that if A and B belong to $\mathscr{B}(\mathbb{R})$, then $A \times B$ belongs to $\mathscr{B}(\mathbb{R}^2)$. Since $\mathscr{B}(\mathbb{R}) \times \mathscr{B}(\mathbb{R})$ is the σ-algebra generated by the collection of all such rectangles $A \times B$, it must be included in $\mathscr{B}(\mathbb{R}^2)$. Thus $\mathscr{B}(\mathbb{R}) \times \mathscr{B}(\mathbb{R}) = \mathscr{B}(\mathbb{R}^2)$. □

Let us introduce some terminology and notation. Suppose that X and Y are sets and that E is a subset of $X \times Y$. Then for each x in X and each y in Y the *sections* E_x and E^y are the subsets of Y and X given by

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

and

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

If f is a function on $X \times Y$, then the *sections* f_x and f^y are the functions on Y and X given by

$$f_x(y) = f(x,y)$$

and

$$f^y(x) = f(x,y).$$

Lemma 5.1.2. *Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces.*

(a) *If E is a subset of $X \times Y$ that belongs to $\mathscr{A} \times \mathscr{B}$, then each section E_x belongs to $\mathscr{B}$ and each section E^y belongs to $\mathscr{A}$.*
(b) *If f is an extended real-valued (or a complex-valued) $\mathscr{A} \times \mathscr{B}$-measurable function on $X \times Y$, then each section f_x is $\mathscr{B}$-measurable and each section f^y is $\mathscr{A}$-measurable.*

Proof. Suppose that x belongs to X, and let $\mathscr{F}$ be the collection of all subsets E of $X \times Y$ such that E_x belongs to $\mathscr{B}$. Then $\mathscr{F}$ contains all rectangles $A \times B$ for which $A \in \mathscr{A}$ and $B \in \mathscr{B}$ (note that $(A \times B)_x$ is either B or $\varnothing$). In particular, $X \times Y \in \mathscr{F}$. Furthermore, the identities $(E^c)_x = (E_x)^c$ and $(\cup_n E_n)_x = \cup_n((E_n)_x)$ imply that $\mathscr{F}$ is closed under complementation and under the formation of countable unions; thus $\mathscr{F}$ is a σ-algebra. It follows that $\mathscr{F}$ includes the σ-algebra $\mathscr{A} \times \mathscr{B}$ and hence that E_x belongs to $\mathscr{B}$ whenever E belongs to $\mathscr{A} \times \mathscr{B}$. A similar argument shows that E^y belongs to $\mathscr{A}$ whenever E belongs to $\mathscr{A} \times \mathscr{B}$. With this part (a) is proved.

Part (b) follows from part (a) and the identities $(f_x)^{-1}(D) = (f^{-1}(D))_x$ and $(f^y)^{-1}(D) = (f^{-1}(D))^y$. □

Proposition 5.1.3. *Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces. If E belongs to the σ-algebra $\mathscr{A} \times \mathscr{B}$, then the function $x \mapsto \nu(E_x)$ is $\mathscr{A}$-measurable and the function $y \mapsto \mu(E^y)$ is $\mathscr{B}$-measurable.*

Proof. First suppose that the measure ν is finite. Let $\mathscr{F}$ be the class of those sets E in $\mathscr{A} \times \mathscr{B}$ for which the function $x \mapsto \nu(E_x)$ is $\mathscr{A}$-measurable (Lemma 5.1.2 implies that E_x belongs to $\mathscr{B}$, and hence that $\nu(E_x)$ is defined). If $A \in \mathscr{A}$ and $B \in \mathscr{B}$, then $\nu((A \times B)_x) = \nu(B)\chi_A(x)$, and so the rectangle $A \times B$ belongs to $\mathscr{F}$. In particular, the space $X \times Y$ belongs to $\mathscr{F}$. Note that if E and F are sets in $\mathscr{A} \times \mathscr{B}$ such that $E \subseteq F$, then $\nu((F-E)_x) = \nu(F_x) - \nu(E_x)$, and that if $\{E_n\}$ is an increasing sequence of sets in $\mathscr{A} \times \mathscr{B}$, then $\nu((\cup_n E_n)_x) = \lim_n \nu((E_n)_x)$; it follows that $\mathscr{F}$ is closed under the formation of proper differences and under the formation of unions of increasing sequences of sets. Thus $\mathscr{F}$ is a d-system (see Sect. 1.6). Since the family of rectangles with measurable sides is closed under the formation of finite intersections (note that

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)),$$

Theorem 1.6.2 implies that $\mathscr{F} = \mathscr{A} \times \mathscr{B}$. Thus $x \mapsto \nu(E_x)$ is measurable for each E in $\mathscr{A} \times \mathscr{B}$.

Now suppose that ν is σ-infinite, and let $\{D_n\}$ be a sequence of disjoint subsets of Y that belong to $\mathscr{B}$, have finite measure under ν, and satisfy $\cup_n D_n = Y$. Define finite measures $\nu_1, \nu_2, \dots$ on $\mathscr{B}$ by letting $\nu_n(B) = \nu(B \cap D_n)$. According to what we have just proved, for each n the function $x \mapsto \nu_n(E_x)$ is $\mathscr{A}$-measurable; since $\nu(E_x) = \sum_n \nu_n(E_x)$ holds for each x, the measurability of $x \mapsto \nu(E_x)$ follows. The function $y \mapsto \mu(E^y)$ can be treated similarly, and so the proof is complete. □

Theorem 5.1.4. *Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces. Then there is a unique measure $\mu \times \nu$ on the σ-algebra $\mathscr{A} \times \mathscr{B}$ such that*

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

holds for each A in $\mathscr{A}$ and B in $\mathscr{B}$. Furthermore, the measure under $\mu \times \nu$ of an arbitrary set E in $\mathscr{A} \times \mathscr{B}$ is given by

$$(\mu \times \nu)(E) = \int_X \nu(E_x)\,\mu(dx) = \int_Y \mu(E^y)\,\nu(dy). \tag{1}$$

The measure $\mu \times \nu$ is called the *product* of μ and ν.

Proof. The measurability of $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ for each E in $\mathscr{A} \times \mathscr{B}$ follows from Proposition 5.1.3. Thus we can define functions $(\mu \times \nu)_1$ and $(\mu \times \nu)_2$ on $\mathscr{A} \times \mathscr{B}$ by $(\mu \times \nu)_1(E) = \int_X \nu(E_x)\,\mu(dx)$ and $(\mu \times \nu)_2(E) = \int_Y \mu(E^y)\,\nu(dy)$. It is clear that $(\mu \times \nu)_1(\varnothing) = (\mu \times \nu)_2(\varnothing) = 0$. If $\{E_n\}$ is a sequence of disjoint sets in $\mathscr{A} \times \mathscr{B}$, if $E = \cup_n E_n$, and if $x \in X$, then $\{(E_n)_x\}$ is a sequence of disjoint sets in $\mathscr{B}$ such that $E_x = \cup_n((E_n)_x)$ and hence such that $\nu(E_x) = \sum_n \nu((E_n)_x)$; thus Corollary 2.4.2 implies that

$$(\mu \times \nu)_1(E) = \int_X \nu(E_x)\,\mu(dx) = \sum_n \int_X \nu((E_n)_x)\,\mu(dx) = \sum_n (\mu \times \nu)_1(E_n),$$

and so $(\mu \times \nu)_1$ is countably additive. A similar argument shows that $(\mu \times \nu)_2$ is countably additive. It is easy to check that if $A \in \mathscr{A}$ and $B \in \mathscr{B}$, then

$$(\mu \times \nu)_1(A \times B) = \mu(A)\nu(B) = (\mu \times \nu)_2(A \times B).$$

Hence $(\mu \times \nu)_1$ and $(\mu \times \nu)_2$ are measures on $\mathscr{A} \times \mathscr{B}$ that have the required values on the rectangles with measurable sides.

The uniqueness of $\mu \times \nu$ follows from Corollary 1.6.4. Thus $(\mu \times \nu)_1 = (\mu \times \nu)_2$, and Eq. (1) holds for each E in $\mathscr{A} \times \mathscr{B}$. □

Example 5.1.5. Let us look again at the space $\mathbb{R}^2$. We have already shown that $\mathscr{B}(\mathbb{R}^2) = \mathscr{B}(\mathbb{R}) \times \mathscr{B}(\mathbb{R})$. Let λ_1 be Lebesgue measure on the Borel subsets of $\mathbb{R}$, and let λ_2 be Lebesgue measure on the Borel subsets of $\mathbb{R}^2$. Each rectangle in $\mathbb{R}^2$ of the form $(a,b] \times (c,d]$ is assigned the same value, namely $(b-a)(d-c)$, by λ_2 and by $\lambda_1 \times \lambda_1$; thus Proposition 1.4.3 or Corollary 1.6.4 implies that $\lambda_2 = \lambda_1 \times \lambda_1$. With this we have a second construction of Lebesgue measure on $\mathbb{R}^2$. □

Exercises

1. Use the results of Sect. 5.1 to give another solution to Exercise 1.4.1.
2. Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces, and let E belong to $\mathscr{A} \times \mathscr{B}$. Show that if μ-almost every section E_x has measure zero under ν, then ν-almost every section E^y has measure zero under μ.
3. Show that every $(d-1)$-dimensional hyperplane in $\mathbb{R}^d$ has zero d-dimensional Lebesgue measure (a $(d-1)$-dimensional hyperplane is a set that has the form $\{x \in \mathbb{R}^d : \sum_i a_i x_i = b\}$ for some b in $\mathbb{R}$ and some nonzero element $(a_1, \dots, a_d)$ of $\mathbb{R}^d$).
4. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces.
   (a) Use Proposition 2.6.2 to show that for each y in Y the function $x \mapsto (x,y)$ is measurable with respect to $\mathscr{A}$ and $\mathscr{A} \times \mathscr{B}$ and that for each x in X the function $y \mapsto (x,y)$ is measurable with respect to $\mathscr{B}$ and $\mathscr{A} \times \mathscr{B}$.
   (b) Use part (a) to give another proof of Lemma 5.1.2. (See Proposition 2.6.1.)
5. Let $\mathscr{M}_1$ be the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}$, and let $\mathscr{M}_2$ be the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}^2$. Show that $\mathscr{M}_2 \neq \mathscr{M}_1 \times \mathscr{M}_1$. (Hint: Which subsets of $\mathbb{R}$ can arise as sections of sets in $\mathscr{M}_2$?)
6. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces, and let K be a kernel from $(X, \mathscr{A})$ to $(Y, \mathscr{B})$ such that $K(x,Y)$ is finite for each x in X (see Exercise 2.4.7).
   (a) Show that the formula $(x,E) \mapsto K(x,E_x)$ defines a kernel from $(X, \mathscr{A})$ to $(X \times Y, \mathscr{A} \times \mathscr{B})$.
   (b) Show that if μ is a measure on $(X, \mathscr{A})$, then
   $$E \mapsto \int K(x,E_x)\,\mu(dx)$$
   defines a measure on $\mathscr{A} \times \mathscr{B}$.
   (c) How can the existence of the product of a pair of finite measures be deduced from parts (a) and (b) of this exercise?

7. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces. Show that if $C \in \mathscr{A} \times \mathscr{B}$, then the collection of sections $\{C_x : x \in X\}$ has at most the cardinality of the continuum. (Hint: See Exercise 1.1.7. Show that if

$$C \in \sigma(\{A_n \times B_n : n = 1, 2, \dots\})$$

and if x_1 and x_2 belong to exactly the same A_n's, then $C_{x_1} = C_{x_2}$. Next use the function $x \mapsto \{\chi_{A_n}(x)\}$ to map X into $\{0,1\}^{\mathbb{N}}$, and note that $\{0,1\}^{\mathbb{N}}$ has the cardinality of the continuum (see A.8).)
8. Show that if the cardinality of X is larger than that of the continuum and if $\mathscr{A}$ is a σ-algebra on X, then the diagonal in $X \times X$ (that is, the set $\{(x_1, x_2) \in X \times X : x_1 = x_2\}$) does not belong to $\mathscr{A} \times \mathscr{A}$. (Hint: Use Exercise 7.)

5.2 Fubini's Theorem

The following two theorems enable one to evaluate integrals with respect to product measures by evaluating iterated integrals.

Proposition 5.2.1 (Tonelli's Theorem). *Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces, and let $f : X \times Y \to [0, +\infty]$ be $\mathscr{A} \times \mathscr{B}$-measurable. Then*

(a) *the function $x \mapsto \int_Y f_x \, d\nu$ is $\mathscr{A}$-measurable and the function $y \mapsto \int_X f^y \, d\mu$ is $\mathscr{B}$-measurable, and*
(b) *f satisfies*

$$\int_{X \times Y} f \, d(\mu \times \nu) = \int_X \left(\int_Y f_x \, d\nu \right) \mu(dx) = \int_Y \left(\int_X f^y \, d\mu \right) \nu(dy). \quad (1)$$

Note that the functions f_x and f^y are nonnegative and measurable (Lemma 5.1.2); thus the expression $\int_Y f_x \, d\nu$ is defined for each x in X and the expression $\int_X f^y \, d\mu$ is defined for each y in Y. Note also that (1) can be reformulated as

$$\begin{aligned} \int_{X \times Y} f(x,y) \, d(\mu \times \nu)(x,y) &= \int_X \left(\int_Y f(x,y) \, \nu(dy) \right) \mu(dx) \\ &= \int_Y \left(\int_X f(x,y) \, \mu(dx) \right) \nu(dy). \end{aligned}$$

Proof. First suppose that E belongs to $\mathscr{A} \times \mathscr{B}$ and that f is the characteristic function of E. Then the sections f_x and f^y are the characteristic functions of the sections E_x and E^y, and so the relations $\int f_x \, d\nu = \nu(E_x)$ and $\int f^y \, d\mu = \mu(E^y)$ hold for each x and y. Thus Proposition 5.1.3 and Theorem 5.1.4 imply that conclusions (a) and (b) hold if f is a characteristic function. The additivity and homogeneity of the integral now imply that they hold for nonnegative simple $\mathscr{A} \times \mathscr{B}$-measurable functions, and Proposition 2.1.5, Proposition 2.1.8, and Theorem 2.4.1 imply that they hold for arbitrary nonnegative $\mathscr{A} \times \mathscr{B}$-measurable functions. □

Note that (1) is applicable to each nonnegative $\mathscr{A} \times \mathscr{B}$-measurable function, integrable or not; thus one can often determine whether an $\mathscr{A} \times \mathscr{B}$-measurable function f is integrable by using Proposition 5.2.1 to calculate $\int |f|\, d(\mu \times \nu)$.

Theorem 5.2.2 (Fubini's Theorem). *Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces, and let $f\colon X \times Y \to [-\infty, +\infty]$ be $\mathscr{A} \times \mathscr{B}$-measurable and $\mu \times \nu$-integrable. Then*

(a) *for μ-almost every x in X the section f_x is ν-integrable and for ν-almost every y in Y the section f^y is μ-integrable,*

(b) *the functions I_f and J_f defined by*

$$I_f(x) = \begin{cases} \int_Y f_x \, d\nu & \text{if } f_x \text{ is } \nu\text{-integrable,} \\ 0 & \text{otherwise} \end{cases}$$

and

$$J_f(y) = \begin{cases} \int_X f^y \, d\mu & \text{if } f^y \text{ is } \mu\text{-integrable,} \\ 0 & \text{otherwise} \end{cases}$$

belong to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$ and $\mathscr{L}^1(Y, \mathscr{B}, and \nu, \mathbb{R})$, respectively, and

(c) *the relation*

$$\int_{X \times Y} f \, d(\mu \times \nu) = \int_X I_f \, d\mu = \int_Y J_f \, d\nu$$

holds.

Note that part (c) of this theorem is just a precise way of rephrasing equation (1) in the case where f is integrable but not necessarily nonnegative.

Proof. Let f^+ and f^- be the positive and negative parts of f. Lemma 5.1.2 implies that the sections f_x, $(f^+)_x$, and $(f^-)_x$ are $\mathscr{B}$-measurable, and Proposition 5.2.1 implies that the functions $x \mapsto \int (f^+)_x \, d\nu$ and $x \mapsto \int (f^-)_x \, d\nu$ are $\mathscr{A}$-measurable and μ-integrable and hence that they are finite μ-almost everywhere (Corollary 2.3.14). Thus f_x is ν-integrable for almost every x. Let N be the set of those x for which $\int (f^+)_x \, d\nu = +\infty$ or $\int (f^-)_x \, d\nu = +\infty$. Then N belongs to $\mathscr{A}$, and $I_f(x)$ is equal to 0 if $x \in N$ and is equal to $\int (f^+)_x \, d\nu - \int (f^-)_x \, d\nu$ otherwise; consequently I_f belongs to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$. Propositions 5.2.1 and 2.3.9 now imply that

$$\begin{aligned} \int f \, d(\mu \times \nu) &= \int f^+ \, d(\mu \times \nu) - \int f^- \, d(\mu \times \nu) \\ &= \int \left(\int (f^+)_x \, d\nu \right) \mu(dx) - \int \left(\int (f^-)_x \, d\nu \right) \mu(dx) \\ &= \int I_f \, d\mu. \end{aligned}$$

Similar arguments apply to the functions f^y and J_f, and so the proof is complete. □

Of course we can deal with a complex-valued function on $X \times Y$ by separating it into its real and imaginary parts.

We briefly sketch the theory of products of a finite number of measure spaces. Let $(X_1, \mathscr{A}_1, \mu_1), \ldots, (X_n, \mathscr{A}_n, \mu_n)$ be σ-finite measure spaces. Then

$$\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$$

is defined to be the σ-algebra on $X_1 \times \cdots \times X_n$ generated by the sets of the form $A_1 \times \cdots \times A_n$, where $A_i \in \mathscr{A}_i$ for $i = 1, \ldots, n$. It is not hard to check that if $1 \le k < n$ and if we make the usual identification of $(X_1 \times \cdots \times X_k) \times (X_{k+1} \times \cdots \times X_n)$ with $X_1 \times \cdots \times X_n$, then

$$(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k) \times (\mathscr{A}_{k+1} \times \cdots \times \mathscr{A}_n) = \mathscr{A}_1 \times \cdots \times \mathscr{A}_n.$$

Thus we can use Theorem 5.1.4 (applied $n-1$ times) to construct a measure $\mu_1 \times \cdots \times \mu_n$ on $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ that satisfies

$$(\mu_1 \times \cdots \times \mu_n)(A_1 \times \cdots \times A_n) = \mu_i(A_1) \cdots \mu_n(A_n)$$

whenever $A_i \in \mathscr{A}_i$ for $i = 1, \ldots, n$. Corollary 1.6.4 implies that the measure $\mu_1 \times \cdots \times \mu_n$ is unique. Integrals with respect to $\mu_1 \times \cdots \times \mu_n$ can be evaluated by repeated applications of Proposition 5.2.1 or Theorem 5.2.2.

Exercises

1. Let λ be Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, let μ be counting measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, and let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the characteristic function of the line $\{(x,y) \in \mathbb{R}^2 : y = x\}$. Show that

$$\int\int f(x,y)\,\mu(dy)\,\lambda(dx) \neq \int\int f(x,y)\,\lambda(dx)\,\mu(dy).$$

2. Suppose that $f\colon \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$f(x,y) = \begin{cases} 1 & \text{if } x \ge 0 \text{ and } x \le y < x+1, \\ -1 & \text{if } x \ge 0 \text{ and } x+1 \le y < x+2, \\ 0 & \text{otherwise.} \end{cases}$$

   Show that $\int\int f(x,y)\,\lambda(dy)\,\lambda(dx) \neq \int\int f(x,y)\,\lambda(dx)\,\lambda(dy)$. Why does this not contradict Theorem 5.2.2?

3. (a) Let μ be a measure on $(X, \mathscr{A})$. Show that if μ is σ-finite, then there are finite measures $\mu_1, \mu_2, \ldots$ on $(X, \mathscr{A})$ such that $\mu = \sum_n \mu_n$.
   (b) Show by example that the converse of part (a) does not hold.

(c) Let $(X,\mathscr{A},\mu)$ and $(Y,\mathscr{B},\nu)$ be measure spaces, and let $f\colon X\times Y\to[0,+\infty]$ be $\mathscr{A}\times\mathscr{B}$-measurable. Show that if μ and ν are sums of series of finite measures, then the functions $x\mapsto\int f(x,y)\,\nu(dy)$ and $y\mapsto\int f(x,y)\,\mu(dx)$ are measurable, and

$$\int\int f(x,y)\,\nu(dy)\,\mu(dx)=\int\int f(x,y)\,\mu(dx)\,\nu(dy).$$

4. Let $(X,\mathscr{A})$ and $(Y,\mathscr{B})$ be measurable spaces, let μ_1 and μ_2 be finite measures on $(X,\mathscr{A})$, and let ν_1 and ν_2 be finite measures on $(Y,\mathscr{B})$. Show that if $\mu_2\ll\mu_1$ and $\nu_2\ll\nu_1$, then $\mu_2\times\nu_2\ll\mu_1\times\nu_1$. How are the various Radon–Nikodym derivatives related? (Hint: Do both parts at once, showing that $\mu_2\times\nu_2$ can be computed by integrating an appropriate function with respect to $\mu_1\times\nu_1$.)
5. Let $(X,\mathscr{A})$, $(Y,\mathscr{B})$, K, and μ be as in Exercise 5.1.6, and let ν be the measure on $(X\times Y,\mathscr{A}\times\mathscr{B})$ defined in part (b) of that exercise. Show that if $f\colon X\times Y\to[0,+\infty]$ is $\mathscr{A}\times\mathscr{B}$-measurable, then
   (a) $x\mapsto\int f(x,y)\,K(x,dy)$ is $\mathscr{A}$-measurable, and
   (b) $\int f\,d\nu=\int\int f(x,y)\,K(x,dy)\,\mu(dx)$.
6. Let $(X,\mathscr{A},\mu)$ and $(Y,\mathscr{B},\nu)$ be σ-finite measure spaces, let $(\mathscr{A}\times\mathscr{B})_{\mu\times\nu}$ be the completion of $\mathscr{A}\times\mathscr{B}$ under $\mu\times\nu$, and let $\overline{\mu\times\nu}$ be the completion of $\mu\times\nu$.

   (a) Suppose that $f\colon X\times Y\to[0,+\infty]$ is $(\mathscr{A}\times\mathscr{B})_{\mu\times\nu}$-measurable. Show that f_x is $\mathscr{B}_\nu$-measurable for μ-almost every x in X and that f^y is $\mathscr{A}_\mu$-measurable for ν-almost every y in Y. Show also that if $\int f(x,y)\,\overline{\nu}(dy)$ is defined to be 0 whenever f_x is not $\mathscr{B}_\nu$-measurable and $\int f(x,y)\,\overline{\mu}(dx)$ is defined to be 0 whenever f^y is not $\mathscr{A}_\mu$-measurable, then

   $$\int f\,d(\overline{\mu\times\nu})=\int\int f(x,y)\,\overline{\nu}(dy)\,\overline{\mu}(dx)=\int\int f(x,y)\,\overline{\mu}(dx)\,\overline{\nu}(dy).$$

   (Hint: See Proposition 2.2.5.)
   (b) State and prove an analogous modification of Theorem 5.2.2.

5.3 Applications

We begin by noting a couple of easy-to-derive consequences of the theory of product measures.

Example 5.3.1. Let $(X,\mathscr{A},\mu)$ be a σ-finite measure space, let λ be Lebesgue measure on $(\mathbb{R},\mathscr{B}(\mathbb{R}))$, and let $f\colon X\to[0,+\infty]$ be $\mathscr{A}$-measurable. Let E be defined by

$$E=\{(x,y)\in X\times\mathbb{R}:0\le y<f(x)\};$$

in other words, E is the “region under the graph of f.” Then E belongs to $\mathscr{A} \times \mathscr{B}(\mathbb{R})$ (check this), and so its measure under $\mu \times \lambda$ can be computed using Theorem 5.1.4. On the one hand,

$$(\mu \times \lambda)(E) = \int_X \lambda(E_x)\,\mu(dx) = \int_X f(x)\,\mu(dx),$$

while on the other,

$$(\mu \times \lambda)(E) = \int_{\mathbb{R}} \mu(E^y)\,\lambda(dy) = \int_0^{\infty} \mu(\{x \in X : f(x) > y\})\,\lambda(dy).$$

Thus we have the often useful relation

$$\int_X f(x)\,\mu(dx) = \int_0^{\infty} \mu(\{x \in X : f(x) > y\})\,dy.$$ □

Example 5.3.2. Next we use Fubini’s theorem to derive a familiar result about double series. Let $\sum_{m,n} a_{m,n}$ be a double series, and let μ be counting measure on $\mathbb{N}$ (more precisely, on the σ-algebra of all subsets of $\mathbb{N}$). The series $\sum_{m,n} a_{m,n}$ is absolutely convergent if and only if the function $(m,n) \mapsto a_{m,n}$ is $\mu \times \mu$-integrable. Thus Fubini’s theorem implies that if $\sum_{m,n} a_{m,n}$ is absolutely convergent, then $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n}$; in other words, the order of summation can be reversed for absolutely convergent series. See also Exercise 3. □

Let us consider one version of integration by parts. Another version will be discussed in Sect. 6.3.

Proposition 5.3.3. *Let F, G: $\mathbb{R} \to \mathbb{R}$ be bounded nondecreasing right-continuous functions that vanish*[1] *at $-\infty$, let μ_F and μ_G be the measures they induce on $\mathscr{B}(\mathbb{R})$, and let a and b be real numbers such that $a < b$. Then*

$$\int_{[a,b]} \frac{F(x) + F(x-)}{2}\,\mu_G(dx) + \int_{[a,b]} \frac{G(x) + G(x-)}{2}\,\mu_F(dx) \qquad (1)$$

$$= F(b)G(b) - F(a-)G(a-).$$

Proof. Let S be the square $[a,b] \times [a,b]$, and let T_1 and T_2 be the triangular regions consisting of those points (x,y) in S for which $x \geq y$ and for which $x < y$, respectively. We compute the measure of S under $\mu_F \times \mu_G$ in two ways. On the one hand,

$$(\mu_F \times \mu_G)(S) = \mu_F([a,b])\mu_G([a,b]) = (F(b) - F(a-))(G(b) - G(a-)). \qquad (2)$$

[1] In other words, $\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} G(x) = 0$.

On the other hand, the relation $S = T_1 \cup T_2$ and Theorem 5.1.4 imply that

$$(\mu_F \times \mu_G)(S) = \int_{[a,b]} \mu_G([a,x])\, \mu_F(dx) + \int_{[a,b]} \mu_F([a,y))\, \mu_G(dy). \tag{3}$$

If we replace the "dummy variable" y in this equation with x and express $\mu_G([a,x])$ and $\mu_F([a,x))$ in terms of G and F, then the right-hand side of (3) becomes

$$\int_{[a,b]} (G(x) - G(a-))\, \mu_F(dx) + \int_{[a,b]} (F(x-) - F(a-))\, \mu_G(dx).$$

Equating this to the expression on the right side of (2) and using a little algebra gives

$$\int_{[a,b]} G(x)\, \mu_F(dx) + \int_{[a,b]} F(x-)\, \mu_G(dx) = F(b)G(b) - F(a-)G(a-).$$

The functions F and G can be interchanged in this identity, yielding

$$\int_{[a,b]} G(x-)\, \mu_F(dx) + \int_{[a,b]} F(x)\, \mu_G(dx) = F(b)G(b) - F(a-)G(a-).$$

These two equations together imply Eq. (1). □

See Exercises 4 and 5 for more information about Eq. (1).

Our last application of Fubini's theorem is to the convolution of functions in $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$.

Proposition 5.3.4. *Let f and g belong to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$. Then*

(a) *for almost every x the function $t \mapsto f(x-t)g(t)$ belongs to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$, and*

(b) *the function $f * g$ defined by*

$$f * g(x) = \begin{cases} \int f(x-t)g(t)\,dt & \text{if } t \mapsto f(x-t)g(t) \text{ is Lebesgue integrable,} \\ 0 & \text{otherwise} \end{cases}$$

belongs to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ and satisfies $\|f * g\|_1 \le \|f\|_1 \|g\|_1$.

Proof. We begin by checking that the function $(x,t) \mapsto f(x-t)g(t)$ is measurable with respect to $\mathscr{B}(\mathbb{R}^2)$, and hence (see Sect. 5.1) with respect to $\mathscr{B}(\mathbb{R}) \times \mathscr{B}(\mathbb{R})$. The function $(x,t) \mapsto f(x-t)$ is the composition of the continuous, and hence Borel measurable, function $(x,t) \mapsto x-t$ with the Borel measurable function f; thus it is Borel measurable (Proposition 2.6.1). A similar argument shows that $(x,t) \mapsto g(t)$ is Borel measurable. Consequently the function $(x,t) \mapsto f(x-t)g(t)$ is Borel measurable.

Thus we can use Proposition 5.2.1 and the translation invariance of Lebesgue measure (see the remarks at the end of Sect. 2.6) to justify the calculation

$$\int |f(x-t)g(t)|\,d(\lambda\times\lambda)(x,t) = \int\int |f(x-t)g(t)|\,\lambda(dx)\,\lambda(dt)$$
$$= \int \|f\|_1 |g(t)|\,\lambda(dt) = \|f\|_1\|g\|_1. \qquad (4)$$

It follows that $(x,t)\mapsto f(x-t)g(t)$ belongs to

$$\mathscr{L}^1(\mathbb{R}^2, \mathscr{B}(\mathbb{R})\times\mathscr{B}(\mathbb{R}), \lambda\times\lambda),$$

and so Fubini's theorem implies that $t\mapsto f(x-t)g(t)$ is integrable for almost every x. Since $|f*g(x)| \le \int |f(x-t)g(t)|\,\lambda(dt)$ holds for each x in $\mathbb{R}$, Fubini's theorem and calculation (4) also imply that $f*g$ belongs to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ and that $\|f*g\|_1 \le \|f\|_1\|g\|_1$. □

The *convolution* of the functions f and g in $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ is the function $f*g$ defined in part (b) of Proposition 5.3.4. Note that if f_1, f_2, g_1, and g_2 belong to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ and if $f_1 = f_2$ and $g_1 = g_2$ hold λ-almost everywhere, then $(f_1 * g_1)(x) = (f_2 * g_2)(x)$ holds at each x in $\mathbb{R}$. Thus convolution, which we have defined as an operator that assigns a function in $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ to each pair of functions in $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$, can be (and usually is) considered as an operator that assigns an element of $L^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ to each pair of elements of $L^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$.

Convolution turns out to be a fundamental operation in harmonic analysis. Although we do not have space to develop its properties in detail, a few are presented in the exercises below. See Chap. 10 for convolutions in probability theory, and see Sect. 9.4 for convolutions in a much more general setting.

Exercises

1. Let μ be a σ-finite measure on $(X, \mathscr{A})$, and let $f, g\colon X \to [0, +\infty]$ be $\mathscr{A}$-measurable functions such that

$$\mu(\{x : f(x) > t\}) \le \mu(\{x : g(x) > t\})$$

   holds for each positive t. Show that $\int f\,d\mu \le \int g\,d\mu$.

2. Let μ be a σ-finite measure on $(X, \mathscr{A})$, let $f\colon X \to [0, +\infty]$ be $\mathscr{A}$-measurable, and let p satisfy $1 \le p < +\infty$. Show that

$$\int f^p\,d\mu = \int_0^\infty p t^{p-1} \mu(\{x : f(x) > t\})\,dt.$$

3. Let $\sum_{m,n} a_{m,n}$ be a double series whose terms are nonnegative. Show that $\sum_m \sum_n a_{m,n} = \sum_n \sum_m a_{m,n}$
   (a) by applying Proposition 5.2.1, and
   (b) by checking directly that $\sum_m \sum_n a_{m,n}$ and $\sum_n \sum_m a_{m,n}$ are both equal to

$$\sup \left\{ \sum_{(m,n)\in F} a_{m,n} : F \text{ is a finite subset of } \mathbb{N} \times \mathbb{N} \right\}.$$

   (Note that we did not assume that the series involved are convergent.)
4. Show that if the functions F and G in Proposition 5.3.3 have no points of discontinuity in common, then Eq. (1) can be replaced with the equation

$$\int_{[a,b]} F(x)\,\mu_G(dx) + \int_{[a,b]} G(x)\,\mu_F(dx) = F(b)G(b) - F(a-)G(a-).$$

5. Show by example that formula (1) of Proposition 5.3.3 cannot in general be replaced with the formula in Exercise 4.
6. Show that if f and g belong to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$, then $f * g = g * f$. (Hint: See the remarks at the end of Sect. 2.6.)
7. Show that if f and g belong to $\mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ and if g is bounded, then $f * g$ is continuous. (Hint: See Exercise 3.4.5.)
8. Suppose that A is a Borel subset of $\mathbb{R}$ that satisfies $0 < \lambda(A) < +\infty$.
   (a) Show that the function $x \mapsto \lambda(A \cap (x + A))$ is continuous and is nonzero throughout some open interval about 0 (of course $x + A$ is the set $\{x + a : a \in A\}$). (Hint: Consider $\chi_{-A} * \chi_A$, where $-A = \{-a : a \in A\}$, and use Exercise 7.)
   (b) Use part (a) to give another proof of Proposition 1.4.10.

Notes

The theory of products of a finite number of σ-finite measure spaces, as given in this chapter, can be found in almost every book on measure and integration. The theory of products of an infinite number of measure spaces of total mass 1, as presented in Sect. 10.6, is needed for the study of probability and can be found in some books on measure theory and in most books on measure-theoretic probability.

The proof of Proposition 1.4.10 indicated in Exercise 5.3.8 was shown to me by Charles Rockland and (independently) by Lee Rubel.

Chapter 6
Differentiation

In this chapter we look at two aspects of the relationship between differentiation and integration. First, in Sect. 6.1, we look at changes of variables in d-dimensional integrals. Such changes of variables occur, for example, when one evaluates an integral over a region in $\mathbb{R}^2$ by converting to polar coordinates. Then, in Sects. 6.2 and 6.3, we look at some deeper aspects of differentiation theory, including the almost everywhere differentiability of monotone functions and of indefinite integrals and the relationship between Radon–Nikodym derivatives and differentiation theory. The Vitali covering theorem is an important tool for this. The discussion of differentiation theory will be resumed when we discuss the Henstock–Kurzweil integral in Appendix H.

6.1 Change of Variable in $\mathbb{R}^d$

In this section we deal with changes of variable in $\mathbb{R}^d$ and with their relation to Lebesgue measure. The main result is Theorem 6.1.7. Let us begin by recalling some definitions.

Let M_d be the set of all d by d matrices with real entries, and let D be a real-valued function on M_d. We will sometimes find it convenient to denote the columns of a d by d matrix A by A_1, A_2, ..., A_d and to write $D(A_1, A_2, \ldots, A_d)$ in place of $D(A)$. The function D is *multilinear* if for each i and each choice of A_j (for $j \neq i$) the map $A_i \mapsto D(A_1, \ldots, A_d)$ is linear, is *alternating* if $D(A) = 0$ holds whenever two of the columns of A are equal, and is a *determinant* if it is multilinear, is alternating, and satisfies $D(I) = 1$ (here I is, of course, the d by d identity matrix).

We need to recall a few basic facts about determinants.

Lemma 6.1.1. *For each positive integer d there is a unique determinant on M_d.*

We follow the standard usage and use $\det(A)$ to denote the determinant of a matrix A.

D.L. Cohn, *Measure Theory: Second Edition*, Birkhäuser Advanced Texts Basler Lehrbücher, DOI 10.1007/978-1-4614-6956-8_6,