
$= |[a,b)|_F + |[b,+\infty)|_F$.
9.$|(-\infty,b)\cup[b,+\infty)|_F = |(-\infty,+\infty)|_F$
$= F_-(+\infty) - F_+(-\infty)$
$= F_-(b) - F_+(-\infty) + F_-(+\infty) - F_-(b)$
$= |(-\infty,b)|_F + |[b,+\infty)|_F$.

The remaining nine cases follow similarly.

(2). Let $E \in \mathcal{B}_0$. If E is written in two ways $E = \bigcup_{n=1}^{j} I_n = \bigcup_{m=1}^{k} J_m$ as the disjoint union of finitely many intervals, we can combine these two partitions into a common refinement $E = \bigcup_{1 \le n \le j, 1 \le m \le k} I_n \cap J_m$. Then, by finite additivity of the F-volume, we have

$$\mu_0(E) = \sum_{1 \le n \le j; 1 \le m \le k} |I_n \cap J_m|_F = \sum_{n=1}^{j} |I_n|_F = \sum_{m=1}^{k} |J_m|_F.$$

Furthermore, let $E, F \in \mathcal{B}_0$ be disjoint, with $E = \bigcup_{n=1}^{j} I_n$ and $F = \bigcup_{m=1}^{k} J_m$ as disjoint unions. Then by definition of μ_0, we have

$$\mu_0(E \cup F) = \sum_{n=1}^{j} |I_n|_F + \sum_{m=1}^{k} |J_m|_F = \mu_0(E) + \mu_0(F).$$

As such, we see that μ_0 is well-defined.

(3). Let $E \in \mathcal{B}_0$ be an interval. If $E = (a, b)$, let n be large enough that $[a + \frac{1}{n}, b - \frac{1}{n}] \subset (a, b)$. By definition, we have

$$|(a, b)|_F - |[a + \frac{1}{n}, b - \frac{1}{n}]|_F = F_-(b) - F_+(a) - [F_+(b - \frac{1}{n}) - F_-(a + \frac{1}{n})].$$

By the monotone nature of F, $\forall \varepsilon > 0$, $\exists N > 0$ such that

$$F_-(b) - F_+(b - \frac{1}{n}) \le \varepsilon, \quad F_-(a + \frac{1}{n}) - F_+(a) \le \varepsilon$$

for all $n \ge N$. Hence we obtain that

$$|(a, b)|_F - |[a + \frac{1}{n}, b - \frac{1}{n}]|_F \le 2\varepsilon$$

for all $n \ge N$. As $\varepsilon > 0$ is arbitrary, by the Heine-Borel theorem, we obtain the claim when $E = (a, b)$. This implies, by monotonicity, that the claim holds for all bounded intervals E.

Now suppose that E is an unbounded interval. If $E = (a, +\infty)$, then in case that $F_-(+\infty)$ is finite, $\exists M > 0$ sufficiently large such that

$$F_-(+\infty) - F_+(M) \le \varepsilon.$$

By the monotone nature of F, $\exists N > 0$ such that

$$F_-(a + \frac{1}{n}) - F_+(a) \le \varepsilon$$

for all $n \geq N$. Combine these two inequalities, we have

$$|(a, +\infty)|_F - |[a + \frac{1}{n}, M]|_F \leq 2\varepsilon$$

for all $n \geq N$. As $\varepsilon > 0$ is arbitrary, by the Heine-Borel theorem, we obtain the claim when $E = (a, +\infty)$ and $F_-(+\infty)$ is finite. This implies, by monotonicity, that the claim holds when $E = [a, +\infty)$ and $F_-(+\infty)$ is finite. In case that $F_-(+\infty)$ is infinite, $F_+(M)$ can be made arbitrarily large for sufficiently large M, so can $|[a + \frac{1}{n}, M]|_F$, thus we obtain the claim when $E = (a, +\infty)$ or $E = [a, +\infty)$. The same argument gives the claim when $E = (-\infty, a)$ or $E = (-\infty, a]$.

Finally, if $E = (-\infty, +\infty)$ with $F_-(+\infty) = +\infty$ or $F_+(-\infty) = -\infty$ (or both), then $\forall M > 0$, there exists b sufficiently large with $F_+(b) > M$ or a sufficiently small with $F_-(a) < -M$ (or both). In all cases the claim follows. Hence we suppose that both $F_-(+\infty)$ and $F_+(-\infty)$ are finite. By the monotone nature of F, exist $a < b$ such that

$$F_-(+\infty) - F_+(b) \leq \varepsilon, \quad F_-(a) - F_+(-\infty) \leq \varepsilon,$$

so in particular

$$|(-\infty, +\infty)|_F - |[a, b]|_F \leq 2\varepsilon.$$

As $\varepsilon > 0$ is arbitrary, we obtain the full claim, and conclude that

$$\mu_0(E) = \sup_{K \subset E} \mu_0(K)$$

where K ranges over all compact intervals contained in the interval E.

(4). Let E_n be an interval. If $E_n = [a, b]$, then as before, by the monotone nature of F, $\forall \varepsilon > 0$, $\exists N > 0$ such that

$$F_-(b + \frac{1}{n}) - F_+(b) \leq \varepsilon, \quad F_-(a) - F_+(a - \frac{1}{n}) \leq \varepsilon$$

for all $n \geq N$, from which we obtain that

$$|(a - \frac{1}{n}, b + \frac{1}{n})|_F - |[a, b]|_F \leq 2\varepsilon$$

for all $n \geq N$. As $\varepsilon > 0$ is arbitrary, we obtain the claim when $E_n = [a, b]$. This implies, by monotonicity, that the claim holds for all bounded intervals E_n.

Now suppose that E_n is an unbounded interval. If $E_n = [a, +\infty)$, then again by the monotone nature of F, $\forall \varepsilon > 0$, $\exists N > 0$ such that

$$F_-(a) - F_+(a - \frac{1}{n}) \leq \varepsilon$$

for all $n \geq N$. This implies as in (3) that the claim holds when $E_n = [a, +\infty)$. A similar argument applies when $E_n = (-\infty, a]$. Hence we conclude that

$$\mu_0(E_n) = \inf_{U \supset E_n} \mu_0(U)$$

where U ranges over all open intervals containing the interval E_n. $\square$

Exercise 20

Proof. Let $F : \mathbf{R} \to \mathbf{R}$ be a monotone non-decreasing function, and $E \subset \mathbf{R}$ a Borel set. Let $P(E)$ be the property that

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

By Exercise 19, $P(E)$ holds for all intervals E. Clearly $P(\emptyset)$ holds trivially. If E is a bounded interval, then $\mathbf{R} \backslash E = I_1 \cup I_2$ is the disjoint union of two intervals (adopting the convention that intervals can be unbounded). Thus by countable additivity of μ_F, we have

$$\begin{aligned}
\mu_F(I_1 \cup I_2) &= \mu_F(I_1) + \mu_F(I_2) \\
&= \sup_{K_1 \subset I_1, K_1 \text{ compact}} \mu_F(K_1) + \sup_{K_2 \subset I_2, K_2 \text{ compact}} \mu_F(K_2) \\
&= \sup_{K_1 \subset I_1, K_2 \subset I_2, K_1, K_2 \text{ compact}} \mu_F(K_1 \cup K_2) \\
&\leq \sup_{K \subset I_1 \cup I_2, K \text{ compact}} \mu_F(K).
\end{aligned}$$

Combining this with monotonicity, we see that $P(\mathbf{R} \backslash E)$ is true when E is a bounded interval. When E is an unbounded interval, then so is the complement of E, hence we conclude that $P(\mathbf{R} \backslash E)$ is true for all intervals E.

Finally, let $E_1, E_2, \ldots$ be disjoint intervals. By writing any unbounded interval as a countable union of bounded ones, we can further assume that the E_n are bounded. Then, by countable additivity,

$$\sup_{K \subset \bigcup_{n=1}^\infty E_n, K \text{ compact}} \mu_F(K) = \sum_{n=1}^\infty \sup_{K \subset \bigcup_{n=1}^\infty E_n, K \text{ compact}} \mu_F(K \cap E_n).$$

On the other hand, we have

$$\mu_F(\bigcup_{n=1}^\infty E_n) = \sum_{n=1}^\infty \mu_F(E_n) = \sum_{n=1}^\infty \sup_{K_n \subset E_n, K_n \text{ compact}} \mu_F(K_n).$$

For every compact set $K_n \subset E_n$, $K_n \subset \bigcup_{n=1}^{\infty} E_n$ and $K_n = K_n \cap E_n$, and so we conclude that

$$\mu(\bigcup_{n=1}^{\infty} E_n) \leq \sup_{K \subset \bigcup_{n=1}^{\infty} E_n, K \text{ compact}} \mu_F(K).$$

Combining this with monotonicity, we see that $P(\bigcup_{n=1}^{\infty} E_n)$ holds. By remark 4 of note 3, we conclude that $P(E)$ holds for every Borel set E. That is, μ_F is inner regular.

In a similar spirit, we can show that μ_F is outer regular. Let $K \subset \mathbf{R}$. By the Heine-Borel theorem, K is closed and bounded. In particular, there exists $M > 0$ such that $K \subset [-M, M]$, by monotonicity, we have $\mu_F(K) \leq |[-M, M]|_F < \infty$. Thus, the Lebesgue-Stieltjes measure μ_F is a Radon measure on $\mathbf{R}$.

Conversely, let μ be a Radon measure on $\mathbf{R}$. Define $F : \mathbf{R} \to \mathbf{R}$ to be such that

$$F(x) := \begin{cases} \mu((0, x]) & x \geq 0 \\ -\mu((x, 0]) & x < 0 \end{cases}.$$

By construction, F is monotone non-decreasing. We further claim that F is right-continuous. When $x > 0$, by downwards monotone convergence, we have

$$\begin{aligned} F(x) = \mu((0, x]) = \mu(\bigcap_{n=1}^{\infty} (0, x + \frac{1}{n}]) = \lim_{n \to \infty} \mu((0, x + \frac{1}{n}]) = \lim_{n \to \infty} F(x + \frac{1}{n}) \\ = F_+(x). \end{aligned}$$

When $x < 0$, by upwards monotone convergence, we have

$$\begin{aligned} F(x) = -\mu((x, 0]) = -\mu(\bigcup_{n=1}^{\infty} (x + \frac{1}{n}, 0]) = \lim_{n \to \infty} -\mu((x + \frac{1}{n}, 0]) = \lim_{n \to \infty} F(x + \frac{1}{n}) \\ = F_+(x). \end{aligned}$$

When $x = 0$, by definition of F and downwards monotone convergence,

$$F_+(0) = \lim_{n \to \infty} F(\frac{1}{n}) = \lim_{n \to \infty} \mu((0, \frac{1}{n}]) = \mu(\bigcap_{n=1}^{\infty} (0, \frac{1}{n}]) = \mu(\emptyset) = F(0).$$

Now, $\forall x \in \mathbf{R}$, by downwards monotone convergence again, we get

$$\mu(\{x\}) = \mu(\bigcap_{n=1}^{\infty} (x - \frac{1}{n}, x]) = \lim_{n \to \infty} \mu((x - \frac{1}{n}, x]).$$

Observe that

$$\lim_{n \to \infty} \mu((x - \frac{1}{n}, x]) = \begin{cases} F(x) - \lim_{n \to \infty} F(x - \frac{1}{n}) & x \geq 0 \\ \lim_{n \to \infty} -F(x - \frac{1}{n}) + F(x) & x < 0. \end{cases}$$

Combining this with the right-continuity of F, we see that $\forall x \in \mathbf{R}$,

$$\mu(\{x\}) = F_+(x) - F_-(x).$$

Pick $-\infty < a < b < +\infty$. If $a < 0 \leq b$, by additivity, we obtain that

$$\begin{aligned}
\mu([a,b)) &= \mu(\{a\} \cup (a,0] \cup (0,b] \backslash \{b\}) \\
&= F_+(a) - F_-(a) - F(a) + F(b) - (F_+(b) - F_-(b)) \\
&= F_+(a) - F_-(a) - F_+(a) + F_+(b) - F_+(b) + F_-(b) \\
&= F_-(b) - F_-(a).
\end{aligned}$$

Likewise, $\mu((a,b]) = \mu((a,0] \cup (0,b]) = -F(a) + F(b) = F_+(b) - F_+(a)$. And

$$\begin{aligned}
\mu((a,b)) &= \mu((a,0] \cup (0,b] \backslash \{b\}) \\
&= -F(a) + F(b) - (F_+(b) - F_-(b)) \\
&= -F_+(a) + F_+(b) - F_+(b) + F_-(b) \\
&= F_-(b) - F_+(a).
\end{aligned}$$

Finally,

$$\begin{aligned}
\mu([a,b]) &= \mu(\{a\} \cup (a,0] \cup (0,b]) = F_+(a) - F_-(a) - F(a) + F(b) \\
&= F_+(a) - F_-(a) - F_+(a) + F_+(b) \\
&= F_+(b) - F_-(a).
\end{aligned}$$

The same argument applies to the cases $a < b \leq 0$ and $0 \leq a < b$ as well. By theorem 18, we conclude that $\mu = \mu_F$. □

Exercise 21

Proof. Suppose that $\mu_F = \mu_{F'}$, where $F, F' : \mathbf{R} \to \mathbf{R}$ are monotone non-decreasing functions. For any $-\infty < x < y < +\infty$, we have $|(x,y)|_F = |(x,y)|_{F'}$. That is,

$$F_-(y) - F_+(x) = F'_-(y) - F'_+(x).$$

Equivalently, $F_-(y) - F'_-(y) = F_+(x) - F'_+(x)$. Similarly, from $|[x,y]|_F = |[x,y]|_{F'}$, we get

$$F_+(y) - F_-(x) = F'_+(y) - F'_-(x).$$

Equivalently, $F_+(y) - F'_+(y) = F_-(x) - F'_-(x)$. On the other hand, from $|[x,y)|_F = |[x,y)|_{F'}$,

$$F_-(y) - F_-(x) = F'_-(y) - F'_-(x).$$

Equivalently, $F_-(y) - F'_-(y) = F_-(x) - F'_-(x)$. From these equalities, we thus get

$$F_-(y) - F'_-(y) = F_-(x) - F'_-(x) = F_+(x) - F'_+(x) = F_+(y) - F'_+(y).$$

Denote this common value by C, then we conclude that

$$F_+(x) = F'_+(x) + C, \quad F_-(x) = F'_-(x) + C$$

for all $x \in \mathbf{R}$. Conversely, suppose that there exists such a constant C. Then it's easy to verify that $|I|_F = |I|_{F'}$ for every bounded interval $I \subset \mathbf{R}$, and $\mu_F(\{a\}) = \mu_{F'}(\{a\})$ for all $a \in \mathbf{R}$. By theorem 18, we obtain that $\mu_F = \mu_{F'}$. $\square$

Exercise 22

Proof. (1). By definition, if $F(x) = x$, then $\mu_F(I) = |I| = m(I)$ for any bounded interval I, and $\mu_F(\{a\}) = 0 = m(a)$ for all $a \in \mathbf{R}$. By theorem 18, we see that $\mu_F = m$ is the Lebesgue measure.

(2). By theorem 18, it suffices to verify that $m_{F'} = \mu_F$ on the bounded intervals as well as the singleton sets. Let $-\infty < a < b < +\infty$. Then

$$\int_{[a,b]} F'(x)\ dx = F(b) - F(a)$$

by the second fundamental theorem for absolutely continuous functions. Note that $F(x) = F_-(x) = F_+(x)$ for all x by continuity, which implies that

$$\int_{\{a\}} F'(x)\ dx = \int_{[a,a]} F'(x)\ dx = F_+(a) - F_-(a) = \mu_F(\{a\}) = 0$$

for all $a \in \mathbf{R}$. By finite additivity of the integral, we conclude that $m_{F'} = \mu_F$ on every bounded interval, and the claim follows.

As such, to show that for any unsigned Borel measurable $f : \mathbf{R} \to [0, +\infty]$,

$$\int_{\mathbf{R}} f(x)\ d\mu_F(x) = \int_{\mathbf{R}} f(x) F'(x)\ dx,$$

is equivalent to showing that

$$\int_{\mathbf{R}} f(x)\ dm_{F'}(x) = \int_{\mathbf{R}} f(x) F'(x)\ dx.$$

By definition, $f = \sup_n f_n(x)$ is the supremum of an increasing sequence $0 \leq f_1 \leq f_2 \leq \ldots$ of unsigned simple functions. Let $f_n = \sum_{i=1}^{k} c_i 1_{E_i}$, where $c_i \geq 0$ and the E_i are Borel measurable sets. Then

$$\int_{\mathbf{R}} f_n(x)\ dm_{F'}(x) = \sum_{i=1}^{k} c_i \int_{E_i} F'(x)\ dx = \int_{\mathbf{R}} f_n(x)F'(x)\ dx.$$

Since F' is non-negative, $0 \leq f_1 F' \leq f_2 F' \leq \ldots \leq fF'$, hence by the monotone convergence theorem,

$$\begin{aligned}\int_{\mathbf{R}} f(x)\ dm_{F'}(x) &= \lim_{n\to\infty} \int_{\mathbf{R}} f_n(x)\ dm_{F'}(x) = \lim_{n\to\infty} \int_{\mathbf{R}} f_n(x)F'(x)\ dx \\ &= \int_{\mathbf{R}} f(x)F'(x)\ dx,\end{aligned}$$

as desired. $\square$

Exercise 23

Proof. (1). By construction, $H_+(0) = 1,\ H_-(0) = 0$. Hence for any bounded interval $I \subset \mathbf{R}$, we have $\mu_H(I) = \begin{cases} 1 & 0 \in I \\ 0 & 0 \notin I \end{cases}$, which agrees with $\delta_0(I) = 1_I(0)$. Also, it's easy to verify that $\delta_0(\{a\}) = H_+(a) - H_-(a)$ for all $a \in \mathbf{R}$. By theorem 18, we conclude that $\mu_H = \delta_0$.

(2). Let J_n be a basic jump function with point of discontinuity x_n, and $I \subset \mathbf{R}$ be a bounded interval with endpoints $a < b$. If $x_n \in I$, then $a \leq x_n$, so

$$J_{n\pm}(b) - J_{n\pm}(a) = 1 - 0 = 1,$$

i.e. $|I|_{J_n} = 1$. Clearly we have $|I|_{J_n} = 0$ if $x_n \notin I$. Thus

$$|I|_F = \sum_n c_n |I|_{J_n} = \sum_{n, x_n \in I} c_n.$$

On the other hand, by definition of the Dirac measure,

$$\sum_n c_n \delta_{x_n}(I) = \sum_n c_n 1_I(x_n) = \sum_{n, x_n \in I} c_n.$$

As such, $\forall a \in \mathbf{R}$, viewed as the interval, $[a, a]$, we also get

$$\sum_n c_n \delta_{x_n}(\{a\}) = \sum_{n, x_n = a} c_n = F_+(a) - F_-(a).$$

By theorem 18, we conclude that $\mu_F = \sum_n c_n \delta_{x_n}$ as desired. $\square$

Exercise 24

Proof. (1). By definition, F is continuous iff $\mu_F(\{x\}) = F_+(x) - F_-(x) = 0$ for all $x \in \mathbf{R}$.

(2). Extending the Cantor function outside $[0,1]$ by constants, by part (4) of Exercise 84 in note 5, we get $F_+(-\infty) = 0$ and $F_-(+\infty) = 1$. That is, μ_F is a probability measure. By definition, F is constant on the intervals removed in the construction of the Cantor set, of which there are countably many. By countable additivity, we thus have $\mu_F([0,1]\backslash C) = 0$. Clearly, $\mu_F((1,+\infty)) = \mu_F((-\infty,0)) = 0$, and so we get $\mu_F(\mathbf{R}\backslash C) = 0$.

(3). Let $I \subset [0,1]$ be an interval with endpoints $a < b$. By continuity of F, $\mu_F(I) = F(b) - F(a)$ for all bounded intervals I. Note that $\frac{1}{3} \cdot I \subset [0, \frac{1}{3}]$ is an interval, as well as $\frac{1}{3} \cdot I + \frac{2}{3} \subset [\frac{2}{3}, 1]$. Also, by taking $n \to \infty$ in the iterative definition of F_n, we obtain that

$$F(x) := \begin{cases} \frac{1}{2}F(3x) & \text{if } x \in [0,1/3]; \\ \frac{1}{2} & \text{if } x \in (1/3,2/3); \\ \frac{1}{2} + \frac{1}{2}F(3x-2) & \text{if } x \in [2/3,1] \end{cases}$$

Hence $\mu_F(\frac{1}{3} \cdot I) = F(\frac{b}{3}) - F(\frac{a}{3}) = \frac{1}{2}F(b) - \frac{1}{2}F(a) = \frac{1}{2}\mu_F(I)$, and likewise

$$\mu_F(\frac{1}{3} \cdot I + \frac{2}{3}) = F(\frac{b}{3} + \frac{2}{3}) - F(\frac{a}{3} + \frac{2}{3}) = \frac{1}{2}F(\frac{b}{3}) - \frac{1}{2}F(\frac{a}{3}) = \frac{1}{2}\mu_F(I).$$

i.e. the claim holds for all intervals $I \subset [0,1]$.

Let $E \in \mathcal{B}[[0,1]]$ be such that the claim holds for E. Since

$$\frac{1}{3} \cdot ([0,1]\backslash E) = [0, \frac{1}{3}]\backslash \frac{1}{3} \cdot E,$$

and μ_F is countably additive, we get

$$\begin{aligned} \mu_F(\frac{1}{3} \cdot ([0,1]\backslash E)) &= \mu_F([0, \frac{1}{3}]) - \mu_F(\frac{1}{3} \cdot E) = \frac{1}{2}\mu_F([0,1]) - \frac{1}{2}\mu(E) \\ &= \frac{1}{2}\mu_F([0,1]\backslash E). \end{aligned}$$

If $E_1, E_2 \ldots \in \mathcal{B}[[0,1]]$ is a disjoint sequence for which the claim holds for every E_n, then

$$\mu_F(\frac{1}{3} \cdot \bigcup_{n=1}^{\infty} E_n) = \mu_F(\bigcup_{n=1}^{\infty} \frac{1}{3} \cdot E_n) = \sum_{n=1}^{\infty} \mu_F(\frac{1}{3} \cdot E_n) = \frac{1}{2}\mu_F(\bigcup_{n=1}^{\infty} E_n).$$

Also, from the fact that

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

for any sets A and B and any point x, we have

$$\begin{aligned}\mu_F(\frac{1}{3} \cdot ([0,1] \backslash E) + \frac{2}{3}) = \mu_F([\frac{2}{3}, 1] \backslash (\frac{1}{3} \cdot E + \frac{2}{3})) &= \frac{1}{2}\mu_F([0,1]) - \frac{1}{2}\mu_F(E) \\ &= \frac{1}{2}\mu_F([0,1] \backslash E).\end{aligned}$$

And likewise we have

$$\begin{aligned}\mu_F(\frac{1}{3} \cdot \bigcup_{n=1}^{\infty} E_n + \frac{2}{3}) = \mu_F(\bigcup_{n=1}^{\infty} (\frac{1}{3} \cdot E_n + \frac{2}{3})) &= \sum_{n=1}^{\infty} \mu_F(\frac{1}{3} \cdot E_n + \frac{2}{3}) \\ &= \sum_{n=1}^{\infty} \frac{1}{2}\mu_F(E_n) = \frac{1}{2}\mu_F(\bigcup_{n=1}^{\infty} E_n).\end{aligned}$$

Thus, we see that the claim holds for $[0,1] \backslash E$ and $\bigcup_{n=1}^{\infty} E_n$. Clearly the claim holds vacuously for the empty set. By Exercise 14 and remark 4 of note 3, we conclude that the claim holds for every Borel-measurable $E \subset [0,1]$. $\square$

Exercise 25

Proof. By the triangle inequality, we can take $f : [a,b] \to \mathbf{R}$ to be unsigned. As continuous functions on compact intervals are uniformly continuous, exists $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever $|x - y| \leq \delta$. Let $a = t_0 < t_1 < \ldots < t_n = b$ and $t_i^* \in [t_{i-1}, t_i]$ for $1 \leq i \leq n$ be such that $\sup_{1 \leq i \leq n} |t_i - t_{i-1}| \leq \delta$. Let

$$m_i := \inf_{x \in I_i} f(x),\ M_i := \sup_{x \in I_i} f(x),$$

where $I_i := [t_{i-1}, t_i]$, then by monotonicity, we see that

$$m_i(F(b_i) - F(a_i)) \leq \int_{[t_{i-1}, t_i)} f\ dF,\ f(t_i^*)(F(b_i) - F(a_i)) \leq M_i(F(b_i) - F(a_i))$$

for every $1 \leq i \leq n$. In particular, we have

$$|\int_{[t_{i-1}, t_i)} f\ dF - f(t_i^*)(F(b_i) - F(a_i))| \leq \varepsilon(F(b_i) - F(a_i))$$

for every i. By finite additivity, we obtain that

$$|\int_{[a,b)} f\ dF - \sum_{i=1}^{n} f(t_i^*)(F(b_i) - F(a_i))| \leq \varepsilon(F(b) - F(a)).$$

Since F is continuous at b, $\mu_F(\{b\}) = 0$, and thus

$$|\int_{[a,b]} f\ dF - \sum_{i=1}^{n} f(t_i^*)(F(b_i) - F(a_i))| \leq \varepsilon(F(b) - F(a)).$$

As $\varepsilon > 0$ is arbitrary, the claim follows. □

Exercise 26

Proof. Let $\varepsilon > 0$ be arbitrary. By Exercise 25, we can find $\delta > 0$ such that any partition $a = t_0 < t_1 < \ldots < t_n = b$ of norm $\sup_{1 \leq i \leq n} |t_i - t_{i-1}| \leq \delta$ is such that

$$|\sum_{i=1}^{n} F(t_{i-1})(G(t_i) - G(t_{i-1})) - \int_{[a,b]} F\ dG| \leq \varepsilon/2,$$
$$|\sum_{i=1}^{n} G(t_i)(F(t_i) - F(t_{i-1})) - \int_{[a,b]} G\ dF| \leq \varepsilon/2.$$

By the triangle inequality, this implies that

$$|\sum_{i=1}^{n} F(t_i)G(t_i) - F(t_{i-1})G(t_{i-1}) - (\int_{[a,b]} F\ dG + \int_{[a,b]} G\ dF)| \leq \varepsilon.$$

Or equivalently, $|F(b)G(b) - F(a)G(a) - (\int_{[a,b]} F\ dG + \int_{[a,b]} G\ dF)| \leq \varepsilon.$ □

Exercise 27

Proof. (1). We aim to show that every σ-algebra that contains $\pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y)$ contains $A := \{E \times F : E \in \mathcal{B}_X,\ F \in \mathcal{B}_Y\}$ and vice versa. If $E \in \mathcal{B}_X,\ F \in \mathcal{B}_Y$, then $E \times F$ can be written as $\pi_X^{-1}(E) \cap \pi_Y^{-1}(F) = E \times Y \cap X \times F$, which is an element of $\pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y)$. i.e. Every σ-algebra that contains $\pi_X^*(\mathcal{B}_X) \cup \pi_Y^*(\mathcal{B}_Y)$ contains A. Conversely, $E \times Y,\ X \times F \in A$, and we get the result.

(2). By (1), if $E \in \mathcal{B}_X$, then $\pi_X^{-1}(E) = E \times Y \in \mathcal{B}_X \times \mathcal{B}_Y$. Hence π_X is a measurable morphism. Similarly for π_Y. Let $\mathcal{T}$ be a σ-algebra on $X \times Y$ that makes the projection maps π_X, π_Y both measurable morphisms. Then for every $E \in \mathcal{B}_X, F \in \mathcal{B}_Y,\ \pi_X^{-1}(E), \pi_Y^{-1}(F) \in \mathcal{T}$. Hence $\pi_X^{-1}(E) \cap \pi_Y^{-1}(F) = E \times F \in \mathcal{T}$. That is, $\mathcal{T} \supset \{E \times F : E \in \mathcal{B}_X, F \in \mathcal{B}_Y\}$, so $\mathcal{T} \supset \mathcal{B}_X \times \mathcal{B}_Y$.

(3). Define $\phi_x : Y \to X \times Y$ by $\phi_x(y) := (x,y)$. Let $P(F)$ be the property of sets $F \subset X \times Y$ that $\phi_x^{-1}(F) \in \mathcal{B}_Y$. If $A \in \mathcal{B}_X$ and $B \in \mathcal{B}_Y$, then

$$\phi_x^{-1}(A \times B) = \begin{cases} B, & \text{if } x \in A \\ \emptyset, & \text{otherwise.} \end{cases}$$

In particular, $P(A \times B)$ holds for every $A \in \mathcal{B}_X$ and $B \in \mathcal{B}_Y$. Clearly $P(\emptyset)$ holds since $\phi_x^{-1}(\emptyset) = \emptyset \in \mathcal{B}_Y$. If $E_1, E_2, \ldots \subset X \times Y$ are such that $P(E_n)$ is true for all n, then $\phi_x^{-1}(\bigcup_{n=1}^{\infty} E_n) = \bigcup_{n=1}^{\infty} \phi_x^{-1}(E_n) \in \mathcal{B}_Y$, so $P(\bigcup_{n=1}^{\infty} E_n)$ is true. If $P(E)$ is true for some $E \subset X \times Y$, then $\phi_x^{-1}(E^c) = Y \backslash \phi_x^{-1}(E) \in \mathcal{B}_Y$. By remark 4 of note 3 and (1), $P(E)$ is true for all $E \in \mathcal{B}_X \times \mathcal{B}_Y$. That is, the sets $E_x := \{y \in Y : (x,y) \in E\}$ lie in $\mathcal{B}_Y$ for every $x \in X$, and similarly the sets $E^y := \{x \in X : (x,y) \in E\}$ lie in $\mathcal{B}_X$ for every $y \in Y$.

(4). Let $U \subset [0, +\infty]$ be open. Then $E := f^{-1}(U) \in \mathcal{B}_X \times \mathcal{B}_Y$. By (3), we see that

$$E_x = \{y \in Y : (x,y) \in f^{-1}(U)\} = \{y \in Y : f(x,y) \in U\} = f_x^{-1}(U) \in \mathcal{B}_Y.$$

Hence the function $f_x : y \mapsto f(x,y)$ is $\mathcal{B}_Y$-measurable for every $x \in X$. Similarly the function $f^y : x \mapsto f(x,y)$ is $\mathcal{B}_X$-measurable for every $y \in Y$.

(5). Let $G := \{A \times B : A \in \mathcal{B}_X, B \in \mathcal{B}_Y\}$. We define:

$$\Sigma := \{E \in \mathcal{B}_X \times \mathcal{B}_Y : E \in \langle G_0 \rangle \text{ for some } G_0 \subset G \text{ at most countable}\}.$$

Clearly $\Sigma \subset \mathcal{B}_X \times \mathcal{B}_Y$. We claim that Σ is a σ-algebra. First note that $\emptyset \in \Sigma$. Then, if $E \in \Sigma$, then $E \in \langle G_0 \rangle$ for some $G_0 \subset G$ at most countable, and $E^c \in \langle G_0 \rangle$, so $E^c \in \Sigma$. Finally, if $E_1, E_2, \ldots \in \Sigma$, with $E_n \in \langle G_n \rangle$ for some $G_n \subset G$ at most countable for all n, then $\bigcup_{n=1}^{\infty} E_n \subset \langle \{E_n : n \geq 1\} \rangle$, and thus $\bigcup_{n=1}^{\infty} E_n \in \Sigma$. We conclude that Σ is a σ-algebra. As $G \subset \Sigma$, we get $\mathcal{B}_X \times \mathcal{B}_Y \subset \Sigma$ and thus $\Sigma = \mathcal{B}_X \times \mathcal{B}_Y$.

Fix some $E \in \mathcal{B}_X \times \mathcal{B}_Y$, from what we have shown, E lies in the σ-algebra generated by an at most countable family of sets $(A_n \times B_n)_{n \in \mathbf{N}}$ with $A_n \in \mathcal{B}_X$ and $B_n \in \mathcal{B}_Y$. $\forall x \in X$, let

$$\mathcal{R} := \{S \subset X \times Y : S_x \in \langle B_n : n \in \mathbf{N} \rangle\}$$

We claim that $\mathcal{R}$ is again a σ-algebra. Clearly, $\emptyset \in \mathcal{R}$. If $S \in \mathcal{R}$, then $S_x^c := \{y \in Y : (x,y) \in S^c\} = Y \backslash \{y \in Y : (x,y) \in S\} = Y \backslash S_x \in \langle B_n : n \in \mathbf{N} \rangle$. Finally, if $S^1, S^2, \ldots \in \mathcal{R}$, then $(\bigcup_{n=1}^{\infty} S^n)_x := \{y \in Y : (x,y) \in \bigcup_{n=1}^{\infty} S^n\} = \bigcup_{n=1}^{\infty} \{y \in Y : (x,y) \in S^n\} = \bigcup_{n=1}^{\infty} S_x^n \in \langle B_n : n \in \mathbf{N} \rangle$. Therefore, we conclude that $\mathcal{R}$ is a σ-algebra.

Since the family $(A_n \times B_n)_{n \in \mathbf{N}}$ lie in $\mathcal{R}$, we see that $E \in \mathcal{R}$. By construction, this implies that $E_x \in \langle B_n : n \in \mathbf{N} \rangle$. i.e. there exists an at most countable collection of sets (which can depend on E) such that $\{E_x : x \in X\}$ is a subset of the σ-algebra generated by such collection. $\square$

Exercise 28

Proof. (1). Let $T_X := \{\emptyset_X, X\}, T_Y := \{\emptyset_Y, Y\}$ be the trivial σ-algebras on X and Y respectively. By definition, $T_X \times T_Y = \langle\{E \times F : E \in T_X, F \in T_Y\}\rangle$, which equals $\{\emptyset_{X\times Y}, X \times Y\}$, the trivial algebra on $X \times Y$.

(3). Let B_X and B_Y be two finite σ-algebras on two spaces X and Y. By Exercise 4 of note 3, they are both atomic algebras. That is,

$$\mathcal{B}_X = \mathcal{A}((A_i)_{i=1}^n), \mathcal{B}_Y = \mathcal{A}((B_j)_{j=1}^m),$$

where $X = \bigcup_{i=1}^n A_i, Y = \bigcup_{j=1}^m B_j$ are partitions of X and Y into atoms. Hence, $X \times Y = \bigcup_{i,j} A_i \times B_j$ is a partition of $X \times Y$ into atoms of $\mathcal{B}_X \times \mathcal{B}_Y$. By construction, the atomic algebra

$$\mathcal{A}((A_i \times B_j)_{1\le i\le n, 1\le j\le m})$$

is finite, and is itself a σ-algebra by Exercise 10 of note 3. Also, it contains the family of sets $E \times F : E \in \mathcal{B}_X, F \in \mathcal{B}_Y$, and thus contains the product algebra $\mathcal{B}_X \times \mathcal{B}_Y$. Hence, we conclude that $\mathcal{B}_X \times \mathcal{B}_Y$ is finite.

(4). By Exercise 17 of note 3, if E, F are Borel measurable subsets of $\mathbf{R}^d, \mathbf{R}^{d'}$ respectively, then $E \times F$ is a Borel measurable subset of $\mathbf{R}^{d+d'}$. Hence we have

$$\mathcal{B}[\mathbf{R}^d] \times \mathcal{B}[\mathbf{R}^{d'}] \subset \mathcal{B}[\mathbf{R}^{d+d'}]$$

by (1) of Exercise 27. On the other hand, we have $\mathcal{B}[\mathbf{R}^d] = \langle B_d\rangle$, $\mathcal{B}[\mathbf{R}^{d'}] = \langle B_{d'}\rangle$, where B_d and $B_{d'}$ are the boxes of $\mathbf{R}^d$ and $\mathbf{R}^{d'}$ respectively, by Exercise 14 of note 3. As the product of boxes are boxes, $\mathcal{B}[\mathbf{R}^d] \times \mathcal{B}[\mathbf{R}^{d'}]$ contains boxes of $\mathbf{R}^{d+d'}$. Therefore,

$$\mathcal{B}[\mathbf{R}^d] \times \mathcal{B}[\mathbf{R}^{d'}] \supset \mathcal{B}[\mathbf{R}^{d+d'}]$$

We conclude that $\mathcal{B}[\mathbf{R}^d] \times \mathcal{B}[\mathbf{R}^{d'}] = \mathcal{B}[\mathbf{R}^{d+d'}]$, as desired.

(5). Suppose for contradiction that $\mathcal{L}[\mathbf{R}^d] \times \mathcal{L}[\mathbf{R}^{d'}] = \mathcal{L}[\mathbf{R}^{d+d'}]$. Use (2) of Exercise 18 in note 3: The Cartesian product of any set with a point is a null set, even if the first set was not measurable. Let $E \subset \mathbf{R}^d$ be a non-measurable set, and $a \in \mathbf{R}^{d'}$. Then $F = E \times \{a\}$ is a null set. By Exercise 27(3),

$$E = F^a = \{x \in \mathbf{R}^d : (x, a) \in E \times \{a\}\} \in \mathcal{L}[\mathbf{R}^d],$$

which is false. Thus, $\mathcal{L}[\mathbf{R}^d] \times \mathcal{L}[\mathbf{R}^{d'}] \neq \mathcal{L}[\mathbf{R}^{d+d'}]$.

(6). By Exercise 27(1) and Exercise 22 of note 1, $\mathcal{L}[\mathbf{R}^d] \times \mathcal{L}[\mathbf{R}^{d'}] \subset \mathcal{L}[\mathbf{R}^{d+d'}]$. Also, the Lebesgue measure space $(\mathbf{R}^{d+d'}, \mathcal{L}[\mathbf{R}^{d+d'}], m)$ is complete. By definition of completion, we have $\overline{\mathcal{L}[\mathbf{R}^d] \times \mathcal{L}[\mathbf{R}^{d'}]} \subset \mathcal{L}[\mathbf{R}^{d+d'}]$. On the other hand, by Exercise 27 of note 3 and (4),

$$\mathcal{L}[\mathbf{R}^{d+d'}] = \overline{\mathcal{B}[\mathbf{R}^d] \times \mathcal{B}[\mathbf{R}^{d'}]} \subset \overline{\mathcal{L}[\mathbf{R}^d] \times \mathcal{L}[\mathbf{R}^{d'}]}.$$

Hence we conclude that $\overline{\mathcal{L}[\mathbf{R}^d] \times \mathcal{L}[\mathbf{R}^{d'}]} = \mathcal{L}[\mathbf{R}^{d+d'}]$, and we are done.

(8). Suppose that X is at most countable. Let $S \subset X \times Y$. Let S_x be the slices $S_x := \{y \in Y : (x, y) \in S\} \in 2^Y$ for every $x \in X$. Since X is at most countable, the set

$$S = \bigcup_{x \in X} \{x\} \times S_x \in 2^X \times 2^Y$$

as a countable union. Hence we get $2^{X \times Y} = 2^X \times 2^Y$. If Y is at most countable, then taking the slices $S^y := \{x \in X : (x, y) \in S\} \subset X$ for every $y \in Y$ gives the same result. □

Exercise 33

Proof. (1). Let δ_x, δ_y be two Dirac measures on $(X, \mathcal{B}_X)$ and $(Y, \mathcal{B}_Y)$ respectively. We claim that $\delta_x \times \delta_y = \delta_{(x,y)}$. Clearly $\delta_{(x,y)}(E \times F) = 1_{(x,y)}(E \times F) = \delta_x(E)\delta_y(F)$ for every $E \in \mathcal{B}_X$ and $F \in \mathcal{B}_Y$. By proposition 30, we obtain the claim.

(2). Let $\#_X$ and $\#_Y$ be the counting measures on $(X, \mathcal{B}_X), (Y, \mathcal{B}_Y)$ respectively. We claim that $\#_X \times \#_Y = \#_{X \times Y}$. Let $E \times F \subset X \times Y$ be such that $E \in \mathcal{B}_X$ and $F \in \mathcal{B}_Y$. If one of E and F is countable, then $E \times F$ is countable and $\#_{X \times F}(E \times F) = \#_X(E)\#_Y(F) = +\infty$. Otherwise, both E and F are finite and the number of distinct pairs are $\#_{X \times Y}(E \times F) = \#_X(E)\#_Y(F)$. By proposition 30 again, we obtain the claim. □

Exercise 34

Proof. Let $\mathcal{T} := \langle \{E \times F \times G : E \in \mathcal{B}_X, F \in \mathcal{B}_Y, G \in \mathcal{B}_Z\} \rangle$, and define

$$S := \{A \subset X \times Y : A \times G \in \mathcal{T} \text{ for all } G \in \mathcal{B}_Z\}.$$

We claim that $S \supset \mathcal{B}_X \times \mathcal{B}_Y$. Clearly S contains all sets of the form $E \times F$ where $E \in \mathcal{B}_X, F \in \mathcal{B}_Y$. Furthermore, if $A \in S$, then

$$A^c \times G = (X \times Y \times Z \backslash A \times G) \cap (X \times Y \times G) \in \mathcal{T}$$

by the construction of $\mathcal{T}$. i.e. $A^c \in S$. Let $A_1, A_2, \dots$ be such that $A_n \in S$ for all n, then $\bigcup_{n=1}^{\infty} A_n \in S$ since $\bigcup_{n=1}^{\infty} A_n \times G = \bigcup_{n=1}^{\infty} (A_n \times G) \in \mathcal{T}$. Finally, we clearly have $\emptyset_{X \times Y} \in S$ say by writing $\emptyset_{X \times Y} = \emptyset_X \times \emptyset_Y$. By remark 4 of note 3, we conclude that $\mathcal{B}_X \times \mathcal{B}_Y \subset S$. This implies that $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z \subset \mathcal{T}$. Conversely, as $E \times F \times G = (E \times F) \times G$, we clearly have $\mathcal{T} \subset (\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z$, and thus $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{T}$. Running the same argument with $\mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z)$, we conclude that

$$(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z) = \mathcal{T}.$$

As desired. Now, let $\mathcal{B}_0$ be the collection of all finite unions

$$S := (E_1 \times F_1 \times G_1) \cup \ldots \cup (E_k \times F_k \times G_k)$$

of Cartesian products of $\mathcal{B}_X$-measurable sets $E_1, \dots, E_k$, $\mathcal{B}_Y$-measurable sets $F_1, \dots, F_k$, and $\mathcal{B}_Z$-measurable sets $G_1, \dots, G_k$. One can verifies that this is a Boolean algebra. Also, any set in $\mathcal{B}_0$ can be easily decomposed into a disjoint union of product sets $E_1 \times F_1 \times G_1, \dots, E_k \times F_k \times G_k$ of $\mathcal{B}_X$-measurable sets, $\mathcal{B}_Y$-measurable sets, and $\mathcal{B}_Z$-measurable sets. We then define the quantity $\mu_0(S)$ associated with such a disjoint union S by the formula

$$\mu_0(S) := \sum_{j=1}^{k} \mu_X(E_j)\mu_Y(F_j)\mu_Z(G_j)$$

whenever S is the disjoint union of products $E_1 \times F_1 \times G_1, \dots, E_k \times F_k \times G_k$ of $\mathcal{B}_X$-measurable sets, $\mathcal{B}_Y$-measurable sets, and $\mathcal{B}_Z$-measurable sets. One can show that this definition does not depend on exactly how S is decomposed, and gives a finitely additive measure $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$.

Now we show that μ_0 is a pre-measure. It suffices to show that if $S \in \mathcal{B}_0$ is the countable disjoint union of sets $S_1, S_2, \dots \in \mathcal{B}_0$, then $\mu_0(S) = \sum_{n=1}^{\infty} \mu_0(S_n)$. Splitting S up into disjoint product sets, and restricting the S_n to each of these product sets in turn, we may assume without loss of generality (using the finite additivity of μ_0) that $S = E \times F$ for some $E \in \mathcal{B}_X$, $F \in \mathcal{B}_Y$ and $G \in \mathcal{B}_Z$. In a similar spirit, by breaking each S_n up into component product sets and using finite additivity again, we may assume without loss of generality that each S_n takes the form $S_n = E_n \times F_n \times G_n$ for some $E_n \in \mathcal{B}_X$, $F_n \in \mathcal{B}_Y$ and $G_n \in \mathcal{B}_Z$. By definition of μ_0, our objective is now to show that

$$\mu_X(E)\mu_Y(F)\mu(G) = \sum_{n=1}^{\infty} \mu_X(E_n)\mu_Y(F_n)\mu_Z(G_n).$$

To do this, first observe from construction that we have the pointwise identity

$$1_E(x)1_F(y)1_G(z) = \sum_{n=1}^{\infty} 1_{E_n}(x)1_{F_n}(y)1_{G_n}(z).$$

for all $x \in X,\ y \in Y$ and $z \in Z$. We fix $x \in X,\ z \in Z$ and integrate this identity in y (noting that both sides are measurable and unsigned) to conclude that

$$\int_Y 1_E(x)1_F(y)1_E(z)\ d\mu_Y(y) = \int_Y \sum_{n=1}^{\infty} 1_{E_n}(x)1_{F_n}(y)1_{G_n}(z)\ d\mu_Y(y).$$

The left-hand side simplifies to $1_E(x)1_G(z)\mu_Y(F)$. To compute the right-hand side, we use the monotone convergence theorem to interchange the summation and integration, and see that the right-hand side is $\sum_{n=1}^{\infty} 1_{E_n}(x)1_{G_n}(z)\mu_Y(F_n)$, thus

$$1_E(x)1_G(z)\mu_Y(F) = \sum_{n=1}^{\infty} 1_{E_n}(x)1_{G_n}(z)\mu_Y(F_n).$$

for all x and z. Both sides are measurable and unsigned in x, so we may integrate in X and conclude that

$$\int_X 1_E(x)1_G(z)\mu_Y(F)\ d\mu_X(x) = \int_X \sum_{n=1}^{\infty} 1_{E_n}(x)1_{G_n}(z)\mu_Y(F_n)\ d\mu_X(x).$$

The left-hand side here is $1_G(z)\mu_X(E)\mu_Y(F)$. Using monotone convergence as before, the right-hand side simplifies to $\sum_{n=1}^{\infty} 1_{G_n}(z)\mu_X(_n)\mu_Y(F_n)$. Finally, we integrate in Z, and the claim follows.

Now that we have established that μ_0 is a pre-measure, we may apply theorem 14 to extend this measure to a countably additive measure $(\mu_X \times \mu_Y) \times \mu_Z$ on a σ-algebra containing $\mathcal{B}_0$. By Exercise 27(2), $(\mu_X \times \mu_Y) \times \mu_Z$ is a countably additive measure on $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z$, and as it extends μ_0, it will obey

$$\mu(E \times F \times G) = \mu_X(E)\mu_Y(F)\mu_Z(G). \tag{1}$$

Observe from finite additivity that any measure on $(\mathcal{B}_X \times \mathcal{B}_Y) \times \mathcal{B}_Z = \mathcal{B}_X \times (\mathcal{B}_Y \times \mathcal{B}_Z)$ that obeys (1) must extend μ_0, and $\mu_X \times (\mu_Y \times \mu_Z)$ is such a measure, thus we have

$$(\mu_X \times \mu_Y) \times \mu_Z = \mu_X \times (\mu_Y \times \mu_Z)$$

by Exercise 15 (and associativity of the product algebra). $\square$

Exercise 41

Proof. (1). We claim that

$$E = F := \bigcap_{n=1}^{\infty} \bigcup_{q \in \mathbf{Q} \cap [0,1]} [q - \frac{1}{n}, q + \frac{1}{n}]^2.$$

By the denseness of rationals among reals, if $x \in [0,1]$, then $\forall n \geq 1$, there exists $q = q(n,x) \in \mathbf{Q} \cap [0,1]$ such that $x \in [q - \frac{1}{n}, q + \frac{1}{n}]$. In other words, $(x,x) \in \bigcup_{q \in \mathbf{Q} \cap [0,1]} [q - \frac{1}{n}, q + \frac{1}{n}]^2$ for all n, and thus $E \subset F$. Conversely, let $(a,b) \in A$. Then $\forall n \geq 1$, there exists some $q =\in \mathbf{Q} \cap [0,1]$ such that $(a,b) \in [q - \frac{1}{n}, q + \frac{1}{n}]^2$, so $|a - b| \leq \frac{2}{n}$ for all n, from which we conclude that $a = b \in [0,1]$. Therefore, we see that $F \subset E$, and the claim follows.

(2). By definition,

$$\int_X (\int_Y f(x,y)\ d\#(y)) dm(x) = \int_X \#(E_x)\ dm(x) = \int_X 1\ dm(x) = 1.$$

(3). Similarly,

$$\int_Y (\int_X f(x,y)\ dm(x)) d\#(y) = \int_Y m(E^y)\ d\#(y) = \int_Y 0\ d\#(y) = 0.$$

(4). Define $\mu_1 : \mathcal{L}([0,1]) \times 2^{[0,1]} \to [0, +\infty]$ by

$$\mu_1(E) := \int_Y m(E^y)\ d\#(y).$$

We verify that this is a well-defined measure. Let $E_1, E_2, \ldots \in \mathcal{L}([0,1]) \times 2^{[0,1]}$ be a countable sequence of disjoint measurable sets, then $(E_n)^y$ is a countable sequence of disjoint Lebesgue measurable sets for any fixed y. Thus

$$\mu_1(\bigcup_{n=1}^{\infty} E_n) = \int_Y m(\bigcup_{n=1}^{\infty} (E_n)^y)\ d\#(y) = \int_Y \sum_{n=1}^{\infty} m((E_n)^y)\ d\#(y),$$

which is $\sum_{n=1}^{\infty} \mu_1(E_n)$ by Tonelli's theorem for sums and integrals. Moreover,

$$\mu_1(\emptyset_{[0,1]^2}) = \int_Y m(\emptyset^y)\ d\#(y) = 0.$$

Hence μ_1 is indeed a measure. Now, we construct an "elementary product pre-measure" that we will then apply theorem 14 to. Let $\mathcal{B}_0$ be the collection of all finite unions

$$S := (E_1 \times F_1) \cup \ldots \cup (E_k \times F_k)$$

of Cartesian products of $\mathcal{L}([0,1])$-measurable sets $E_1, \ldots, E_k$ and $2^{[0,1]}$-measurable sets $F_1, \ldots, F_k$. It is not difficult to verify that this is a Boolean algebra (though it is not, in general, a σ-algebra). Also, any set in $\mathcal{B}_0$ can be easily decomposed into a disjoint union of product sets $E_1 \times F_1, \ldots, E_k \times F_k$ of $\mathcal{L}([0,1])$-measurable sets and $2^{[0,1]}$-measurable sets (cf. Lemma 2 (and Exercise 2) from the prologue). We then define the quantity $\mu_0(S)$ associated such a disjoint union S by the formula

$$\mu_0(S) := \sum_{j=1}^{k} m(E_j)\#(F_j)$$

whenever S is the disjoint union of products $E_1 \times F_1, \ldots, E_k \times F_k$ of $\mathcal{L}([0,1])$-measurable sets and $2^{[0,1]}$-measurable sets. One can show that this definition does not depend on exactly how S is decomposed, and gives a finitely additive measure $\mu_0 : \mathcal{B}_0 \to [0,+\infty]$ (cf. Exercise 2 from the prologue, and also Exercise 31 from Notes 3). Arguing as in the proof of proposition 30, one can show that μ_0 is a pre-measure. We may then apply Theorem 14 to extend this measure to a countably additive measure $\mu_2 = \mu_X \times \mu_Y$ on a σ-algebra containing $\mathcal{B}_0$. By Exercise 27(2), μ_2 is a countably additive measure on $\mathcal{L}([0,1]) \times 2^{[0,1]}$, and as it extends μ_0, it will obey (6).

Finally, we show that $\mu_1 \neq \mu_2$. As in the proof of Theorem 14, for every $E \in \mathcal{L}([0,1]) \times 2^{[0,1]}$, $\mu_2(E)$ is the quantity

$$\mu_2(E) := \inf\{\sum_{n=1}^{\infty} \mu_0(E_n) : E \subset \bigcup_{n=1}^{\infty} E_n; E_n \in \mathcal{B}_0 \text{ for all } n\}.$$

For the diagonal set $E = \{(x,x) : x \in [0,1]\}$, it is obvious that $\mu_2(E) = 1$, since we have $(x,x) \in E_j \times F_j$ for some $E_j \in \mathcal{L}([0,1])$ and $F_j \in 2^{[0,1]}$ for all $x \in [0,1]$, which forces any countable union of E_n that contains E to equal $[0,1]^2$. On the other hand, we have $\mu_1(E) = 0$, as shown by (3). $\square$

Remark to (4) of Exercise 41: It is tempting to define μ_2 in the same manner as μ_1 to be $\mu_2 := \int_X \#(E_x)\ dm(x)$, but the map $x \to \#(E_x)$ might not be measurable w.r.t $\mathcal{L}([0,1])$. For instance, let $C \subset [0,1]$ be a non-measurable set, and $E = \{(x,x) : x \in [0,1]\}$ the diagonal set, then

$$([0,1] \times C) \cap E = \{(x,x) : x \in C\} \in \mathcal{L}([0,1]) \times 2^{[0,1]}.$$

Denote this set by A. The map $x \to \#(A_x)$ is then the indicator function of C, which is not measurable.

Exercise 44

Proof. Let $(x_{n,m})_{n,m\in\mathbf{N}}$ be a doubly infinite sequence of extended non-negative reals such that $x_{n,m}$ equals $+1$ when $n = m$, -1 when $n = m+1$, and 0 otherwise. Now, partition the unit interval $[0,1]$ into $[\frac{1}{2^{k-1}}, \frac{1}{2^k})$, $k = 1, 2, \ldots$ as countably many sub-intervals of length 2^{-k}. Denote by I_n the n^{th} interval in the x-axis, and J_m the m^{th} interval in the y-axis. Define $f : [0,1]^2 \to \mathbf{R}$ by

$$f(x,y) := \sum_{n,m} \frac{x_{n,m}}{|I_n| \cdot |J_m|} 1_{I_n \times J_m}(x,y).$$

$\forall x \in [0,1],\ f(x,y) = \frac{1}{|I_n|} \sum_m \frac{x_{n,m}}{|J_m|} 1_{J_m}(y)$ for some fixed n. Specifically,

$$g(x) := \int_{[0,1]} f(x,y)\ dy = \begin{cases} 2, & x \in I_1 \\ 0, & x \in I_n \text{ for } n \geq 2 \end{cases}$$

Hence $g(x)$ is integrable, and we have

$$\int_{[0,1]} (\int_{[0,1]} f(x,y)\ dy)\ dx = \int_{[0,1]} g(x)\ dx = 1.$$

$\forall y \in [0,1],\ f(x,y) = \frac{1}{|J_m|} \sum_n \frac{x_{n,m}}{|I_n|} 1_{I_n}(x)$ for some fixed m. Specifically,

$$h(y) := \int_{[0,1]} f(x,y)\ dx = 0.$$

Hence $h(y)$ is integrable, and we have

$$\int_{[0,1]} (\int_{[0,1]} f(x,y)\ dx)\ dy = \int_{[0,1]} h(y)\ dy = 0.$$

The function f as defined is clearly Borel-measurable. But

$$|f(x,y)| = \sum_{n,m} \frac{1}{|I_n| \cdot |J_m|} 1_{I_n \times J_m}(x,y)$$

is not integrable with respect to the product space. □

Exercise 47

Proof. By definition, there exists a sequence $f_n : X \to [0, +\infty]$ of simple functions such that $\int_X f(x)\ d\mu(x) = \lim_{n\to\infty} \int_X f_n(x)\ d\mu(x)$. The areas under simple functions are measurable in $\mathcal{B} \times \mathcal{B}[\mathbf{R}]$, as they can be expressed as sets of the form $\bigcup_{i=1}^k f^{-1}(\{a_i\}) \times \{a_i\}$. Then the set $A_f := \{(x,t) \in X \times \mathbf{R} : 0 \leq t \leq f(x)\}$ is measurable as a countable union of measurable sets.

By Tonelli's theorem, we therefore get

$$
\begin{aligned}
(\mu \times m)(A_f) &= \int_{X\times\mathbf{R}} 1_{A_f}(x,t)\ d\mu \times m(x,t) \\
&= \int_X (\int_{\mathbf{R}} 1_{A_f}(x,t)\ dm(t)) d\mu(x) \\
&= \int_X (\int_{\mathbf{R}} 1_{[0,f(x)]}(t)\ dm(t))\ d\mu(x) \\
&= \int_X f(x)\ d\mu(x),
\end{aligned}
$$

as desired. □

Exercise 48

Proof. Let $g : X \times [0, +\infty] \to [0, +\infty]$ be defined by

$$g(x, \lambda) := 1_{f^{-1}([\lambda,+\infty])}(x).$$

Then g is measurable w.r.t $\mathcal{B} \times \mathcal{B}[\mathbf{R}^+]$, as the level set (here $\gamma \in [0, 1)$)

$$\{(x, \lambda) \in X \times [0, +\infty] : g(x, \lambda) > \gamma\} = \{(x, \lambda) \in X \times [0, +\infty] : f(x) \geq \lambda\}$$

can be expressed as $H^{-1}([0, +\infty])$, for $H : X \times [0, +\infty] \to [0, +\infty]$ given by

$$H(x, \lambda) := f(x) - \lambda.$$

Clearly H is measurable w.r.t $\mathcal{B} \times \mathcal{B}[\mathbf{R}^+]$ as a composition and difference of measurable maps. Therefore g is measurable as well. By Tonelli's theorem, we thus have

$$
\begin{aligned}
\int_{[0,+\infty]} \mu(\{x \in X : f(x) \geq \lambda\})\ d\lambda &= \int_{[0,+\infty]} (\int_X g(x,\lambda)\ d\mu(x))\ d\lambda \\
&= \int_X (\int_{[0,+\infty]} g(x,\lambda)\ d\lambda)\ d\mu(x) \\
&= \int_X (\int_{[0,+\infty]} 1_{[0,f(x)]}(\lambda)\ d\lambda))\ d\mu(x) \\
&= \int_X f(x)\ d\mu(x),
\end{aligned}
$$

as claimed. □

Exercise 49

Proof. By definition, we will need to show that

$$\|f * P_t - f\|_{L^1(\mathbf{R}^d)} = \int_{\mathbf{R}^d} |f * P_t(x) - f(x)| \ dx \to 0$$

as $t \to 0$. We first verify this claim for a dense subclass of f, namely the functions f which are continuous and compactly supported (i.e. they vanish outside of a compact set). By the Lebesgue differentiation theorem, almost every point of f is a Lebesgue point. Then by (4) of Exercise 51 in Note 5, the convolutions $f * P_t$ converge essentially uniformly to f as $t \to 0$. Furthermore, as f is compactly supported, the support of $f * P_t - f$ stays uniformly bounded for t in a bounded set. From this we see that $f * P_t$ also converges to f in L^1 norm as required.

Next, we observe the quantitative estimate

$$\|f * P_t\|_{L^1(\mathbf{R}^d)} \leq \|f\|_{L^1(\mathbf{R}^d)} \tag{2}$$

for any $t > 0$. This follows from the triangle inequality and Tonelli's theorem

$$\begin{aligned}
\int_{\mathbf{R}^d} \left| \int_{\mathbf{R}^d} f(y)P_t(x-y) \ dy \right| \ dx &\leq \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} |f(y)P_t(x-y)| \ dy \ dx \\
&\leq \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} |f(y)P_t(x-y)| \ dx \ dy \\
&= \int_{\mathbf{R}^d} |f(y)| \int_{\mathbf{R}^\mathbf{d}} P_t(x-y) \ dx \ dy
\end{aligned}$$

together with the translation invariance of the Lebesgue integral:

$$\int_{\mathbf{R}^d} P_t(x-y) \ dx = \int_{\mathbf{R}^d} P_t(x) \ dx = 1.$$

Now we put the two ingredients together. Let $f : \mathbf{R}^d \to \mathbf{C}$ be absolutely integrable, and let $\varepsilon > 0$ be arbitrary. Applying Littlewood's second principle to f, we can find a continuous, compactly supported function $g : \mathbf{R}^d \to \mathbf{C}$ such that

$$\int_{\mathbf{R}^d} |f(x) - g(x)| \ dx \leq \varepsilon.$$

Applying (2), and using the commutativity and distributivity of the convolution, we conclude that

$$\int_{\mathbf{R}^d} |(f-g) * P_t(x)| \ dx = \int_{\mathbf{R}^d} |f * P_t(x) - g * P_t(x)| \ dx \leq \varepsilon.$$

By the dense subclass result and the triangle inequality, we thus obtain

$$\begin{aligned}
&\|f * P_t - f\|_{L^1(\mathbf{R}^d)} \\
&= \int_{\mathbf{R}^d} |f * P_t(x) - g * P_t(x) + g * P_t(x) - f(x)|\ dx \\
&\leq \int_{\mathbf{R}^d} |(f-g) * P_t(x)|\ dx + \int_{\mathbf{R}^d} |g * P_t(x) - g(x)|\ dx + \int_{\mathbf{R}^d} |g(x) - f(x)|\ dx \\
&\leq \int_{\mathbf{R}^d} |g * P_t(x) - g(x)|\ dx + 2\varepsilon.
\end{aligned}$$

By the dense subclass result again, we see that the claim follows. □

Exercise 51

Proof. Let $f : X \to Y$ be a Lipschitz continuous function from one metric space (X, d_X) to another (Y, d_Y), and $C > 0$ is such that $d_Y(f(x), f(x')) \leq C d_X(x, x')$ for all $x, x' \in X$. Then for every $\varepsilon > 0$, once we have $d_X(x, x') \leq \varepsilon/C$, then $d_Y(f(x), f(x')) \leq \varepsilon$. i.e. f is uniformly continuous.

As shown in Exercise 87 of Note 5, the function $x \mapsto \sqrt{x}$ is absolutely continuous (and thus uniformly continuous), but not Lipschitz continuous, on the interval $[0, 1]$. □

Exercise 54

Proof. Let $v \in \mathbf{R}^d, v \neq 0$. From the total differentiability of f at x_0,

$$\lim_{r \to 0; r \in \mathbf{R} \setminus \{0\}} \frac{f(x_0 + rv) - f(x_0)}{|rv|} = \lim_{r \to 0; r \in \mathbf{R} \setminus \{0\}} \frac{rv \cdot \nabla f(x_0)}{|rv|}$$

which we rearrange to obtain

$$\frac{1}{|v|} \lim_{r \to 0; r \in \mathbf{R} \setminus \{0\}} \frac{f(x_0 + rv) - f(x_0)}{|r|} = \frac{1}{|v|} \lim_{r \to 0; r \in \mathbf{R} \setminus \{0\}} \frac{rv \cdot \nabla f(x_0)}{|r|}$$

which we simplify as

$$\lim_{r \to 0; r \in \mathbf{R} \setminus \{0\}} \frac{f(x_0 + rv) - f(x_0)}{|r|} = \operatorname{sgn}(r) v \cdot \nabla f(x_0)$$

Thus we see that

$$D_v f(x_0) := \lim_{r \to 0; r \in \mathbf{R} \setminus \{0\}} \frac{f(x_0 + rv) - f(x_0)}{r} = v \cdot \nabla f(x_0).$$

In particular, the partial derivatives $\frac{\partial f}{\partial x_i} f(x_0)$ exist for $i = 1, \ldots, d$. If $v = e_i$ is one of the standard basis vectors of $\mathbf{R}^d$, then

$$e_i \cdot \nabla f(x_0) = \frac{\partial f}{\partial x_i} f(x_0)$$

is the ith coordinate of $\nabla f(x_0)$ by the definition of dot product. $\square$

Exercise 55

Proof. Let $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$. Also, denote by f_{x_i} the partial derivative of f in the e_i direction. By induction, we see that the difference

$$f(x_1 + h_1, \ldots, x_d + h_d) - f(x_1, \ldots, x_d)$$

can be written as

$$\begin{aligned}&f(x_1, \ldots, x_d) \pm f(x_1, x_2 + h_2, \ldots, x_d + h_d) \pm f(x_1, x_2, x_3 + h_3, \ldots, x_d + h_d)\\ &\pm \ldots \pm f(x_1, \ldots, x_{d-1}, x_d + h_d) - f(x_1, \ldots, x_d).\end{aligned}$$

In short, by adding and subtracting the $2(d-2)$ terms

$$f(x_1, \ldots, x_{i-1}, x_i + h_i, \ldots, x_d + h_d)$$

for $2 \leq i \leq d$. Collect the terms into d pairs, by the mean value theorem, we get

$$\begin{aligned}&f_{x_1}(\xi_1, x_2 + h_2, \ldots, x_d + h_d)h_1 + f_{x_2}(x_1, \xi_2, x_3 + h_3, \ldots, x_d + h_d)h_2\\ &+ f_{x_3}(x_1, x_2, \xi_3, x_4 + h_4, \ldots, x_d + h_d)h_3 + \ldots\\ &+ f_{x_d}(x_1 + h_1, \ldots, x_{d-1} + h_{d-1}, \xi_d)h_d,\end{aligned}$$

where $\xi_i \in (x_i, x_i + h_i)$ for all $1 \leq i \leq d$. In conclusion, we have

$$\sum_{i=1}^{d} f_{x_i}(x_1, \ldots, x_{i-1}, \xi_i, x_{i+1} + h_{i+1}, \ldots, x_d + h_d)h_i$$

Insert $h \cdot \nabla f(x) = h_1 f_{x_1}(x) + \ldots + h_d f_{x_d}(x)$ and collect like terms, we get an upper bound for

$$|f(x + h) - f(x) - h \cdot \nabla f(x)|$$

as follows:

$$\sum_{i=1}^{d} |f_{x_i}(x_1, \ldots, x_{i-1}, \xi_i, x_{i+1} + h_{i+1}, \ldots, x_d + h_d) - f_{x_i}(x)||h_i|.$$

As the partial derivatives are continuous at $x = (x_1, \ldots, x_d)$, there exists $\delta > 0$ such that when $|h| \leq \delta$,

$$|f_{x_i}(x_1, \ldots, x_{i-1}, \xi_i, x_{i+1} + h_{i+1}, \ldots, x_d + h_d) - f_{x_i}(x)| \leq \frac{\varepsilon}{d}$$

for all i. In particular, this implies that

$$|f(x+h) - f(x) - h \cdot \nabla f(x)| \leq \varepsilon |h|$$

whenever $|h| \leq \delta$. Or equivalently, that

$$\lim_{h \to 0; h \to \mathbf{R}^d \setminus \{0\}} \frac{f(x+h) - f(x) - h \cdot \nabla f(x)}{|h|} = 0.$$

That is, f is totally differentiable at x. $\square$

Exercise 56

Proof. Let $x, v \in \mathbf{R}^2$. By direct computation we get

$$\frac{f(x+hv) - f(x)}{h} = \frac{1}{h}\left(\frac{(x_1 + hv_1)(x_2 + hv_2)^2}{(x_1 + hv_1)^2 + (x_2 + hv_2)^2} - \frac{x_1 x_2^2}{x_1^2 + x_2^2}\right).$$

First assume $x \neq 0$. For simplicity let $N(x) := x_1 x_2^2,\ D(x) := x_1^2 + x_2^2$. Then

$$f(x) = \frac{N(x)}{D(x)},\ f(x+hv) = \frac{N(x+hv)}{D(x+hv)}.$$

Now we expand these terms. For the numerator:

$$\begin{aligned} N(x+hv) &= (x_1 + hv_1)(x_2 + hv_2)^2 \\ &= x_1 x_2^2 + h(v_1 x_2^2 + 2x_1 x_2 v_2) + h^2(2x_2 v_1 v_2 + x_1 v_2^2) + h^3 v_1 v_2^2 \\ &= N(x) + hA + h^2 B + h^3 C, \end{aligned}$$

where $A = v_1 x_2^2 + 2x_1 x_2 v_2, B = 2x_2 v_1 v_2 + x_1 v_2^2, C = v_1 v_2^2$. For the denominator:

$$\begin{aligned} D(x+hv) &= (x_1 + hv_1)^2 + (x_2 + hv_2)^2 \\ &= D(x) + h(2x_1 v_1 + 2x_2 v_2) + h^2(v_1^2 + v_2^2) \\ &= D(x) + hE + h^2 F, \end{aligned}$$

where $E = 2x_1 v_1 + 2x_2 v_2, F = v_1^2 + v_2^2$. Now we insert these expressions:

$$\frac{f(x+hv) - f(x)}{h} = \frac{1}{h}\left(\frac{N + hA + h^2 B + h^3 C}{D + hF + h^2 E} - \frac{N}{D}\right) \overset{h \to 0}{=} \frac{DA - NE}{D^2}$$

after some algebraic simplifications. Substitute the x and v, we see that

$$D_v f(x) = \frac{v_1 x_2^2(x_2^2 - x_1^2) + 2x_1^3 x_2 v_2}{(x_1^2 + x_2^2)^2}, \text{ for } x \neq 0.$$

If $x = 0$, we can easily calculate

$$\frac{f(hv) - f(0)}{h} = \frac{v_1 v_2^2}{v_1^2 + v_2^2}$$

for all v. Hence it follows that

$$D_v f(0) = \frac{v_1 v_2^2}{v_1^2 + v_2^2}.$$

Hence, the directional derivatives $D_v f(x)$ exist for all $x, v \in \mathbf{R}^2$. On the other hand, we have

$$\frac{\partial f}{\partial x_1}(0) = \frac{\partial f}{\partial x_2}(0) = 0$$

by what we have just shown. Therefore, by condition and (10) we have

$$\frac{f(0 + h) - f(0) - h \cdot \nabla f(x_0)}{|h|} = \frac{f(h)}{|h|}.$$

Take $h = (t, t) \in \mathbf{R}^2 \backslash \{0\}$, then $f(t) = t/2$, while $|h| = \sqrt{2}t$. In particular, one has $\frac{f(h)}{|h|} = \frac{1}{2\sqrt{2}} \neq 0$ for all t, and so f is not totally differentiable at the origin $(0, 0)$. □

Remark on the proof of Rademacher differentiation theorem:

1. As f is continuous, the function $g_h(x_0) := \frac{f(x_0 + hv) - f(v)}{h}$ is continuous for each $h \in \mathbf{Q}\backslash\{0\}$, and thus Borel-measurable by Definition 8 of Note 3. Hence the function

$$r(x_0) := \lim \sup_{h \to 0; r \in \mathbf{Q}\backslash\{0\}} g_h(x_0) - \lim \inf_{h \to 0; r \in \mathbf{Q}\backslash\{0\}} g_h(x_0)$$

is Borel-measurable (It is not difficult to adapt the measurability results for unsigned limsup/liminf to signed limsup/liminf. One could also use exp and log to move between the unsigned and signed cases, although this is a somewhat strange way to proceed.)

2. Let $T : X \to Y$ be a linear map between an n-dimensional normed space X and an m-dimensional normed space Y, then T is Lipschitz. i.e. $\exists \lambda \in F$ s.t

$$\frac{\|T(x) - T(y)\|}{\|x - y\|} \leq \lambda$$

for all $x, y \in X$. By linearity of T, this reduces to

$$\|T(\frac{x-y}{\|x-y\|})\| \leq \lambda.$$

Let $z := \frac{x-y}{\|x-y\|} = \sum_{i=1}^{n} a_i e_i$, where $\{e_i\}$ is an orthonormal basis, and assume $x \neq y$, we see that

$$\|T(z)\| \leq (|a_1| + \ldots + |a_n|) \max_{1 \leq i \leq n} \|T(e_i)\|.$$

By the Cauchy-Schwarz inequality, we obtain that

$$\sum_{i=1}^{n} |a_i| = \sum_{i=1}^{n} |a_i| \cdot 1 \leq (\sum_{i=1}^{n} a_i^2)^{1/2} (\sum_{i=1}^{n} 1^2)^{1/2} = \|z\|\sqrt{n} = \sqrt{n}.$$

Hence the claim holds for $\lambda := \sqrt{n} \cdot \max_{1 \leq i \leq n} \|T(e_i)\|$.

3. Now, let $T : \mathbf{R}^d \to \mathbf{R}^d$ be an invertible linear transformation such that $T(v) = e_1$, and define $g(x) := f(Tx)$. Note that by (2) we have

$$|g(x) - g(y)| = |f(Tx) - f(Ty)| \leq C|T(x) - T(y)| \leq C\lambda|x - y|,$$

hence g is Lipschitz continuous. Furthermore, by construction we have

$$E_v = T^{-1}(\{y_0 \in \mathbf{R}^d : D_{e_1} g(y_0) \text{ does not exist}\})$$

By Exercise 8 of Note 1, E_v is a null set if $\{y_0 \in \mathbf{R}^d : D_{e_1} g(y_0) \text{ does not exist}\}$ is. Thus we may assume without loss of generality that v is the basis vector e_1.

4. By condition, $E^{y_0} = \{x_0 \in \mathbf{R} : x \mapsto f(x, y_0) \text{ is not differentiable at } x_0\}$ is a null set in $\mathbf{R}$ by the Lipschitz differentiation theorem. Applying Tonelli's theorem for sets, we conclude that

$$m^d(E_{e_1}) = m^{d-1} \times m(E_{e_1}) = \int_{\mathbf{R}^{d-1}} m(E^{y_0}) \; dm^{d-1}(y_0) = 0.$$

i.e. E_{e_1} is a null set as required.

5. Let $g : \mathbf{R}^d \to \mathbf{R}$ be continuously differentiable and compactly supported, which means that the partial derivatives exist and are all continuous. Let $a = (a_1, \dots, a_d), b = (b_1, \dots, b_d) \in \mathbf{R}^d$. For each $1 \leq i \leq d$, define the functions $g_i : \mathbf{R} \to \mathbf{R}$ by

$$g_i(x) := g(b_1, \dots, b_{i-1}, x, a_{i+1}, \dots, a_d).$$

Then by telescoping series and the triangle inequality, we have

$$|g(a) - g(b)| \leq \sum_{i=1}^{d} |g_i(b_i) - g_i(a_i)|.$$

Applying the second FTC, this implies that

$$|g(a) - g(b)| \leq \sum_{i=1}^{d} |\int_{[a_i,b_i]} g_i'(x)\ dx|.$$

Since g is compactly supported, so are the partial derivatives, we may thus assume that $a, b \in \operatorname{supp}(g)$. As the $g_i'(x)$ are continuous and compactly supported, they are bounded. Let $K := \max_{1 \leq i \leq d} \|g_i'\|_\infty$, we obtain that

$$|g(a) - g(b)| \leq K \sum_{i=1}^{d} |a_i - b_i|,$$

from which we get $|g(a) - g(b)| \leq K|a - b|$, and thus g is Lipschitz.

Equivalently, we specialize the d-dimensional function to one-dimensional lines, and then apply the one-dimensional FTC on each such line. That is, we specialize g to any line $\{x + tv : t \in \mathbf{R}\}$ by considering the one-dimensional function $t \mapsto g(x + tv)$.

Consider the function $f : \mathbf{R} \to \mathbf{R}$ given by $f(t) := g(x + th)$ for some $x, h \in \mathbf{R}^d$. By the chain rule, we have

$$f'(t) = \nabla g(x + ht) \cdot h$$

which is bounded by the fact that each component of ∇g is bounded, hence by the second FTC, we have

$$|g(x+h) - g(x)| = |\int_{[0,1]} \nabla g(x + ht) \cdot h\ dt| \leq \int_{[0,1]} |\nabla g(x + ht)\ \cdot h|\ dt.$$

By Cauchy-Schwarz inequality, $|\nabla g(x+ht)\ \cdot h| \leq |\nabla g(x+ht)||h| \leq M|h|$, where $M := |\nabla g|_\infty < \infty$.

6. By Exercise 1.5.10 of “Analysis 2”, the unit sphere B is compact, and thus totally bounded. Hence $\forall \varepsilon > 0$, there exists $n \in \mathbf{N}$ such that

$$B \subset \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon/2).$$

Since the set $\mathbf{Q}^d$ is dense in $\mathbf{R}^d$, we can pick any vector $v^{(i)}$ from $\mathbf{Q}^d \cap B(x^{(i)}, \varepsilon/2)$ for each $1 \leq i \leq n$. Define $V_\varepsilon := \{v^{(i)} : 1 \leq i \leq n\}$. Then

$$|u - v^{(i)}| \leq |u - x^{(i)}| + |x^{(i)} - v^{(i)}| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

given that $u \in B(x^{(i)}, \varepsilon/2)$. Clearly, V_ε depends only on ε (and on d).

(7). Why do we restrict the possible choices of v to a finite subset V_ε of $\mathbf{Q}^d$? - So that h will only depend on ε rather than on both ε and v in the bound $|F(h)|/|h| \leq (C+1)\varepsilon$. The discretization here is needed to obtain uniformity in v.

Exercise 58

Proof. Suppose that $F \neq 0$ almost everywhere. By the Lebesgue differentiation theorem, there must be some $x_0 \in \mathbf{R}^d$ such that $F(x_0) \neq 0$ and

$$\lim_{r \to 0} \frac{1}{m(B(x_0, r))} \int_{B(x_0,r)} F(y)\ dy = F(x_0).$$

In particular, for every $\varepsilon > 0$ we can pick r small enough such that

$$|\int_{B(x_0,r)} F(y)\ dy - m(B(x_0, r))F(x_0)| < \varepsilon m(B(x_0, r))$$

Set $\varepsilon := |F(x_0)/2|$, and fix the corresponding r, we deduce that

$$|\int_{B(x_0,r)} F(y)\ dy| > m(B(x_0, r))|F(x_0)|/2 > 0.$$

Pick $g(y) := 1_{B(x_0,r)}(y)$, a compactly supported, continuously differentiable function, but

$$|\int_{\mathbf{R}^d} F(y)g(y)\ dy| = |\int_{B(x_0,r)} F(y)\ dy| > 0,$$

a contradiction. The claim then follows. $\square$

Exercise 59

Proof. (1). By definition, $\mathcal{B}_A = \prod_{\beta\in A}\mathcal{B}_\beta := \langle\bigcup_{\beta\in A}\pi_\beta^*(\mathcal{B}_\beta)\rangle$. Note that

$$\forall E_\beta \in \mathcal{B}_\beta,\ \pi_\beta^{-1}(E_\beta) \in \pi_\beta^*(\mathcal{B}_\beta) \in \mathcal{B}_A.$$

Hence $\pi_\beta : X_A \to X_\beta$ is indeed a measurable morphism. Let $\mathcal{A}$ be any σ-algebra on X_A that also has this property. Then $\mathcal{A}$ contains $\pi_\beta^*(\mathcal{B}_\beta)$ for all $\beta \in A$, and thus contains $\mathcal{B}_A$ by the definition of generated algebra.

(2). By definition, the partial projection $\pi_B : X_A \to X_B$ is given by

$$\pi_B((x_\alpha)_{\alpha\in A}) := (x_\alpha)_{\alpha\in B},$$

where the target space X_B is endowed with the product σ-algebra

$$\mathcal{B}_B = \prod_{\beta\in B}\mathcal{B}_\beta := \langle\bigcup_{\beta\in B}\pi^*_{\{\beta\}\leftarrow B}(\mathcal{B}_\beta)\rangle.$$

Let $P(E)$ be the property of sets $E \subset X_B$ that $\pi_B^{-1}(E) \in \mathcal{B}_A$. For clarity, write π_B as $\pi_{B\leftarrow A}$. For any $\beta \in B$ and $E_\beta \in \mathcal{B}_\beta$, observe that

$$\pi^{-1}_{B\leftarrow A}(\pi^{-1}_{\{\beta\}\leftarrow B}(E_\beta)) = (\pi_{\{\beta\}\leftarrow B}\circ\pi_{B\leftarrow A})^{-1}(E_\beta) = \pi_\beta^{-1}(E_\beta) \in \mathcal{B}_A$$

by (1) and the composition law. Thus P holds on the generating family of $\mathcal{B}_B$. Since $\mathcal{B}_A$ is a σ-algebra, it can be shown that $P(\emptyset_{X_B})$ holds, $P(E^c)$ holds whenever $P(E)$ holds, and if $E_1, E_2, \ldots \subset X_B$ are such that $P(E_n)$ holds for all n, then $P(\bigcup_n E_n)$ holds as well. Hence we conclude that $\pi_B^{-1}(E) \in \mathcal{B}_A$ for all $E \in \mathcal{B}_B$. i.e. π_B is a measurable morphism.

(3). If A is at most countable, we can take $B = A$, and the claim is trivial. Hence we assume that A is uncountable. Let

$$\mathcal{F} := \bigcup_{\beta\in A}\pi_\beta^*(\mathcal{B}_\beta) = \{\pi_\beta^{-1}(E_\beta) : E_\beta \in \mathcal{B}_\beta, \beta \in A\},$$

and consider the set

$$\mathcal{B}_C := \bigcup_{\mathcal{F}_0\subset\mathcal{F},\mathcal{F}_0 \text{ countable}}\langle\mathcal{F}_0\rangle$$

Clearly, $\mathcal{B}_C$ is a σ-algebra, and $\mathcal{B}_C = \mathcal{B}_A$. In particular, $\forall E_A \in \mathcal{B}_A,\ E_A \in \langle\mathcal{F}_0\rangle$ for some countable $\mathcal{F}_0 \subset \mathcal{F}$. Fix this $\mathcal{F}_0$, and write it as

$$\{\pi_\beta^{-1}(E_{\beta,n}) : E_{\beta,n} \in \mathcal{B}_\beta, \beta \in B, n \in \mathbf{N}\}$$

for some countable $B \subset A$, where we let $E_{\beta,n} := \emptyset$ if necessary. From the composition law $\pi_{\{\beta\}\leftarrow B} \circ \pi_B = \pi_\beta$, we see that

$$\mathcal{F}_0 = \{\pi_B^{-1}(\pi_{\{\beta\}\leftarrow B}^{-1}(E_{\beta,n})) : E_{\beta,n} \in \mathcal{B}_\beta, \beta \in B, n \in \mathbf{N}\}.$$

For simplicity, we write $\pi_{\{\beta\}\leftarrow B}^{-1}(E_{\beta,n}) := F_{\beta,n}$. Hence we have:

$$\mathcal{F}_0 = \{\pi_B^{-1}(F_{\beta,n}) : \beta \in B, n \in \mathbf{N}\}.$$

Since $\pi_{\{\beta\}\leftarrow B}$ are measurable morphisms, $F_{\beta,n} \in \mathcal{B}_B$. Now we consider the set

$$\mathcal{A} = \{G \in \langle \mathcal{F}_0 \rangle : G = \pi_B^{-1}(E) \text{ for some } E \in \mathcal{B}_B\}.$$

We claim that $\mathcal{A} = \langle \mathcal{F}_0 \rangle$, from which the result follows. Let $P(G)$ be the property of sets $G \subset X_A$ that $G = \pi_B^{-1}(E)$ for some $E \in \mathcal{B}_B$. By construction, P holds on the generating family $\mathcal{F}_0$. If $G = \pi_B^{-1}(E)$ with $E \in \mathcal{B}_B$, then $X_A \backslash G = \pi_B^{-1}(X_B \backslash E)$, $X_B \backslash E \in \mathcal{B}_B$. Let $G_1, G_2, \ldots \subset X_A$ be such that $G_n = \pi_B^{-1}(E_n)$ with $E_n \in \mathcal{B}_B$ for all n, then $\bigcup_n G_n = \pi_B^{-1}(\bigcup_n E_n)$, $\bigcup_n E_n \in \mathcal{B}_B$. Finally, we have $\emptyset_{X_A} = \pi_B^{-1}(\emptyset_{X_B})$. We conclude that $\mathcal{A}$ is a σ-algebra, so $\mathcal{A} = \langle \mathcal{F}_0 \rangle$, as desired.

(4). Assume A is uncountable. By the denseness of $\mathbf{Q}$ in $\mathbf{R}$ and the fact that every open set in $\mathbf{R}$ can be written as the union of at most countably many disjoint open intervals, we consider the rational intervals $I_1, I_2, \ldots \subset [0, +\infty]$. By (3), we have

$$\forall n \geq 1,\ f^{-1}(I_n) = \pi_{B_n}^{-1}(E_n)$$

for an at most countable set $B_n \subset A$ and some $E_n \in \mathcal{B}_{B_n}$. Let $B := \bigcup_{n=1}^\infty B_n$, then B is at most countable. Now, for every open set $U \subset [0, +\infty]$, there exists an at most countable subset $Z(U) \subset \mathbf{N}$ such that $U = \bigcup_{n \in Z(U)} I_n$, hence

$$f^{-1}(U) = \bigcup_{n \in Z(U)} f^{-1}(I_n) = \bigcup_{n \in Z(U)} \pi_{B_n}^{-1}(E_n) = \bigcup_{n \in Z(U)} \pi_B^{-1}(\pi_{B_n \leftarrow B}^{-1}(E_n)).$$

We conclude that f is measurable w.r.t the pullback algebra $\pi_B^*(\mathcal{B}_B)$.

Next, we show that f is constant on each fiber of π_B. Let $x, x' \in \pi_B^{-1}(y)$ for some $y \in X_B$. It suffices to show that $f(x) \in I_n$ iff $f(x') \in I_n$ for all n. By construction, $f(x) \in I_n$ iff

$$x \in \pi_{B_n}^{-1}(E_n) = \pi_B^{-1}(\pi_{B_n \leftarrow B}^{-1}(E_n)) = \pi_B^{-1}(F_n)$$

where $F_n := \pi_{B_n \leftarrow B}^{-1}(E_n) \in \mathcal{B}_B$. This occurs iff $y \in F_n$, iff $f(x') \in I_n$, as required. Let $f_B : X_B \to [0, +\infty]$ be s.t $f_B(y) := f(x)$ for all $x \in \pi_B^{-1}(y)$, we see that f_B is well-defined, with $f = f_B \circ \pi_B$.

Finally, we show that f_B is $\mathcal{B}_B$-measurable. Let $U \subset [0, +\infty]$ be open. By definition,

$$f_B^{-1}(U) = \{y \in X_B : f_B(y) \in U\} = \pi_B(\{x \in X_A : f(x) \in U\}) = \pi_B(f^{-1}(U)).$$

By what we have shown before, there exists some $E_U \in \mathcal{B}_B$ ($E_U = \bigcup_{n \in Z(U)} F_n$) such that

$$f^{-1}(U) = \pi_B^{-1}(E_U).$$

In particular, this implies that $f_B^{-1}(U) = \pi_B \circ \pi_B^{-1}(E_U) = E_U \in \mathcal{B}_B$. i.e. f_B is $\mathcal{B}_B$-measurable. Thus a measurable function can only depend on at most countably many of the coordinates.

(5). Let $A = \{\beta_1, \beta_2, \ldots\}$ be countable. Since $\prod_{n \in \mathbf{N}} E_{\beta_n} = \bigcap_{n \in \mathbf{N}} \pi_{\beta_n}^{-1}(E_{\beta_n})$ for every $E_{\beta_n} \in \mathcal{B}_{\beta_n}$, $\mathcal{B}_A$ contains the σ-algebra generated by the product sets. Conversely, $\pi_{\beta_n}^{-1}(E_{\beta_n}) = \prod_{i=1}^{n-1} X_{\beta_i} \times E_{\beta_n} \times \prod_{j=n+1}^{\infty} X_{\beta_j}$, hence the σ-algebra generated by the product sets must contain $\mathcal{B}_A$. The claim thus follows.

(6). Pick $E_\beta \in \mathcal{B}_\beta$ for all $\beta \in A$ (valid since the $\mathcal{B}_\alpha$ are all non-trivial), and set $E_A := \prod_{\beta \in A} E_\beta$. By construction, E_A is in the product σ-algebra. If $E_A \in \mathcal{B}_A$, by (3) there exists an at most countable set $B \subset A$ and $E_B \in \mathcal{B}_B$ such that $E_A = \pi_B^{-1}(E_B)$. But then $\forall \alpha \neq B$ (of which there are countably many), we have

$$\pi_\alpha(E_A) = E_\alpha \neq X_\alpha = \pi_\alpha(\pi_B^{-1}(E_B)),$$

a contradiction. That is, $E_A \notin \mathcal{B}_A$, and therefore $\mathcal{B}_A$ is not the σ-algebra generated by sets $\prod_{\beta \in A} E_\beta$ with $E_\beta \in \mathcal{B}_\beta$ for all $\beta \in A$.

(7). For any $\beta \in A$ and $E_\beta \in \mathcal{B}_\beta$, let $E = \pi_\beta^{-1}(E_\beta)$. Observe that

$$E_{x_{A \setminus B}, B} = \begin{cases} \pi_{\{\beta\} \leftarrow B}^{-1}(E_\beta), \text{ if } \beta \in B \\ X_B, \text{ if } \beta \in A \setminus B \text{ and } (x_{A \setminus B})_\beta \in E_\beta \\ \emptyset_B, \text{if } \beta \in A \setminus B \text{ and } (x_{A \setminus B})_\beta \notin E_\beta. \end{cases}$$

Hence $E_{x_{A \setminus B}, B} \in \mathcal{B}_B$. Let $P(E)$ be the property of $E \subset X_A$ that $E_{x_{A \setminus B}, B} \in \mathcal{B}_B$. We just verified P on $\bigcup_{\beta \in A} \pi_\beta^*(\mathcal{B}_\beta)$. Given $E_{x_{A \setminus B}, B} \in \mathcal{B}_B$, we have

$$E^c_{x_{A \setminus B}, B} = X_B \setminus E_{x_{A \setminus B}, B} \in \mathcal{B}_B$$

by the definition of $E_{x_{A \setminus B}, B}$. If $E^n_{x_{A \setminus B}, B} \in \mathcal{B}_B$ for all $n \in \mathbf{N}$, then

$$(\bigcup_n E^n)_{x_{A \setminus B}, B} = \bigcup_n E^n_{x_{A \setminus B}, B} \in \mathcal{B}_B.$$

Lastly, we have $(\emptyset_{X_A})_{x_{A \setminus B}, B} = \emptyset_{X_B} \in \mathcal{B}_B$. We thus conclude that $P(E)$ is true for all $E \in \mathcal{B}_A$.

(8). For every open set $U \subset [0, +\infty]$, $f^{-1}_{x_{A\setminus B}, B}(U)$ is

$$\{x_B \in X_B : f(x_B, x_{A\setminus B}) \in U\} = \{x_B \in X_B : (x_B, x_{A\setminus B}) \in f^{-1}(U)\},$$

which is $\mathcal{B}_B$-measurable by (7), as $f^{-1}(U) \in \mathcal{B}_A$. $\square$

Remark on the proof of the Kolmogorov extension theorem

1. By (13), we have

$$\mu_B(E_B) = (\pi_{B \leftarrow B\cup B'})_* \mu_{B\cup B'}(E_B) = \mu_{B\cup B'}(\pi^{-1}_{B\leftarrow B\cup B'}(E_B))$$

Substitute $E_B = \pi_{B\leftarrow B\cup B'}(E_{B\cup B'})$, we get $\mu_B(E_B) = \mu_{B\cup B'}(E_{B\cup B'})$.

2. The reason why we may assume that the B_N are increasing in N by enlarging each B_N as necessary, is because the composition law allows us to pull back to a larger coordinates set without altering F_N:

$$F_N = \pi^{-1}_{B_N}(G_N) = \pi^{-1}_{B_{N'}}(\pi^{-1}_{B_N \leftarrow B_{N'}}(G_N))$$

for any $B_{N'} \supset B_N$ and $N' > N$. The decreasing nature of F_N then gives

$$\pi^{-1}_{B_{N+1}}(\pi^{-1}_{B_N \leftarrow B_{N+1}}(G_N)) = F_N \supset F_{N+1} = \pi^{-1}_{B_{N+1}}(G_{N+1}).$$

Hence we get the inclusion $G_{N+1} \subset \pi^{-1}_{B_N \leftarrow B_{N+1}}(G_N)$.

3. Since the projection maps are continuous, the preimages $\pi^{-1}_{N' \leftarrow N}(K_{N'})$ are closed, as $\pi^{-1}_{N\leftarrow N}(K_N) = K_N$ is compact, if we set $K'_N := \bigcap_{N'=1}^{N} \pi^{-1}_{B_{N'} \leftarrow B_N}(K_{N'})$, then K'_N is a closed subset of a compact set, hence compact. To get the following lower bound, observe that

$$G_N \setminus \bigcup_{N'=1}^{N} \pi^{-1}_{B'_N \leftarrow B_N}(G_{N'} \setminus K_{N'}) \subset \bigcap_{N'=1}^{N} \pi^{-1}_{B'_N \leftarrow B_N}(K_{N'}).$$

Indeed, from the inclusion $G_{N+1} \subset \pi^{-1}_{B_N \leftarrow B_{N+1}}(G_N)$, we get

$$G_N \subset \pi^{-1}_{B_{N'} \leftarrow B_N}(G_{N'})$$

for all $1 \leq N' \leq N$. Hence $\pi^{-1}_{B'_N \leftarrow B_N}(G_{N'} \backslash K_{N'}) \supset G_N \backslash \pi^{-1}_{B'_N \leftarrow B_N}(K_{N'})$. So

$$\bigcup_{N'=1}^{N} \pi^{-1}_{B'_N \leftarrow B_N}(G_{N'} \backslash K_{N'}) \supset G_N \backslash \bigcap_{N'=1}^{N} \pi^{-1}_{B'_N \leftarrow B_N}(K_{N'}).$$

It follows from $\mu_0(F_N) > \varepsilon$ that

$$\mu_{B_N}(K'_N) \geq \varepsilon - \sum_{N'=1}^{N} \mu_{B_{N'}}(G_{N'} \backslash K_{N'}) \geq \varepsilon - \sum_{N'=1}^{N} \varepsilon/2^{N'+1} \geq \varepsilon/2.$$

4. By construction, we have

$$\begin{aligned} K'_{N+1} &= \bigcap_{N'=1}^{N} \pi^{-1}_{B_N \leftarrow B_{N+1}}(\pi^{-1}_{B_{N'} \leftarrow B_N}(K_{N'})) \cap K_{N+1} \\ &= \pi^{-1}_{B_N \leftarrow B_{N+1}}(\bigcap_{N'=1}^{N} \pi^{-1}_{B_{N'} \leftarrow B_N}(K_{N'})) \cap K_{N+1} \\ &= \pi^{-1}_{B_N \leftarrow B_{N+1}}(K'_N) \cap K_{N+1} \end{aligned}$$

which is a subset of $\pi^{-1}_{B_N \leftarrow B_{N+1}}(K'_N)$.

5. Since the H_N are decreasing in N, $\forall N_0 \leq N$, we see that

$$x_N \in H_N \subset H_{N_0} = \pi^{-1}_{B_{N_0}}(K'_{N_0}),$$

and thus $\pi_{B_{N_0}}(x_N) \in K'_{N_0}$. Applying the Heine-Borel theorem repeatedly, we may obtain nested subsequences $x_{N_{j,m}}$ for $m = 1, 2, \ldots$ and $j = 1, 2, \ldots$ such that for each $j = 1, 2, \ldots$ (indicating the jth subsequence of x_N), the sequence $m \mapsto \pi_{B_j}(x_{N_{j,m}})$ converges (ensuring $N_{j,m} \geq j$ for all j and m).

6. Concerning the diagonalisation trick, as we are picking subsequences, $N_{j,m} \leq N_{m,m}$ for $j \leq m$, thus for each j, $\pi_{B_j}(x_{N_{m,m}})$ converges to a limit as $m \to \infty$. Note that any σ-algebra that contains $\mathcal{B}_0$ will contain $\mathcal{B}_A$ since it contains all sets of the form $\pi_\beta^{-1}(E_\beta)$ for $\beta \in A$ and $E_\beta \in \mathcal{B}_\beta$, so Exercise 15 applies.

When constructing a product probability measure, μ_B is defined on X_B, for sets in $\mathcal{B}_B = \langle \bigcup_{\alpha \in B} \pi^*_{\alpha \leftarrow X_B}(\mathcal{B}_\alpha) \rangle$, and the X_B are locally compact, σ-compact spaces since each X_α is so, and finite products preserve local compactness and σ-compactness.