Chapter 8

Product Spaces

In Chapter 1 we defined Lebesgue measure m_1 on $\mathbb{R}$. The starting point was that we wanted the Lebesgue measure of an interval to be its length. Similarly, when constructing Lebesgue measure on $\mathbb{R}^n$, m_n, one will want the Lebesgue measure of an n-box

$$S = \{(x_1, ..., x_n) : a_i \leq x_i \leq b_i\}, \text{ where } a_i, b_i, \in \mathbb{R}, a_i < b_i \text{ for } i = 1, ..., n,$$

to be the product of its sidelengths, $m_n(S) = (b_1 - a_1) \cdot ... \cdot (b_n - a_n)$. In this chapter we will use m_1 and an inductive procedure to define m_n.

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be measure spaces with positive measures μ and λ. We want to define a measure $\mu \times \lambda$ on a σ-algebra over the cartesian product $X \times Y$ that has the property that

$$(\mu \times \lambda)(A \times B) = \mu(A)\lambda(B), A \in \mathcal{A}, B \in \mathcal{B} \tag{8.1}$$

i.e. the measure of rectangles is the product of the measure of the sides. One could proceed as in Chapter 1 and define an outer measure on $\mathcal{P}(X \times Y)$ by

$$(\mu \times \lambda)^*(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i)\lambda(B_i) : \ A_i \in \mathcal{A}, B_i \in \mathcal{B}, E \subseteq \bigcup_{i=1}^{\infty} A_i \times B_i \right\}$$

and then hope to obtain a measure $\mu \times \lambda$ on a σ-algebra that contains the rectangles and satisfies (8.1). After this one still needs a way to calculate the measure of more general sets $E \subseteq X \times Y$. For this one will want to use iterated integrals and one has to show:

$$\begin{aligned}\mu \times \lambda(E) &= \int_{X\times Y} \chi_E d(\mu\times\lambda)\\ &= \int_Y \left(\int_X \chi_E(x,y)d\mu(x)\right) d\lambda(y)\\ &= \int_X \left(\int_Y \chi_E(x,y)d\lambda(y)\right) d\mu(x)\end{aligned}$$

This process will indeed work for σ-finite measures μ and λ, but it turns out it is easier to start by showing the equality of the last two iterated integrals, and then use the above string of identities as the definition for $\mu \times \lambda$. This is what we will do in this chapter. The principal tools will be the Monotone Class Theorem and the Monotone Convergence Theorem.

8.1 Product σ-algebras and the Monotone Class Theorem

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A *measurable rectangle* is a set of the form

$$A \times B = \{(x, y) : x \in A \text{ and } y \in B\},$$

where $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Of course, if $X = Y = \mathbb{R}$, then we can take A and B to be intervals and note that $A \times B$ is an ordinary rectangle in $\mathbb{R}^2$. However, the example where A equals the rationals and B the Cantor set shows that one should be careful with an oversimplified mental image of what a measurable rectangle is.

The collection $\mathcal{E} \subseteq X \times Y$ of *elementary sets* is comprised of sets which are finite unions of mutually disjoint measurable rectangles, i.e. $Q \in \mathcal{E}$ if and only if there are $A_1, ..., A_n \in \mathcal{A}$, $B_1, ..., B_n \in \mathcal{B}$ with

$$Q = \bigcup_{i=1}^{n} A_i \times B_i$$

and $(A_i \times B_i) \cap (A_j \times B_j) = \emptyset$ for all $i \neq j$. $\mathcal{A} \times \mathcal{B}$ is the smallest σ-algebra containing all the measurable rectangles.

If $\mathcal{A} = \mathcal{B} = \mathcal{B}(\mathbb{R})$ are the Borel sets of $\mathbb{R}$, then it is an exercise to verify that $\mathcal{A} \times \mathcal{B} = \mathcal{B}(\mathbb{R}^2)$, the Borel sets in $\mathbb{R}^2$. Similarly, $\mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^k) = \mathcal{B}(\mathbb{R}^{n+k})$.

Definition 8.1. A monotone class $\mathcal{M} \subseteq \mathcal{P}(X \times Y)$ is a collection of sets with the following two properties:

(a) If $E_1 \subseteq E_2 \subseteq ...$, $E_i \in \mathcal{M}$ for all $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} E_i \in \mathcal{M}$.

(b) If $F_1 \supseteq F_2 \supseteq ...$, $F_i \in \mathcal{M}$ for all $i \in \mathbb{N}$, then $\bigcap_{i=1}^{\infty} F_i \in \mathcal{M}$.

Note that any σ-algebra is a monotone class, and that any intersection of monotone classes in $\mathcal{P}(X \times Y)$ is a monotone class. Thus, for any collection of subsets of $X \times Y$ there is a smallest monotone class containing it.

The following theorem is a special case of what is called the Monotone Class Theorem.

Theorem 8.2. *$\mathcal{A} \times \mathcal{B}$ is the smallest monotone class which contains the elementary sets $\mathcal{E}$.*

Proof. Let $\mathcal{M}$ be the smallest monotone class which contains $\mathcal{E}$. Since $\mathcal{E} \subseteq \mathcal{A} \times \mathcal{B}$ and since $\mathcal{A} \times \mathcal{B}$ is a monotone class it is clear that $\mathcal{M} \subseteq \mathcal{A} \times \mathcal{B}$, and we will conclude the proof by showing that $\mathcal{M}$ is a σ-algebra. Since $X \times Y \in \mathcal{E} \subseteq \mathcal{M}$ and since $\mathcal{M}$ is a monotone class it suffices to show that $\mathcal{M}$ is closed under taking complements and finite unions, because a countable infinite union can be written as the union of an increasing sequence of sets each of which is a finite union of the given sets. By use of de Morgan's laws we see that this is equivalent to the following two statements:

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

(b) If $E, F \in \mathcal{M}$, then $E \cap F \in \mathcal{M}$.

We start by verifying that (a) and (b) hold if $\mathcal{M}$ is replaced by the elementary sets $\mathcal{E}$. In fact, we start with (b). If $P, Q \in \mathcal{E}$, then $P = \bigcup_{i=1}^{n}(A_i \times B_i)$, $Q = \bigcup_{j=1}^{m}(C_j \times D_j)$, where each of the families of measurable rectangles $\{A_i \times B_i\}_i$ and $\{C_j \times D_j\}_j$ consists of mutually disjoint sets. Then

$$\begin{aligned} P \cap Q &= \bigcup_{i,j} ((A_i \times B_i) \cap (C_j \times D_j)) \\ &= \bigcup_{i,j} ((A_i \cap C_j) \times (B_i \cap D_j)) \in \mathcal{E}, \end{aligned}$$

so (b) holds for $\mathcal{E}$. We further note that this inductively implies that whenever $P_1, P_2, ..., P_n \in \mathcal{E}$, then $\bigcap_{i=1}^{n} P_i \in \mathcal{E}$. If $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then

$$(A \times B)^c = (X \times B^c) \cup (A^c \times B) \in \mathcal{E}.$$

Hence, if $P = \bigcup_{i=1}^{n}(A_i \times B_i) \in \mathcal{E}$, then $P^c = \bigcap_{i=1}^{n}(A_i \times B_i)^c \in \mathcal{E}$.

Now let

$$\mathcal{N}_1 = \{P \in \mathcal{M} : P^c \in \mathcal{M}\}.$$

Then by the above $\mathcal{E} \subseteq \mathcal{N}_1 \subseteq \mathcal{M}$. We will show that $\mathcal{N}_1 = \mathcal{M}$ by showing that $\mathcal{N}_1$ is a monotone class.

For $i = 1, 2, ...$ let $P_i \in \mathcal{N}_1$ with $P_1 \subseteq P_2 \subseteq ...$, then $P = \bigcup_{i=1}^{\infty} P_i \in \mathcal{M}$ since $\mathcal{M}$ is a monotone class. Furthermore, the definition of $\mathcal{N}_1$ implies that $P_i^c \in \mathcal{M}$ for each i. Of course, we have $P_1^c \supseteq P_2^c \supseteq ...$, hence $P^c = \bigcap_{i=1}^{\infty} P_i^c \in \mathcal{M}$, because $\mathcal{M}$ is a monotone class. This means that $P \in \mathcal{N}_1$. Similarly one proves that if $P_i \in \mathcal{N}_1$ with $P_1 \supseteq P_2 \supseteq ...$, then $\bigcap_{i=1}^{\infty} P_i \in \mathcal{N}_1$. Thus $\mathcal{N}_1$ is a monotone class with $\mathcal{E} \subseteq \mathcal{N}_1 \subseteq \mathcal{M}$, hence $\mathcal{N}_1 = \mathcal{M}$. This implies that (a) holds.

We will verify (b) in two steps. Let

$$\mathcal{N}_2 = \{P \in \mathcal{M} : P \cap Q \in \mathcal{M} \text{ for all } Q \in \mathcal{E}\}.$$

Then we see that $\mathcal{E} \subseteq \mathcal{N}_2 \subseteq \mathcal{M}$ and as above one easily shows that $\mathcal{N}_2$ is a monotone class, thus $\mathcal{N}_2 = \mathcal{M}$. Finally we set

$$\mathcal{N}_3 = \{P \in \mathcal{M} : P \cap Q \in \mathcal{M} \text{ for all } q \in \mathcal{M}\}.$$

Then since $\mathcal{N}_2 = \mathcal{M}$ we have $\mathcal{E} \subseteq \mathcal{N}_3 \subseteq \mathcal{M}$, and again one easily verifies that $\mathcal{N}_3$ is a monotone class, thus $\mathcal{N}_3 = \mathcal{M}$ and (b) holds. □

8.2 Product Measures and Fubini's Theorem

If $E \in \mathcal{A} \times \mathcal{B}$ and $x \in X$, then we define the *x-section* of E by

$$E_x = \{y \in Y : (x, y) \in E\}.$$

Similarly for $y \in Y$ we set

$$E^y = \{x \in X : (x, y) \in E\}.$$

Lemma 8.3. *If $E \in \mathcal{A} \times \mathcal{B}$, then for all $x \in X$ and all $y \in Y$ we have*

$$E_x \in \mathcal{B} \text{ and } E^y \in \mathcal{A}.$$

Proof. Let Ω be the collection of all $E \in \mathcal{A} \times \mathcal{B}$ such that $E_x \in \mathcal{B}$ for each $x \in X$. We will show that Ω is a σ-algebra which contains all the measurable rectangles, hence $\Omega = \mathcal{A} \times \mathcal{B}$.

Clearly, if $A \in A$, $B \in \mathcal{B}$, then $(A \times B)_x = B \in \mathcal{B}$ if $x \in A$ and $(A \times B)_x = \emptyset \in \mathcal{B}$ if $x \notin A$. Hence $A \times B \in \Omega$. Thus in particular $X \times Y \in \Omega$. Next we show that if $E \in \Omega$, then $E^c \in \Omega$. If $E \in \Omega$ and $x \in X$, then $E_x \in \mathcal{B}$ and

$$\begin{aligned}(E^c)_x &= \{y \in Y : (x,y) \in E^c\} \\ &= \{y \in Y : (x,y) \notin E\} \\ &= \{y \in Y : y \notin E_x\} = (E_x)^c \in \mathcal{B}.\end{aligned}$$

Finally, if $E_1, E_2, ... \in \Omega$, then $E_{1,x}, E_{2,x}, ... \in \mathcal{B}$ and hence $\bigcup_{i=1}^{\infty} E_{i,x} \in \mathcal{B}$ for all $x \in X$. One checks that $\left(\bigcup_{i=1}^{\infty} E_i\right)_x = \bigcup_{i=1}^{\infty} E_{i,x}$. Thus Ω is a σ-algebra containing all measurable rectangles, hence $\Omega = \mathcal{A} \times \mathcal{B}$.

Similarly, if $E \in \mathcal{A} \times \mathcal{B}$, $y \in Y$, then $E^y \in \mathcal{A}$. $\square$

We will consider measurable functions $f : X \times Y \to Z$, where $Z = [0, \infty]$ or $Z = \mathbb{C}$. Because we are now dealing with three σ-algebras, we will indicate which of these σ-algebras the word measurable refers to.

One can construct $\mathcal{A}\times\mathcal{B}$-measurable functions from $\mathcal{A}$- and/or $\mathcal{B}$-measurable functions as follows: If $u : X \to Z$ is $\mathcal{A}$-measurable, then $f(x, y) = u(x)$ for all $(x, y) \in X \times Y$ defines an $\mathcal{A} \times \mathcal{B}$-measurable function. Indeed, if $W \subseteq Z$ is open, then $f^{-1}(W) = u^{-1}(W) \times Y$ is a measurable rectangle in $\mathcal{A} \times \mathcal{B}$. It follows that $g(x, y) = \sum_{i=1}^{n} u_i(x)v_i(y)$ is $\mathcal{A} \times \mathcal{B}$-measurable, whenever $u_1, ..., u_n$ are $\mathcal{A}$-measurable and $v_1, ..., v_n$ are $\mathcal{B}$-measurable.

For the following theorems we will be interested in the opposite process, i.e. we will start with an $\mathcal{A} \times \mathcal{B}$-measurable function and we will see that the x- and y-sections are measurable.

If $f : X \times Y \to Z$, then for all $x \in X$ we define a function $f_x : Y \to Z$ by $f_x(y) = f(x, y)$, $y \in Y$. Similarly, for $y \in Y$ we set $f^y : X \to Z, f^y(x) = f(x, y)$.

Lemma 8.4. *Let $Z = [0, \infty]$ or $Z = \mathbb{C}$. If $f : X \times Y \to Z$ is $\mathcal{A} \times \mathcal{B}$-measurable, then for all $x \in X$ and all $y \in Y$ we have f_x is a $\mathcal{B}$-measurable function on Y and f^y is an $\mathcal{A}$-measurable function on X.*

Proof. Let $E \in \mathcal{A}\times\mathcal{B}$ and $f = \chi_E$. For $x \in X$ we have $f_x(y) = 1$ if and only if $(x, y) \in E$, hence $f_x = \chi_{E_x}$, and this is $\mathcal{B}$-measurable by Lemma 8.3. It follows that f_x is $\mathcal{B}$-measurable, whenever f is an $\mathcal{A} \times \mathcal{B}$-measurable simple function on $X \times Y$. If $f : X \times Y \to [0, \infty]$ is an $\mathcal{A} \times \mathcal{B}$-measurable function, then $f(x, y)$ is the pointwise limit of a nondecreasing sequence $s_n(x, y)$ of nonnegative measurable simple functions. Hence $f_x(y) = \lim_{n\to\infty} s_{n,x}(y)$

is a pointwise limit of $\mathcal{B}$-measurable functions. It follows that such f_x is $\mathcal{B}$-measurable.

Finally, if $f : X \times Y \to \mathbb{C}$ is $\mathcal{A} \times \mathcal{B}$-measurable, then $f = f_1^+ - f_1^- + i(f_2^+ - f_2^-)$, where $f_j^\pm$ are $\mathcal{A} \times \mathcal{B}$-measurable and nonnegative for $j = 1, 2$. Hence $f_{j,x}^\pm$ are $\mathcal{B}$-measurable, and so is $f_x = f_{1,x}^+ - f_{1,x}^- + i(f_{2,x}^+ - f_{2,x}^-)$. The proof that f^y is $\mathcal{A}$-measurable is exactly the same. □

Theorem 8.5. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be σ-finite measure spaces.*
(a) Let $E \in \mathcal{A} \times \mathcal{B}$. Define

$$\varphi : X \to [0, \infty], \;\; \psi : Y \to [0, \infty]$$

by $\varphi(x) = \lambda(E_x)$ and $\psi(y) = \mu(E^y)$, $x \in X, y \in Y$.

Then φ is $\mathcal{A}$-measurable, ψ is $\mathcal{B}$-measurable, and

$$\int_X \varphi d\mu = \int_Y \psi d\lambda.$$

(b) If $f : X \times Y \to [0, \infty]$ is $\mathcal{A} \times \mathcal{B}$ measurable, if

$$\varphi(x) = \int_Y f_x d\lambda \;\; \text{and} \;\; \psi(y) = \int_X f^y d\mu,$$

then φ is $\mathcal{A}$-measurable, ψ is $\mathcal{B}$-measurable, and

$$\int_X \varphi d\mu = \int_Y \psi d\lambda.$$

Of course, (a) is just the special case of (b) where $f = \chi_E$. (b) is a version of Tonelli's Theorem.

It follows from Lemma 8.4 that f_x is $\mathcal{B}$-measurable for all $x \in X$. Hence $\varphi(x) = \int_Y f_x d\lambda$ is well-defined, and

$$\varphi(x) = \int_Y f(x, y) d\lambda(y).$$

Thus

$$\int_X \varphi(x) d\mu(x) = \int_X \left(\int_Y f(x, y) d\lambda(y) \right) d\mu(x).$$

Hence we see that the conclusion of the theorem is that two iterated integrals coincide

$$\int_X \left(\int_Y f d\lambda \right) d\mu = \int_Y \left(\int_X f d\mu \right) d\lambda$$

for every nonnegative $\mathcal{A} \times \mathcal{B}$-measurable function f.

Proof. We will first prove (a) and (b) when μ and λ are finite measures, and then we will use the same trick that we used in the proof of the Lebesgue-Radon-Nikodym Theorem to extend the theorem to the σ-finite case.

Thus suppose that μ and λ are finite measures and let $\mathcal{M}$ be the collection of sets in $\mathcal{A} \times \mathcal{B}$ such that all conclusions of part (a) of the theorem hold. We will show that $\mathcal{E} \subseteq \mathcal{M}$ and that $\mathcal{M}$ is a monotone class, hence by the Monotone Class Theorem it will follow that $\mathcal{M} = \mathcal{A} \times \mathcal{B}$.

If $E = A \times B$ for some $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then one checks that $\varphi(x) = \chi_A(x)\lambda(B)$ and this is $\mathcal{A}$-measurable. Similarly, $\psi(y) = \chi_B(y)\mu(A)$ is $\mathcal{B}$-measurable and $\int_X \varphi d\mu = \mu(A)\lambda(B) = \int_Y \psi d\lambda$. Hence $E = A \times B \in \mathcal{M}$.

Next if $E = \bigcup_{i=1}^n E_i$ is a finite union of mutually disjoint measurable rectangles E_i, then $E_x = \bigcup_{i=1}^n E_{i,x}$ is a finite union of mutually disjoint sets $E_{i,x} \in \mathcal{B}$. Hence

$$\varphi(x) = \lambda(E_x) = \sum_{i=1}^n \lambda(E_{i,x}) = \sum_{i=1}^n \varphi_i(x),$$

where each $\varphi_i(x) = \lambda(E_{i,x})$ is $\mathcal{A}$-measurable since E_i is a measurable rectangle. Thus φ is $\mathcal{A}$-measurable and

$$\int_X \varphi d\mu = \sum_{i=1}^n \int_X \varphi_i d\mu.$$

Similarly, ψ is $\mathcal{B}$-measurable and

$$\int_Y \psi d\lambda = \sum_{i=1}^n \int_Y \psi_i d\lambda,$$

where $\psi_i(y) = \mu(E_i^y)$. Since we already showed that the theorem holds for the measurable rectangles E_i, we know that for each i we have $\int_X \varphi_i d\mu = \int_Y \psi_i d\lambda$, and hence we conclude that $\int_X \varphi d\mu = \int_Y \psi d\lambda$. Thus $\mathcal{E} \subseteq \mathcal{M}$.

Next we show that $\mathcal{M}$ is a monotone class. Let

$$P_1 \subseteq P_2 \subseteq ..., \quad P_i \in \mathcal{M} \text{ for all } i = 1, 2, ...$$

We must show that $P = \bigcup_{i=1}^\infty P_i \in \mathcal{M}$.

For $x \in X$ and $y \in Y$ we set $\varphi(x) = \lambda(P_x)$ and $\psi(y) = \mu(P^y)$. The hypothesis on the P_i implies that for each $x \in X$ we have $P_{1,x} \subseteq P_{2,x} \subseteq ...$ and hence $\{\varphi_i(x)\}_i = \{\lambda(P_{i,x})\}_i$ forms a nondecreasing sequence of functions.

Since $P_i \in \mathcal{M}$ we also have that each φ_i is $\mathcal{A}$-measurable. The continuity property of the measure λ implies that

$$\lim_{i\to\infty} \varphi_i(x) = \lim_{i\to\infty} \lambda(P_{i,x}) = \lambda(\bigcup_{i=1}^{\infty} P_{i,x}) = \lambda((\bigcup_{i=1}^{\infty} P_i)_x) = \varphi(x).$$

Thus φ is $\mathcal{A}$-measurable. Similarly

$$\psi(y) = \lim_{i\to\infty} \psi_i(y), \quad \psi_i(y) = \mu(P_i^y)$$

is $\mathcal{B}$-measurable. Since $P_i \in \mathcal{M}$ we have $\int_X \varphi_i d\mu = \int_Y \psi_i d\lambda$ for each i, hence the Monotone Convergence Theorem implies

$$\int_X \varphi d\mu = \lim_{i\to\infty} \int_X \varphi_i d\mu = \lim_{i\to\infty} \int_Y \psi_i d\lambda = \int_Y \psi d\lambda.$$

This shows that $P \in \mathcal{M}$.

If $P_1 \supseteq P_2 \supseteq \ldots,\ \ P_i \in \mathcal{M}$ for all $i = 1, 2, \ldots$, then we can use almost the same argument to show that $P = \bigcap_{i=1}^{\infty} P_i \in \mathcal{M}$. The difference is that we have nonincreasing sequences φ_i and ψ_i. Thus at the very end we use the Dominated Convergence Theorem instead of the Monotone Convergence Theorem. That is where we use that μ and λ are finite measures. This concludes the proof of (a) in the case when μ and λ are finite measures.

Next we will verify (b) for finite measures. Linearity implies that (b) holds for all $\mathcal{A} \times \mathcal{B}$-measurable nonnegative simple functions. The general case follows, because the nondecreasing limit of measurable functions is measurable and we can write any nonnegative measurable function as a pointwise limit of a nondecreasing sequence of measurable simple functions and then use the Monotone Convergence Theorem.

Finally we will prove the theorem for σ-finite measures μ and λ. By Lemma 7.14 there are functions $w \in L^1(\mu), v \in L^1(\lambda)$ such that $0 < w(x), v(y) \leq 1$ for all x, y. Then $d\mu_1 = w d\mu$ and $d\lambda_1 = v d\lambda$ define finite measures.

If $f : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ is $\mathcal{A} \times \mathcal{B}$ measurable, then

$$g(x, y) = \frac{f(x, y)}{w(x)v(y)}$$

is $\mathcal{A} \times \mathcal{B}$-measurable, and

$$\varphi(x) = \int_Y f(x, y) d\lambda(y) = w(x) \int_Y g(x, y) d\lambda_1(y)$$

is $\mathcal{A}$ measurable since λ_1 is a finite measure. It also follows that

$$\int_X \varphi(x)d\mu(x) = \int_X \left(\int_Y g(x,y)d\lambda_1(y)\right) d\mu_1(x).$$

By symmetry we obtain that ψ is $\mathcal{B}$-measurable and

$$\int_Y \psi(y)d\lambda(y) = \int_Y \left(\int_X g(x,y)d\mu_1(x)\right) d\lambda_1(y).$$

Hence the theorem follows from the case of finite measures. □

Definition 8.6. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be σ-finite measure spaces. For $E \in \mathcal{A} \times \mathcal{B}$ define

$$(\mu \times \lambda)(E) = \int_X \varphi d\mu = \int_Y \psi d\lambda,$$

where $\varphi(x) = \lambda(E_x)$ and $\psi(y) = \mu(E^y)$ as in the previous theorem.

One easily verifies that this defines a positive measure on $\mathcal{A} \times \mathcal{B}$. In fact, it is trivial that $\mu \times \lambda(\emptyset) = 0$ and the countable additivity follows from part (b) of the previous theorem and the Beppo Levi Theorem. Note also that $\mu \times \lambda$ is a σ-finite measure and that

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

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Theorem 8.7 (Fubini-Tonelli Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be σ-finite measure spaces, and let f be an $\mathcal{A} \times \mathcal{B}$-measurable function on $X \times Y$.*

(a) If $0 \le f \le \infty$, then

$$\int_X \left(\int_Y f d\lambda\right) d\mu = \int_{X \times Y} f d(\mu \times \lambda) = \int_Y \left(\int_X f d\mu\right) d\lambda. \tag{8.2}$$

(b) If f is complex-valued and if

$$\int_X \left(\int_Y |f| d\lambda\right) d\mu < \infty,$$

then $f \in L^1(\mu \times \lambda)$.

(c) If $f \in L^1(\mu \times \lambda)$, then $f_x \in L^1(\lambda)$ for $[\mu]$a.e. $x \in X$, $f^y \in L^1(\mu)$ for $[\lambda]$ a.e. $y \in Y$. Thus the functions $\varphi(x) = \int_Y f_x d\lambda$ and $\psi(y) = \int_X f^y d\mu$ are defined a.e.. They satisfy $\varphi \in L^1(\mu)$ and $\psi \in L^1(\lambda)$ and

$$\int_X \varphi d\mu = \int_{X\times Y} f d(\mu \times \lambda) = \int_Y \psi d\lambda,$$

i.e. (8.2) holds.

Thus the order of integration may be reversed for $\mathcal{A} \times \mathcal{B}$ measurable functions f whenever $f \geq 0$ (by (a)) and also whenever one of the iterated integrals for $|f|$ is finite (by (b) and (c)).

Proof. (a) In Theorem 8.5 (b) we already proved the equality of the terms on the left and on the right in equation (8.2). That the term in the middle also equals the other two follows from the definition of $\mu \times \lambda$ and the Monotone Convergence Theorem just as (b) followed from (a) in Theorem 8.5.

(b) follows immediately from (a).

(c) It is enough to verify this for real-valued functions $f \in L^1(\mu \times \lambda)$. If f is real-valued, then (a) applies to f^+ and f^-. Let φ_1 and φ_2 correspond to f^+ and f^- as φ corresponds to f. $f \in L^1(\mu \times \lambda)$ and $f^+ \leq |f|$ implies $\varphi_1 \in L^1(\mu)$ by (a). Thus $\varphi_1(x) < \infty$ for a.e. x $[\mu]$. Similarly, $\varphi_2 \in L^1(\mu)$ and $\varphi_2(x) < \infty$ a.e. The identities $f_x = f_x^+ - f_x^-$ and $|f_x| = f_x^+ + f_x^-$ imply that $\varphi(x) = \varphi_1(x) - \varphi_2(x)$ and $f_x \in L^1(\lambda)$, whenever both $\varphi_1(x) < \infty$ and $\varphi_2(x) < \infty$ which happens a.e. $[\mu]$. Equation (8.2) holds for f^+ and f^- hence it holds for f. The remainder of the statement follows analogously. □

Example 8.8. Let $f(x, y) = \frac{x^2-y^2}{(x^2+y^2)^2}$ if $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$, then f is Borel measurable on $\mathbb{R}^2$ (verify). We want to integrate f over $[0, 1] \times [0, 1]$ with respect to Lebesgue measure. Note that for $x \neq 0$ we have

$$\begin{aligned}
\varphi(x) &= \int_0^1 f(x, y)dy = \int_0^1 \frac{1}{x^2+y^2} - \frac{2y^2}{(x^2+y^2)^2} dy \\
&= \int_0^1 \frac{1}{x^2+y^2} dy + \int_0^1 y \frac{d}{dy}(x^2+y^2)^{-1} dy \\
&= \int_0^1 \frac{1}{x^2+y^2} dy + [y \frac{1}{x^2+y^2}]_{y=0}^{y=1} - \int_0^1 \frac{1}{x^2+y^2} dy \\
&= \frac{1}{1+x^2}.
\end{aligned}$$

Hence $\int_0^1 \varphi(x)dx = \frac{\pi}{4}$. On the other hand since $f(x,y) = -f(y,x)$ it is clear that reversing the order of integration should yield $-\frac{\pi}{4}$.

This does not contradict Fubini's Theorem since one can show that

$$\int_0^1 \int_0^1 |f(x,y)|dydx = \infty.$$

Example 8.9. Let $X = Y = [0,1]$, $\mu =$ Lebesgue measure on X, $\lambda =$ counting measure on Y, and $f(x,y) = 1$ if $x = y$ and $f(x,y) = 0$ otherwise. Then $\psi(y) = \int_X f(x,y)d\mu(x) = 0$ for all $y \in Y$ and $\varphi(x) = \int_Y f(x,y)d\lambda(y) = 1$ for all $x \in X$. Thus

$$\int_Y \left(\int_X f d\mu \right) d\lambda = 0 \neq 1 = \int_X \left(\int_Y f d\lambda \right) d\mu.$$

This does not contradict Fubini's theorem since λ is not a σ-finite measure.

If one of the measures is counting measure on a countable set, then the Fubini-Tonelli Theorem reduces to the Beppo Levi Theorem and we don't need to assume that the second measure is σ-finite.

We also remark that it is possible to construct examples of two measures and a function f such that the two iterated integrals exist and don't equal one another, because f is not jointly measurable (i.e. measurable with respect to the product σ-algebra).

8.3 Completion of product measure spaces and Lebesgue measure on $\mathbb{R}^n$

Even if $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ are complete measures, it may happen that $(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \lambda)$ is not complete. For example, let $(X, \mathcal{A}, \mu) = (Y, \mathcal{B}, \lambda) = ([0,1], \mathcal{L}_1, m_1)$ be Lebesgue measure on $[0,1]$. Then $E = [0,1] \times \{0\}$ is a measurable rectangle with $(m_1 \times m_1)(E) = m_1([0,1])m_1(\{0\}) = 0$. Now let $G \subseteq [0,1]$ be a non-Lebesgue-measurable set. Then $F = G \times \{0\} \subseteq E$ is a subset of a set of measure 0, yet $F \notin \mathcal{L}_1 \times \mathcal{L}_1$, because if it were, then by Lemma 8.3 $G = F^0 \in \mathcal{L}_1$.

One can extend Fubini's Theorem to the completion of product measures. We omit the details, but mention that one needs to be a little careful. Indeed, if $(\mathcal{A} \times \mathcal{B})^*$ denotes the completion of $\mathcal{A} \times \mathcal{B}$ with respect to some measure $\mu \times \lambda$, then one wants to prove a version of Fubini's Theorem for $(\mathcal{A} \times \mathcal{B})^*$-measurable functions f. The key observation is that there is an $\mathcal{A} \times \mathcal{B}$-measurable function g that agrees with f a.e. $[\mu \times \lambda]$.

By an inductive process one can easily extend the definition of product measures and Fubini's theorem to the product of n measures. The completion of $(\mathbb{R}^n, \mathcal{L}_1 \times ... \times \mathcal{L}_1, m_1 \times ... \times m_1)$ is called n-dimensional Lebesgue measure and is denoted by $(\mathbb{R}^n, \mathcal{L}_n, m_n)$. Thus, if $a, b \in \mathbb{R}^n$, say $a = (a_1, ..., a_n), b = (b_1, ..., b_n)$ with $a_i < b_i$ for all $i = 1, ..., n$, if

$$[a, b] = \{(x_1, ..., x_n) \in \mathbb{R}^n : a_i \leq x_i \leq b_i \text{ for all } i = 1, ..., n\},$$

then $m_n([a, b]) = \prod_{i=1}^n (b_i - a_i)$, i.e. the Lebesgue measure of an n-box is the product of the sidelengths.

Recall that $\mathcal{B}(\mathbb{R}^n)$ denotes the σ-algebra of Borel measurable sets. We already mentioned that $\mathcal{B}(\mathbb{R}^{n+k}) = \mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^k)$, thus $\mathcal{B}(\mathbb{R}^n) \subseteq \mathcal{L}_1 \times ... \times \mathcal{L}_1 \subseteq \mathcal{L}_n$. Hence all Borel measurable functions on $\mathbb{R}^n$ are Lebesgue measurable.

In Chapter 4 we mentioned that any positive measure on the Borel sets of $\mathbb{R}^n$ which is finite on the compact sets is automatically regular. Since every compact subset of $\mathbb{R}^n$ is a subset of an n box, this applies to m_n. In fact, as for $n = 1$ one can prove that for every $E \in \mathcal{L}_n$, there is an F_σ-set $F \subseteq E$ and a G_δ-set $G \supseteq E$ with $m_n(G \setminus F) = 0$. Thus m_n is a Radon measure on the Borel sets.

For further properties of m_n we refer to Rudin's textbook.

8.4 Convolutions

Theorem 8.10. *Let $f, g \in L^1(\mathbb{R})$. Then*

$$\int_{-\infty}^{\infty} |f(x-y)g(y)| dy < \infty \text{ for a.e. } x \in \mathbb{R}$$

and

$$h(x) = \int_{-\infty}^{\infty} f(x-y)g(y) dy \in L^1(\mathbb{R})$$

with $\|h\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}$

The function h is called the convolution of f and g, one writes $h = f * g$. Note that in general $f, g \in L^1(\mathbb{R})$ implies $fg \in L^2(\mathbb{R})$, but one cannot conclude that $fg \in L^1(\mathbb{R})$.

Proof. There exist Borel measurable functions f_0 and g_0 such that $f_0(x) = f(x)$ and $g_0(x) = g(x)$ a.e.. The integrals in the theorem are unaffected for every $x \in \mathbb{R}$, if we replace f, g with f_0, g_0. Thus we will assume that both f and g are Borel measurable. Set $\varphi(x,y) = x - y$ and $\psi(x,y) = y$, then φ and ψ are Borel measurable on $\mathbb{R}^2$, hence

$$K(x,y) = f(\varphi(x,y))g(\psi(y) = f(x-y)g(y)$$

is Borel measurable $\mathbb{R}^2 \to \mathbb{C}$. Hence K is $\mathcal{L}_1 \times \mathcal{L}_1$-measurable and thus $\int_{-\infty}^{\infty} |K(x,y)|dy$ is $\mathcal{L}_1$-measurable and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |K(x,y)|dydx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |K(x,y)|dxdy.$$

Thus by Tonelli's Theorem

$$\begin{aligned}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} |f(x-y)g(y)|dy\right)dx &= \int_{-\infty}^{\infty} |g(y)| \int_{-\infty}^{\infty} |f(x-y)|dxdy \\ &= \int_{-\infty}^{\infty} |g(y)| \|f\|_{L^1} dy \\ &= \|f\|_{L^1}\|g\|_{L^1} < \infty.\end{aligned}$$

Note that $\int_{-\infty}^{\infty} |f(x-y)|dx = \|f\|_{L^1}$ by the translation invariance of Lebesgue measure. Thus $K \in L^1(\mathbb{R}^2)$ and we conclude $\int_{-\infty}^{\infty} |f(x-y)g(y)|dy < \infty$ for a.e. $x \in \mathbb{R}$. For each $x \in \mathbb{R}$ where the previous expression is finite $h(x)$ is well defined and by Fubini's theorem h is measurable and satisfies $\|h\|_{L^1} \le \|f\|_{L^1}\|g\|_{L^1}$. □

It is easy to see that $f*g = g*f$. Furthermore, if $1 \le p \le \infty$ if $f \in L^1(\mathbb{R})$, $g \in L^p(\mathbb{R})$), then it turns out that $f * g \in L^p(\mathbb{R})$ and $\|f * g\|_p \le \|f\|_1\|g\|_p$. In fact, for $p = \infty$ a stronger conclusion is true.

Theorem 8.11. *If $f \in L^1(\mathbb{R})$, $g \in L^\infty(\mathbb{R})$, then $f * g$ is bounded and continuous.*

Proof. The measurability of all expressions follows as in the proof of Theorem 8.10. We have

$$\int_{-\infty}^{\infty} |f(x-y)g(y)|dy \le \|g\|_\infty \int_{-\infty}^{\infty} |f(x-y)|dy = \|g\|_\infty\|f\|_1.$$

This implies $h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy$ is defined for all $x \in \mathbb{R}$ and

$$|h(x+t) - h(x)| \le \int_{-\infty}^{\infty} |(f(x+t-y) - f(x-y))g(y)|dy \le \|g\|_\infty\|f_t - f\|_1$$

where we have written $f_t(x) = f(x+t)$. Thus the continuity of h follows from Exercise 5.11. □

8.5 Exercises

1. Show that $\mathcal{B}(\mathbb{R}^{n+k}) = \mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^k)$.

2. Let $(X, \mathcal{A}, \mu)$ and $(y, \mathcal{B}, \lambda)$ be σ-finite measure spaces. Suppose that ν is a positive measure on $\mathcal{A} \times \mathcal{B}$ such that $\nu(A \times B) = \mu(A)\lambda(B)$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Show that $\nu = \mu \times \lambda$ on $\mathcal{A} \times \mathcal{B}$.

3. Show that Lebesgue measure on $\mathbb{R}^n$ is translation invariant, i.e. if $E \in \mathcal{L}_n$ and if $x \in \mathbb{R}^n$, then $E + x \in \mathcal{L}_n$ and $m_n(E + x) = m_n(E)$.

4. Let $1 \le p < \infty$, let $f \in L^1(\mathbb{R})$, $g \in L^p(\mathbb{R}))$. Show that $f * g \in L^p(\mathbb{R})$ and $\|f * g\|_p \le \|f\|_1 \|g\|_p$.

5. Let f, g be non-negative real measurable functions on $[0, \infty)$. Let $1 < p \le 2$ and $q = \frac{p}{p-1}$. If $g \in L^p([0, \infty), m_1)$ and $\int_0^\infty f(t) t^{\frac{1}{q}-1} dt < \infty$, show that

$$\int_0^\infty \int_0^x \frac{g(x)}{x} f(t) dt dx \le \|g\|_p (q-1)^{-\frac{1}{q}} \int_0^\infty f(t) t^{\frac{1}{q}-1} dt.$$

6. Use Fubini's Theorem and the relation

$$\frac{1}{x} = \int_0^\infty e^{-xt} dt$$

to prove that

$$\lim_{M \to \infty} \int_0^M \frac{\sin x}{x} dx = \frac{\pi}{2}.$$

Similarly show that

$$\int_0^\infty \frac{e^{-x}}{x} \sin x dx = \frac{\pi}{4}.$$

7. A function $F : \mathbb{R} \to \mathbb{C}$ is said to be in $L^1_{loc}(\mathbb{R})$, if F is Lebesgue-measurable and $\int_K |F(t)| dt < \infty$ for every compact set $K \subseteq \mathbb{R}$.

Let μ be a finite Borel measure on $\mathbb{R}$, $g \in L^2(\mu)$, $f \in L^2(\mathbb{R})$. Show that

$$\int_{\mathbb{R}} |g(x) f(t-x)| d\mu(x) < \infty \text{ for a.e. } t \ [m_1],$$

and that

$$F(t) = \int_{\mathbb{R}} g(x) f(t-x) d\mu(x)$$

is well-defined a.e. with $F \in L^1_{loc}(\mathbb{R})$.

8. Let μ be a σ-finite measure on $(X, \mathcal{A})$ and let $f : X \to [0, \infty]$ be measurable. Show that

$$\int_X f d\mu = \int_0^\infty \mu(\{x \in X : f(x) > t\}) dt.$$