CHAPTER 5

DECOMPOSITION OF MEASURES

Introduction

In this section a version of the fundamental theorem of calculus for Lebesgue integrals will be proved. Moreover, the concept of differentiating a measure with respect to another measure will be developped. A very important result in this chapter is the so called Radon-Nikodym Theorem.

5.1. Complex Measures

Let $(X, \mathcal{M})$ be a measurable space. Recall that if $A_n \subseteq X$, $n \in \mathbf{N}_+$, and $A_i \cap A_j = \phi$ if $i \neq j$, the sequence $(A_n)_{n \in \mathbf{N}_+}$ is called a disjoint denumerable collection. The collection is called a measurable partition of A if $A = \cup_{n=1}^{\infty} A_n$ and $A_n \in \mathcal{M}$ for every $n \in \mathbf{N}_+$.

A complex function μ on $\mathcal{M}$ is called a complex measure if

$$\mu(A) = \Sigma_{n=1}^{\infty} \mu(A_n)$$

for every $A \in \mathcal{M}$ and measurable partition $(A_n)_{n=1}^{\infty}$ of A. Note that $\mu(\phi) = 0$ if μ is a complex measure. A complex measure is said to be a real measure if it is a real function. The reader should note that a positive measure need not be a real measure since infinity is not a real number. If μ is a complex measure $\mu = \mu_{\mathrm{Re}} + i\mu_{\mathrm{Im}}$, where μ_{Re} =Re μ and μ_{Im} =Im μ are real measures.

If $(X, \mathcal{M}, \mu)$ is a positive measure and $f \in L^1(\mu)$ it follows that

$$\lambda(A) = \int_A f d\mu, \ A \in \mathcal{M}$$

is a real measure and we write $d\lambda = f d\mu$.

A function $\mu : \mathcal{M} \to [-\infty, \infty]$ is called a signed measure measure if

(a) $\mu : \mathcal{M} \to]-\infty, \infty]$ or $\mu : \mathcal{M} \to [-\infty, \infty[$
(b) $\mu(\phi) = 0$
and
(c) for every $A \in \mathcal{M}$ and measurable partition $(A_n)_{n=1}^{\infty}$ of A,

$$\mu(A) = \Sigma_{n=1}^{\infty} \mu(A_n)$$

where the latter sum converges absolutely if $\mu(A) \in \mathbf{R}$.

Here $-\infty - \infty = -\infty$ and $-\infty + x = -\infty$ if $x \in \mathbf{R}$. The sum of a positive measure and a real measure and the difference of a real measure and a positive measure are examples of signed measures and it can be proved that there are no other signed measures (see Folland $[F]$). Below we concentrate on positive, real, and complex measures and will not say more about signed measures here.

Suppose μ is a complex measure on $\mathcal{M}$ and define for every $A \in \mathcal{M}$

$$\mid \mu \mid (A) = \sup \Sigma_{n=1}^{\infty} \mid \mu(A_n) \mid,$$

where the supremum is taken over all measurable partitions $(A_n)_{n=1}^{\infty}$ of A. Note that $\mid \mu \mid (\phi) = 0$ and

$$\mid \mu \mid (A) \geq \mid \mu(B) \mid \text{ if } A, B \in \mathcal{M} \text{ and } A \supseteq B.$$

The set function $\mid \mu \mid$ is called the total variation of μ or the total variation measure of μ. It turns out that $\mid \mu \mid$ is a positive measure. In fact, as will shortly be seen, $\mid \mu \mid$ is a finite positive measure.

Theorem 5.1.1. *The total variation $\mid \mu \mid$ of a complex measure is a positive measure.*

PROOF. Let $(A_n)_{n=1}^{\infty}$ be a measurable partition of A.

For each n, suppose $a_n <\mid \mu \mid (A_n)$ and let $(E_{kn})_{k=1}^{\infty}$ be a measurable partition of A_n such that

$$a_n < \Sigma_{k=1}^{\infty} \mid \mu(E_{kn}) \mid .$$

Since $(E_{kn})_{k,n=1}^{\infty}$ is a partition of A it follows that

$$\Sigma_{n=1}^{\infty} a_n < \Sigma_{k,n=1}^{\infty} \mid \mu(E_{kn}) \mid \leq \mid \mu \mid (A).$$

Thus

$$\Sigma_{n=1}^{\infty} \mid \mu \mid (A_n) \leq \mid \mu \mid (A).$$

To prove the opposite inequality, let $(E_k)_{k=1}^{\infty}$ be a measurable partition of A. Then, since $(A_n \cap E_k)_{n=1}^{\infty}$ is a measurable partition of E_k and $(A_n \cap E_k)_{k=1}^{\infty}$ a measurable partition of A_n,

$$\Sigma_{k=1}^{\infty} \mid \mu(E_k) \mid = \Sigma_{k=1}^{\infty} \mid \Sigma_{n=1}^{\infty} \mu(A_n \cap E_k) \mid$$

$$\leq \Sigma_{k,n=1}^{\infty} \mid \mu(A_n \cap E_k) \mid \leq \Sigma_{n=1}^{\infty} \mid \mu \mid (A_n)$$

and we get

$$\mid \mu \mid (A) \leq \Sigma_{n=1}^{\infty} \mid \mu \mid (A_n).$$

Thus

$$\mid \mu \mid (A) = \Sigma_{n=1}^{\infty} \mid \mu \mid (A_n).$$

Since $\mid \mu \mid (\phi) = 0$, the theorem is proved.

Theorem 5.1.2. *The total variation $\mid \mu \mid$ of a complex measure μ is a finite positive measure.*

PROOF. Since

$$\mid \mu \mid \leq \mid \mu_{\mathrm{Re}} \mid + \mid \mu_{\mathrm{Im}} \mid$$

there is no loss of generality to assume that μ is a real measure.

Suppose $\mid \mu \mid (E) = \infty$ for some $E \in \mathcal{M}$. We first prove that there exist disjoint sets $A, B \in \mathcal{M}$ such that

$$A \cup B = E$$

and

$$| \mu(A) |> 1 \text{ and } | \mu | (B) = \infty.$$

To this end let $c = 2(1+ | \mu(E) |)$ and let $(E_k)_{k=1}^{\infty}$ be a measurable partition of E such that

$$\Sigma_{k=1}^{n} | \mu(E_k) |> c$$

for some sufficiently large n. There exists a subset N of $\{1, ..., n\}$ such that

$$| \Sigma_{k\in N}\mu(E_k) |> \frac{c}{2}.$$

Set $A = \cup_{k\in N} E_k$ and $B = E \setminus A$. Then $| \mu(A) |> \frac{c}{2} \geq 1$ and

$$| \mu(B) |=| \mu(E) - \mu(A) |$$

$$\geq| \mu(A) | - | \mu(E) |> \frac{c}{2} - | \mu(E) |= 1.$$

Since $\infty =| \mu | (E) =| \mu | (A)+ | \mu | (B)$ we have $| \mu | (A) = \infty$ or $| \mu | (B) = \infty$. If $| \mu | (B) < \infty$ we interchange A and B and have $| \mu(A) |> 1$ and $| \mu | (B) = \infty$.

Suppose $| \mu | (X) = \infty$. Set $E_0 = X$ and choose disjoint sets $A_0, B_0 \in \mathcal{M}$ such that

$$A_0 \cup B_0 = E_0$$

and

$$| \mu(A_0) |> 1 \text{ and } | \mu | (B_0) = \infty.$$

Set $E_1 = B_0$ and choose disjoint sets $A_1, B_1 \in \mathcal{M}$ such that

$$A_1 \cup B_1 = E_1$$

and

$$| \mu(A_1) |> 1 \text{ and } | \mu | (B_1) = \infty.$$

By induction, we find a measurable partition $(A_n)_{n=0}^{\infty}$ of the set $A =_{def} \cup_{n=0}^{\infty} A_n$ such that $| \mu(A_n) |> 1$ for every n. Now, since μ is a complex measure,

$$\mu(A) = \Sigma_{n=0}^{\infty}\mu(A_n).$$

But this series cannot converge, since the general term does not tend to zero as $n \to \infty$. This contradiction shows that $| \mu |$ is a finite positive measure.

If μ is a real measure we define

$$\mu^+ = \frac{1}{2}(|\ \mu\ | + \mu)$$

and

$$\mu^- = \frac{1}{2}(|\ \mu\ | - \mu).$$

The measures μ^+ and μ^- are finite positive measures and are called the positive and negative variations of μ, respectively . The representation

$$\mu = \mu^+ - \mu^-$$

is called the Jordan decomposition of μ.

Exercises

1. Suppose $(X, \mathcal{M}, \mu)$ is a positive measure space and $d\lambda = fd\mu$, where $f \in L^1(\mu)$. Prove that $d \mid \lambda \mid = \mid f \mid d\mu$.

2. Suppose λ, μ, and ν are real measures defined on the same σ-algebra and $\lambda \leq \mu$ and $\lambda \leq \nu$. Prove that

$$\lambda \leq \min(\mu, \nu)$$

where

$$\min(\mu, \nu) = \frac{1}{2}(\mu + \nu - \mid \mu - \nu \mid).$$

3. Suppose $\mu : \mathcal{M} \to \mathbf{C}$ is a complex measure and $f, g : X \to \mathbf{R}$ measurable functions. Show that

$$\mid \mu(f \in A) - \mu(g \in A) \mid \leq \mid \mu \mid (f \neq g)$$

for every $A \in \mathcal{R}$.

5.2. The Lebesque Decomposition and the Radon-Nikodym Theorem

Let μ be a positive measure on $\mathcal{M}$ and λ a positive or complex measure on $\mathcal{M}$. The measure λ is said to be absolutely continuous with respect to μ (abbreviated $\lambda << \mu$) if $\lambda(A) = 0$ for every $A \in \mathcal{M}$ for which $\mu(A) = 0$. If we define

$$\mathcal{Z}_\lambda = \{A \in \mathcal{M};\ \lambda(A) = 0\}$$

it follows that $\lambda << \mu$ if and only if

$$\mathcal{Z}_\mu \subseteq \mathcal{Z}_\lambda.$$

For example, $\gamma_n << v_n$ and $v_n << \gamma_n$.

The measure λ is said to be concentrated on $E \in \mathcal{M}$ if $\lambda = \lambda^E$, where $\lambda^E(A) =_{def} \lambda(E \cap A)$ for every $A \in \mathcal{M}$. This is equivalent to the hypothesis that $A \in \mathcal{Z}_\lambda$ if $A \in \mathcal{M}$ and $A \cap E = \phi$. Thus if $E_1, E_2 \in \mathcal{M}$, where $E_1 \subseteq E_2$, and λ is concentrated on E_1, then λ is concentrated on E_2. Moreover, if $E_1, E_2 \in \mathcal{M}$ and λ is concentrated on both E_1 and E_2, then λ is concentrated on $E_1 \cap E_2$. Two measures λ_1 and λ_2 are said to be mutually singular (abbreviated $\lambda_1 \perp \lambda_2$) if there exist disjoint measurable sets E_1 and E_2 such that λ_1 is concentrated on E_1 and λ_2 is concentrated on E_2.

Theorem 5.2.1. *Let μ be a positive measure and λ, λ_1, and λ_2 complex measures.*

(i) *If $\lambda_1 << \mu$ and $\lambda_2 << \mu$, then $(\alpha_1\lambda_1 + \alpha_2\lambda_2) << \mu$ for all complex numbers α_1 and α_2.*

(ii) *If $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$, then $(\alpha_1\lambda_1 + \alpha_2\lambda_2) \perp \mu$ for all complex numbers α_1 and α_2.*

(iii) *If $\lambda << \mu$ and $\lambda \perp \mu$, then $\lambda = 0$.*

(iv) *If $\lambda << \mu$, then $\mid \lambda \mid << \mu$.*

PROOF. The properties (i) and (ii) are simple to prove and are left as exercises.

To prove (iii) suppose $E \in \mathcal{M}$ is a μ-null set and $\lambda = \lambda^E$. If $A \in \mathcal{M}$, then $\lambda(A) = \lambda(A \cap E)$ and $A \cap E$ is a μ-null set. Since $\lambda << \mu$ it follows that $A \cap E \in Z_\lambda$ and, hence, $\lambda(A) = \lambda(A \cap E) = 0$. This proves (iii)

To prove (iv) suppose $A \in \mathcal{M}$ and $\mu(A) = 0$. If $(A_n)_{n=1}^\infty$ is measurable partition of A, then $\mu(A_n) = 0$ for every n. Since $\lambda << \mu$, $\lambda(A_n) = 0$ for every n and we conclude that $\mid \lambda \mid (A) = 0$. This proves (vi).

Theorem 5.2.2. *Let μ be a positive measure on $\mathcal{M}$ and λ a complex measure on $\mathcal{M}$. Then the following conditions are equivalent:*

(a) $\lambda << \mu$.

(b) *To every $\varepsilon > 0$ there corresponds a $\delta > 0$ such that $\mid \lambda(E) \mid < \varepsilon$ for all $E \in \mathcal{M}$ with $\mu(E) < \delta$.*

If λ is a positive measure, the implication $(a) \Rightarrow (b)$ in Theorem 5.2.2 is, in general, wrong. To see this take $\mu = \gamma_1$ and $\lambda = v_1$. Then $\lambda << \mu$ and if we choose $A_n = [n, \infty[$, $n \in \mathbf{N}_+$, then $\mu(A_n) \to 0$ as $n \to \infty$ but $\lambda(A_n) = \infty$ for each n.

PROOF. (a)$\Rightarrow$(b). If (b) is wrong there exist an $\varepsilon > 0$ and sets $E_n \in \mathcal{M}$, $n \in \mathbf{N}_+$, such that $\mid \lambda(E_n) \mid \geq \varepsilon$ and $\mu(E_n) < 2^{-n}$. Set

$$A_n = \cup_{k=n}^\infty E_k \text{ and } A = \cap_{n=1}^\infty A_n.$$

Since $A_n \supseteq A_{n+1} \supseteq A$ and $\mu(A_n) < 2^{-n+1}$, it follows that $\mu(A) = 0$ and using that $\mid \lambda \mid (A_n) \geq \mid \lambda(E_n) \mid$, Theorem 1.1.2 (f) implies that

$$\mid \lambda \mid (A) = \lim_{n \to \infty} \mid \lambda \mid (A_n) \geq \varepsilon.$$

This contradicts that $\mid \lambda \mid << \mu$.

(b)$\Rightarrow$(a). If $E \in \mathcal{M}$ and $\mu(E) = 0$ then to each $\varepsilon > 0$, $\mid \lambda(E) \mid < \varepsilon$, and we conclude that $\lambda(E) = 0$. The theorem is proved.

Theorem 5.2.3. *Let μ be a σ-finite positive measure and λ a real measure on $\mathcal{M}$.*

(a) **(The Lebesgue Decomposition of λ)** *There exists a unique pair of real measures λ_a and λ_s on$\mathcal{M}$ such that*

$$\lambda = \lambda_a + \lambda_s, \ \lambda_a << \mu, \text{ and } \lambda_s \perp \mu.$$

If λ is a finite positive measure, λ_a and λ_s are finite positive measures.

(b) **(The Radon-Nikodym Theorem)** *There exits a unique $g \in L^1(\mu)$ such that*

$$d\lambda_a = g d\mu.$$

If λ is a finite positive measure, $g \geq 0$ a.e. $[\mu]$.

The proof of Theorem 5.2.3 is based on the following

Lemma 5.2.1. *Let (X, M, μ) be a finite positive measure space and suppose $f \in L^1(\mu)$.*

(a) *If $a \in \mathbf{R}$ and*

$$\int_E f d\mu \leq a\mu(E), \text{ all } E \in \mathcal{M}$$

then $f \leq a$ a.e. $[\mu]$.

(b) *If $b \in \mathbf{R}$ and*

$$\int_E f d\mu \geq b\mu(E), \text{ all } E \in \mathcal{M}$$

then $f \geq b$ a.e. $[\mu]$.

PROOF. (a) Set $g = f - a$ so that

$$\int_E g d\mu \leq 0, \ \textit{all } E \in \mathcal{M}.$$

Now choose $E = \{g > 0\}$ to obtain

$$0 \geq \int_E g d\mu = \int_X \chi_E g d\mu \geq 0$$

as $\chi_E g \geq 0$ a.e. $[\mu]$. But then Example 2.1.2 yields $\chi_E g = 0$ a.e. $[\mu]$ and we get $E \in Z_\mu$. Thus $g \leq 0$ a.e. $[\mu]$ or $f \leq a$ a.e. $[\mu]$.

Part (b) follows in a similar way as Part (a) and the proof is omitted here.

PROOF. Uniqueness: (a) Suppose $\lambda_a^{(k)}$ and $\lambda_s^{(k)}$ are real measures on $\mathcal{M}$ such that

$$\lambda = \lambda_a^{(k)} + \lambda_s^{(k)},\ \lambda_a^{(k)} << \mu, \text{ and } \lambda_s^{(k)} \perp \mu$$

for $k = 1, 2$. Then

$$\lambda_a^{(1)} - \lambda_a^{(2)} = \lambda_s^{(2)} - \lambda_s^{(1)}$$

and

$$\lambda_a^{(1)} - \lambda_a^{(2)} << \mu \text{ and } \lambda_a^{(1)} - \lambda_a^{(2)} \perp \mu.$$

Thus by applying Theorem 5.2.1, $\lambda_a^{(1)} - \lambda_a^{(2)} = 0$ and $\lambda_a^{(1)} = \lambda_a^{(2)}$. From this we conclude that $\lambda_s^{(1)} = \lambda_s^{(2)}$.

(b) Suppose $g_k \in L^1(\mu), k = 1, 2$, and

$$d\lambda_a = g_1 d\mu = g_2 d\mu.$$

Then $hd\mu = 0$ where $h = g_1 - g_2$. But then

$$\int_{\{h>0\}} hd\mu = 0$$

and it follows that $h \leq 0$ a.e. $[\mu]$. In a similar way we prove that $h \geq 0$ a.e. $[\mu]$. Thus $h = 0$ in $L^1(\mu)$, that is $g_1 = g_2$ in $L^1(\mu)$.

Existence: The beautiful proof that follows is due to von Neumann.

First suppose that μ and λ are finite positive measures and set $\nu = \lambda + \mu$. Clearly, $L^1(\lambda) \supseteq L^1(\nu) \supseteq L^2(\nu)$. Moreover, if $f : X \to \mathbf{R}$ is measurable

$$\int_X |\, f \,|\, d\lambda \leq \int_X |\, f \,|\, d\nu \leq \sqrt{\int_X f^2 d\nu} \sqrt{\nu(X)}$$

and from this we conclude that the map

$$f \to \int_X f d\lambda$$

is a continuous linear functional on $L^2(\nu)$. Therefore, in view of Theorem 4.2.2, there exists a $g \in L^2(\nu)$ such that

$$\int_X f d\lambda = \int_X fg d\nu \text{ all } f \in L^2(\nu).$$

Suppose $E \in \mathcal{M}$ and put $f = \chi_E$ to obtain

$$0 \leq \lambda(E) = \int_E g d\nu$$

and, since $\nu \geq \lambda$,

$$0 \leq \int_E g d\nu \leq \nu(E).$$

But then Lemma 5.2.1 implies that $0 \leq g \leq 1$ a.e. $[\nu]$. Therefore, without loss of generality we can assume that $0 \leq g(x) \leq 1$ for all $x \in X$ and, in addition, as above

$$\int_X f d\lambda = \int_X fg d\nu \text{ all } f \in L^2(\nu)$$

that is

$$\int_X f(1-g) d\lambda = \int_X fg d\mu \text{ all } f \in L^2(\nu).$$

Put $A = \{0 \leq g < 1\}$, $S = \{g = 1\}$, $\lambda_a = \lambda^A$, and $\lambda_s = \lambda^S$. Note that $\lambda = \lambda^A + \lambda^S$. The choice $f = \chi_S$ gives $\mu(S) = 0$ and hence $\lambda_s \perp \mu$. Moreover, the choice

$$f = (1 + ... + g^n)\chi_E$$

where $E \in \mathcal{M}$, gives

$$\int_E (1 - g^{n+1}) d\lambda = \int_E (1 + ... + g^n) g d\mu.$$

By letting $n \to \infty$ and using monotone convergence

$$\lambda(E \cap A) = \int_E h d\mu.$$

where

$$h = \lim_{n \to \infty} (1 + ... + g^n) g.$$

Since h is non-negative and

$$\lambda(A) = \int_X h d\mu$$

it follows that $h \in L^1(\mu)$. Moreover, the construction above shows that $\lambda = \lambda_a + \lambda_s$.

In the next step we assume that μ is a σ-finite positive measure and λ a finite positive measure. Let $(X_n)_{n=1}^{\infty}$ be a measurable partition of X such that $\mu(X_n) < \infty$ for every n. Let n be fixed and apply Part (a) to the pair μ^{X_n} and λ^{X_n} to obtain finite positive measures $(\lambda^{X_n})_a$ and $(\lambda^{X_n})_s$ such that

$$\lambda^{X_n} = (\lambda^{X_n})_a + (\lambda^{X_n})_s,\ (\lambda^{X_n})_a << \mu^{X_n},\ \text{and } (\lambda^{X_n})_s \perp \mu^{X_n}$$

and

$$d(\lambda^{X_n})_a = h_n d\mu^{X_n} \text{ (or } (\lambda^{X_n})_a = h_n \mu^{X_n})$$

where $0 \leq h_n \in L^1(\mu^{X_n})$. Without loss of generality we can assume that $h_n = 0$ off X_n and that $(\lambda^{X_n})_s$ is concentrated on $A_n \subseteq X_n$ where $A_n \in \mathcal{Z}_\mu$. In particular, $(\lambda^{X_n})_a = h_n\mu$. Now

$$\lambda = h\mu + \Sigma_{n=1}^{\infty}(\lambda^{X_n})_s$$

where

$$h = \Sigma_{n=1}^{\infty} h_n$$

and

$$\int_X h d\mu \leq \lambda(X) < \infty.$$

Thus $h \in L^1(\mu)$. Moreover, $\lambda_s =_{def} \Sigma_{n=1}^{\infty}(\lambda^{X_n})_s$ is concentrated on $\cup_{n=1}^{\infty} A_n \in \mathcal{Z}_\mu$. Hence $\lambda_s \perp \mu$.

Finally if λ is a real measure we apply what we have already proved to the positive and negative variations of λ and we are done.

Example 5.2.1. Let λ be Lebesgue measure in the unit interval and μ the counting measure in the unit interval restricted to the class of all Lebesgue measurable subsets of the unit interval. Clearly, $\lambda << \mu$. Suppose there is an

$h \in L^1(\mu)$ such that $d\lambda = h d\mu$. We can assume that $h \geq 0$ and the Markov inequality implies that the set $\{h \geq \varepsilon\}$ is finite for every $\varepsilon > 0$. But then

$$\lambda(h \in]0,1]) = \lim_{n\to\infty} \lambda(h \geq 2^{-n}) = 0$$

and it follows that $1 = \lambda(h = 0) = \int_{\{h=0\}} h d\mu = 0$, which is a contradiction.

Corollary 5.2.1. *Suppose μ is a real measure. Then there exists*

$$h \in L^1(\mid \mu \mid)$$

such that $\mid h(x) \mid = 1$ for all $x \in X$ and

$$d\mu = hd \mid \mu \mid .$$

PROOF. Since $\mid \mu(A) \mid \leq \mid \mu \mid (A)$ for every $A \in \mathcal{M}$, the Radon-Nikodym Theorem implies that $d\mu = hd \mid \mu \mid$ for an appropriate $h \in L^1(\mid \mu \mid)$. But then $d \mid \mu \mid = \mid h \mid d \mid \mu \mid$ (see Exercise 1 in Chapter 5.1). Thus

$$\mid \mu \mid (E) = \int_E \mid h \mid d \mid \mu \mid, \text{ all } E \in \mathcal{M}$$

and Lemma 5.2.1 yields $h = 1$ a.e. $[\mid \mu \mid]$. From this the theorem follows at once.

Theorem 5.2.4. (Hahn's Decomposition Theorem) *Suppose μ is a real measure. There exists an $A \in \mathcal{M}$ such that*

$$\mu^+ = \mu^A \text{ and } \mu^- = -\mu^{A^c}.$$

PROOF. Let $d\mu = hd \mid \mu \mid$ where $\mid h \mid = 1$. Note that $hd\mu = d \mid \mu \mid$. Set $A = \{h = 1\}$. Then

$$d\mu^+ = \frac{1}{2}(d \mid \mu \mid + d\mu) = \frac{1}{2}(h+1)d\mu = \chi_A d\mu$$

and

$$d\mu^- = d\mu^+ - d\mu = (\chi_A - 1)d\mu = -\chi_{A^c} d\mu.$$

The theorem is proved.

If a real measure λ is absolutely continuous with respect to a σ-finite positive measure μ, the Radon-Nikodym Theorem says that $d\lambda = f d\mu$ for an approprite $f \in L^1(\mu)$. We sometimes write

$$f = \frac{d\lambda}{d\mu}$$

and call f the Radon-Nikodym derivate of λ with respect to μ.

Exercises

1. Suppose μ and $\nu_n, n \in$ **N**, are positive measures defined on the same σ-algebra and set $\theta = \Sigma_{n=0}^{\infty} \nu_n$. Prove that
   a) $\theta \perp \mu$ if $\nu_n \perp \mu$, all $n \in \mathbf{N}$.
   b) $\theta << \mu$ if $\nu_n << \mu$, all $n \in \mathbf{N}$.

2. Suppose μ is a real measure and $\mu = \lambda_1 - \lambda_2$, where λ_1 and λ_2 are finite positive measures. Prove that $\lambda_1 \geq \mu^+$ and $\lambda_2 \geq \mu^-$.

3. Let λ_1 and λ_2 be mutually singular complex measures on the same σ-algebra. Show that $| \lambda_1 | \perp | \lambda_2 |$.

4. Let $(X, \mathcal{M}, \mu)$ be a σ-finite positive measure space and suppose λ and τ are two probability measures defined on the σ-algebra $\mathcal{M}$ such that $\lambda << \mu$ and $\tau << \mu$. Prove that

$$\sup_{A \in \mathcal{M}} | \lambda(A) - \tau(A) | = \frac{1}{2} \int_X | \frac{d\lambda}{d\mu} - \frac{d\tau}{d\mu} | \, d\mu.$$

5.3. The Wiener Maximal Theorem and the Lebesgue Differentiation Theorem

We say that a Lebesgue measurable function f in $\mathbf{R}^n$ is locally Lebesgue integrable and belongs to the class $L^1_{loc}(m_n)$ if $f\chi_K \in L^1(m_n)$ for each compact subset K of $\mathbf{R}^n$. In a similar way $f \in L^1_{loc}(v_n)$ if f is a Borel function such that $f\chi_K \in L^1(v_n)$ for each compact subset K of $\mathbf{R}^n$. If $f \in L^1_{loc}(m_n)$, we define the average $A_r f(x)$ of f on the open ball $B(x,r)$ as

$$A_r f(x) = \frac{1}{m_n(B(x,r))} \int_{B(x,r)} f(y)dy.$$

It follows from dominated convergence that the map $(x,r) \to A_r f(x)$ of $\mathbf{R}^n \times]0,\infty[$ into $\mathbf{R}$ is continuous. The Hardy-Littlewood maximal function f^* is, by definition, $f^* = \sup_{r>0} A_r \mid f \mid$ or, stated more explicitly,

$$f^*(x) = \sup_{r>0} \frac{1}{m_n(B(x,r))} \int_{B(x,r)} \mid f(y) \mid dy, \ x \in \mathbf{R}^n.$$

The function $f^* : (\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n)) \to ([0,\infty], \mathcal{R}_{0,\infty})$ is measurable since

$$f^* = \sup_{\substack{r>0 \\ r\in\mathbf{Q}}} A_r \mid f \mid .$$

Theorem 5.3.1. (Wiener's Maximal Theorem) *There exists a positive constant $C = C_n < \infty$ such that for all $f \in L^1(m_n)$,*

$$m_n(f^* > \alpha) \leq \frac{C}{\alpha} \parallel f \parallel_1 \text{ if } \alpha > 0.$$

The proof of the Wiener Maximal Theorem is based on the following remarkable result.

Lemma 5.3.1. *Let $\mathcal{C}$ be a collection of open balls in $\mathbf{R}^n$ and set $V = \cup_{B\in\mathcal{C}} B$. If $c < m_n(V)$ there exist pairwise disjoint $B_1, ..., B_k \in \mathcal{C}$ such that*

$$\Sigma_{i=1}^k m_n(B_i) > 3^{-n} c.$$

PROOF. Let $K \subseteq V$ be compact with $m_n(K) > c$, and suppose $A_1, ..., A_p \in \mathcal{C}$ cover K. Let B_1 be the largest of the $A_i's$ (that is, B_1 has maximal radius), let B_2 be the largest of the $A_i's$ which are disjoint from B_1, let B_3 be the largest of the $A_i's$ which are disjoint from $B_1 \cup B_2$, and so on until the process stops after k steps. If $B_i = B(x_i, r_i)$ put $B_i^* = B(x_i, 3r_i)$. Then $\cup_{i=1}^k B_i^* \supseteq K$ and

$$c < \Sigma_{i=1}^k m_n(B_i^*) = 3^n \Sigma_{i=1}^k m_n(B_i).$$

The lemma is proved.

PROOF OF THEOREM 5.3.1. Set

$$E_\alpha = \{f^* > \alpha\}.$$

For each $x \in E_\alpha$ choose an $r_x > 0$ such that $A_{r_x} \mid f \mid (x) > \alpha$. If $c < m_n(E_\alpha)$, by Lemma 5.3.1 there exist $x_1, ..., x_k \in E_\alpha$ such that the balls $B_i = B(x_i, r_{x_i})$, $i = 1, ..., k$, are mutually disjoint and

$$\Sigma_{i=1}^k m_n(B_i) > 3^{-n} c.$$

But then

$$c < 3^n \Sigma_{i=1}^k m_n(B_i) < \frac{3^n}{\alpha} \Sigma_{i=1}^k \int_{B_i} \mid f(y) \mid dy \leq \frac{3^n}{\alpha} \int_{\mathbf{R}^n} \mid f(y) \mid dy.$$

The theorem is proved.

Theorem 5.3.2. *If* $f \in L^1_{loc}(m_n)$,

$$\lim_{r \to 0} \frac{1}{m_n(B(x,r))} \int_{B(x,r)} f(y) dy = f(x) \text{ a.e. } [m_n].$$

PROOF. Clearly, there is no loss of generality to assume that $f \in L^1(m_n)$. Suppose $g \in C_c(\mathbf{R}^n) =_{def} \{f \in C(\mathbf{R}^n);\ f(x) = 0 \text{ if } \mid x \mid \text{ large enough}\}$. Then

$$\lim_{r \to 0} A_r g(x) = g(x) \text{ all } x \in \mathbf{R}^n.$$

Since $A_r f - f = A_r(f-g) - (f-g) + A_r g - g$,

$$\overline{\lim_{r\to 0}} \mid A_r f - f \mid \leq (f-g)^* + \mid f-g \mid .$$

Now, for fixed $\alpha > 0$,

$$m_n(\overline{\lim_{r\to 0}} \mid A_r f - f \mid > \alpha)$$

$$\leq m_n((f-g)^* > \frac{\alpha}{2}) + m_n(\mid f-g \mid > \frac{\alpha}{2})$$

and the Wiener Maximal Theorem and the Markov Inequality give

$$m_n(\overline{\lim_{r\to 0}} \mid A_r f - f \mid > \alpha)$$

$$\leq (\frac{2C}{\alpha} + \frac{2}{\alpha}) \parallel f - g \parallel_1 .$$

Remembering that $C_c(\mathbf{R}^n)$ is dense in $L^1(m_n)$, the theorem follows at once.

If $f \in L^1_{loc}(m_n)$ we define the so called Lebesgue set L_f to be

$$L_f = \left\{ x;\ \lim_{r\to 0} \frac{1}{m_n(B(x,r))} \int_{B(x,r)} \mid f(y) - f(x) \mid dy = 0 \right\}.$$

Note that if q is real and

$$E_q = \left\{ x;\ \lim_{r\to 0} \frac{1}{m_n(B(x,r))} \int_{B(x,r)} \mid f(y) - q \mid dy = \mid f(x) - q \mid \right\}$$

then $m_n(\cup_{q\in\mathbf{Q}} E_q^c) = 0$. If $x \in \cap_{q\in\mathbf{Q}} E_q$,

$$\overline{\lim_{r\to 0}} \frac{1}{m_n(B(x,r))} \int_{B(x,r)} \mid f(y) - f(x) \mid dy \leq 2 \mid f(x) - q \mid$$

for all rational numbers q and it follows that $m_n(L_f^c) = 0$.

A family $\mathcal{E}_{x,\alpha} = (E_{x,r})_{r>0}$ of Borel sets in $\mathbf{R}^n$ is said to shrink nicely to a point x in $\mathbf{R}^n$ if $E_{x,r} \subseteq B(x,r)$ for each r and there is a positive constant α, independent of r, such that $m_n(E_{x,r}) \geq \alpha m_n(B(x,r))$.

Theorem 5.3.3. (The Lebesgue Differentiation Theorem) *Suppose* $f \in L^1_{loc}(m_n)$ *and* $x \in L_f$. *Then*

$$\lim_{r \to 0} \frac{1}{m_n(E_{x,r})} \int_{E_{x,r}} \mid f(y) - f(x) \mid dy = 0$$

and

$$\lim_{r \to 0} \frac{1}{m_n(E_{x,r})} \int_{E_{x,r}} f(y)dy = f(x).$$

PROOF. The result follows from the inequality

$$\frac{1}{m_n(E_{x,r})} \int_{E_{x,r}} \mid f(y) - f(x) \mid dy \leq \frac{1}{\alpha m_n(B(x,r))} \int_{B(x,r)} \mid f(y) - f(x) \mid dy.$$

Theorem 5.3.4. *Suppose* λ *is a real or positive measure on* $\mathcal{R}_n$ *and suppose* $\lambda \perp v_n$. *If* λ *is a positive measure it is assumed that* $\lambda(K) < \infty$ *for every compact subset of* $\mathbf{R}^n$. *Then*

$$\lim_{r \to 0} \frac{\lambda(E_{x,r})}{v_n(E_{x,r})} = 0 \text{ a.e. } [v_n]$$

If $E_{x,r} = B(x,r)$ and λ is the counting measure $c_{\mathbf{Q}^n}$ restricted to $\mathcal{R}_n$ then $\lambda \perp v_n$ but the limit in Theorem 5.3.4 equals plus infinity for all $x \in \mathbf{R}^n$. The hypothesis "$\lambda(K) < \infty$ for every compact subset of $\mathbf{R}^n$" in Theorem 5.3.4 is not superflous.

PROOF. Since $\mid \lambda(E) \mid \leq \mid \lambda \mid (E)$ if $E \in \mathcal{R}_n$, there is no restriction to assume that λ is a positive measure (cf. Theorem 3.1.4). Moreover, since

$$\frac{\lambda(E_{x,r})}{v_n(E_{x,r})} \leq \frac{\lambda(B(x,r))}{\alpha v_n(B(x,r))}$$

it can be assumed that $E_{x,r} = B(x,r)$. Note that the function $\lambda(B(\cdot,r))$ is Borel measurable for fixed $r > 0$ and $\lambda(B(x,\cdot))$ left continuous for fixed $x \in \mathbf{R}^n$.

Suppose $A \in \mathcal{Z}_\lambda$ and $v_n = (v_n)^A$. Given $\delta > 0$, it is enough to prove that $F \in \mathcal{Z}_{v_n}$ where

$$F = \left\{ x \in A;\ \overline{\lim_{r \to 0}} \frac{\lambda(B(x,r))}{m_n(B(x,r))} > \delta \right\}$$

To this end let $\varepsilon > 0$ and use Theorem 3.1.3 to get an open $U \supseteq A$ such that $\lambda(U) < \varepsilon$. For each $x \in F$ there is an open ball $B_x \subseteq U$ such that

$$\lambda(B_x) > \delta v_n(B_x).$$

If $V = \cup_{x \in F} B_x$ and $c < v_n(V)$ we use Lemma 5.3.1 to obtain $x_1, ..., x_k$ such that $B_{x_1}, ..., B_{x_k}$ are pairwise disjoint and

$$c < 3^n \Sigma_{i=1}^k v_n(B_{x_i}) < 3^n \delta^{-1} \Sigma_{i=1}^k \lambda(B_{x_i})$$

$$\leq 3^n \delta^{-1} \lambda(U) < 3^n \delta^{-1} \varepsilon.$$

Thus $v_n(V) \leq 3^n \delta^{-1} \varepsilon$. Since $V \supseteq F \in \mathcal{R}_n$ and $\varepsilon > 0$ is arbitrary, $v_n(F) = 0$ and the theorem is proved.

Corollary 5.3.1. *Suppose $F : \mathbf{R} \to \mathbf{R}$ is an increasing function. Then $F'(x)$ exists for almost all x with respect to linear measure.*

PROOF. Let D be the set of all points of discontinuity of F. Suppose $-\infty < a < b < \infty$ and $\varepsilon > 0$. If $a < x_1 < ... < x_n < b$, where $x_1, ..., x_n \in D$ and

$$F(x_k+) - F(x_k-) \geq \varepsilon,\ \ k = 1, ..., n$$

then

$$n\varepsilon \leq \Sigma_{k=1}^n (F(x_k+) - F(x_k-)) \leq F(b) - F(a).$$

Thus $D \cap [a, b]$ is at most denumerable and it follows that D is at most denumerable. Set $H(x) = F(x+) - F(x)$, $x \in \mathbf{R}$, and let $(x_j)_{j=0}^N$ be an enumeration of the members of the set $\{H > 0\}$. Moreover, for any $a > 0$,

$$\sum_{|x_j| < a} H(x_j) \leq \sum_{|x_j| < a} (F(x_j+) - F(x_j-))$$

$$\leq F(a) - F(-a) < \infty.$$

Now, if we introduce

$$\nu(A) = \Sigma_1^N H(x_j)\delta_{x_j}(A),\ \ A \in \mathcal{R}$$

then ν is a positive measure such that $\nu(K) < \infty$ for each compact subset K of $\mathbf{R}$. Furthermore, if h is a non-zero real number,

$$\mid \frac{1}{h}(H(x+h) - H(x) \mid \leq \frac{1}{h}(H(x+h) + H(x)) \leq 4\frac{1}{4 \mid h \mid}\nu(B(x, 2 \mid h \mid)$$

and Theorem 5.3.4 implies that $H'(x) = 0$ a.e. $[v_1]$. Therefore, without loss of generality it may be assumed that F is right continuous and, in addition, there is no restriction to assume that $F(+\infty) - F(-\infty) < \infty$.

By Section 1.6 F induces a finite positive Borel measure μ such that

$$\mu(]x, y]) = F(y) - F(x) \text{ if } x < y.$$

Now consider the Lebesgue decomposition

$$d\mu = f dv_1 + d\lambda$$

where $f \in L^1(v_1)$ and $\lambda \perp v_1$. If $x < y$,

$$F(y) - F(x) = \int_x^y f(t)dt + \lambda(]x, y])$$

and the previous two theorems imply that

$$\lim_{y \downarrow x} \frac{F(y) - F(x)}{y - x} = f(x) \text{ a.e. } \ [v_1]$$

If $y < x$,

$$F(x) - F(y) = \int_y^x f(t)dt + \lambda(]y, x])$$

and we get

$$\lim_{y \uparrow x} \frac{F(y) - F(x)}{y - x} = f(x) \text{ a.e. } \ [v_1]\,.$$

The theorem is proved.

Exercises

1. Suppose $F : \mathbf{R} \to \mathbf{R}$ is increasing and let $f \in L^1_{loc}(v_1)$ be such that $F'(x) = f(x)$ a.e. $[v_1]$. Prove that

$$\int_x^y f(t)dt \leq F(y) - F(x) \text{ if } -\infty < x \leq y < \infty.$$

5.4. Absolutely Continuous Functions and Functions of Bounded Variation

Throughout this section a and b are reals with $a < b$ and to simplify notation we set $m_{a,b} = m_{|[a,b]}$. If $f \in L^1(m_{a,b})$ we know from the previous section that the function

$$(If)(x) =_{def} \int_a^x f(t)dt, \;\; a \leq x \leq b$$

has the derivative $f(x)$ a.e. $[m_{a,b}]$, that is

$$\frac{d}{dx} \int_a^x f(t)dt = f(x) \text{ a.e. } [m_{a,b}].$$

Our next main task will be to describe the range of the linear map I.

A function $F : [a, b] \to \mathbf{R}$ is said to be absolutely continuous if to every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\Sigma_{i=1}^n \mid b_i - a_i \mid < \delta \text{ implies } \Sigma_{i=1}^n \mid F(b_i) - F(a_i) \mid < \varepsilon$$

whenever $]a_1, b_1[, ...,]a_n, b_n[$ are disjoint open subintervals of $[a, b]$. It is obvious that an absolutely continuous function is continuous. It can be proved that the Cantor function is not absolutely continuous.

Theorem 5.4.1. *If $f \in L^1(m_{a,b})$, then If is absolutely continuous.*

PROOF. There is no restriction to assume $f \geq 0$. Set

$$d\lambda = f dm_{a,b}.$$

By Theorem 5.2.2, to every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\lambda(A) < \varepsilon$ for each Lebesgue set A in $[a, b]$ such that $m_{a,b}(A) < \delta$. Now restricting A to be a finite disjoint union of open intervals, the theorem follows.

Suppose $-\infty \leq \alpha < \beta \leq \infty$ and $F :]\alpha, \beta[\to \mathbf{R}$. For every $x \in]\alpha, \beta[$ we define

$$T_F(x) = \sup \Sigma_{i=1}^n \mid F(x_i) - F(x_{i-1}) \mid$$

where the supremum is taken over all positive integers n and all choices $(x_i)_{i=0}^n$ such that

$$\alpha < x_0 < x_1 < ... < x_n = x.$$

The function $T_F :]\alpha, \beta[\to [0, \infty]$ is called the total variation of F. Note that T_F is increasing. If T_F is a bounded function, F is said to be of bounded variation. A bounded increasing function on $\mathbf{R}$ is of bounded variation. Therefore the difference of two bounded increasing functions on $\mathbf{R}$ is of bounded variation. Interestingly enough, the converse is true. In the special case $]\alpha, \beta[= \mathbf{R}$ we write $F \in BV$ if F is of bounded variation.

Theorem 5.4.2. *Suppose $F \in BV$.*

(a) *The functions $T_F + F$ and $T_F - F$ are increasing and*

$$F = \frac{1}{2}(T_F + F) - \frac{1}{2}(T_F - F).$$

In particular, F is differentiable almost everywhere with respect to linear measure.

(b) *If F is right continuous, then so is T_F.*

PROOF. (a) Let $x < y$ and $\varepsilon > 0$. Choose $x_0 < x_1 < ... < x_n = x$ such that

$$\Sigma_{i=1}^n \mid F(x_i) - F(x_{i-1}) \mid \geq T_f(x) - \varepsilon.$$

Then

$$T_F(y) + F(y)$$

$$\geq \Sigma_{i=1}^{n} \mid F(x_i) - F(x_{i-1}) \mid + \mid F(y) - F(x) \mid + (F(y) - F(x)) + F(x)$$
$$\geq T_F(x) - \varepsilon + F(x)$$

and, since $\varepsilon > 0$ is arbitrary, $T_F(y) + F(y) \geq T_F(x) + F(x)$. Hence $T_F + F$ is increasing. Finally, replacing F by $-F$ it follows that the function $T_F - F$ is increasing.

(b) If $c \in \mathbf{R}$ and $x > c$,

$$T_f(x) = T_F(c) + \sup \Sigma_{i=1}^{n} \mid F(x_i) - F(x_{i-1}) \mid$$

where the supremum is taken over all positive integers n and all choices $(x_i)_{i=0}^{n}$ such that

$$c = x_0 < x_1 < ... < x_n = x.$$

Suppose $T_F(c+) > T_F(c)$ where $c \in \mathbf{R}$. Then there is an $\varepsilon > 0$ such that

$$T_F(x) - T_F(c) > \varepsilon$$

for all $x > c$. Now, since F is right continuous at the point c, for fixed $x > c$ there exists a partition

$$c < x_{11} < ... < x_{1n_1} = x$$

such that

$$\Sigma_{i=2}^{n_1} \mid F(x_{1i}) - F(x_{1i-1}) \mid > \varepsilon.$$

But

$$T_F(x_{11}) - T_F(c) > \varepsilon$$

and we get a partition

$$c < x_{21} < ... < x_{2n_2} = x_{11}$$

such that

$$\Sigma_{i=2}^{n_2} \mid F(x_{2i}) - F(x_{2i-1}) \mid > \varepsilon.$$

Summing up we have got a partition of the interval $[x_{21}, x]$ with

$$\Sigma_{i=2}^{n_2} \mid F(x_{2i}) - F(x_{2i-1}) \mid + \Sigma_{i=2}^{n_1} \mid F(x_{1i}) - F(x_{1i-1}) \mid > 2\varepsilon.$$

By repeating the process the total variation of F becomes infinite, which is a contradiction. The theorem is proved.

Theorem 5.4.3. *Suppose $F : [a,b] \to \mathbf{R}$ is absolutely continuous. Then there exists a unique $f \in L^1(m_{a,b})$ such that*

$$F(x) = F(a) + \int_a^x f(t)dt, \; a \leq x \leq b.$$

In particular, the range of the map I equals the set of all real-valued absolutely continuous maps on $[a,b]$.

PROOF. Set $F(x) = F(a)$ if $x \leq a$ and $F(x) = F(b)$ if $x \geq b$. There exists a $\delta > 0$ such that

$$\Sigma_{i=1}^n \mid b_i - a_i \mid < \delta \text{ implies } \Sigma_{i=1}^n \mid F(b_i) - F(a_i) \mid < 1$$

whenever $]a_1, b_1[, ...,]a_n, b_n[$ are disjoint subintervals of $[a,b]$. Let p be the least positive integer such that $a + p\delta \geq b$. Then $T_F \leq p$ and $F \in BV$. Let $F = G - H$, where $G = \frac{1}{2}(T_F + F)$ and $H = \frac{1}{2}(T_F - F)$. There exist finite positive Borel measures λ_G and λ_H such that

$$\lambda_G(]x,y]) = G(y) - G(x), \; x \leq y$$

and

$$\lambda_H(]x,y]) = H(y) - H(x), \; x \leq y.$$

If we define $\lambda = \lambda_G - \lambda_H$,

$$\lambda(]x,y]) = F(y) - F(x), \; x \leq y.$$

Clearly,

$$\lambda(]x,y[) = F(y) - F(x), \; x \leq y$$

since F is continuous.

Our next task will be to prove that $\lambda << v_1$. To this end, suppose $A \in \mathcal{R}$ and $v_1(A) = 0$. Now choose $\varepsilon > 0$ and let $\delta > 0$ be as in the definition of the absolute continuity of F on $[a,b]$. For each $k \in \mathbf{N}_+$ there exists an open set

$V_k \supseteq A$ such that $v_1(V_k) < \delta$ and $\lim_{k\to\infty} \lambda(V_k) = \lambda(A)$. But each fixed V_k is a disjoint union of open intervals $(]a_i, b_i[)_{i=1}^{\infty}$ and hence

$$\Sigma_{i=1}^{n} \mid b_i - a_i \mid < \delta$$

for every n and, accordingly from this,

$$\Sigma_{i=1}^{\infty} \mid F(b_i) - F(a_i) \mid \leq \varepsilon$$

and

$$\mid \lambda(V_k) \mid \leq \Sigma_{i=1}^{\infty} \mid \lambda(]a_i, b_i[) \mid \leq \varepsilon.$$

Thus $\mid \lambda(A) \mid \leq \varepsilon$ and since $\varepsilon > 0$ is arbitrary, $\lambda(A) = 0$. From this $\lambda << v_1$ and the theorem follows at once.

Suppose $(X, \mathcal{M}, \mu)$ is a positive measure space. From now on we write $f \in L^1(\mu)$ if there exist a $g \in \mathcal{L}^1(\mu)$ and an $A \in \mathcal{M}$ such that $A^c \in \mathcal{Z}_\mu$ and $f(x) = g(x)$ for all $x \in A$. Furthermore, we define

$$\int_X f d\mu = \int_X g d\mu$$

(cf the discussion in Section 2). Note that $f(x)$ need not be defined for every $x \in X$.

Corollary 5.4.1. *A function $f : [a, b] \to \mathbf{R}$ is absolutely continuous if and only if the following conditions are true:*

(i) $f'(x)$ exists for $m_{a,b}$-almost all $x \in [a, b]$

(ii) $f' \in L^1(m_{a,b})$

(iii) $f(x) = f(a) + \int_a^x f'(t)dt$, all $x \in [a, b]$.

Exercises

1. Suppose $f : [0, 1] \to \mathbf{R}$ satisfies $f(0) = 0$ and

$$f(x) = x^2 \sin \frac{1}{x^2} \text{ if } 0 < x \leq 1.$$

Prove that f is differentiable everywhere but f is not absolutely continuous.

2. Suppose α is a positive real number and f a function on $[0,1]$ such that $f(0) = 0$ and $f(x) = x^\alpha \sin\frac{1}{x}$, $0 < x \leq 1$. Prove that f is absolutely continuous if and only if $\alpha > 1$.

3. Suppose $f(x) = x\cos(\pi/x)$ if $0 < x < 2$ and $f(x) = 0$ if $x \in \mathbf{R}\setminus\,]0,2[$. Prove that f is not of bounded variation on $\mathbf{R}$.

4 A function $f : [a,b] \to \mathbf{R}$ is a Lipschitz function, that is there exists a positive real number C such that

$$\mid f(x) - f(y) \mid \leq C \mid x - y \mid$$

for all $x, y \in [a,b]$. Show that f is absolutely continuous and $\mid f'(x) \mid \leq C$ a.e. $[m_{a,b}]$.

5. Suppose $f : [a,b] \to \mathbf{R}$ is absolutely continuous. Prove that

$$T_g(x) = \int_a^x \mid f'(t) \mid dt,\ a < x < b$$

if g is the restriction of f to the open interval $]a,b[$.

6. Suppose f and g are real-valued absolutely continuous functions on the compact interval $[a,b]$. Show that the function $h = \max(f,g)$ is absolutely continuous and $h' \leq \max(f', g')$ a.e. $[m_{a,b}]$.

$\downarrow\downarrow\downarrow$

5.5. Conditional Expectation

Let $(\Omega, \mathcal{F}, P)$ be a probability space and suppose $\xi \in L^1(P)$. Moreover, suppose $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra and set

$$\mu(A) = P[A], \ A \in \mathcal{G}$$

and

$$\lambda(A) = \int_A \xi dP, \ A \in \mathcal{G}.$$

It is trivial that $\mathcal{Z}_\mu = \mathcal{Z}_P \cap \mathcal{G} \subseteq \mathcal{Z}_\lambda$ and the Radon-Nikodym Theorem shows there exists a unique $\eta \in L^1(\mu)$ such that

$$\lambda(A) = \int_A \eta d\mu \text{ all } A \in \mathcal{G}$$

or, what amounts to the same thing,

$$\int_A \xi dP = \int_A \eta dP \text{ all } A \in \mathcal{G}.$$

Note that η is $(\mathcal{G}, \mathcal{R})$-measurable. The random variable η is called the conditional expectation of ξ given $\mathcal{G}$ and it is standard to write $\eta = E[\xi \mid \mathcal{G}]$.

A sequence of σ-algebras $(\mathcal{F}_n)_{n=1}^\infty$ is called a filtration if

$$\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}.$$

If $(\mathcal{F}_n)_{n=1}^\infty$ is a filtration and $(\xi_n)_{n=1}^\infty$ is a sequence of real valued random variables such that for each n,

(a) $\xi_n \in L^1(P)$
(b) ξ_n is $(\mathcal{F}_n, \mathcal{R})$-measurable
(c) $E\left[\xi_{n+1} \mid \mathcal{F}_n\right] = \xi_n$

then $(\xi_n, \mathcal{F}_n)_{n=1}^\infty$ is called a martingale. There are very nice connections between martingales and the theory of differentiation (see e.g Billingsley $[B]$ and Malliavin $[M]$).

↑↑↑