Chapter 6

Constructing measures

So far we have been taking measures for granted; except for a few, almost trivial examples we have not shown that they exist. In this chapter we shall develop a powerful technique for constructing measures with the properties we want, e.g. the Lebesgue measures and the coin tossing measure described in Section 5.1.

When we construct a measure, we usually start by knowing how we want it to behave on a family of simple sets: For the Lebesgue measure on $\mathbb{R}$, we want the intervals (a, b) to have measure $b - a$, and for the coin tossing measure we know what we want the measure of a cylinder set to be. The art is to extend such "pre-measures" to full-blown measures.

We shall use a three step procedure to construct measures. We start with a collection $\mathcal{R}$ of subsets of our space X and a function $\rho : \mathcal{R} \to \overline{\mathbb{R}}_+$. The idea is that the sets R in $\mathcal{R}$ are the sets we "know" the size $\rho(R)$ of – if, e.g., we want to construct the Lebesgue measure, $\mathcal{R}$ could be the collection of all open intervals, and ρ would then be given by $\rho((a, b)) = b - a$. From $\mathcal{R}$ and ρ, we first construct an "outer measure" μ^* which assigns a "size" $\mu^*(A)$ to *all* subsets A of X. The problem with μ^* is that it usually fails to be countably additive; i.e. the crucial equality

$$\mu^*(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu^*(A_n)$$

doesn't hold for all disjoint sequences $\{A_n\}$ of subsets of X. The second step in our procedure is therefore to identify a σ-algebra $\mathcal{A}$ such that countable additivity holds when the disjoint sets A_n belong to $\mathcal{A}$. The restriction of μ^* to $\mathcal{A}$ will then be our desired measure μ. The final step in the procedure is to check that μ really is an extension of ρ, i.e. that $\mathcal{R} \subset \mathcal{A}$ and $\mu(R) = \rho(R)$ for all $R \in \mathcal{R}$. This is not always the case, but requires special properties of $\mathcal{R}$ and ρ.

We begin by constructing outer measures.

6.1 Outer measure

To construct outer measures, we don't need to require much of $\mathcal{R}$ and ρ:

Definition 6.1.1 *Assume that X is a nonempty set and that $\mathcal{R}$ is a collection of subsets of X such that $\emptyset, X \in \mathcal{R}$. Throughout this section, we assume that $\rho : \mathcal{R} \to \overline{\mathbb{R}}_+$ is a function such that $\rho(\emptyset) = 0$.*

Assume that B is a subset of X. A *covering* of X is a countable collection $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$ of sets from $\mathcal{R}$ such that

$$B \subset \bigcup_{n\in\mathbb{N}} C_n$$

We define the *size* of the covering to be

$$|\mathcal{C}| = \sum_{n=1}^{\infty} \rho(C_n)$$

We are now ready to define the *outer measure μ^* generated by $\mathcal{R}$ and ρ*: For all $B \subset X$, we set

$$\mu^*(B) = \inf\{|\mathcal{C}| \,:\, \mathcal{C} \text{ is a covering of } B\}$$

We see why μ^* is called an *outer* measure; it is obtained by approximating sets from the *outside* by unions of sets in $\mathcal{R}$.

The essential properties of outer measures are not hard to establish:

Proposition 6.1.2 *The outer measure μ^* satisfies:*

(i) $\mu^*(\emptyset) = 0$.

(ii) *(Monotonicity) If $B \subset C$, then $\mu^*(B) \leq \mu^*(C)$.*

(iii) *(Countable subadditivity) If $\{B_n\}_{n\in\mathbb{N}}$ is a sequence of subsets of $\mathbb{R}^d$, then*

$$\mu^*(\bigcup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} \mu^*(B_n)$$

Proof: (i) Since $\mathcal{C} = \{\emptyset, \emptyset, \emptyset, \ldots\}$ is a covering of $\emptyset$ and $\rho(\emptyset) = 0$, we get $\mu^*(\emptyset) = 0$.

(ii) Since any covering of C is a covering of B, we have $\mu^*(B) \leq \mu^*(C)$.

(iii) If $\mu^*(B_n) = \infty$ for some $n \in \mathbb{N}$, there is nothing to prove, and we may hence assume that $\mu^*(B_n) < \infty$ for all n. Let $\epsilon > 0$ be given. For each $n \in \mathbb{N}$, we can find a covering $C_1^{(n)}, C_2^{(n)}, \ldots$ of B_n such that

$$\sum_{k=1}^{\infty} \rho(C_k^{(n)}) < \mu^*(B_n) + \frac{\epsilon}{2^n}$$

The collection $\{C_k^{(n)}\}_{k,n\in\mathbb{N}}$ of all sets in all the coverings is a countable covering of $\bigcup_{n=1}^{\infty} B_n$, and

$$\sum_{k,n\in\mathbb{N}} \rho(C_k^{(n)}) = \sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty} \rho(C_k^{(n)})\right) \leq \sum_{n=1}^{\infty}\left(\mu^*(B_n) + \frac{\epsilon}{2^n}\right) = \sum_{n=1}^{\infty} \mu^*(B_n) + \epsilon$$

(if you feel unsure about these manipulations, take a look at Exercise 5). This means that

$$\mu^*(\bigcup_{n\in\mathbb{N}} B_n) \leq \sum_{n=1}^{\infty} \mu^*(B_n) + \epsilon$$

and since ϵ is an arbitrary, positive number, we must have

$$\mu^*(\bigcup_{n\in\mathbb{N}} B_n) \leq \sum_{n=1}^{\infty} \mu^*(B_n)$$

□

Remark: Note that property (iii) in the proposition above also holds for finite sums; i.e.

$$\mu^*(\bigcup_{n=1}^{N} B_n) \leq \sum_{n=1}^{N} \mu^*(B_n)$$

(to see this, just apply (iii) to the sequence $B_1, B_2, \ldots, B_N, \emptyset, \emptyset, \ldots$). In particular, we always have $\mu^*(A \cup B) \leq \mu^*(A) + \mu^*(B)$.

We have now completed the first part of our program: We have constructed the outer measure and described its fundamental properties. The next step is to define the measurable sets, to prove that they form a σ-algebra, and show that μ^* is a measure when we restrict it to this σ-algebra.

Exercises for Section 6.1

1. Show that $\mu^*(R) \leq \rho(R)$ for all $R \in \mathcal{R}$.
2. Let $X = \{1, 2\}$, $\mathcal{R} = \{\emptyset, \{1\}, \{1, 2\}\}$, and define $\rho : \mathcal{R} \to \overline{\mathbb{R}}$ by $\rho(\emptyset) = 0$, $\rho(\{1\}) = 2$, $\rho(\{1, 2\}) = 1$. Show that $\mu^*(\{1\}) < \rho(\{1\})$.
3. Assume that $X = \mathbb{R}$, and let $\mathcal{R}$ consist of $\emptyset$, X, plus all open intervals (a, b), where $a, b \in \mathbb{R}$. Define $\rho : \mathcal{R} \to \overline{\mathbb{R}}$ by $\rho(\emptyset) = 0$, $\rho(\mathbb{R}) = \infty$, $\rho((a, b)) = b - a$.

a) Show that if $I = [c, d]$ is a closed and bounded interval, and $\mathcal{C} = \{C_n\}$ is a covering of I, then there is a finite number $C_{i_1}, C_{i_2}, \ldots, C_{i_n}$ of sets from $\mathcal{C}$ that covers I (i.e., such that $I \subset C_{i_1} \cup C_{i_2} \cup \ldots \cup C_{i_n}$). (*Hint:* Compactness.)

b) Show that $\mu^*([c, d]) = \rho([c, d]) = d - c$ for all closed and bounded intervals.

4. Assume that $\mathcal{R}$ is a σ-algebra and that ρ is a measure on $\mathcal{R}$. Let $(X, \bar{\mathcal{R}}, \bar{\mu})$ be the completion of $(X, \mathcal{R}, \mu)$. Show that $\mu^*(A) = \bar{\mu}(A)$ for all $A \in \bar{\mathcal{R}}$.

5. Let $\{a_{n,k}\}_{n,k\in\mathbb{N}}$ be a collection of nonnegative, real numbers, and let A be the supremum over all finite sums of distinct elements in this collection, i.e.

$$A = \sup\{\sum_{i=1}^{I} a_{n_i,k_i} \ : \ I \in \mathbb{N} \text{ and all pairs } (n_1, k_1), \ldots, (n_I, k_I) \text{ are different}\}$$

a) Assume that $\{b_m\}_{m\in\mathbb{N}}$ is a sequence which contains each element in the set $\{a_{n,k}\}_{n,k\in\mathbb{N}}$ exactly ones. Show that $\sum_{m=1}^{\infty} b_m = A$.

b) Show that $\sum_{n=1}^{\infty} (\sum_{k=1}^{\infty} a_{n,k}) = A$.

c) Comment on the proof of Proposition 6.1.2(iii).

6.2 Measurable sets

The definition of a measurable set is not difficult, but it is rather mysterious in the sense that it is not easy to see why it captures the essence of measurability.

Definition 6.2.1 *A subset E of X is called* measurable *if*

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) = \mu^*(A)$$

for all $A \subset X$. The collection of all measurable sets is denoted by $\mathcal{M}$.

This approach to measurability was introduced by the Greek mathematician Constantin Carathéodory (1873-1950) and replaced more cumbersome (but easier to understand) earlier approaches (see Exercise 6.3.4 for one such). As already mentioned, it is not at all easy to explain why it captures the intuitive notion of measurability. The best explanation I can offer, is that the reason why some sets are impossible to measure (and hence non-measurable), is that they have very irregular boundaries. The definition above says that a set is measurable if we can use it to cut any other set in two parts without introducing any further irregularities – hence all parts of its boundary must be reasonably regular. I admit that this explanation is rather vague, and a better argument may simply be to show that the definition works. So let us get started.

Let us first of all make a very simple observation. Since $A = (A \cap E) \cup (A \cap E^c)$, subadditivity (recall Proposition 6.1.2(iii)) tells us that we always have

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \geq \mu^*(A)$$

Hence to prove that a set is measurable, we only need to prove that

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \mu^*(A)$$

Our first result is easy.

Lemma 6.2.2 *If E has outer measure 0, then E is measurable. In particular, $\emptyset \in \mathcal{M}$.*

Proof: If E has outer measure 0, so has $A \cap E$ since $A \cap E \subset E$. Hence

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) = \mu^*(A \cap E^c) \leq \mu^*(A)$$

for all $A \subset X$. □

Next we have:

Proposition 6.2.3 *$\mathcal{M}$ satisfies:*[1]

(i) $\emptyset \in \mathcal{M}$.

(ii) *If $E \in \mathcal{M}$, then $E^c \in \mathcal{M}$.*

(iii) *If $E_1, E_2, \ldots, E_n \in \mathcal{M}$, then $E_1 \cup E_2 \cup \ldots \cup E_n \in \mathcal{M}$.*

(iv) *If $E_1, E_2, \ldots, E_n \in \mathcal{M}$, then $E_1 \cap E_2 \cap \ldots \cap E_n \in \mathcal{M}$.*

Proof: We have already proved (i), and (ii) is obvious from the definition of measurable sets. Since $E_1 \cup E_2 \cup \ldots \cup E_n = (E_1^c \cap E_2^c \cap \ldots \cap E_n^c)^c$ by De Morgan's laws, (iii) follows from (ii) and (iv). Hence it only remains to prove (iv).

To prove (iv) is suffices to prove that if $E_1, E_2 \in \mathcal{M}$, then $E_1 \cap E_2 \in \mathcal{M}$ as we can then add more sets by induction. If we first use the measurability of E_1, we see that for any set $A \subset \mathbb{R}^d$

$$\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c)$$

Using the measurability of E_2, we get

$$\mu^*(A \cap E_1) = \mu^*((A \cap E_1) \cap E_2) + \mu^*((A \cap E_1) \cap E_2^c)$$

[1] As you probably know from Chapter 5, a family of sets satisfying (i)-(iii) is usually called an *algebra*. As (iv) is a consequence of (ii) and (iii) using one of De Morgan's laws, the proposition simply states that $\mathcal{M}$ is an algebra, but in the present context (iv) is in many ways the most important property.

Combining these two expressions, we have

$$\mu^*(A) = \mu^*((A \cap (E_1 \cap E_2)) + \mu^*((A \cap E_1) \cap E_2^c) + \mu^*(A \cap E_1^c)$$

Observe that (draw a picture!)

$$(A \cap E_1 \cap E_2^c) \cup (A \cap E_1^c) = A \cap (E_1 \cap E_2)^c$$

and hence

$$\mu^*(A \cap E_1 \cap E_2^c) + \mu^*(A \cap E_1^c) \geq \mu^*(A \cap (E_1 \cap E_2)^c)$$

Putting this into the expression for $\mu^*(A)$ above, we get

$$\mu^*(A) \geq \mu^*((A \cap (E_1 \cap E_2)) + \mu^*(A \cap (E_1 \cap E_2)^c)$$

which means that $E_1 \cap E_2 \in \mathcal{M}$. □

We would like to extend parts (iii) and (iv) in the proposition above to countable unions and intersection. For this we need the following lemma:

Lemma 6.2.4 *If $E_1, E_2, \ldots, E_n$ is a disjoint collection of measurable sets, then*

$$\mu^*(A \cap (E_1 \cup E_2 \cup \ldots \cup E_n)) = \mu^*(A \cap E_1) + \mu^*(A \cap E_2) + \ldots + \mu^*(A \cap E_n)$$

Proof: It suffices to prove the lemma for two sets E_1 and E_2 as we can then extend it by induction. Using the measurability of E_1, we see that

$$\mu^*(A \cap (E_1 \cup E_2)) = \mu^*((A \cap (E_1 \cup E_2)) \cap E_1) + \mu^*(A \cap (E_1 \cup E_2)) \cap E_1^c) =$$

$$= \mu^*(A \cap E_1) + \mu^*(A \cap E_2)$$

□

We can now prove that $\mathcal{M}$ is closed under countable unions.

Lemma 6.2.5 *If $A_n \in \mathcal{M}$ for each $n \in \mathbb{N}$, then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$.*

Proof: Define a new sequence $\{E_n\}$ of sets by $E_1 = A_1$ and

$$E_n = A_n \cap (E_1 \cup E_2 \cup \ldots \cup E_{n-1})^c$$

for $n > 1$, and note that $E_n \in \mathcal{M}$ since $\mathcal{M}$ is an algebra. The sets $\{E_n\}$ are disjoint and have the same union as $\{A_n\}$ (make a drawing!), and hence it suffices to prove that $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{M}$, i.e.

$$\mu^*(A) \geq \mu^*\big(A \cap \bigcup_{n=1}^{\infty} E_n\big) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

Since $\bigcup_{n=1}^{N} E_n \in \mathcal{M}$ for all $N \in \mathbb{N}$, we have:

$$\mu^*(A) = \mu^*\big(A \cap \bigcup_{n=1}^{N} E_n\big) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{N} E_n\big)^c\big) \geq$$

$$\geq \sum_{n=1}^{N} \mu^*(A \cap E_n) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

where we in the last step have used the lemma above plus the observation that $\big(\bigcup_{n=1}^{\infty} E_n\big)^c \subset \big(\bigcup_{n=1}^{N} E_n\big)^c$. Since this inequality holds for all $N \in \mathbb{N}$, we get

$$\mu^*(A) \geq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

By sublinearity, we have $\sum_{n=1}^{\infty} \mu^*(A \cap E_n) \geq \mu^*(\bigcup_{n=1}^{\infty}(A \cap E_n)) = \mu^*(A \cap \bigcup_{n=1}^{\infty}(E_n))$, and hence

$$\mu^*(A) \geq \mu^*\big(A \cap \bigcup_{n=1}^{\infty} E_n\big) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

$\square$

Lemmas 6.2.3 and 6.2.5 tell us that $\mathcal{M}$ is a σ-algebra. We may restrict μ^* to $\mathcal{M}$ to get a function

$$\mu : \mathcal{M} \to \overline{\mathbb{R}}_+$$

defined by[2]

$$\mu(A) = \mu^*(A) \qquad \text{for all } A \in \mathcal{M}$$

We can now complete the second part of our program:

Theorem 6.2.6 *$\mathcal{M}$ is a σ-algebra, and μ is a complete measure on $\mathcal{M}$.*

Proof: We already know that $\mathcal{M}$ is a σ-algebra, and if we prove that μ is a measure, the completeness will follow from Lemma 6.2.2. As we already know that $\mu(\emptyset) = \mu^*(\emptyset) = 0$, we only need to prove that

$$\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n) \tag{6.2.1}$$

for all disjoint sequences $\{E_n\}$ of sets from $\mathcal{M}$

[2]Since μ is just the restriction of μ^* to a smaller domain, it may seem a luxury to introduce a new symbol for it, but in some situations it is important to be able to distinguish quickly between μ and μ^*.

By Proposition 6.1.2(iii), we already know that

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

To get the opposite inequality, we use Lemma 6.2.4 with $A = X$ to see that

$$\sum_{n=1}^{N} \mu(E_n) = \mu(\bigcup_{n=1}^{N} E_n) \leq \mu(\bigcup_{n=1}^{\infty} E_n)$$

Since this holds for all $N \in \mathbb{N}$, we must have

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

Hence we have both inequalities, and (6.2.1) is proved. □

We have now completed the second part of our program — we have shown how we can turn an outer measure μ^* into a measure μ by restricting it to the measurable sets. This is in an interesting result in itself, but we still have some work to do – we need to compare the measure μ to the original function ρ.

Exercises for Section 6.2

1. Explain in detail why 6.2.3(iii) follows from (ii) and (iv).
2. Carry out the induction step in the proof of Proposition 6.2.3(iv).
3. Explain the equality $(A \cap E_1 \cap E_2^c) \cup (A \cap E_1^c) = A \cap (E_1 \cap E_2)^c$ in the proof of Lemma 6.2.3.
4. Carry out the induction step in the proof of Lemma 6.2.4.
5. Explain why the sets E_n in the proof of Lemma 6.2.5 are disjoint and have the same union as the sets A_n. Explain in detail why the sets E_n belong to $\mathcal{M}$.

6.3 Carathéodory's Theorem

In the generality we have have been working so far, there isn't much that can be said about the relationship between the original set function ρ and the measure μ generated via the outer measure construction. Since $R, \emptyset, \emptyset, \ldots$ is a covering of R, we always have $\mu^*(R) \leq \rho(R)$ for all $R \in \mathcal{R}$, but if we want equality instead of inequality, we need to impose conditions on $\mathcal{R}$ and ρ:

Definition 6.3.1 *$\mathcal{R}$ is called an* algebra of sets *on X if the following conditions are satisfied:*

(i) $\emptyset \in \mathcal{R}$

(ii) *If* $R \in \mathcal{R}$, *then* $R^c \in \mathcal{R}$.

(iii) *If* $R, S \in \mathcal{R}$, *then* $R \cup S \in \mathcal{R}$

If $\mathcal{R}$ *is an algebra of sets, the function* $\rho : \mathcal{R} \to \overline{\mathbb{R}}_+$ *is called a* premeasure *if the following conditions are satisfied:*

(iv) $\rho(\emptyset) = 0$

(v) *If* $\{R_n\}_{n\in\mathbb{N}}$ *is a disjoint sequence of sets in* $\mathcal{R}$ *whose union happens to belong to* $\mathcal{R}$, *then*[3]

$$\rho(\bigcup_{n\in\mathbb{N}} R_n) = \sum_{n=1}^{\infty} \rho(R_n)$$

We are now ready for the main result of this section.

Theorem 6.3.2 (Carathéodory's Extension Theorem) *Assume that* $\mathcal{R}$ *is an algebra and that* ρ *is a premeasure on* $\mathcal{R}$. *Then there exists a complete measure* μ *extending* ρ, *i.e., a measure* μ *defined on a* σ*-algebra* $\mathcal{M}$ *containing* $\mathcal{R}$ *such that* $\mu(R) = \rho(R)$ *for all* $R \in \mathcal{R}$.

Proof: We know that the outer measure construction generates a σ-algebra $\mathcal{M}$ of measurable sets and a complete measure μ on $\mathcal{M}$, and it suffices to show that all sets in $\mathcal{R}$ are measurable and that $\mu(R) = \rho(R)$ for all $R \in \mathcal{R}$.

Let us first prove that any set $R \in \mathcal{R}$ is measurable, i.e. that

$$\mu^*(A) \geq \mu^*(A \cap R) + \mu^*(A \cap R^c)$$

for any set $A \subset X$. Given $\epsilon > 0$, we can find a covering $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$ of A such that $\sum_{n=1}^{\infty} \rho(C_n) < \mu^*(A) + \epsilon$. Note that $\{C_n \cap R\}$ and $\{C_n \cap R^c\}$ are coverings of $A \cap R$ and $A \cap R^c$, respectively, and hence

$$\mu^*(A) + \epsilon > \sum_{n=1}^{\infty} \rho(C_n) = \sum_{n=1}^{\infty} \big(\rho(C_n \cap R) + \rho(C_n \cap R^c)\big) =$$

$$= \sum_{n=1}^{\infty} \rho(C_n \cap R) + \sum_{n=1}^{\infty} \rho(C_n \cap R^c) \geq \mu^*(A \cap R) + \mu(A \cap R^c)$$

Since $\epsilon < 0$ is arbitrary, this means that $\mu^*(A) \geq \mu^*(A \cap R) + \mu^*(A \cap R^c)$, and hence R is measurable.

It remains to prove that $\mu(R) = \mu^*(R) = \rho(R)$ for all $R \in \mathcal{R}$. As we already know that $\mu^*(R) \leq \rho(R)$, it suffices to prove the opposite inequality.

[3] In general, there is no reason why such a union should belong to $\mathcal{R}$, but it does happen, e.g. when all but a finite number of the sets R_n equals $\emptyset$

For any $\epsilon > 0$, there is a covering $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$ of R such that $\sum_{n=1}^{\infty} \rho(C_n) < \mu^*(R) + \epsilon$. The sets $C'_n = R \cap (C_n \setminus \bigcup_{k=1}^{n-1} C_k)$ are disjoint elements of $\mathcal{R}$ whose union is R, and hence by condition (v) in the definition above

$$\rho(R) = \sum_{n=1}^{\infty} \rho(C'_n) \leq \sum_{n=1}^{\infty} \rho(C_n) < \mu^*(R) + \epsilon$$

Since $\epsilon > 0$ is arbitrary, this means that $\rho(R) \leq \mu^*(R)$, and since we already have the opposite inequality, we have proved that $\mu^*(R) = \rho(R)$. □

In general, there is more than one measure extending a given premeasure ρ, but for most spaces that occur in practice, there isn't too much freedom (but see Exercise 2 for an extreme case):

Proposition 6.3.3 *Let ρ be a premeasure on an algebra $\mathcal{R}$ and let $(X, \mathcal{M}, \mu)$ be the measure space obtained by the outer measure construction. Assume that $\mathcal{B}$ is a σ-algebra containing $\mathcal{A}$ and that ν is a measure on $\mathcal{B}$ such that $\nu(R) = \rho(R)$ for all $R \in \mathcal{R}$. Then $\nu(A) \leq \mu(A)$ for all $A \in \mathcal{M} \cap \mathcal{B}$, with equality whenever $\mu(A) < \infty$.*

Proof: Assume $A \in \mathcal{M} \cap \mathcal{B}$ and let $\mathcal{C} = \{C_n\}$ be a covering of A. Then

$$\nu(A) \leq \nu(\bigcup_{n\in\mathbb{N}} C_n) \leq \sum_{n=1}^{\infty} \nu(C_n) = \sum_{n=1}^{\infty} \rho(C_n) = |\mathcal{C}|$$

and since $\mu(A) = \mu^*(A) = \inf\{|\mathcal{C}| : \mathcal{C}$ is a covering of $A\}$, we see that $\nu(A) \leq \mu(A)$.

Now assume that $\mu(A) < \infty$. There clearly exists a covering $\mathcal{C} = \{C_n\}$ of A such that $|\mathcal{C}| < \infty$. Replacing C_n by $C_n \setminus (C_1 \cup \ldots C_{n-1})$ if necessary, we may assume that the sets C_n are disjoint, and hence $\mu(C) = \nu(C) = \sum_{n=1}^{\infty} \rho(C_n) < \infty$. We thus have

$$\nu(A) + \nu(C \setminus A) = \nu(C) = \mu(C) = \mu(A) + \mu(C \setminus A)$$

and since we already know that $\nu(A) \leq \mu(A)$ and $\nu(C \setminus A) \leq \mu(C \setminus A)$, this is only possible if $\nu(A) = \mu(A)$ (and $\nu(C \setminus A) = \mu(C \setminus A)$). □

Theoretically, Carathéodory's theorem is a very natural and satisfactory result, but it is a little inconvenient to use in practice as we seldom start with a premeasure defined on an algebra of sets – experience shows that what we usually start from is something slightly weaker called a *semi-algebra*. We shall now extend Carathéodory's result to deal with this situation, and we begin with the definition of a semi-algebra.

Definition 6.3.4 *Let X be a non-empty set and $\mathcal{R}$ a non-empty collection of subsets of X. We call $\mathcal{R}$ a* semi-algebra *if the following conditions are satisfied:*

(i) If $R, S \in \mathcal{R}$, then $R \cap S \in \mathcal{R}$.

(ii) If $R \in \mathcal{R}$, then R^c is a a disjoint union of sets in $\mathcal{R}$.

Starting with a semi-algebra, it is not hard to build an algebra:

Lemma 6.3.5 *Assume that $\mathcal{R}$ is a semi-algebra on a set X and let $\mathcal{A}$ consist of $\emptyset$ plus all finite, disjoint unions of sets in $\mathcal{R}$. Then $\mathcal{A}$ is the algebra generated by $\mathcal{R}$.*

Proof: As all sets in $\mathcal{A}$ clearly have to be in any algebra containing $\mathcal{R}$, we only need to show that $\mathcal{A}$ is an algebra, and for this it suffices to show that $\mathcal{A}$ is closed under complements and finite intersections.

Let us start with the intersections. Assume that A, B are two nonempty sets in $\mathcal{A}$, i.e. that

$$A = R_1 \cup R_2 \cup \ldots \cup R_n$$

$$B = Q_1 \cup Q_2 \cup \ldots \cup Q_m$$

are disjoint unions of sets from $\mathcal{R}$. Then

$$A \cap B = \bigcup_{i,j} (R_i \cap Q_j)$$

is clearly a disjoint, finite union of sets in $\mathcal{R}$, and hence $A \cap B \in \mathcal{A}$. By induction, we see that any finite intersection of sets in $\mathcal{A}$ is in $\mathcal{A}$.

Turning to complements, we first observe that $X = \emptyset^c$ belongs to $\mathcal{A}$. To see this, pick a nonempty set $R \in \mathcal{R}$. Since $\mathcal{R}$ is a semi-algebra, $R^c = R_1 \cup \ldots \cup R_n$ for disjoint sets $R_1, \ldots, R_n \in \mathcal{R}$, and hence

$$X = R \cup R_1 \cup \ldots \cup R_n$$

which proves that $X \in \mathcal{A}$. Assume next that $A = R_1 \cup R_2 \cup \ldots \cup R_n$ is a disjoint union of sets in $\mathcal{R}$. Then by one of De Morgan's laws

$$A^c = R_1^c \cap R_2^c \cap \ldots \cap R_n^c$$

Since $\mathcal{R}$ is a semi-algebra, each R_i^c is a disjoint union of sets in $\mathcal{R}$ and hence belongs to $\mathcal{A}$. Since we have already proved that $\mathcal{A}$ is closed under finite intersections, $A^c \in \mathcal{A}$. □

Here is the promised extension of Carathéodory's theorem to semi-algebras:

Theorem 6.3.6 (Carathéodory's Theorem for Semi-Algebras) *Let $\mathcal{R}$ be a semi-algebra of subsets of X and assume that $\lambda : \mathcal{R} \to \overline{\mathbb{R}}_+$ is a function such that*

(i) *if a set $R \in \mathcal{R}$ is a disjoint, finite union $R = \bigcup_{i=1}^n R_i$ of sets in $\mathcal{R}$, then*

$$\lambda(R) = \sum_{i=1}^n \lambda(R_i)$$

(ii) *If a set $R \in \mathcal{R}$ is a disjoint, countable union $R = \bigcup_{i\in\mathbb{N}} R_i$ of sets in $\mathcal{R}$, then*

$$\lambda(R) = \sum_{i=1}^\infty \lambda(R_i)$$

Then λ has an extension to a complete measure on a σ-algebra containing $\mathcal{R}$.[4]

Before we turn to the proof of the theorem, it is convenient to prove the following lemma.

Lemma 6.3.7 *Assume that $\mathcal{R}$ and λ are as in the theorem above.*

(i) *Assume that a set $A \subset X$ can be written as disjoint, finite unions of sets in $\mathcal{R}$ in two different ways: $A = \bigcup_{i=1}^n R_i$ and $A = \bigcup_{j=1}^m S_j$. Then*

$$\sum_{i=1}^n \lambda(R_i) = \sum_{j=1}^m \lambda(S_j)$$

(ii) *Assume that a set $A \subset X$ can be written as disjoint unions of sets in $\mathcal{R}$ in a finite and a countable way: $A = \bigcup_{i=1}^n R_i$ and $A = \bigcup_{j\in\mathbb{N}} S_j$. Then*

$$\sum_{i=1}^n \lambda(R_i) = \sum_{j=1}^\infty \lambda(S_j)$$

Proof: We only prove (ii). The proof of (i) is similar, the only difference being that we in one place use condition (i) of the theorem instead of condition (ii).

To prove (ii), observe that since $\mathcal{R}$ is a semi-algebra, the intersections $R_i \cap S_j$ belong to $\mathcal{R}$, and hence by condition (ii) in the theorem

$$\lambda(R_i) = \sum_{j=1}^\infty \lambda(R_i \cap S_j)$$

Similarly by condition (i) in the theorem

$$\lambda(S_j) = \sum_{i=1}^n \lambda(R_i \cap S_j)$$

[4] If $\emptyset \in \mathcal{R}$ and $\lambda(\emptyset) = 0$, (i) is just a special case of (ii), but since we haven't assumed that $\emptyset \in \mathcal{R}$, we have to treat finite and infinite unions separately.

This means that

$$\sum_{i=1}^{n}\lambda(R_i)=\sum_{i=1}^{n}\sum_{j=1}^{\infty}\lambda(R_i\cap S_j)=\sum_{j=1}^{\infty}\sum_{i=1}^{n}\lambda(R_i\cap S_j)=\sum_{j=1}^{\infty}\lambda(S_j)$$

which is what we wanted to prove. □

Proof of Carathéodory's theorem for semi-algebras: Let $\mathcal{A}$ be the algebra generated by $\mathcal{R}$. The plan is to extend λ to a premeasure ρ on $\mathcal{A}$, and then use the version of Carathéodory's theorem that we have already proved.

Since any non-empty set A in $\mathcal{A}$ is a finite, disjoint union $A=\bigcup_{i=1}^{n}R_i$ of sets in $\mathcal{R}$, we define $\rho:\mathcal{A}\to\overline{\mathbb{R}}_+$ by putting $\rho(\emptyset)=0$ and

$$\rho(A)=\sum_{i=1}^{n}\lambda(R_i)$$

(ρ is well-defined as part (i) of the lemma tells us that if $A=\bigcup_{j=1}^{m}S_j$ is another way of writing A as a disjoint union of sets in $\mathcal{R}$, then $\sum_{i=1}^{n}\lambda(R_i)=\sum_{j=1}^{m}\lambda(S_j)$). To use Carathéodory's Extension Theorem 6.3.2, we need to show that ρ is a premeasure, i.e. that

$$\rho(A)=\sum_{n=1}^{\infty}\rho(A_n)$$

whenever A and $A_1,A_2,\ldots$ are in $\mathcal{A}$ and A is the disjoint union of $A_1,A_2,\ldots$. Since A and the A_n's are in $\mathcal{A}$, they can be written as finite, disjoint unions of sets in $\mathcal{R}$:

$$A=\bigcup_{j=i}^{M}R_j$$

and

$$A_n=\bigcup_{k=1}^{N_n}S_{n,k}$$

Since $A=\bigcup_{n\in\mathbb{N}}A_n=\bigcup_{n,k}S_{n,k}$, part (ii) of the lemma tells us that

$$\sum_{j=1}^{M}\lambda(R_j)=\sum_{n,k}\lambda(S_{n,k})$$

The rest is easy: By definition of ρ

$$\rho(A)=\sum_{j=1}^{M}\lambda(R_j)=\sum_{n,k}\lambda(S_{n,k})=\sum_{n=1}^{\infty}\sum_{k=1}^{N_n}\lambda(S_{n,k})=\sum_{n=1}^{\infty}\rho(A_n)$$

We have now proved that ρ is a premeasure on the algebra $\mathcal{A}$, and by the original Carathéodory's theorem 6.3.2, there is a complete measure on a σ-algebra containing $\mathcal{A}$ which extends ρ (and hence λ). $\square$

We now have the machinery we need to construct measures, and in the next section we shall use it to construct Lebesgue measure on $\mathbb{R}$.

Exercises for Section 6.3

1. Prove part (i) of Lemma 6.3.7.
2. A measure space $(X, \mathcal{A}, \mu)$ is called *σ-finite* if there there is a disjoint family $\{A_n\}_{n\in\mathbb{N}}$ of sets in $\mathcal{A}$ such that $X = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.
   a) Show that $(X, \mathcal{A}, \mu)$ is σ-finite if and only if there is an increasing family $\{E_n\}_{n\in\mathbb{N}}$ of sets in $\mathcal{A}$ such that $X = \bigcup_{n\in\mathbb{N}} E_n$ and $\mu(E_n) < \infty$ for all $n \in \mathbb{N}$.
   b) Show that the Lebesgue measure is σ-finite.
   c) Show that if the measure μ in Proposition 6.3.3 is σ-finite, then $\nu(A) = \mu(A)$ for all $A \in \mathcal{M}\cap\mathcal{B}$. Hence μ has a unique extension to the σ-algebra generated by $\mathcal{R}$.
3. Let $X = \mathbb{Q}$ and let $\mathcal{R}$ be the collection of subsets of X consisting of
   (i) All bounded, half-open, rational intervals $(r, s]_{\mathbb{Q}} = \{q \in \mathbb{Q} \mid r < q \leq s\}$ where $r, s \in \mathbb{Q}$.
   (ii) All unbounded, half-open, rational intervals $(-\infty, s]_{\mathbb{Q}} = \{q \in \mathbb{Q} \mid q \leq s\}$ and $(r, \infty)_{\mathbb{Q}} = \{q \in \mathbb{Q} \mid r < q\}$ where $r, s \in \mathbb{Q}$.

   a) Show that $\mathcal{R}$ is a semi-algebra.
   b) Show that the σ-algebra $\mathcal{M}$ generated by $\mathcal{R}$ is the collection of all subset of X.

   Define $\rho : \mathcal{R} \to \overline{\mathbb{R}}_+$ by $\rho(\emptyset) = 0$, $\rho(R) = \infty$ otherwise.

   c) Show that ρ is a satisfies the conditions of Carathéodory's theorem for semi-algebras.
   d) Show that there are infinitely many extensions of ρ to a measure ν on $\mathcal{M}$. (*Hint:* Which values may $\nu(\{q\})$ have?)
4. In this problem we shall take a look at a different (and perhaps more intuitive) approach to measurability. We assume that $\mathcal{R}$ is an algebra and that ρ is a premeasure on $\mathcal{R}$. We also assume that $\rho(X) < \infty$. Define the *inner measure* of a subset E of X by
$$\mu_*(E) = \mu^*(X) - \mu^*(E^c)$$
   We call a set E *$*$-measurable* if $\mu^*(E) = \mu_*(E)$. (Why is this a natural condition?)
   a) Show that $\mu_*(E) \leq \mu^*(E)$ for all $E \subset X$.

b) Show that if E is measurable, then E is $*$-measurable. (*Hint:* Use Definition 6.2.1 with $A = X$.)

c) Show that if E is $*$-measurable, then for every $\epsilon > 0$ there are measurable sets D, F such that $D \subset E \subset F$ and

$$\mu^*(D) > \mu^*(E) - \frac{\epsilon}{2} \qquad \text{and} \qquad \mu^*(F) < \mu^*(E) + \frac{\epsilon}{2}$$

d) Show that $\mu^*(F \backslash D) < \epsilon$ and explain that $\mu^*(F \backslash E) < \epsilon$ and $\mu^*(E \backslash D) < \epsilon$.

e) Explain that for any set $A \subset X$

$$\mu^*(A \cap F) = \mu^*(A \cap D) + \mu^*(A \cap (F \setminus D)) < \mu^*(A \cap D) + \epsilon$$

and

$$\mu^*(A \cap D^c) = \mu^*(A \cap F^c) + \mu^*(A \cap (F \setminus D)) < \mu^*(A \cap F^c) + \epsilon$$

and use this to show that $\mu^*(A \cap D) > \mu^*(A \cap E) - \epsilon$ and $\mu^*(A \cap F^c) > \mu^*(A \cap E^c) - \epsilon$.

f) Explain that for any set $A \subset X$

$$\mu^*(A) \geq \mu^*(A \cap (F^c \cup D)) =$$

$$= \mu^*(A \cap F^c) + \mu^*(A \cap D) \geq \mu^*(A \cap E^c) + \mu^*(A \cap E) - 2\epsilon$$

and use it to show that if E is $*$-measurable, then E is measurable. The notions of measurable and $*$-measurable hence coincide when $\mu^*(X)$ is finite.

6.4 Lebesgue measure on $\mathbb{R}$

In this section we shall use the theory in the previous section to construct the Lebesgue measure on $\mathbb{R}$. Essentially the same argument can be used to construct Lebesgue measure on $\mathbb{R}^d$ for $d > 1$, but as the geometric considerations become a little more complicated, we shall restrict ourselves to the one dimensional case. In Section 6.6 we shall see how we can obtain higher dimensional Lebesgue measure by a different method.

Recall that the one dimensional Lebesgue measure is a generalization of the notion of length: We know how long intervals are and want to extend this notion of size to a full-blown measure. For technical reasons, it is convenient to work with half-open intervals $(a, b]$.

Definition 6.4.1 *In this section, $\mathcal{R}$ consists of the following subsets of $\mathbb{R}$:*

(i) All finite, half-open intervals $(a, b]$, where $a, b \in \mathbb{R}$, $a < b$.

(ii) All infinite intervals of the form $(-\infty, b]$ and (a, ∞) where $a, b \in \mathbb{R}$.

The advantage of working with half-open intervals become clear when we check what happens when we take intersections and complements:

Proposition 6.4.2 *$\mathcal{R}$ is a semi-algebra of sets.*

Proof: I leave it to the reader to check the various cases. The crucial observation is that the complement of a half-open interval is either a half-open interval or the union of two half-open intervals. □

We define $\lambda : \mathcal{R} \to \overline{\mathbb{R}}_+$ simply by letting $\lambda(R)$ be the length of the interval R. To use Carathéodory's theorem for semi-algebras, we first need to check that the main condition is satisfied.

Lemma 6.4.3 *If a set $R \in \mathcal{R}$ is a disjoint, countable union $R = \bigcup_{i\in\mathbb{N}} R_i$ of sets in $\mathcal{R}$, then*

$$\lambda(R) = \sum_{i=1}^{\infty} \lambda(R_i)$$

The same holds for finite unions: If $R \in \mathcal{R}$ is a disjoint, finite union $R = R_1 \cup R_2 \cup \ldots \cup R_n$ of sets in $\mathcal{R}$, then $\lambda(R) = \lambda(R_1) + \lambda(R_2) + \ldots \lambda(R_n)$.

Proof: The case for finite unions is easy and left to the reader. For the countable case, note that for any finite N, the nonoverlapping intervals $R_1, R_2, \ldots, R_N$ can be ordered from left to right, and obviously make up a part of R in a such a way that $\lambda(R) \geq \sum_{n=1}^{N} \lambda(R_n)$. Since this holds for all finite N, we must have $\lambda(R) \geq \sum_{n=1}^{\infty} \lambda(R_n)$.

The opposite inequality is more subtle, and we need to use a compactness argument. Let us first assume that R is a finite interval $(a, b]$. Given an $\epsilon > 0$, we extend each interval $R_n = (a_n, b_n]$ to an open interval $\hat{R}_n = (a_n, b_n + \frac{\epsilon}{2^n})$. These open intervals cover the compact interval $[a + \epsilon, b]$, and by Theorem 2.6.6 there is a finite subcover $\hat{R}_{n_1}, \hat{R}_{n_2}, \ldots, \hat{R}_{n_k}$. Since these finite intervals cover an interval of length $b - a - \epsilon$, we clearly have $\sum_{j=1}^{k} \lambda(\hat{R}_{n_j}) \geq (b - a - \epsilon)$. But since $\lambda(\hat{R}_n) = \lambda(R_n) + \frac{\epsilon}{2^n}$, this means that $\sum_{j=1}^{k} \lambda(R_{n_j}) \geq (b - a - 2\epsilon)$. Consequently, $\sum_{n=1}^{\infty} \lambda(R_n) \geq b - a - 2\epsilon = \lambda(R) - 2\epsilon$, and since ϵ is an arbitrary, positive number, we must have $\sum_{n=1}^{\infty} \lambda(R_n) \geq \lambda(R)$.

It remains to see what happens if R is an interval of infinite length, i.e. when $\lambda(R) = \infty$. For each $n \in \mathbb{N}$, the intervals $\{R_i \cap (-n, n]\}_{i\in\mathbb{N}}$ are disjoint and have union $R \cap (-n, n]$. By what we have already proved, $\lambda(R \cap (-n, n]) = \sum_{i=1}^{\infty} \lambda(R_i \cap (-n, n]) \leq \sum_{i=1}^{\infty} \lambda(R_i)$. Since $\lim_{n\to\infty} \lambda(R \cap (-n, n]) = \infty$, we see that $\sum_{i=1}^{\infty} \lambda(R_i) = \infty = \lambda(R)$. □

We have reached our goal: The lemma above tells us that we can apply Carathéodory's theorem for semi-algebras to the semi-algebra $\mathcal{R}$ and the function λ. The resulting measure μ is the *Lebesgue measure on* $\mathbb{R}$. Let us sum up the essential points.

Theorem 6.4.4 *The Lebesgue measure μ is the unique, completed Borel measure on $\mathbb{R}$ such that $\mu(I) = |I|$ for all intervals I.*

Proof: By construction, the μ-measure of all half-open intervals is equal to their length, and by continuity of measure (Proposition 5.1.5), the same must hold for open and closed intervals. By Proposition 6.3.3, if ν is any other completed, Borel measure with the same property, $\mu(A) = \nu(A)$ for all Borel sets such that $\mu(A) < \infty$. But the same must hold for a general Borel set A as we can slice it up into pieces of finite measure:

$$\mu(A) = \sum_{n\in\mathbb{Z}} \mu\big(A \cap (n, n+1]\big) = \sum_{n\in\mathbb{Z}} \nu\big(A \cap (n, n+1]\big) = \nu(A)$$

□

The next result tells us that measurable sets can be approximated arbitrarily well by open and closed set.

Proposition 6.4.5 *Assume that $A \subset \mathbb{R}$ is a measurable set. For each $\epsilon > 0$, there is an open set $G \supset A$ such that $\mu(G \setminus A) < \epsilon$, and a closed set $F \subset A$ such that $\mu(A \setminus F) < \epsilon$.*

Proof: We begin with the open sets. Assume first that A has finite measure. Then for every $\epsilon > 0$ there is a covering $\{C_n\}$ of A by half-open rectangles $C_n = (a_n, b_n]$ such that

$$\sum_{n=1}^{\infty} |C_n| < \mu(A) + \frac{\epsilon}{2}$$

If we replace the half-open intervals $C_n = (a_n, b_n]$ by the open intervals $B_n = (a_n, b_n + \frac{\epsilon}{2^{n+1}})$, we get an open set $G = \bigcup_{n=1}^{\infty} B_n$ containing A with

$$\mu(G) \le \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \left(|C_n| + \frac{\epsilon}{2^{n+1}}\right) < \mu(A) + \epsilon$$

and hence

$$\mu(G \setminus A) = \mu(G) - \mu(A) < \epsilon$$

by Lemma 5.1.4c).

If $\mu(A)$ is infinite, we slice A into pieces of finite measure $A_n = A \cap (n, n+1]$ for all $n \in \mathbb{Z}$, and use what we have already proved to find an open set G_n such that $A_n \subset G_n$ and $\mu(G_n \setminus A_n) < \frac{\epsilon}{2^{n+2}}$. Then $G = \bigcup_{n\in\mathbb{Z}} G_n$ is an open set which contains A, and since $G \setminus A \subset \bigcup_{n\in\mathbb{Z}} (G_n \setminus A_n)$, we get

$$\mu(G \setminus A) \le \sum_{n\in\mathbb{Z}} \mu(G_n \setminus A_n) < \sum_{n\in\mathbb{Z}} \frac{\epsilon}{2^{n+2}} < \epsilon,$$

proving the statement about approximation by open sets.

To prove the statement about closed sets, just note that if we apply the first part of the theorem to A^c, we get an open set $G \supset A^c$ such that $\mu(G \setminus A^c) < \epsilon$. This means that $F = G^c$ is a closed set such that $F \subset A$, and since $A \setminus F = G \setminus A^c$, we have $\mu(A \setminus F) < \epsilon$. □

Another important property of the Lebesgue measure is that it is *translation invariant* — if we move a set a distance to the left or to the right, it keeps its measure. To formulate this mathematically, let E be a subset of $\mathbb{R}$ and a a real number, and write

$$E + a = \{e + a \,|\, e \in E\}$$

for the set we obtain by moving all points in E a distance a.

Proposition 6.4.6 *If $E \subset \mathbb{R}$ is measurable, so is $E + a$ for all $a \in \mathbb{R}$, and $\mu(E + a) = \mu(E)$.*

Proof: We shall leave this to the reader (see Exercise 7). The key observation is that $\mu^*(E + a) = \mu^*(E)$ holds for outer measure since intervals keep their length when we translate them. □

One of the reasons why we had to develop the rather complicated machinery of σ-algebras, is that we can not in general expect to define a measure on *all* subsets of our measure space X — some sets are just so complicated that they are nonmeasurable. We shall now take a look at such a set. Before we begin, we need to modify the notion of translation so that it works on the interval $[0, 1)$. If $x, y \in [0, 1)$, we first define

$$x \dot{+} y = \begin{cases} x + y & \text{if } x + y \in [0,1) \\ \\ x + y - 1 & \text{otherwise} \end{cases}$$

If $E \subset [0, 1)$ and $y \in [0, 1)$, let

$$E \dot{+} y = \{e \dot{+} y \,|\, e \in E\}$$

Note that $E \dot{+} y$ is the set obtained by first translating E by y and then moving the part that sticks out to the right of $[0, 1)$ one unit to left so that it fills up the empty part of $[0, 1)$. It follows from translation invariance that $E \dot{+} y$ is measurable if E is, and that $\mu(E) = \mu(E \dot{+} y)$.

Example 1: A nonmeasurable set. We start by introducing an equivalence relation $\sim$ on the interval $[0, 1)$:

$$x \sim y \Longleftrightarrow x - y \text{ is rational}$$

Next, we let E be a set obtained by picking one element from each equivalence class.[5] We shall work with the sets $E \dot{+} q$ for all rational numbers q in the interval $[0,1)$, i.e., for all $q \in \hat{\mathbb{Q}} = \mathbb{Q} \cap [0,1)$.

First observe that if $q_1 \neq q_2$, then $(E \dot{+} q_1) \cap (E \dot{+} q_2) = \emptyset$. If not, we could write the common element x as both $x = e_1 \dot{+} q_1$ and $x = e_2 \dot{+} q_2$ for some $e_1, e_2 \in E$. The equality $e_1 \dot{+} q_1 = e_2 \dot{+} q_2$, implies that $e_1 - e_2$ is rational, and by definition of E this is only possible if $e_1 = e_2$. Hence $q_1 = q_2$, contradicting the assumption that $q_1 \neq q_2$.

The next observation is that $[0,1) = \bigcup_{q\in\hat{\mathbb{Q}}}(E \dot{+} q)$. To see this, pick an arbitrary $x \in [0,1)$. By definition, there is an e in E that belongs to the same equivalence class as x, i.e. such that $q = x - e$ is rational. If $q \in [0,1)$, then $x \in E \dot{+} q$, if $q < 0$ (the only other possibility), we have $x \in E \dot{+} (q+1)$ (check this!). In either case, $x \in \bigcup_{q\in\hat{\mathbb{Q}}}(E \dot{+} q)$.

Assume for contradiction that E is Lebesgue measurable. Then, as already observed, $E \dot{+} q$ is Lebesgue measurable with $\mu(E \dot{+} q) = \mu(E)$ for $q \in \hat{\mathbb{Q}}$. But by countable additivity

$$\mu([0,1)) = \sum_{q\in\hat{\mathbb{Q}}} \mu(E \dot{+} q)$$

and since $\mu([0,1)) = 1$, this is impossible — a sum of countable many, equal, nonnegative numbers is either ∞ (if the numbers are positive) or 0 (if the numbers are 0).

Note that this argument works not only for the Lebesgue measure, but for any (non-zero) translation invariant measure on $\mathbb{R}$. This means that it is impossible to find a translation invariant measure on $\mathbb{R}$ that makes all sets measurable. ♣

The existence of nonmeasurable sets is not a surprise – there is no reason to expect that all sets should be measurable – but it is a nuisance which complicates many arguments. As we have seen in this chapter, the hard part is often to find the right class of measurable sets and prove that it is a σ-algebra.

Exercises for Section 6.4

1. Complete the proof of Proposition 6.4.2.
2. Prove the "finite case" of Lemma 6.4.3.
3. In the proof of Lemma 6.4.3, explain why $\sum_{n=1}^{\infty} |I_n| \geq b - a - 2\epsilon = |I| - 2\epsilon$.
4. Explain that $A \setminus F = G \setminus A^c$ at the end of the proof of Proposition 6.4.5.

[5]Here we are using a principle from set theory called the Axiom of Choice which allows us to make a new set by picking one element from each set in an infinite family. It has been proved that it is not possible to construct a nonmeasurable subset of $\mathbb{R}$ without using a principle of this kind.

5. A subset of $\mathbb{R}$ is called a $\mathcal{G}_\delta$-set if it is the intersection of countably many open sets, and it is called a $\mathcal{F}_\sigma$-set if it is union of countably many closed set.
   a) Explain why all $\mathcal{G}_\delta$- and $\mathcal{F}_\sigma$-sets are measurable.
   b) Show that if $A \subset \mathbb{R}$ is measurable, there is a $\mathcal{G}_\delta$-set G such that $A \subset G$ and $\mu(G \setminus A) = 0$.
   c) Show that if $A \subset \mathbb{R}$ is measurable, there is a $\mathcal{F}_\sigma$-set F such that $F \subset A$ and $\mu(A \setminus F) = 0$.
6. Assume that $A \subset \mathbb{R}$ is a measurable set with finite measure. Show that for every $\epsilon > 0$, there is a compact set $K \subset A$ such that $\mu(A \setminus K) < \epsilon$.
7. Prove Proposition 6.4.6 by:
   a) Showing that if $E \subset \mathbb{R}$ and $a \in \mathbb{R}$, then $\mu^*(E + a) = \mu(E)$.
   b) Showing that if $E \subset \mathbb{R}$ is measurable, then so is $E + a$ for any $a \in \mathbb{R}$.
   c) Explaining how to obtain the proposition from a) and b).
8. Check that the equivalence relation in Example 1 really is an equivalence relation.
9. If A is a subset of $\mathbb{R}$ and r is a positive, real number, let
$$rA = \{ra \mid a \in A\}$$
   Show that if A is measurable, then so is rA and $\mu(rA) = r\mu(A)$.
10. Use Proposition 6.4.6 to prove that that if $E \subset [0, 1)$ is Lebesgue measurable, then $E \dot{+} y$ is Lebesgue measurable with $\mu(E \dot{+} y) = \mu(E)$ or all $y \in [0, 1)$.

6.5 The coin tossing measure

As a second example of how to construct measures, we shall construct the natural probability measure on the space of infinite sequences of unbiased coin tosses (recall Example 8 of Section 5.1). Be aware that this section is rather sketchy – it is more like a structured sequence of exercises than an ordinary section. The results here will not be used in the sequel.

Let us recall the setting. The underlying space Ω consists of all infinite sequences
$$\omega = (\omega_1, \omega_2, \ldots, \omega_n, \ldots)$$
where each ω_n is either H (for heads) or T (for tails). If $\mathbf{a} = a_1, a_2, \ldots, a_n$ is a *finite* sequence of H's and T's, we let
$$\mathcal{C}_{\mathbf{a}} = \{\omega \in \Omega \mid \omega_1 = a_1, \omega_2 = a_2, \ldots, \omega_n = a_n\}$$
and call it the *cylinder set* generated by $\mathbf{a}$. We call n the *length* of $\mathcal{C}_{\mathbf{a}}$. Let $\mathcal{R}$ be the collection of all cylinder sets (of all possible lengths).

Lemma 6.5.1 *$\mathcal{R}$ is a semi-algebra.*

Proof: The intersection of two cylinder sets $\mathcal{C}_{\mathbf{a}}$ and $\mathcal{C}_{\mathbf{b}}$ is either equal to one of them (if one of the sequences $\mathbf{a}$, $\mathbf{b}$ is an extension of the other) or empty. The complement of a cylinder set is the union of all other cylinder sets of the same length. $\square$

We define a function $\lambda : \mathcal{R} \to [0, 1]$ by putting

$$\lambda(\mathcal{C}_{\mathbf{a}}) = \frac{1}{2^n}$$

where n is the length of $\mathcal{C}_{\mathbf{a}}$. (There are 2^n cylinder sets of length n and they correspond to 2^n equally probable events).

We first check that λ behaves the right way under finite unions:

Lemma 6.5.2 *If $A_1, A_2, \ldots, A_N$ are disjoint sets in $\mathcal{R}$ whose union $\bigcup_{n=1}^N A_n$ belongs to $\mathcal{R}$, then*

$$\lambda(\bigcup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \lambda(A_n)$$

Proof: Left to the reader (see Exercise 1). $\square$

To use Carathéodory's theorem for semi-algebras we must extend the result above to countable unions, i.e., we need to show that if $\{A_n\}$ is a disjoint sequence of cylinder sets whose union is a cylinder set, then

$$\lambda(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{i=1}^{\infty} \lambda(A_n)$$

The next lemma tells us that this condition is trivially satisfied because the situation never occurs – a cylinder set is *never* a disjoint, countable union of infinitely many cylinder sets! As this is the difficult part of the construction, I spell out the details. The argument is actually a compactness argument in disguise, and corresponds to Lemma 6.4.3 in the construction of the Lebesgue measure.

Before we begin, we need some notation and terminology. For each $k \in \mathbb{N}$, we write $\omega \sim_k \hat{\omega}$ if $\omega_i = \hat{\omega}_i$ for $i = 1, 2, \ldots, k$, i.e., if the first k coin tosses in the sequences ω and $\hat{\omega}$ are equal. We say that a subset A of Ω is *determined at $k \in \mathbb{N}$* if whenever $\omega \sim_k \hat{\omega}$ then either both ω and $\hat{\omega}$ belong to A or none of them do (intuitively, this means that you can decide whether ω belongs to A by looking at the k first coin tosses). A cylinder set of length k is obviously determined at time k. We say that a set A is *finally determined* if it is determined at some $k \in \mathbb{N}$.

Lemma 6.5.3 *A cylinder set A is never an infinite, countable, disjoint union of cylinder sets.*

Proof: Assume for contradiction that $\{A_n\}_{n\in\mathbb{N}}$ is a disjoint sequence of cylinder sets whose union $A = \bigcup_{n\in\mathbb{N}} A_n$ is also a cylinder set. Since the A_n's are disjoint and nonempty, the set $B_N = A \setminus \bigcup_{n=1}^{N} A_n$ is nonempty for all $N \in \mathbb{N}$. Note that the sets B_N are finitely determined since A and the A_n's are.

Since the sets B_N are decreasing and nonempty, there must either be strings $\omega = (\omega_1, \omega_2, \ldots)$ starting with $\omega_1 = H$ in all the sets B_N or strings starting with $\omega_1 = T$ in all the sets B_N (or both, in which case we just choose H). Call the appropriate symbol $\hat{\omega}_1$. Arguing in the same way, we see that there must either be strings starting with $(\hat{\omega}_1, H)$ or strings starting with $(\hat{\omega}_1, T)$ in all the sets B_N (or both, in which case we just choose H). Call the appropriate symbol $\hat{\omega}_2$. Continuing in this way, we get an infinite sequence $\hat{\omega} = (\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3, \ldots)$ such that for each $k \in \mathbb{N}$, there is a sequence starting with $(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3, \ldots, \hat{\omega}_k)$ in each B_N. Since B_N is finitely determined, this means that $\hat{\omega} \in B_N$ for all N (just choose k so large that B_N is determined at k). But this implies that $\hat{\omega} \in A \setminus \bigcup_{n\in\mathbb{N}} A_n$, which is a contradiction since $A = \bigcup_{n\in\mathbb{N}} A_n$ by assumption. □

We are now ready for the main result:

Theorem 6.5.4 *There is a complete measure P on Ω such that $P(A) = \lambda(A)$ for all finitely determined sets A. The measure is unique on the σ-algebra of measurable sets.*

Proof: The three lemmas above give us exactly what we need to apply Carathéodory's theorem for semi-algebras. The uniqueness follows from Proposition 6.3.3. The details are left to the reader. □

Exercises for Section 6.5

1. Prove Lemma 6.5.2 (*Hint:* If the cylinder sets $A_1, A_2, \ldots, A_N$ have length $K_1, K_2, \ldots, K_N$, respectively, then this is really a statement about finite sequences of length $K = \max\{K_1, K_2, \ldots, K_N\}$.)
2. Fill in the details in the proof of Theorem 6.5.4.
3. Let $\{B_n\}_{n\in\mathbb{N}}$ be a decreasing sequence of nonempty, finitely determined sets. Show that $\bigcap_{n\in\mathbb{N}} B_n \neq \emptyset$.
4. Let $E = \{\omega \in \Omega \,|\, \omega \text{ contains infinitely many } H\text{'s}\}$. Show that E is not finitely determined.
5. Let $H = 1$, $T = 0$. Show that

$$E = \{\omega \in \Omega \,|\, \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \omega_i = \frac{1}{2}\}$$

is not finitely determined.

6. Show that a set is in the algebra generated by $\mathcal{R}$ if and only if it is finitely determined. Show that if E is a measurable set, then there is for any $\epsilon > 0$ a finitely determined set D such that $P(E \triangle D) < \epsilon$

7. (You should do Exercise 6 before you attempt this one.) Assume that I is a nonempty subset of $\mathbb{N}$. Define an equivalence relation $\sim_I$ on Ω by

$$\omega \sim_I \hat{\omega} \Longleftrightarrow \omega_i = \hat{\omega}_i \text{ for all } i \in I$$

We say that a set $B \subset \Omega$ is *I-determined* if whenever $\omega \sim_I \hat{\omega}$, then either ω and $\hat{\omega}$ are both in B or both in B^c (intuitively, this means that we can determine whether ω belongs to B by looking at the coin tosses ω_i where $i \in I$).

a) Let $\mathcal{A}$ be the σ-algebra of all measurable sets, and define

$$\mathcal{A}_I = \{A \in \mathcal{A} \mid A \text{ is } I\text{-determined}\}$$

Show that $\mathcal{A}_I$ is a σ-algebra.

b) Assume that $A \in \mathcal{A}_I$ and that C is a finitely determined set such that $A \subset C$. Show that there is a finitely determined $\hat{C} \in \mathcal{A}_I$ such that $A \subset \hat{C} \subset C$.

c) Assume that $A \in \mathcal{A}_I$. Show that for any $\epsilon > 0$, there is a finitely determined $B \in \mathcal{A}_I$ such that $P(A \triangle B) < \epsilon$.

d) Assume that $I, J \subset \mathbb{N}$ are disjoint. Show that if $B \in \mathcal{A}_I$ and $D \in \mathcal{A}_J$ are finitely generated, then $P(B \cap D) = P(B)P(D)$. In the language of probability theory, B and D are *independent events.* (*Hint:* This is just finite combinatorics and has nothing to do with measures. Note that finitely determined sets are in the algebra generated by $\mathcal{R}$, and hence their measures are given directly in terms of λ.)

e) We still assume that $I, J \subset \mathbb{N}$ are disjoint. Show that if $A \in \mathcal{A}_I$ and $C \in \mathcal{A}_J$, then $P(A \cap C) = P(A)P(C)$. (*Hint:* Combine c) and d).)

f) Let $I_n = \{n, n+1, n+2, \ldots\}$. A set $E \subset \Omega$ is called a *tail event* if $E \in I_n$ for all $n \in \mathbb{N}$. Show that the sets E in Exercises 4 and 5 are tail events.

g) Assume that E is a tail event and that A is finitely generated. Show that $P(A \cap E) = P(A)P(E)$.

h) Assume that E is a tail event, and let E_n be finitely determined sets such that $P(E \triangle E_n) < \frac{1}{n}$ (such sets exist by Exercise 6). Show that $P(E \cap E_n) \to P(E)$ as $n \to \infty$.

i) Show that on the other hand $P(E \cap E_n) = P(E_n)P(E) \to P(E)^2$. Conclude that $P(E) = P(E)^2$, which means that $P(E) = 0$ or $P(E) = 1$. We have proved *Kolmogorov's 0-1-law*: A tail event can only have probability 0 or 1.

6.6 Product measures

In calculus you learned how to compute double integrals by iterated integration: If $R = [a, b] \times [c, d]$ is a rectangle in the plane, then

$$\iint_R f(x, y)\, dxdy = \int_c^d \left[\int_a^b f(x, y)\, dx\right] dy = \int_a^b \left[\int_c^d f(x, y)\, dy\right] dx$$

There is a similar result in measure theory that we are now going to look at. Starting with two measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$, we shall first construct a product measure space $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \times \nu)$ and then prove that

$$\int f\, d(\mu \times \nu) = \int \left[\int f(x, y)\, d\mu(x)\right] d\nu(y) = \int \left[\int f(x, y)\, d\nu(y)\right] d\mu(x)$$

(I am putting in x's and y's to indicate which variables we are integrating with respect to). The guiding light in the construction of the product measure $\mu \times \nu$ is that we want it to satisfy the natural product rule

$$\mu \times \nu(A \times B) = \mu(A)\nu(B)$$

for all $A \in \mathcal{A}$ and all $B \in \mathcal{B}$ (think of the formula for the area of an ordinary rectangle).

As usual, we shall apply Carathéodory's theorem for semi-algebras. If we define *measurable rectangles* to be subsets of $X \times Y$ the form $R = A \times B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we first observe that the class of all such sets form a semi-algebra.

Lemma 6.6.1 *The collection $\mathcal{R}$ of measurable rectangles is a semi-algebra.*

Proof: Observe first that since

$$(R_1 \times S_1) \cap (R_2 \times S_2) = (R_1 \cap R_2) \times (S_1 \cap S_2)$$

the intersection of two measurable rectangles is a measurable rectangle. Moreover, since

$$(A \times B)^c = (A^c \times B^c) \cup (A^c \times B) \cup (A \times B^c) \tag{6.6.1}$$

the complement of measurable rectangle is a finite, disjoint union of measurable rectangles, and hence $\mathcal{R}$ is a semi-algebra. □

The next step is to define a function $\lambda : \mathcal{R} \to \overline{\mathbb{R}}_+$ by

$$\lambda(A \times B) = \mu(A)\nu(B)$$

To use Carathéodory's theorem for semi-algebras, we need to show that if

$$A \times B = \bigcup_{n \in \mathbb{N}} (C_n \times D_n)$$

is a disjoint union, then $\lambda(A \times B) = \sum_{n=1}^{\infty} \lambda(C_n \times D_n)$, or in other words

$$\mu(A)\nu(B) = \sum_{n=1}^{\infty} \mu(C_n)\nu(D_n) \tag{6.6.2}$$

(note that in this case, $\emptyset \in \mathcal{R}$ and there is no need to treat finite sums separately). Observe that since $\mu(C) = \int \mathbf{1}_C(x) d\mu(x)$ and $\nu(D) = \int \mathbf{1}_D(y) d\nu(y)$, we have

$$\mu(C)\nu(D) = \int \mathbf{1}_C(x) d\mu(x) \int \mathbf{1}_D(y) d\nu(x) =$$

$$= \int \left[\int \mathbf{1}_C(x) \mathbf{1}_D(y) d\mu(x) \right] d\nu(x) = \int \left[\int \mathbf{1}_{C \times D}(x, y) d\mu(x) \right] d\nu(x)$$

for any two sets $C \in \mathcal{A}$ and $D \in \mathcal{B}$. If $A \times B = \bigcup_{n \in \mathbb{N}} (C_n \times D_n)$ is a disjoint union, the Monotone Convergence Theorem 5.5.6 thus tells us that

$$\sum_{n=1}^{\infty} \mu(C_n)\nu(D_n) = \lim_{N \to \infty} \sum_{n=1}^{N} \mu(C_n)\nu(D_n) =$$

$$= \lim_{N \to \infty} \sum_{n=1}^{N} \int \left[\int \mathbf{1}_{C_n \times D_n}(x, y) d\mu(x) \right] d\nu(x)$$

$$= \lim_{N \to \infty} \int \left[\int \sum_{n=1}^{N} \mathbf{1}_{C_n \times D_n}(x, y) d\mu(x) \right] d\nu(x)$$

$$= \int \left[\lim_{N \to \infty} \int \sum_{n=1}^{N} \mathbf{1}_{C_n \times D_n}(x, y) d\mu(x) \right] d\nu(x)$$

$$= \int \left[\int \lim_{N \to \infty} \sum_{n=1}^{N} \mathbf{1}_{C_n \times D_n}(x, y) d\mu(x) \right] d\nu(x)$$

$$= \int \left[\int \sum_{n=1}^{\infty} \mathbf{1}_{C_n \times D_n}(x, y) d\mu(x) \right] d\nu(x)$$

$$= \int \left[\int \mathbf{1}_{C \times D}(x, y) d\mu(x) \right] d\nu(x) = \mu(C)\nu(D)$$

which proves equation (6.6.2).

We are now ready to prove the main theorem. Remember that a measure space is σ-finite if it is a countable union of sets of finite measure.

Theorem 6.6.2 *Assume that $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ are two measure spaces and let $\mathcal{A}\otimes\mathcal{B}$ be the σ-algebra generated by the measurable rectangles $A\times B$, $A\in\mathcal{A}$, $B\in\mathcal{B}$. Then there exists a measure $\mu\times\nu$ on $\mathcal{A}\otimes\mathcal{B}$ such that*

$$\mu\times\nu(A\times B)=\mu(A)\nu(B)\qquad \textit{for all } A\in\mathcal{A}, B\in\mathcal{B}$$

If μ and ν are σ-finite, this measure is unique and is called the product measure *of μ and ν.*

Proof: The existence follows from Theorem 6.3.6 as formula (6.6.2) guarantees that conditions (i) and (ii) hold (as $\emptyset=\emptyset\times\emptyset\in\mathcal{R}$, we need not treat finite and countable sums separately in this case). The uniqueness follows from Proposition 6.3.3 (see also Exercise 6.3.1c)). □

Product measures can be used to construct Lebesgue measure in higher dimension. If μ is Lebesgue measure on $\mathbb{R}$, the completion of the product $\mu\times\mu$ is the Lebesgue measure on $\mathbb{R}^2$. To get Lebesgue measure on $\mathbb{R}^3$, take a new product $(\mu\times\mu)\times\mu$ and complete it. Continuing in this way, we get Lebesgue measure in all dimensions.

Exercises for Section 6.6

1. Check that formula $(R_1\times S_1)\cap(R_2\times S_2)=(R_1\cap R_2)\times(S_1\cap S_2)$ in the proof of Lemma 6.6.1 is correct.
2. Check that formula (6.6.1) is correct and that the union is disjoint.
3. Assume that μ is the counting measure on $\mathbb{N}$. Show that $\mu\times\mu$ is counting measure on $\mathbb{N}^2$.
4. Show that any open set in $\mathbb{R}^d$ is a countable union of open boxes of the form

$$(a_1,b_1)\times(a_2,b_2)\times\ldots\times(a_d,b_d)$$

   where $a_1<b_1, a_2<b_2,\ldots,a_d<b_d$ (this can be used to show that the Lebesgue measure on $\mathbb{R}^d$ is a completed Borel measure).
5. In this problem we shall generalize Proposition 6.4.5 from $\mathbb{R}$ to $\mathbb{R}^2$. Let μ be the Lebesgue integral on $\mathbb{R}$ and let $\lambda=\mu\times\mu$.
   a) Show that if D,E are open sets in $\mathbb{R}$, then $D\times E$ is open in $\mathbb{R}^2$.
   b) Assume that $A\times B$ is a measurable rectangle with $\mu(A),\mu(B)<\infty$. Show that for any $\epsilon>0$ there are open sets $D,E\subset\mathbb{R}$ such that $A\times B\subset D\times E$ and $\lambda(E\times D)-\lambda(A\times B)<\epsilon$.
   c) Assume that $Z\subset\mathbb{R}^2$ is measurable with $\lambda(Z)<\infty$. Show that for any $\epsilon>0$, there is an open set $G\subset\mathbb{R}^2$ such that $Z\subset G$ and $\lambda(G)-\lambda(Z)<\epsilon$. Explain why this mean that $\lambda(G\setminus Z)<\epsilon$.
   d) Assume that $Z\subset\mathbb{R}^2$ is measurable. Show that for any $\epsilon>0$, there is an open set $G\subset\mathbb{R}^2$ such that $Z\subset G$ and $\lambda(G\setminus Z)<\epsilon$.
   e) Assume that $Z\subset\mathbb{R}^2$ is measurable. Show that for any $\epsilon>0$, there is a closed set $F\subset\mathbb{R}^2$ such that that $Z\supset F$ and $\lambda(Z\setminus F)<\epsilon$.

6.7 Fubini's Theorem

In this section we shall see how we can integrate with respect to a product measure; i.e. we shall prove the formulas

$$\int f\, d(\mu \times \nu) = \int \left[\int f(x,y)\, d\mu(x)\right] d\nu(y) = \int \left[\int f(x,y)\, d\nu(y)\right] d\mu(x) \quad (6.7.1)$$

mentioned in the previous section. As one might expect, these formulas do not hold for all measurable functions, and part of the challenge is to find the right conditions. We shall prove two theorems; one (Tonelli's Theorem) which only works for nonnegative functions, but doesn't need any additional conditions; and one (Fubini's Theorem) which works for functions taking both signs, but which has integrability conditions that might be difficult to check. Often the two theorems are used in combination – we use Tonelli's Theorem to show that the conditions for Fubini's Theorem are satisfied.

We have to begin with some technicalities. For formula (6.7.1) to make sense, the functions $x \mapsto f(x,y)$ and $y \mapsto f(x,y)$ we get by fixing one of the variables in the function $f(x,y)$, have to be measurable. To simplify notation, we write $f^y(x)$ for the function $x \mapsto f(x,y)$ and $f_x(y)$ for the function $y \mapsto f(x,y)$. Similarly for subsets of $X \times Y$:

$$E^y = \{x \in X \mid (x,y) \in E\} \quad \text{and} \quad E_x = \{y \in Y \mid (x,y) \in E\}$$

These sets and functions are called *sections* of f and E, respectively (make a drawing).

Lemma 6.7.1 *Assume that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure spaces, and let $\mathcal{A} \otimes \mathcal{B}$ be the product σ-algebra.*

(i) For any $E \in \mathcal{A} \otimes \mathcal{B}$, we have $E^y \in \mathcal{A}$ and $E_x \in \mathcal{B}$ for all $y \in Y$ and $x \in X$.

(ii) For any $\mathcal{A} \otimes \mathcal{B}$-measurable $f : X \times Y \to \overline{\mathbb{R}}$, the sections f^y and f_x are $\mathcal{A}$- and $\mathcal{B}$-measurable, respectively, for all $y \in Y$ and $x \in X$.

Proof: I only prove the lemma for the y-sections, and leave the x-sections to the readers.

(i) Since $\mathcal{A} \otimes \mathcal{B}$ is the smallest σ-algebra containing the measurable rectangles, it clearly suffices to show that

$$\mathcal{C} = \{E \subset X \times Y \mid E^y \in \mathcal{A}\}$$

is a σ-algebra containing the measurable rectangles. That the measurable rectangles are in $\mathcal{C}$, follows from the observation

$$(A \times B)^y = \begin{cases} A & \text{if } y \in B \\ \\ \emptyset & \text{if } y \notin B \end{cases}$$

To show that $\mathcal{C}$ is closed under complements, just note that if $E \in \mathcal{C}$, then $E^y \in \mathcal{A}$, and hence $(E^c)^y = (E^y)^c \in \mathcal{A}$ (check this!) which means that $E^c \in \mathcal{C}$. Similarly for countable unions: If for each $n \in \mathbb{N}$, $E_n \in \mathcal{C}$, then $(E_n)^y \in \mathcal{A}$ for all n, and hence $\bigcup_{n\in\mathbb{N}}(E_n)^y = (\bigcup_{n\in\mathbb{N}} E_n)^y \in \mathcal{A}$ (check this!) which means that $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{C}$.

(ii) We need to check that $(f^y)^{-1}(I) \in \mathcal{A}$ for all intervals of the form $[-\infty, r)$. But this follows from (i) and the measurability of f since

$$(f^y)^{-1}(I) = (f^{-1}(I))^y$$

(check this!) $\square$

There is another measurability problem in formula (6.7.1): We need to know that the integrated sections

$$y \mapsto \int f(x,y)\, d\mu(x) = \int f^y(x)\, d\mu(x)$$

and

$$x \mapsto \int f(x,y)\, d\nu(y) = \int f_x(y)\, d\nu(y)$$

are measurable. This is a more complicated question, and we shall need a quite useful and subtle result known as the Monotone Class Theorem. A *monotone class* of subsets of a set Z is a collection $\mathcal{M}$ of subsets of Z that is closed under increasing countable unions and decreasing countable interesections. More precisely:

(i) If $E_1 \subset E_2 \subset \ldots \subset E_n \subset \ldots$ are in $\mathcal{M}$, then $\bigcup_{n\in\mathbb{N}} \in \mathcal{M}$

(ii) If $E_1 \supset E_2 \supset \ldots \supset E_n \supset \ldots$ are in $\mathcal{M}$, then $\bigcap_{n\in\mathbb{N}} \in \mathcal{M}$

All σ-algebras are monotone classes, but a monotone class need not be a σ-algebra. If $\mathcal{R}$ is a collection of subsets of Z, there is (by the usual argument) a smallest monotone class containing $\mathcal{R}$. It is called the *monotone class generated by* $\mathcal{R}$.

Theorem 6.7.2 (Monotone Class Theorem) *Assume that Z is a nonempty set and that $\mathcal{A}$ is an algebra of subsets of Z. Then σ-algebra and the monotone class generated by $\mathcal{A}$ coincide.*

Proof: Let $\mathcal{C}$ be the σ-algebra and $\mathcal{M}$ the monotone class generated by $\mathcal{A}$. Since all σ-algebras are monotone classes, we must have $\mathcal{M} \subset \mathcal{C}$.

To prove the opposite inclusion, we show that $\mathcal{M}$ is a σ-algebra. Observe that it suffices to prove that $\mathcal{M}$ is an algebra as closure under countable unions will then take care of itself: If $\{E_n\}$ is a sequence from $\mathcal{M}$, the sets $F_n = E_1 \cup E_2 \cup \ldots \cup E_n$ are in $\mathcal{M}$ since $\mathcal{M}$ is an algebra, and hence

$\bigcup_{n\in\mathbb{N}} E_n = \bigcup_{n\in\mathbb{N}} F_n \in \mathcal{M}$ since $\mathcal{M}$ is closed under increasing countable unions.

To prove that $\mathcal{M}$ is an algebra, we use a trick. For each $M \in \mathcal{M}$ define

$$\mathcal{M}(M) = \{F \in \mathcal{M} \mid F \setminus M, F \setminus M, F \cap M \in \mathcal{M}\}$$

It is not hard to check that since $\mathcal{M}$ is a monotone class, so is $\mathcal{M}(M)$. Note also that by symmetry, $N \in \mathcal{M}(M) \iff M \in \mathcal{M}(N)$.

Our aim is to prove that $\mathcal{M}(M) = \mathcal{M}$ for all $M \in \mathcal{M}$. This would mean that the intersection and difference between any two sets in $\mathcal{M}$ are in $\mathcal{M}$, and since $Z \in \mathcal{M}$ (because $Z \in \mathcal{A} \subset \mathcal{M}$), we may conclude that $\mathcal{M}$ is an algebra.

To show that $\mathcal{M}(M) = \mathcal{M}$ for all $M \in \mathcal{M}$, pick a set $A \in \mathcal{A}$. Since $\mathcal{A}$ is an algebra, we have $\mathcal{A} \subset \mathcal{M}(A)$. Since $\mathcal{M}$ is the smallest monotone class containing $\mathcal{A}$, this means that $\mathcal{M}(A) = \mathcal{M}$, and hence $M \in \mathcal{M}(A)$ for all $M \in \mathcal{M}$. By symmetry, $A \in \mathcal{M}(M)$ for all $M \in \mathcal{M}$. Since A was an arbitrary element in $\mathcal{A}$, we have $\mathcal{A} \subset \mathcal{M}(M)$ for all $M \in \mathcal{M}$, and using the minimality of $\mathcal{M}$ again, we see that $\mathcal{M}(M) = \mathcal{M}$ for all $M \in \mathcal{M}$. □

The advantage of the Monotone Class Theorem is that it is often much easier to prove that families are closed under monotone unions and intersections than under arbitrary unions and intersections, especially when there's a measure involved in the definition. The next lemma is a typical case. Note the σ-finiteness condition; Exercise 9 shows that the result does not always hold without it.

Lemma 6.7.3 *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two σ-finite measure spaces, and assume that $E \subset X \times Y$ is $\mathcal{A} \otimes \mathcal{B}$-measurable. Then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are $\mathcal{A}$- and $\mathcal{B}$-measurable, respectively, and*

$$\int \nu(E_x)\, d\mu(x) = \int \mu(E^y)\, d\nu(y) = \mu \times \nu(E)$$

Proof: We shall prove the part about the x-sections E_x and leave the (similar) proof for y-sections to the readers. We shall first carry out the proof for finite measure spaces, i.e. we assume that $\mu(X), \nu(Y) < \infty$.

Let

$$\mathcal{C} = \{E \subset X \times Y \mid x \mapsto \nu(E_x) \text{ is } \mathcal{A}\text{-measurable and } \textstyle\int \nu(E_x)\, d\mu(x) = \mu \times \nu(E)\}$$

If we can show that $\mathcal{C}$ is a monotone class containing the algebra generated by the measurable rectangles, Monotone Class Theorem will tell us that $\mathcal{C}$ is a σ-algebra, and hence contains $\mathcal{A} \otimes \mathcal{B}$ (which is the *smallest* σ-algebra containing the measurable rectangles). This obviously suffices to prove the finite case of the theorem.

To show that any measurable rectangle $E = A \times B$ belongs to $\mathcal{C}$, just observe that

$$\nu(E_x) = \begin{cases} \nu(B) & \text{if } x \in A \\ \\ 0 & \text{if } x \notin A \end{cases}$$

and that $\int \nu(E_x)\, d\mu(x) = \int_A \nu(B)\, d\mu = \mu(A)\nu(B) = \mu \times \nu(E)$.

A set F in the algebra generated by the measurable rectangles, is a disjoint union $F = \bigcup_{i=1}^n E_i$ of measurable rectangles, and since $\nu(F_x) = \sum_{i=1}^n \nu((E_i)_x)$, the function $x \mapsto \nu(F_x)$ is $\mathcal{A}$-measurable (a sum of measurable functions is measurable) and $\int \nu(F_x)\, d\mu(x) = \int \sum_{i=1}^n \nu((E_i)_x) d\mu(x) = \sum_{i=1}^n \mu \times \nu(E_i) = \mu \times \nu(F)$. Hence $F \in \mathcal{C}$.

To show that $\mathcal{C}$ is monotone class, assume that $\{E_n\}$ is an increasing sequence of sets in $\mathcal{C}$. Let $E = \bigcup_{n\in\mathbb{N}} E_n$, and note that $E_x = \bigcup_{n=1}^\infty (E_n)_x$. By continuity of measure, $\nu(E_x) = \lim_{n\to\infty} \nu((E_n)_x)$, and hence $x \mapsto \nu(E_x)$ is measurable as the limit of a sequence of measurable functions. Moreover, by the Monotone Convergence Theorem and continuity of measures,

$$\int \nu(E_x)\, d\mu(x) = \int \lim_{n\to\infty} \nu((E_n)_x)\, d\mu(x)) = \lim_{n\to\infty} \int \nu((E_n)_x)\, d\mu(x)) =$$

$$= \lim_{n\to\infty} \mu \times \nu(E_n) = \mu \times \nu(E)$$

This means that $E \in \mathcal{C}$.

We must also check monotone intersections. Assume that $\{E_n\}$ is a decreasing sequence of sets in $\mathcal{C}$. Let $E = \bigcap_{n\in\mathbb{N}} E_n$, and note that $E_x = \bigcap_{n=1}^\infty (E_n)_x$. By continuity of measure (here we are using that ν is a finite measure), $\nu(E_x) = \lim_{n\to\infty} \nu((E_n)_x)$, and hence $x \mapsto \nu(E_x)$ is measurable. Moreover, using the Dominated Convergenge Theorem (since the measure space is finite, we can use the function that is constant 1 as the dominating function), we see that

$$\int \nu(E_x)\, d\mu(x) = \int \lim_{n\to\infty} \nu((E_n)_x)\, d\mu(x)) = \lim_{n\to\infty} \int \nu((E_n)_x)\, d\mu(x)) =$$

$$= \lim_{n\to\infty} \mu \times \nu(E_n) = \mu \times \nu(E)$$

and hence $E \in \mathcal{C}$.

Since this shows that $\mathcal{C}$ is a monotone class and hence a σ-algebra containing $\mathcal{A}\times\mathcal{B}$, we have proved the lemma for finite measure spaces. To extend it to σ-finite spaces, let $\{X_n\}$ and $\{Y_n\}$ be increasing sequence of subsets of X and Y of finite measure such that $X = \bigcup_{n\in\mathbb{N}} X_n$ and $Y = \bigcup_{n\in\mathbb{N}} Y_n$. If E is a $\mathcal{A} \otimes \mathcal{B}$-measurable subset of $X \times Y$, it follows from what we have already proved that $x \mapsto \nu((E \cap (X_n \times Y_n)_x)$ is measurable and

$$\int \nu((E \cap (X_n \times Y_n)_x)\, d\mu(x) = \mu \times \nu(E \cap (X_n \times Y_n))$$

The lemma for σ-finite spaces now follows from the Monotone Convergence Theorem and continuity of measures. $\Box$

Remark: The proof above illustrates a useful technique. To prove that all sets in the σ-algebra $\mathcal{F}$ generated by a family $\mathcal{R}$ satisfies a certain property P, we prove

(i) All sets in $\mathcal{R}$ has property P.

(ii) The sets with property P form a σ-algebra $\mathcal{G}$.

Then $\mathcal{F} \subset \mathcal{G}$ (since $\mathcal{F}$ is the *smallest* σ-algebra containing $\mathcal{R}$), and hence all sets in $\mathcal{F}$ has property P.

We are now ready to prove the first of our main theorems.

Theorem 6.7.4 (Tonelli's Theorem) *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two σ-finite measure spaces, and assume that $f : X \times Y \to \overline{\mathbb{R}}_+$ is a nonnegative, $\mathcal{A} \otimes \mathcal{B}$-measurable function. Then the functions $x \mapsto \int f(x,y)\, d\nu(y)$ and $y \mapsto \int f(x,y) d\mu(x)$ are $\mathcal{A}$- and $\mathcal{B}$-measurable, respectively, and*

$$\int f\, d(\mu \times \nu) = \int \left[\int f(x,y)\, d\nu(y) \right] d\mu(x) = \int \left[\int f(x,y)\, d\mu(x) \right] d\nu(y)$$

Proof: We prove the first equality and leave the second to the reader. Notice first that if $f = \mathbf{1}_E$ is an indicator function, then by the lemma

$$\int \mathbf{1}_E\, d(\mu \times \nu) = \mu \times \nu(E) = \int \nu(E_x)\, d\mu(x) = \int \left[\int \mathbf{1}_E(x,y)\, d\nu(y) \right] d\mu(x)$$

which proves the theorem for indicator functions. By linearity, it also holds for nonnegative, simple functions.

For the general case, let $\{f_n\}$ be an increasing sequence of nonnegative simple functions converging pointwise to f. The functions $x \mapsto \int f_n(x,y)\, d\nu(y)$ increase to $x \mapsto \int f(x,y)\, d\nu(y)$ by the Monotone Convergence Theorem. Hence the latter function is measurable (as the limit of a sequence of measurable functions), and using the Monotone Convergence Theorem again, we get

$$\int f(x,y)\, d(\mu \times \nu) = \lim_{n \to \infty} \int f_n(x,y)\, d(\mu \times \nu) =$$

$$= \lim_{n \to \infty} \int \left[\int f_n(x,y)\, d\nu(y) \right] d\mu(x) = \int \left[\lim_{n \to \infty} \int f_n(x,y)\, d\nu(y) \right] d\mu(x) =$$

$$= \int \left[\int \lim_{n \to \infty} f_n(x,y)\, d\nu(y) \right] d\mu(x) = \int \left[\int f(x,y)\, d\nu(y) \right] d\mu(x)$$

$\Box$

Fubini's Theorem is now an easy application of Tonelli's Theorem.

Theorem 6.7.5 (Fubini's Theorem) *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two σ-finite measure spaces, and assume that $f : X \times Y \to \overline{\mathbb{R}}$ is $\mu \times \nu$-integrable. Then the functions f_x and f^y are integrable for almost all x and y, and the integrated functions $x \mapsto \int f(x, y)\, d\nu(y)$ and $y \mapsto \int f(x, y) d\mu(x)$ are μ- and ν-integrable, respectively. Moreover,*

$$\int f\, d(\mu \times \nu) = \int \left[\int f(x, y)\, d\nu(y)\right] d\mu(x) = \int \left[\int f(x, y)\, d\mu(x)\right] d\nu(y)$$

Proof: Since f is integrable, it splits as the difference $f = f_+ - f_-$ between two nonnegative, integrable functions. Applying Tonelli's Theorem to $|f|$, we get

$$\int \left[\int |f(x, y)|\, d\nu(y)\right] d\mu(x) = \int \left[\int |f(x, y)|\, d\mu(x)\right] d\nu(y) =$$

$$= \int |f|\, d(\mu \times \nu) < \infty$$

which implies the integrability statements for f_x, f^y, $x \mapsto \int f(x, y)\, d\nu(y)$ and $y \mapsto \int f(x, y) d\mu(x)$. Applying Tonelli's Theorem to f_+ and f_- separately, we get

$$\int f_+\, d(\mu \times \nu) = \int \left[\int f_+(x, y)\, d\nu(y)\right] d\mu(x) = \int \left[\int f_+(x, y)\, d\mu(x)\right] d\nu(y)$$

and

$$\int f_-\, d(\mu \times \nu) = \int \left[\int f_-(x, y)\, d\nu(y)\right] d\mu(x) = \int \left[\int f_-(x, y)\, d\mu(x)\right] d\nu(y)$$

and subtracting the second from the first, we get Fubini's Theorem. □

Remark: The integrability condition in Fubini's Theorem is occasionally a nuisance: The natural way to show that f is integrable, is by calculating the iterated integrals, but this presupposes that f is integrable! The solution is often first to apply Tonelli's Theorem to the absolute value $|f|$, and use the iterated integrals there to show that $\int |f|\, d(\mu \times \nu$ is finite. This means that f is integrable, and we are ready to apply Fubini's Theorem.

Even when the original measures μ and ν are complete, the product measure $\mu \times \nu$ rarely is. A natural response is to take it's completion $\overline{\mu \times \nu}$ (as we did with higher dimensional Lebesgue measures), but the question is if Fubini's theorem still holds. This is not obvious since there are more $\overline{\mu \times \nu}$-measurable functions than $\mu \times \nu$-measurable ones, but fortunately the answer is yes. I just state the result without proof (see Excercise 11).

Theorem 6.7.6 *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two complete, σ-finite measure spaces, and assume that $f : X \times Y \to \overline{\mathbb{R}}$ is $\overline{\mu \times \nu}$-measurable, where $\overline{\mu \times \nu}$ is the completion of the product measure. Then the functions f_x and f^y are measurable for almost all x and y, and the integrated functions $x \mapsto \int f(x,y)\, d\nu(y)$ and $y \mapsto \int f(x,y) d\mu(x)$ are measurable as well. Moreover*

(i) ***(Tonelli's Theorem for Completed Measures)*** *If f is nonnegative,*

$$\int f\, d\overline{\mu \times \nu} = \int \left[\int f(x,y)\, d\nu(y)\right] d\mu(x) = \int \left[\int f(x,y)\, d\mu(x)\right] d\nu(y)$$

(ii) ***(Fubini's Theorem for Completed Measures)*** *If f is integrable, the functions f_x and f^y are integrable for almost all x and y, and the integrated functions $x \mapsto \int f(x,y)\, d\nu(y)$ and $y \mapsto \int f(x,y) d\mu(x)$ are μ- and ν-integrable, respectively. Moreover,*

$$\int f\, d\overline{\mu \times \nu} = \int \left[\int f(x,y)\, d\nu(y)\right] d\mu(x) = \int \left[\int f(x,y)\, d\mu(x)\right] d\nu(y)$$

Exercises for Section 6.7

1. Show that $(E^c)^y = (E^y)^c$ (here c i referring to complements).
2. Show that $(\bigcup_{n\in\mathbb{N}} E_n)^y = \bigcup_{n\in\mathbb{N}} E_n^y$.
3. Show that $(f^y)^{-1}(I) = (f^{-1}(I))^y$.
4. Show that
$$\mathcal{M} = \{M \subset \mathbb{R} \,|\, 0 \in M\}$$
is a monotone class, but not a σ-algebra.
5. Show that the sets $\mathcal{M}(M)$ in the proof of the Monotone Class Theorem really are monotone classes.
6. In this problem μ Lebesgue measure on $\mathbb{R}$, while ν is counting measure on $\mathbb{N}$. Let $\lambda = \mu \times \nu$ be the product measure and let
$$f : \mathbb{R} \times \mathbb{N} \to \mathbb{R}$$
be given by
$$f(x,n) = \frac{1}{1 + (2^n x)^2}$$
Compute $\int f\, d\lambda$. Remember that $\int \frac{1}{1+u^2}\, du = \arctan u + C$.
7. Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be defined by $f(x,y) = \frac{x-y}{(x+y)^3}$ (the expression doesn't make sense for $x = y = 0$, and you may give the function whatever value you want at that point). Show by computing the integrals that
$$\int_0^1 \left[\int_0^1 f(x,y)\, dx\right] dy = -\frac{1}{2}$$

and

$$\int_0^1 \left[\int_0^1 f(x, y)\, dy \right] dx = \frac{1}{2}$$

(you may want to use that $f(x, y) = \frac{1}{x^2+y^2} - \frac{2y}{(x+y)^3}$ in the first integral and argue by symmetry in the second one). Let μ be the Lebesgue measure on $[0, 1]$ and $\lambda = \mu \times \mu$. Is f integrable with respect to λ?

8. Define $f : [0, \infty) \times [0, \infty)$ by $f(x, y) = xe^{-x^2(1+y^2)}$.

   a) Show by performing the integrations that

   $$\int_0^\infty \left[\int_0^\infty f(x, y)\, dx \right] dy = \frac{\pi}{4}$$

   b) Use Tonelli's theorem to show that $\int_0^\infty \left[\int_0^\infty f(x, y)\, dy \right] dx = \frac{\pi}{4}$

   c) Make the substitution $u = xy$ in the inner integral of

   $$\int_0^\infty \left[\int_0^\infty f(x, y)\, dy \right] dx = \int_0^\infty \left[\int_0^\infty xe^{-x^2(1+y^2)})\, dy \right] dx$$

   and show that $\int_0^\infty \left[\int_0^\infty f(x, y)\, dy \right] dx = \left(\int_0^\infty e^{-u^2} du \right)^2$.

   d) Conclude that $\int_0^\infty e^{-u^2} du = \frac{\sqrt{\pi}}{2}$.

9. Let $X = Y = [0, 1]$, let μ be the Lebesgue measure on X (this is just the restriction of the Lebesgue measure on $\mathbb{R}$ to $[0, 1]$) and let ν be the counting measure. Let $E = \{(x, y) \in X \times Y \,|x = y\}$, and show that $\int \nu(E_x)\, d\mu(x)$, $\int \mu(E^y)\, d\nu(y)$, and $\mu \times \nu(E)$ are all different (compare lemma 6.7.3).

10. Assume that $X = Y = \mathbb{N}$ and that $\mu = \nu$ is the counting measure. Let $f : X \times Y \to \mathbb{R}$ be defined by

    $$f(x, y) = \begin{cases} 1 & \text{if } x = y \\ -1 & \text{if } x = y + 1 \\ 0 & \text{otherwise} \end{cases}$$

    Show that $\int f\, d\mu{\times}\nu = \infty$, but that the iterated integrals $\int \left[\int f(x, y)\, d\mu(x) \right] d\nu(u)$ and $\int \left[\int f(x, y)\, d\nu(y) \right] d\mu(x)$ are both finite, but unequal.

11. In this exercise, we shall sketch the proof of Theorem 6.7.6, and we assume that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are as in that theorem.

    a) Assume that $E \in \mathcal{A} \otimes \mathcal{B}$ and that $\mu \times \nu(E) = 0$. Show that $\mu(E^y) = 0$ for ν-almost all y and that $\nu(E_x) = 0$ for μ-almost all x.

    b) Assume that N is an $\mu \times \nu$-null set, i.e. there is an $E \in \mathcal{A} \otimes \mathcal{B}$ such that $N \subset E$ and $\mu \times \nu(E) = 0$. Show that for ν-almost all y, N^y is μ-measurable and $\mu(N^y) = 0$. Show also that for μ-almost all x, N_x is ν-measurable and $\nu(N_x) = 0$. (Here you need to use that the original measure spaces are complete).

c) Assume that $D \subset X \times Y$ is in the completion of $\mathcal{A} \otimes \mathcal{B}$ with respect to $\mu \times \nu$. Show that for ν-almost all y, D^y is μ-measurable, and that for μ-almost all x, D_x is ν-measurable (use that by Theorem 5.2.5 D can be written as a disjoint union $D = E \cup N$, where $E \in \mathcal{A} \otimes \mathcal{B}$ and N is a null set).

d) Let D be as above. Show that the functions $x \mapsto \nu(D_x)$ and $y \mapsto \mu(D^y)$ are $\mathcal{A}$- and $\mathcal{B}$-measurable, respectively (define $\nu(D_x)$ and $\mu(D^y)$ arbitrarily on the sets of measure zero where D_x and D^y fail to be measurable), and show that

$$\int \nu(D_x)\, d\mu(x) = \int \mu(D^y)\, d\nu(y) = \overline{\mu \times \nu}(D)$$

e) Prove Theorem 6.7.6 (this is just checking that the arguments that got us from Lemma 6.7.3 to Theorems 6.7.4 and 6.7.5, still works in the new setting).