Course: Theory of Probability I
Term: Fall 2013
Instructor: Gordan Zitkovic

Lecture 5
Theorems of Fubini-Tonelli and Radon-Nikodym

Products of measure spaces

We have seen that it is possible to define products of arbitrary collections of measurable spaces - one generates the σ-algebra on the product by all finite-dimensional cylinders. The purpose of the present section is to extend that construction to products of measure spaces, i.e., to define products of measures.

Let us first consider the case of two measure spaces $(S, \mathcal{S}, \mu_S)$ and $(T, \mathcal{T}, \mu_T)$. If the measures are stripped, the product $S \times T$ is endowed with the product σ-algebra $\mathcal{S} \otimes \mathcal{T} = \sigma(\{A \times B : A \in \mathcal{S}, B \in \mathcal{T}\})$. The family $\mathcal{P} = \{A \times B : A \in \mathcal{S}, B \in \mathcal{T}\}$ serves as a good starting point towards the creation of the product measure $\mu_S \otimes \mu_T$. Indeed, if we interpret of the elements in $\mathcal{P}$ as rectangles of sorts, it is natural to define

$$(\mu_S \otimes \mu_T)(A \times B) = \mu_S(A)\mu_T(B).$$

The family $\mathcal{P}$ is a π-system (why?), but not necessarily an algebra, so we cannot use Theorem 2.9 (Caratheodory's extension theorem) to define an extension of $\mu_S \otimes \mu_T$ to the whole $\mathcal{S} \otimes \mathcal{T}$. It it not hard, however, to enlarge $\mathcal{P}$ a little bit, so that the resulting set is an algebra, but that the measure $\mu_S \otimes \mu_T$ can still be defined there in a natural way. Indeed, consider the smallest algebra that contains $\mathcal{P}$. It is easy to see that it must contain the family $\mathcal{S}$ defined by

$$\mathcal{A} = \{\cup_{k=1}^n A_k \times B_k : n \in \mathbb{N}, A_k \in \mathcal{S}, B_k \in \mathcal{T}, k = 1, \dots, n\}.$$

Problem 5.1. Show that $\mathcal{A}$ is, in fact, an algebra and that each element $C \in \mathcal{A}$ can be written in the form

$$C = \cup_{k=1}^n A_k \times B_k,$$

for $n \in \mathbb{N}, A_k \in \mathcal{S}, B_k \in \mathcal{T}, k = 1, \dots, n$, such that $A_1 \times B_1, \dots, A_n \times B_n$ are *pairwise disjoint*.

The problem above allows us to extend the definition of the set function $\mu_S \otimes \mu_T$ to the entire $\mathcal{A}$ by

$$(\mu_S \otimes \mu_T)(C) = \sum_{k=1}^{n} \mu_S(A_k)\mu_T(B_k),$$

where $C = \cup_{k=1}^{n} A_k \times B_k$ for $n \in \mathbb{N}, A_k \in \mathcal{S}, B_k \in \mathcal{T}, k = 1, \dots, n$ is a representation of C with pairwise disjoint $A_1 \times B_1, \dots, A_n \times B_n$.

At this point, we could attempt to show that the so-defined set function is σ-additive on $\mathcal{A}$ and extend it using the Caratheodory extension theorem. This is indeed possible - under the additional assumption of σ-finiteness - but we will establish the existence of product measures as a side-effect in the proof of Fubini's theorem below.

Lemma 5.1 (Sections of measurable sets are measurable). *Let C be an $\mathcal{S} \otimes \mathcal{T}$-measurable subset of $S \times T$. For each $x \in S$ the **section** $C_x = \{y \in T : (x, y) \in C\}$ is measurable in $\mathcal{T}$.*

Proof. In the spirit of most of the measurability arguments seen so far in these notes, let $\mathcal{C}$ denote the family of all $C \in \mathcal{S} \times \mathcal{T}$ such that C_x is $\mathcal{T}$-measurable for each $x \in S$. Clearly, the "rectangles" $A \times B$, $A \in \mathcal{S}$, $B \in \mathcal{T}$ are in $\mathcal{A}$ because their sections are either equal to $\emptyset$ or B, for each $x \in S$. Remember that the set of all rectangles generates $\mathcal{S} \otimes \mathcal{T}$. The proof of the theorem will, therefore, be complete once it is established that $\mathcal{C}$ is a σ-algebra. This easy exercise is left to the reader. □

Problem 5.2. Show that an analogous result holds for measurable functions, i.e., show that if $f : S \times T \to \bar{\mathbb{R}}$ is a $\mathcal{S} \otimes \mathcal{T}$-measurable function, then the function $x \mapsto f(x, y_0)$ is $\mathcal{S}$-measurable for each $y_0 \in T$, and the function $y \mapsto f(x_0, y)$ is $\mathcal{T}$-measurable for each $y_0 \in T$.

Proposition 5.2 (A simple Cavalieri's principle). *Let μ_S and μ_T be finite measures. For $C \in \mathcal{S} \otimes \mathcal{T}$, define the functions $\varphi_C : T \to [0, \infty)$ and $\psi_C : S \to [0, \infty)$ by*

$$\varphi_C(y) = \mu_S(C_y),\ \psi_C(x) = \mu_T(C_x).$$

Then,

$$\varphi_C \in \mathcal{L}^0_+(\mathcal{T}),\ \psi_C \in \mathcal{L}^0_+(\mathcal{S}) \text{ and } \int \varphi_C \, d\mu_T = \int \psi_C \, d\mu_S. \tag{5.1}$$

Proof. Note that, by Problem 5.2, the function $x \mapsto \mathbf{1}_C(x, y)$ is $\mathcal{S}$-measurable for each $y \in T$. Therefore,

$$\int \mathbf{1}_C(\cdot, y) \, d\mu_S = \mu_S(C_y) = \varphi_C(y), \tag{5.2}$$

and the function φ_C is well-defined.

Let $\mathcal{C}$ denote the family of all sets in $\mathcal{S}\otimes\mathcal{T}$ such that (5.1) holds. First, observe that $\mathcal{C}$ contains all rectangles $A\times B$, $A\in\mathcal{S}$, $B\in\mathcal{T}$, i.e., it contains a π-system which generates $\mathcal{S}\otimes\mathcal{T}$. Therefore, by the π-λ Theorem (Theorem 2.12), it will be enough to show that $\mathcal{C}$ is a λ-system. We leave the details to the reader[1] □

[1] Use representation (5.2) and the monotone convergence theorem. Where is the finiteness of the measures used?

Proposition 5.3 (Simple Cavalieri holds for σ-finite measures)**.** *The conclusion of Proposition 5.2 continues to hold if we assume that μ_S and μ_T are only σ-finite.*

Proof (*). Thanks to σ-finiteness, there exists pairwise disjoint sequences $\{A_n\}_{n\in\mathbb{N}}$ and $\{B_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$ and $\mathcal{T}$, respectively, such that $\cup_n A_n = S$, $\cup_m B_m = T$ and $\mu_S(A_n)<\infty$ and $\mu_S(B_m)<\infty$, for all $m,n\in\mathbb{N}$.

For $m,n\in\mathbb{N}$, define the set-functions μ_S^n and μ_T^m on $\mathcal{S}$ and $\mathcal{T}$ respectively by

$$\mu_S^n(A)=\mu_S(A_n\cap A),\ \mu_T^m(B)=\mu_T(B_m\cap B).$$

It is easy to check that all μ_S^n and μ_T^m, $m,n\in\mathbb{N}$ are finite measures on $\mathcal{S}$ and $\mathcal{T}$, respectively. Moreover, $\mu_S(A)=\sum_{n=1}^{\infty}\mu_S^n(A)$, $\mu_T(B)=\sum_{m=1}^{\infty}\mu_T^m(B)$. In particular, if we set $\varphi_C^n(y)=\mu_S^n(C_y)$ and $\psi_C^m(x)=\mu_T^m(C_x)$, for all $x\in S$ and $y\in S$, we have

$$\varphi_C(y)=\mu_S(C_y)=\sum_{n=1}^{\infty}\mu_S^n(C_y)=\sum_{n=1}^{\infty}\varphi_C^n(y),\ \text{and}$$
$$\psi_C(x)=\mu_T(C_x)=\sum_{m=1}^{\infty}\mu_T^m(C_x)=\sum_{m=1}^{\infty}\psi_C^m(x),$$

for all $x\in S$, $y\in T$.

We can apply the conclusion of Proposition 5.2 to all pairs $(S,\mathcal{S},\mu_S^n)$ and $(T,\mathcal{T},\mu_T^m)$, $m,n\in\mathbb{N}$, of finite measure spaces to conclude that all elements of the sums above are measurable functions and that so are φ_C and ψ_C.

Similarly, the sequences of non-negative functions $\sum_{i=1}^{n}\varphi_C^n(y)$ and $\sum_{i=1}^{m}\psi_C^i(x)$ are non-decreasing and converge to φ_C and ψ_C. Therefore, by the monotone convergence theorem,

$$\int\varphi_C\,d\mu_T=\lim_n\sum_{i=1}^{n}\int\varphi_C^i\,d\mu_T,\ \text{and}\ \int\psi_C\,d\mu_S=\lim_n\sum_{i=1}^{n}\int\psi_C^i\,d\mu_S.$$

On the other hand, we have $\int\varphi_C^n d\mu_T^m=\int\psi_C^m\,d\mu_S^n$, by Proposition 5.2. Summing over all $n\in\mathbb{N}$ we have

$$\int\varphi_C\,d\mu_T^m=\sum_{n\in\mathbb{N}}\int\psi_C^m\,d\mu_S^n=\int\psi_C^m\,d\mu_S,$$

where the last equality follows from the fact (see Problem 5.3 below) that

$$\sum_{n\in\mathbb{N}} \int f\, d\mu^n_S = \int f\, d\mu_S,$$

for all $f \in \mathcal{L}^0_+$. Another summation - this time over $m \in \mathbb{N}$ - completes the proof [2]. $\square$

[2] The argument of the proof above uncovers the fact that integration is a bilinear operation, i.e., that the mapping

$$(f, \mu) \to \int f\, d\mu,$$

is linear in both arguments.

Problem 5.3. Let $\{A_n\}_{n\in\mathbb{N}}$ be a measurable partition of S, and let the measure μ^n be defined by $\mu^n(A) = \mu(A \cap A_n)$ for all $A \in \mathcal{S}$. Show that for $f \in \mathcal{L}^0_+$, we have

$$\int f\, d\mu = \sum_{n\in\mathbb{N}} \int f\, d\mu^n.$$

Proposition 5.4 (Finite products of measure spaces)**.** *Let $(S_i, \mathcal{S}_i, \mu_i)$, $i = 1, \dots, n$ be σ-finite measure spaces. There exists a unique measure[3] - denoted by $\mu_1 \otimes \dots \otimes \mu_n$ - on the product space $(S_1 \times \dots \times S_n, \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n)$ with the property that*

$$(\mu_1 \otimes \dots \otimes \mu_n)(A_1 \times \dots \times A_n) = \mu_1(A_1) \dots \mu_n(A_n),$$

for all $A_i \in \mathcal{S}_i$, $i = 1, \dots, n$. Such a measure is necessarily σ-finite.

[3] The measure $\mu_1 \otimes \dots \otimes \mu_n$ is called the **product measure**, and the measure space $(S_1 \times \dots \times S_n, \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n, \mu_1 \otimes \dots \otimes \mu_n)$ the **product (measure space)** of measure spaces $(S_1, \mathcal{S}_1, \mu_1)$, $\dots$, $(S_n, \mathcal{S}_n, \mu_n)$.

Proof. To simplify the notation, we assume that $n = 2$ - the general case is very similar. For $C \in \mathcal{S}_1 \otimes \mathcal{S}_2$, we define

$$(\mu_1 \otimes \mu_2)(C) = \int_{S_2} \varphi_C\, d\mu_2,$$

where $\varphi_C(y) = \mu_1(C_y)$ and $C_y = \{x \in S_1 : (x, y) \in C\}$. It follows from Proposition 5.3 that $\mu_1 \otimes \mu_2$ is well-defined as a map from $\mathcal{S}_1 \otimes \mathcal{S}_2$ to $[0, \infty]$. Also, it is clear that $(\mu_1 \otimes \mu_2)(A \times B) = \mu_1(A)\mu_2(B)$, for all $A \in \mathcal{S}_1$, $B \in \mathcal{S}_2$. It remains to show that $\mu_1 \otimes \mu_2$ is a measure. We start with a pairwise disjoint sequence $\{C_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}_1 \otimes \mathcal{S}_2$. For $y \in S_2$, the sequence $\{(C_n)_y\}_{n\in\mathbb{N}}$ is also pairwise disjoint, and so, with $C = \cup_n C_n$, we have

$$\varphi_C(y) = \mu_1(C_y) = \sum_{n\in\mathbb{N}} \mu_2\Big((C_n)_y\Big) = \sum_{n\in\mathbb{N}} \varphi_{C_n}(y),\ \forall\, y \in S_2.$$

Therefore, by the monotone convergence theorem (see Problem 3.13 for details) we have

$$(\mu_1 \otimes \mu_2)(C) = \int_{S_2} \varphi_C\, d\mu_2 = \sum_{n\in\mathbb{N}} \int_{S_2} \varphi_{C_n}\, d\mu = \sum_{n\in\mathbb{N}} (\mu_1 \otimes \mu_2)(C_n).$$

Finally, let $\{A_n\}_{n\in\mathbb{N}}$, $\{B_n\}_{n\in\mathbb{N}}$ be sequences in $\mathcal{S}_1$ and $\mathcal{S}_2$ (respectively) such that $\mu_1(A_n) < \infty$ and $\mu_2(B_n) < \infty$ for all $n \in \mathbb{N}$ and $\cup_n A_n = S_1$, $\cup_n B_n = S_2$. Define $\{C_n\}_{n\in\mathbb{N}}$ as an enumeration of the countable family $\{A_i \times B_j : i, j \in \mathbb{N}\}$ in $\mathcal{S}_1 \otimes \mathcal{S}_2$. Then $(\mu_1 \otimes \mu_2)(C_n) < \infty$ and all $n \in \mathbb{N}$ and $\cup_n C_n = S_1 \times S_2$. Therefore, $\mu_1 \otimes \mu_2$ is σ-finite. $\square$

Now that we know that product measures exist, we can state and prove the important theorem which, when applied to integrable functions bears the name of Fubini, and when applied to non-negative functions, of Tonelli. We state it for both cases simultaneously (i.e., on $\mathcal{L}^{0-1}$) in the case of a product of two measure spaces. An analogous theorem for finite products can be readily derived from it. When the variable or the underlying measure space of integration needs to be specified, we write $\int_S f(x)\,\mu(dx)$ for the Lebesgue integral $\int f\,d\mu$.

Theorem 5.5 (Fubini, Tonelli)**.** *Let $(S,\mathcal{S},\mu_S)$ and $(T,\mathcal{T},\mu_T)$ be two σ-finite measure spaces. For $f \in \mathcal{L}^{0-1}(S\times T)$ we have*

$$\begin{aligned}\int f\,d(\mu_S\otimes\mu_T) &= \int_S \left(\int_T f(x,y)\,\mu_T(dy)\right)\,\mu_S(dx)\\ &= \int_T \left(\int_S f(x,y)\,\mu_S(dx)\right)\,\mu_T(dy)\end{aligned} \tag{5.3}$$

Proof. All the hard work has already been done. We simply need to crank the Standard Machine. Let $\mathcal{H}$ denote the family of all functions in $\mathcal{L}^0_+(S\times T)$ with the property that (5.3) holds. Proposition 5.3 implies that $\mathcal{H}$ contains the indicators of all elements of $\mathcal{S}\otimes\mathcal{T}$. Linearity of all components of (5.3) implies that $\mathcal{H}$ contains all simple functions in $\mathcal{L}^0_+$, and the approximation theorem 3.10 implies that the whole $\mathcal{L}^0_+$ is in $\mathcal{H}$. Finally, the extension to $\mathcal{L}^{0-1}$ follows by additivity. □

Since f^- is always in $\mathcal{L}^{0-1}$, we have the following corollary

Corollary 5.6. *For $f\in\mathcal{L}^0(S\times T)$, we have*

$$f\in\mathcal{L}^{0-1}(S\times T) \text{ if and only if } \int_S\left(\int_T f^-(x,y)\,\mu_T(dy)\right)\,\mu_S(dx)<\infty.$$

Example 5.7. The assumption of σ-finiteness cannot be left out of the statement of Theorem 5.5. Indeed, let $(S,\mathcal{S},\mu)=([0,1],\mathcal{B}([0,1]),\lambda)$ and $(T,\mathcal{T},\nu)=([0,1],2^{[0,1]},\gamma)$, where γ is the counting measure on $2^{[0,1]}$, so that $(T,\mathcal{T},\nu)$ fails the σ-finite property. Define $f\in\mathcal{L}^0(S\times T)$ (why it is product-measurable?) by

$$f(x,y)=\begin{cases}1, & x=y,\\ 0, & x\neq y.\end{cases}$$

Then

$$\int_S f(x,y)\,\mu(dx)=\lambda(\{y\})=0,$$

and so

$$\int_T\int_S f(x,y)\,\mu(dx)\,\nu(dy)=\int_{[0,1]}0\,\gamma(dy)=0.$$

On the other hand,

$$\int_T f(x,y)\,\nu(dy) = \gamma(\{x\}) = 1,$$

and so

$$\int_S \int_T f(x,y)\,\nu(dy)\,\mu(dx) = \int_{[0,1]} 1\,\lambda(dx) = 1.$$

Example 5.8. The integrability of either f^+ or f^- for $f \in \mathcal{L}^0(S \times T)$ is (essentially) necessary for validity of Fubini's theorem, even if all iterated integrals exist. Here is what can go wrong. Let $(S, \mathcal{S}, \mu) = (T, \mathcal{T}, \nu) = (\mathbb{N}, 2^{\mathbb{N}}, \gamma)$, where γ is the counting measure. Define the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ by

$$f(n,m) = \begin{cases} 1, & m = n, \\ -1, & m = n+1, \\ 0, & \text{otherwise} \end{cases}$$

Then

$$\int_T f(n,m)\,\gamma(dm) = \sum_{m \in \mathbb{N}} f(n,m) = 0 + \cdots + 0 + 1 + (-1) + 0 + \cdots = 0,$$

and so

$$\int_S \int_T f(n,m)\,\gamma(dm)\,\gamma(dn) = 0.$$

On the other hand,

$$\int_S f(n,m)\,\gamma(dn) = \sum_{n \in \mathbb{N}} f(n,m) = \begin{cases} 1, & m = 1, \\ 0, & m > 1, \end{cases}$$

i.e.,

$$\int_S f(n,m)\,\gamma(dn) = \mathbf{1}_{\{m=1\}}.$$

Therefore,

$$\int_T \int_S f(n,m)\,\gamma(dn)\,\gamma(dm) = \int_T \mathbf{1}_{\{m=1\}}\,\gamma(dm) = 1.$$

If you think that using the counting measure is cheating, convince yourself that it is not hard to transfer this example to the setup where $(S, \mathcal{S}, \mu) = (T, \mathcal{T}, \nu) = ([0,1], \mathcal{B}([0,1]), \lambda)$.

The existence of the product measure gives us an easy access to the Lebesgue measure on higher-dimensional Euclidean spaces. Just as λ on $\mathbb{R}$ measures the "length" of sets, the Lebesgue measure on $\mathbb{R}^2$ will measure "area", the one on $\mathbb{R}^3$ "volume", etc. Its properties are collected in the following problem:

Problem 5.4. For $n \in \mathbb{N}$, show the following statements:

1. There exists a unique measure λ (note the notation overload) on $\mathcal{B}(\mathbb{R}^n)$ with the property that

$$\lambda\Big([a_1, b_1) \times \cdots \times [a_n, b_n)\Big) = (b_1 - a_1) \dots (b_n - a_n),$$

   for all $a_1 < b_1, \dots, a_n < b_n$ in $\mathbb{R}$.

2. The measure λ on $\mathbb{R}^n$ is invariant with respect to all isometries[4] of $\mathbb{R}^n$.
   Note: Feel free to use the following two facts without proof:
   (a) A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry if and only if there exists $x_0 \in \mathbb{R}^n$ and an orthogonal linear transformation $O : \mathbb{R}^n \to \mathbb{R}^n$ such that $f(x) = x_0 + Ox$.
   (b) Let O be an orthogonal linear transformation. Then R_1 and OR_1 have the same Lebesgue measure, where R_1 denotes the unit rectangle $[0, 1) \times \cdots \times [0, 1)$.

[4] An **isometry** of $\mathbb{R}^n$ is a map $f : \mathbb{R}^n \to \mathbb{R}^n$ with the property that $d(x, y) = d(f(x), f(y))$ for all $x, y \in \mathbb{R}^n$. It can be shown that the isometries of $\mathbb{R}^3$ are precisely translations, rotations, reflections and compositions thereof.

Note: probably the least painful way to prove this fact is by using the change-of-variable formula for the Riemann integral.

The Radon-Nikodym Theorem

We start the discussion of the Radon-Nikodym theorem with a simple observation:

Problem 5.5 (Integral as a measure)**.** For a function $f \in \mathcal{L}^0([0, \infty])$, we define the set-function $\nu : \mathcal{S} \to [0, \infty]$ by

$$\nu(A) = \int_A f \, d\mu. \tag{5.4}$$

1. Show that ν is a measure.
2. Show that $\mu(A) = 0$ implies $\nu(A) = 0$, for all $A \in \mathcal{S}$.
3. Show that the following two properties are equivalent
   - $\mu(A) = 0$ if and only if $\nu(A) = 0$, $A \in \mathcal{S}$, and
   - $f > 0$, a.e.

Definition 5.9 (Absolute continuity, etc.)**.** Let μ, ν be measures on the measurable space $(S, \mathcal{S})$. We say that

1. ν is **absolutely continuous** with respect to μ - denoted by $\nu \ll \mu$ - if $\nu(A) = 0$, whenever $\mu(A) = 0$, $A \in \mathcal{S}$.
2. μ and ν are **equivalent** if $\nu \ll \mu$ and $\mu \ll \nu$, i.e., if $\mu(A) = 0 \Leftrightarrow \nu(A) = 0$, for all $A \in \mathcal{S}$,

3. μ and ν are **(mutually) singular** - denoted by $\mu \perp \nu$ - if there exists $D \in \mathcal{S}$ such that $\mu(D) = 0$ and $\nu(D^c) = 0$.

Problem 5.6. Let μ and ν be measures with ν finite and $\nu \ll \mu$. Show that for each $\varepsilon > 0$ there exists $\delta > 0$ such that for each $A \in \mathcal{S}$, we have $\mu(A) \leq \delta \Rightarrow \nu(A) \leq \varepsilon$. Show that the assumption that ν is finite is necessary.

Problem 5.5 states that the prescription (5.4) defines a measure on $\mathcal{S}$ which is absolutely continuous with respect to μ. What is surprising is that the converse also holds under the assumption of σ-finiteness: all absolutely continuous measures on $\mathcal{S}$ are of that form. That statement (and more) is the topic of this section. Since there is more than one measure in circulation, we use the convention that *a.e.* always uses the notion of the null set as defined by the measure μ.

Theorem 5.10 (The Lebesgue decomposition)**.** *Let $(S, \mathcal{S})$ be a measurable space and let μ and ν be two σ-finite measures on $\mathcal{S}$. Then there exists a unique decomposition $\nu = \nu_a + \nu_s$, where*

1. $\nu_a \ll \mu$,
2. $\nu_s \perp \mu$.

Furthermore, there exists an a.e.-unique function $h \in \mathcal{L}^0_+$ such that

$$\nu_a(A) = \int_A h \, d\mu.$$

The proof is based on the following particular case of the Riesz representation theorem

Proposition 5.11. *Let $(S, \mathcal{S}, \mu)$ be a measure space, and let $L : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a linear map with the property that*

$$|Lf| \leq C||f||_{\mathcal{L}^2} \text{ for all } f \in \mathcal{L}^2$$

for some constant $C \geq 0$. Then, there exists an a.e.-unique element $g \in \mathcal{L}^2$ such that

$$Lf = \int fg \, d\mu, \text{ for all } f \in \mathcal{L}^2.$$

Proof. If $L = 0$, the statement clearly holds with $g = 0$, so we assume $Lh \neq 0$ for some $h \in \mathcal{L}^2$. The set

$$\operatorname{Nul} L = \{f \in \mathcal{L}^2 : Lf = 0\}$$

is a nonempty and closed (why?) subspace of $\mathcal{L}^2$ with $h \notin \operatorname{Nul} L$. Therefore, there exists a sequence $\{f_n\}_{n\in\mathbb{N}}$ in $\operatorname{Nul} L$ such that

$$a_n = ||f_n - h|| \to M = \inf_{f \in \operatorname{Nul} L} ||f - h|| > 0.$$

It is easy to check that

$$\begin{aligned} ||f_n - f_m||^2 &= -||f_n + f_m - 2h||^2 + 2||f_n - h||^2 + 2||f_m - h||^2 \\ &\leq -4M^2 + 2a_n^2 + 2a_m^2 \to 0 \text{ as } m, n \to \infty, \end{aligned}$$

and, hence, that $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}^2$. By completeness (Proposition 4.13) there exists $\hat{f} \in \mathcal{L}^2$ such that $f_n \to \hat{f}$ in $\mathcal{L}^2$. It follows immediately that $||\hat{g}|| = M$, where $\hat{g} = \hat{f} - h$. Therefore,

$$||\hat{g}||^2 \leq ||\hat{g} + \varepsilon f||^2, \text{ for all } \varepsilon > 0, f \in \operatorname{Nul} L,$$

and so

$$\varepsilon||f||^2 + 2\int f\hat{g} \geq 0, \text{ for all } \varepsilon > 0, f \in \operatorname{Nul} L,$$

which, in turn, immediately implies that[5]

$$\int f\hat{g}\, d\mu = 0, \text{ for all } f \in \operatorname{Nul} L.$$

In words, $\hat{g}$ is orthogonal to $\operatorname{Nul} L$. Starting from the identity

$$Lf = L\alpha h, \text{ where } \alpha = Lf/Lh,$$

we deduce that $f - \alpha h \in \operatorname{Nul} L$ and, so,

$$\int (f - \alpha h)\hat{g}\, d\mu = 0, \text{ i.e., } \int f\hat{g}\, d\mu = \tfrac{Lf}{Lh}\int h\hat{g}\, d\mu.$$

Therefore, noting that h cannot be orthogonal to $\hat{g}$, we get

$$Lf = \int fg\, d\mu \text{ where } g = \tfrac{Lh}{\int h\hat{g}\, d\mu}\hat{g}. \qquad \square$$

Proof of Theorem 5.10. Uniqueness: Suppose that $\nu_a^1 + \nu_s^1 = \nu = \nu_a^2 + \nu_s^2$ are two decompositions satisfying (1) and (2) in the statement. Let D^1 and D^2 be as in the definition of mutual singularity applied to the pairs μ, ν_s^1 and μ, ν_s^2, respectively. Set $D = D^1 \cup D^2$, and note that $\mu(D) = 0$ and $\nu_s^1(D^c) = \nu_s^2(D^c) = 0$. For any $A \in \mathcal{S}$, we have $\mu(A \cap D) = 0$ and so, thanks to absolute continuity,

$$\nu_a^1(A \cap D) = \nu_a^2(A \cap D) = 0$$

so that $\nu_s^1(A \cap D) = \nu_s^2(A \cap D) = \nu(A \cap D)$. By singularity, we have

$$\nu_s^1(A \cap D^c) = \nu_s^2(A \cap D^c) = 0,$$

and, consequently, $\nu_a^1(A \cap D^c) = \nu_a^2(A \cap D^c) = \nu(A \cap D^c)$. Finally,

$$\nu_a^1(A) = \nu_a^1(A \cap D) + \nu_a^1(A \cap D^c) = \nu_a^2(A \cap D) + \nu_a^2(A \cap D^c) = \nu^2(A),$$

and, similarly, $\nu_s^1 = \nu_s^2$.

[5] Compare this argument to Problem 4.9

To establish the uniqueness of the function f with the property that $\nu_a(A) = \int_A f\, d\mu$ for all $A \in \mathcal{S}$, we assume that there are two such functions, f_1 and f_2, say. Define the sequence $\{B_n\}_{n\in\mathbb{N}}$ by

$$B_n = \{f_1 \geq f_2\} \cap C_n,$$

where $\{C_n\}_{n\in\mathbb{N}}$ is a pairwise-disjoint sequence in $\mathcal{S}$ with the property that $\nu(C_n) < \infty$, for all $n \in \mathbb{N}$ and $\cup_n C_n = S$. Then, with $g_n = f_1 \mathbf{1}_{B_n} - f_2 \mathbf{1}_{B_n} \in \mathcal{L}^1_+$ we have

$$\int g_n\, d\mu = \int_{B_n} f_1\, d\mu - \int_{B_n} f_2\, d\mu = \nu_a(B_n) - \nu_a(B_n) = 0.$$

By Problem 3.9, we have $g_n = 0$, a.e., i.e., $f_1 = f_2$, a.e., on B_n, for all $n \in \mathbb{N}$, and so $f_1 = f_2$, a.e., on $\{f_1 \geq f_2\}$. A similar argument can be used to show that $f_1 = f_2$, a.e., on $\{f_1 < f_2\}$, as well.

Existence: Assume, first, that μ and ν are probability measures and set $\varphi = \mu + \nu$. We define

$$Lf = \int f\, d\nu \text{ for } f \in \mathcal{L}^2(\varphi),$$

and note that,

$$|Lf| \leq \int |f|\, d\nu \leq \int |f|\, d\varphi \leq ||f||_{\mathcal{L}^2(\varphi)},$$

Proposition 5.11 can be applied to yield the existence of $g \in \mathcal{L}^2(\varphi)$ such that $Lf = \int fg\, d\varphi$, for all $f \in \mathcal{L}^2(\varphi)$. Since $Lf \geq 0$, for $f \geq 0$, φ-a.e., we conclude (why?) that $g \geq 0$, φ-a.e. By approximation (via the monotone convergence theorem) we then have

$$\int F\, d\nu = \int Fg\, d\nu + \int Fg d\mu, \text{ for all } F \in \mathcal{L}^0_+(\varphi). \tag{5.5}$$

If we plug $F = \mathbf{1}_{\{g \geq 1+\varepsilon\}}$ in (5.5), we obtain

$$\nu(g \geq 1+\varepsilon) \geq (1+\varepsilon)\varphi(g \geq 1+\varepsilon),$$

and conclude that $g \leq 1$, φ-a.e. It also follows, by taking $F = \mathbf{1}_D$, where $D = \{g = 1\}$, that $\mu(D) = 0$. The function $h = \frac{g}{1-g}\mathbf{1}_{\{g<1\}}$ is nonnegative, measurable and finitely valued φ-a.e., so we can use $F = \frac{f}{1-g}\mathbf{1}_{D^c}$, for $f \in \mathcal{L}^0_+(\varphi)$ in (5.5) to obtain

$$\int f \mathbf{1}_{D^c}\, d\nu = \int fh\, d\mu.$$

It follows immediately that

$$\nu(A) = \int \mathbf{1}_A h\, d\mu + \nu(A \cap D),$$

which provides the required decomposition

$$\nu_a(A) = \int \mathbf{1}_A h\, d\mu,\ \nu_s(A) = \nu(A \cap D). \qquad \square$$

Corollary 5.12 (Radon-Nikodym)**.** *Let μ and ν be σ-finite measures on $(S, \mathcal{S})$ with $\nu \ll \mu$. Then there exists $f \in \mathcal{L}^0_+$ such that*

$$\nu(A) = \int_A f\, d\mu, \text{ for all } A \in \mathcal{S}. \tag{5.6}$$

For any other $g \in \mathcal{L}^0_+$ with the same property, we have $f = g$, a.e.

Any function f for which (5.6) holds is called the **Radon-Nikodym derivative** of ν with respect to μ and is denoted by $f = \frac{d\nu}{d\mu}$, a.e. [6]

[6] The Radon-Nikodym derivative $f = \frac{d\nu}{d\mu}$ is defined only up to a.e.-equivalence, and there is no canonical way of picking a representative defined for all $x \in S$. For that reason, we usually say that a function $f \in \mathcal{L}^0_+$ is a **version** of the Radon-Nikodym derivative of ν with respect to μ if (5.6) holds. Moreover, to stress the fact that we are talking about a whole class of functions instead of just one, we usually write

$$\frac{d\nu}{d\mu} \in \mathbb{L}^0_+ \text{ and not } \frac{d\nu}{d\mu} \in \mathcal{L}^0_+.$$

We often neglect this fact notationally, and write statements such as "*If $f \in \mathcal{L}^0_+$ and $f = \frac{d\nu}{d\mu}$ then ...*". What we really mean is that the statement holds regardless of the particular *representative* f of the Radon-Nikodym derivative we choose. Also, when we write $\frac{d\nu}{d\mu} = \frac{d\rho}{d\mu}$, we mean that they are equal as elements of $\mathbb{L}^0_+$, i.e., that there exists $f \in \mathcal{L}^0_+$, which is both a version of $\frac{d\nu}{d\mu}$ and a version of $\frac{d\rho}{d\mu}$.

Problem 5.7. Let μ, ν and ρ be σ-finite measures on $(S, \mathcal{S})$. Show that

1. If $\nu \ll \mu$ and $\rho \ll \mu$, then $\nu + \rho \ll \mu$ and
$$\frac{d\nu}{d\mu} + \frac{d\rho}{d\mu} = \frac{d(\nu+\rho)}{d\mu}.$$
2. If $\nu \ll \mu$ and $f \in \mathcal{L}^0_+$, then
$$\int f\, d\nu = \int g\, d\mu \text{ where } g = f\frac{d\nu}{d\mu}.$$
3. If $\nu \ll \mu$ and $\rho \ll \nu$, then $\rho \ll \mu$ and
$$\frac{d\nu}{d\mu}\frac{d\rho}{d\nu} = \frac{d\rho}{d\mu}.$$
Note: Make sure to pay attention to the fact that different measure give rise to different families of null sets, and, hence, to different notions of *almost everywhere*.
4. If $\mu \sim \nu$, then
$$\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1}.$$

Problem 5.8. Let $\mu_1, \mu_2, \nu_1, \nu_2$ be σ-finite measures with μ_1 and ν_1, as well as μ_2 and ν_2, defined on the same measurable space. If $\nu_1 \ll \mu_1$ and $\nu_2 \ll \mu_2$, show that $\nu_1 \otimes \nu_2 \ll \mu_1 \otimes \mu_2$.

Example 5.13. Just like in the statement of Fubini's theorem, the assumption of σ-finiteness cannot be omitted. Indeed, take $(S, \mathcal{S}) = ([0,1], \mathcal{B}([0,1]))$ and consider the Lebesgue measure λ and the counting measure γ on $(S, \mathcal{S})$. Clearly, $\lambda \ll \gamma$, but there is no $f \in \mathcal{L}^0_+$ such that $\lambda(A) = \int_A f\, d\gamma$. Indeed, suppose that such f exists and set $D_n = \{x \in S : f(x) > 1/n\}$, for $n \in \mathbb{N}$, so that $D_n \nearrow \{f > 0\} = \{f \neq 0\}$. Then

$$1 \geq \lambda(D_n) = \int_{D_n} f\, d\gamma \geq \int_{D_n} \tfrac{1}{n}\, d\gamma = \tfrac{1}{n} \# D_n,$$

and so $\#D_n \leq n$. Consequently, the set $\{f > 0\} = \cup_n D_n$ is countable. This leads to a contradiction since the Lebesgue measure does not "charge" countable sets, and so

$$1 = \lambda([0,1]) = \int f \, d\gamma = \int_{\{f>0\}} f \, d\gamma = \lambda(\{f > 0\}) = 0.$$

Additional Problems

Problem 5.9 (Area under the graph of a function)**.** For $f \in \mathcal{L}^0_+$, let $H = \{(x,r) \in S \times [0,\infty) : f(x) \geq r\}$ be the "region under the graph" of f. Show that $\int f \, d\mu = (\mu \otimes \lambda)(H)$.

Note: The equality in this problem is consistent with our intuition that the value of the integral $\int f \, d\mu$ corresponds to the "area under the graph of f".

Problem 5.10 (A layered representation)**.** Let ν be a measure on $\mathcal{B}([0,\infty))$ such that $N(u) = \nu([0,u)) < \infty$, for all $u \in \mathbb{R}$. Let $(S, \mathcal{S}, \mu)$ be a σ-finite measure space. For $f \in \mathcal{L}^0_+(S)$, show that

1. $\int N \circ f \, d\mu = \int_{[0,\infty)} \mu(\{f > u\}) \, \nu(du)$.
2. for $p > 0$, we have $\int f^p \, d\mu = p \int_{[0,\infty)} u^{p-1} \mu(\{f > u\}) \, \lambda(du)$, where λ is the Lebesgue measure.

Problem 5.11 (A useful integral)**.**

1. Show that $\int_0^\infty \left|\frac{\sin x}{x}\right| \, dx = \infty$.

Hint: Find a function laying below $\left|\frac{\sin x}{x}\right|$ which is easier to integrate.

2. For $a > 0$, let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} e^{-xy} \sin(x), & 0 \leq x \leq a, \; 0 \leq y, \\ 0, & \text{otherwise.} \end{cases}$$

   Show that $f \in \mathcal{L}^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$, where λ denotes the Lebesgue measure on $\mathbb{R}^2$.

3. Establish the equality

$$\int_0^a \frac{\sin x}{x} \, dx = \frac{\pi}{2} - \cos(a) \int_0^\infty \frac{e^{-ay}}{1+y^2} \, dy - \sin(a) \int_0^\infty \frac{y e^{-ay}}{1+y^2} \, dy.$$

4. Conclude that for $a > 0$, $\left| \int_0^a \frac{\sin(x)}{x} \, dx - \frac{\pi}{2} \right| \leq \frac{2}{a}$, so that

$$\lim_{a\to\infty} \int_0^a \frac{\sin(x)}{x} \, dx = \frac{\pi}{2}.$$

Problem 5.12 (The Cantor measure)**.** Let $(\{-1,1\}^{\mathbb{N}}, \mathcal{B}(\{-1,1\}^{\mathbb{N}}), \mu_C)$ be the coin-toss space. Define the mapping $f : \{-1,1\}^{\mathbb{N}} \to [0,1]$ by

$$f(\mathbf{s}) = \sum_{n\in\mathbb{N}} (1 + s_n) 3^{-n}, \text{ for } \mathbf{s} = (s_1, s_2, \dots).$$

Let δ be the push-forward of μ_C by the map f. It is called the **Cantor measure**.

1. Let d be the metric on $\{-1,1\}^{\mathbb{N}}$ (as given by the equation (1.1)). Show that for $\alpha = \log_3(2)$ and $\boldsymbol{s}^1, \boldsymbol{s}^2 \in \{-1,1\}^{\mathbb{N}}$, we have

$$d(\boldsymbol{s}^1, \boldsymbol{s}^2)^\alpha \leq \left| f(\boldsymbol{s}^2) - f(\boldsymbol{s}^1) \right| \leq 3d(\boldsymbol{s}^1, \boldsymbol{s}^2)^\alpha.$$

2. Show that δ is atom-free, i.e., that $\delta(\{x\}) = 0$, for all $x \in [0,1]$,

3. For a measure μ on the σ-algebra of Borel sets of a topological space X, the **support** of μ is collection of all $x \in X$ with the property that $\mu(O) > 0$ for each open set O with $x \in O$. Describe the support of δ. *Hint:* Guess what it is and prove that your guess is correct. Use the result in (1).

4. Prove that $\delta \perp \lambda$. *Note:* The Cantor measure is an example of a **singular measure**. It has no atoms, but is still singular with respect to the Lebesgue measure.

Problem 5.13 (Joint measurability)**.**

1. Give an example of a function $f : [0,1] \times [0,1] \to [0,1]$ such that $x \mapsto f(x,y)$ and $y \mapsto f(x,y)$ are $\mathcal{B}([0,1])$-measurable functions for each $y \in [0,1]$ and $x \in [0,1]$, respectively, but that f is not $\mathcal{B}([0,1] \times [0,1])$-measurable.

Hint: You can use the fact that there exists a subset of $[0,1]$ which is not Borel measurable.

2. Let $(S, \mathcal{S})$ be a measurable space. A function $f : S \times \mathbb{R} \to \mathbb{R}$ is called a **Caratheodory function** if
   - $x \mapsto f(x,y)$ is $\mathcal{S}$-measurable for each $y \in \mathbb{R}$, and
   - $y \mapsto f(x,y)$ is continuous for each $x \in \mathbb{R}$.

   Show that Caratheodory functions are $\mathcal{S} \otimes \mathcal{B}(\mathbb{R})$-measurable.

Hint: Express a Caratheodory function as limit of a sequence of the form $f_n = \sum_{k \in \mathbb{Z}} g_{n,k}(x) h_{n,k}(r)$, $n \in \mathbb{N}$.