CHAPTER 6

Differentiation

The generalization from elementary calculus of differentiation in measure theory is less obvious than that of integration, and the methods of treating it are somewhat involved.

Consider the fundamental theorem of calculus (FTC) for smooth functions of a single variable. In one direction (FTC-I, say) it states that the derivative of the integral is the original function, meaning that

$$f(x) = \frac{d}{dx} \int_a^x f(y)\, dy. \tag{6.1}$$

In the other direction (FTC-II, say) it states that we recover the original function by integrating its derivative

$$F(x) = F(a) + \int_a^x f(y)\, dy, \qquad f = F'. \tag{6.2}$$

As we will see, (6.1) holds pointwise a.e. provided that f is locally integrable, which is needed to ensure that the right-hand side is well-defined. Equation (6.2), however, does not hold for all continuous functions F whose pointwise derivative is defined a.e. and integrable; we also need to require that F is absolutely continuous. The Cantor function is a counter-example.

First, we consider a generalization of (6.1) to locally integrable functions on $\mathbb{R}^n$, which leads to the Lebesgue differentiation theorem. We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is locally integrable if it is Lebesgue measurable and

$$\int_K |f|\, dx < \infty$$

for every compact subset $K \subset \mathbb{R}^n$; we denote the space of locally integrable functions by $L^1_{\mathrm{loc}}(\mathbb{R}^n)$.

Let

$$B_r(x) = \{y \in \mathbb{R}^n : |y - x| < r\} \tag{6.3}$$

denote the open ball of radius r and center $x \in \mathbb{R}^n$. We denote Lebesgue measure on $\mathbb{R}^n$ by μ and the Lebesgue measure of a ball B by $\mu(B) = |B|$.

To motivate the statement of the Lebesgue differentiation theorem, observe that (6.1) may be written in terms of symmetric differences as

$$f(x) = \lim_{r \to 0^+} \frac{1}{2r} \int_{x-r}^{x+r} f(y)\, dy. \tag{6.4}$$

An n-dimensional version of (6.4) is

$$f(x) = \lim_{r \to 0^+} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\, dy \tag{6.5}$$

where the integral is with respect n-dimensional Lebesgue measure. The Lebesgue differentiation theorem states that (6.5) holds pointwise μ-a.e. for any locally integrable function f.

To prove the theorem, we will introduce the maximal function of an integrable function, whose key property is that it is weak-L^1, as stated in the Hardy-Littlewood theorem. This property may be shown by the use of a simple covering lemma, which we begin by proving.

Second, we consider a generalization of (6.2) on the representation of a function as an integral. In defining integrals on a general measure space, it is natural to think of them as defined on sets rather than real numbers. For example, in (6.2), we would write $F(x) = \nu([a, x])$ where $\nu : \mathcal{B}([a, b]) \to \mathbb{R}$ is a signed measure. This interpretation leads to the following question: if μ, ν are measures on a measurable space X is there a function $f : X \to [0, \infty]$ such that

$$\nu(A) = \int_A f \, d\mu.$$

If so, we regard $f = d\nu/d\mu$ as the (Radon-Nikodym) derivative of ν with respect to μ. More generally, we may consider signed (or complex) measures, whose values are not restricted to positive numbers. The Radon-Nikodym theorem gives a necessary and sufficient condition for the differentiability of ν with respect to μ, subject to a σ-finiteness assumption: namely, that ν is absolutely continuous with respect to μ.

6.1. A covering lemma

We need only the following simple form of a covering lemma; there are many more sophisticated versions, such as the Vitali and Besicovitch covering theorems, which we do not consider here.

Lemma 6.1. *Let $\{B_1, B_2, \ldots, B_N\}$ be a finite collection of open balls in $\mathbb{R}^n$. There is a disjoint subcollection $\{B_1', B_2', \ldots, B_M'\}$ where $B_j' = B_{i_j}$, such that*

$$\mu\left(\bigcup_{i=1}^N B_i\right) \le 3^n \sum_{i=1}^M \left|B_j'\right|.$$

Proof. If B is an open ball, let $\widehat{B}$ denote the open ball with the same center as B and three times the radius. Then

$$|\widehat{B}| = 3^n |B|.$$

Moreover, if B_1, B_2 are nondisjoint open balls and the radius of B_1 is greater than or equal to the radius of B_2, then $\widehat{B}_1 \supset B_2$.

We obtain the subfamily $\{B_j'\}$ by an iterative procedure. Choose B_1' to be a ball with the largest radius from the collection $\{B_1, B_2, \ldots, B_N\}$. Delete from the collection all balls B_i that intersect B_1', and choose B_2' to be a ball with the largest radius from the remaining balls. Repeat this process until the balls are exhausted, which gives $M \le N$ balls, say.

By construction, the balls $\{B_1', B_2', \ldots, B_M'\}$ are disjoint and

$$\bigcup_{i=1}^N B_i \subset \bigcup_{j=1}^M \widehat{B}_j'.$$

It follows that

$$\mu\left(\bigcup_{i=1}^{N} B_i\right) \leq \sum_{i=1}^{M}\left|\widehat{B}'_j\right| = 3^n \sum_{i=1}^{M}\left|B'_j\right|,$$

which proves the result. □

6.2. Maximal functions

The maximal function of a locally integrable function is obtained by taking the supremum of averages of the absolute value of the function about a point. Maximal functions were introduced by Hardy and Littlewood (1930), and they are the key to proving pointwise properties of integrable functions. They also play a central role in harmonic analysis.

Definition 6.2. If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then the maximal function Mf of f is defined by

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|}\int_{B_r(x)} |f(y)|\, dy.$$

The use of centered open balls to define the maximal function is for convenience. We could use non-centered balls or other sets, such as cubes, to define the maximal function. Some restriction on the shapes on the sets is, however, required; for example, we cannot use arbitrary rectangles, since averages over progressively longer and thinner rectangles about a point whose volumes shrink to zero do not, in general, converge to the value of the function at the point, even if the function is continuous.

Note that any two functions that are equal a.e. have the same maximal function.

Example 6.3. If $f : \mathbb{R} \to \mathbb{R}$ is the step function

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

then

$$Mf(x) = \begin{cases} 1 & \text{if } x > 0, \\ 1/2 & \text{if } x \leq 0. \end{cases}$$

This example illustrates the following result.

Proposition 6.4. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then the maximal function Mf is lower semi-continuous and therefore Borel measurable.*

PROOF. The function $Mf \geq 0$ is lower semi-continuous if

$$E_t = \{x : Mf(x) > t\}$$

is open for every $0 < t < \infty$. To prove that E_t is open, let $x \in E_t$. Then there exists $r > 0$ such that

$$\frac{1}{|B_r(x)|}\int_{B_r(x)} |f(y)|\, dy > t.$$

Choose $r' > r$ such that we still have

$$\frac{1}{|B_{r'}(x)|}\int_{B_r(x)} |f(y)|\, dy > t.$$

If $|x' - x| < r' - r$, then $B_r(x) \subset B_{r'}(x')$, so

$$t < \frac{1}{|B_{r'}(x)|}\int_{B_r(x)} |f(y)|\, dy \leq \frac{1}{|B_{r'}(x')|}\int_{B_{r'}(x')} |f(y)|\, dy \leq Mf(x'),$$

It follows that $x' \in E_t$, which proves that E_t is open. $\square$

The maximal function of a function $f \in L^1(\mathbb{R}^n)$ is not in $L^1(\mathbb{R}^n)$, unless $f = 0$, because it decays too slowly at infinity for its integral to converge. To show this, let $a > 0$ and suppose that $|x| \geq a$. Then, by considering the average of $|f|$ at x over a ball of radius $r = 2|x|$ and using the fact that $B_{2|x|}(x) \supset B_a(0)$, we see that

$$\begin{aligned} Mf(x) &\geq \frac{1}{|B_{2|x|}(x)|} \int_{B_{2|x|}(x)} |f(y)|\, dy \\ &\geq \frac{C}{|x|^n} \int_{B_a(0)} |f(y)|\, dy, \end{aligned}$$

where $C > 0$. The function $1/|x|^n$ is not integrable on $\mathbb{R}^n \setminus B_a(0)$, so if Mf is integrable then we must have

$$\int_{B_a(0)} |f(y)|\, dy = 0$$

for every $a > 0$, which implies that $f = 0$ a.e. in $\mathbb{R}^n$.

Moreover, as the following example shows, the maximal function of an integrable function need not even be locally integrable.

Example 6.5. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1/(x \log^2 x) & \text{if } 0 < x < 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

The change of variable $u = \log x$ implies that

$$\int_0^{1/2} \frac{1}{x|\log x|^n}\, dx$$

is finite if and only if $n > 1$. Thus $f \in L^1(\mathbb{R})$ and for $0 < x < 1/2$

$$\begin{aligned} Mf(x) &\geq \frac{1}{2x} \int_0^{2x} |f(y)|\, dy \\ &\geq \frac{1}{2x} \int_0^{x} \frac{1}{y \log^2 y}\, dy \\ &\geq \frac{1}{2x|\log x|} \end{aligned}$$

so $Mf \notin L^1_{\mathrm{loc}}(\mathbb{R})$.

6.3. Weak-L^1 spaces

Although the maximal function of an integrable function is not integrable, it is not much worse than an integrable function. As we show in the next section, it belongs to the space weak-L^1, which is defined as follows

Definition 6.6. The space weak-$L^1(\mathbb{R}^n)$ consists of measurable functions

$$f : \mathbb{R}^n \to \mathbb{R}$$

such that there exists a constant C, depending on f but not on t, with the property that for every $0 < t < \infty$

$$\mu\left\{x \in \mathbb{R}^n : |f(x)| > t\right\} \leq \frac{C}{t}.$$

An estimate of this form arises for integrable function from the following, almost trivial, Chebyshev inequality.

Theorem 6.7 (Chebyshev's inequality). *Suppose that $(X, \mathcal{A}, \mu)$ is a measure space. If $f : X \to \mathbb{R}$ is integrable and $0 < t < \infty$, then*

$$\mu\left(\{x \in X : |f(x)| > t\}\right) \leq \frac{1}{t}\|f\|_{L^1}. \tag{6.6}$$

Proof. Let $E_t = \{x \in X : |f(x)| > t\}$. Then

$$\int |f|\, d\mu \geq \int_{E_t} |f|\, d\mu \geq t\mu(E_t),$$

which proves the result. □

Chebyshev's inequality implies immediately that if f belongs to $L^1(\mathbb{R}^n)$, then f belongs to weak-$L^1(\mathbb{R}^n)$. The converse statement is, however, false.

Example 6.8. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \frac{1}{x}$$

for $x \neq 0$ satisfies

$$\mu\{x \in \mathbb{R} : |f(x)| > t\} = \frac{2}{t},$$

so f belongs to weak-$L^1(\mathbb{R})$, but f is not integrable or even locally integrable.

6.4. Hardy-Littlewood theorem

The following Hardy-Littlewood theorem states that the maximal function of an integrable function is weak-L^1.

Theorem 6.9 (Hardy-Littlewood). *If $f \in L^1(\mathbb{R}^n)$, there is a constant C such that for every $0 < t < \infty$*

$$\mu\left(\{x \in \mathbb{R}^n : Mf(x) > t\}\right) \leq \frac{C}{t}\|f\|_{L^1}$$

where $C = 3^n$ depends only on n.

Proof. Fix $t > 0$ and let

$$E_t = \{x \in \mathbb{R}^n : Mf(x) > t\}.$$

By the inner regularity of Lebesgue measure

$$\mu(E_t) = \sup\{\mu(K) : K \subset E_t \text{ is compact}\}$$

so it is enough to prove that

$$\mu(K) \leq \frac{C}{t}\int_{\mathbb{R}^n} |f(y)|\, dy.$$

for every compact subset K of E_t.

If $x \in K$, then there is an open ball B_x centered at x such that

$$\frac{1}{|B_x|}\int_{B_x} |f(y)|\, dy > t.$$

Since K is compact, we may extract a finite subcover $\{B_1, B_2, \dots, B_N\}$ from the open cover $\{B_x : x \in K\}$. By Lemma 6.1, there is a finite subfamily of disjoint balls $\{B_1', B_2', \dots, B_M'\}$ such that

$$\begin{aligned}\mu(K) &\le \sum_{i=1}^{N} |B_i| \\ &\le 3^n \sum_{j=1}^{M} |B_j'| \\ &\le \frac{3^n}{t} \sum_{j=1}^{M} \int_{B_j'} |f|\, dx \\ &\le \frac{3^n}{t} \int |f|\, dx,\end{aligned}$$

which proves the result with $C = 3^n$. □

6.5. Lebesgue differentiation theorem

The maximal function provides the crucial estimate in the following proof.

Theorem 6.10. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then for a.e. $x \in \mathbb{R}^n$*

$$\lim_{r\to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\, dy\right] = f(x).$$

Moreover, for a.e. $x \in \mathbb{R}^n$

$$\lim_{r\to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy\right] = 0.$$

Proof. Since

$$\left|\frac{1}{|B_r(x)|} \int_{B_r(x)} f(y) - f(x)\, dy\right| \le \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy,$$

we just need to prove the second result. We define $f^* : \mathbb{R}^n \to [0, \infty]$ by

$$f^*(x) = \limsup_{r\to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy\right].$$

We want to show that $f^* = 0$ pointwise a.e.

If $g \in C_c(\mathbb{R}^n)$ is continuous, then given any $\epsilon > 0$ there exists $\delta > 0$ such that $|g(x) - g(y)| < \epsilon$ whenever $|x - y| < \delta$. Hence if $r < \delta$

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy < \epsilon,$$

which implies that $g^* = 0$. We prove the result for general f by approximation with a continuous function.

First, note that we can assume that $f \in L^1(\mathbb{R}^n)$ is integrable without loss of generality; for example, if the result holds for $f\chi_{B_k(0)} \in L^1(\mathbb{R}^n)$ for each $k \in \mathbb{N}$ except on a set E_k of measure zero, then it holds for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ except on $\bigcup_{k=1}^{\infty} E_k$, which has measure zero.

Next, observe that since

$$|f(y) + g(y) - [f(x) + g(x)]| \leq |f(y) - f(x)| + |g(y) - g(x)|$$

and $\limsup(A + B) \leq \limsup A + \limsup B$, we have

$$(f + g)^* \leq f^* + g^*.$$

Thus, if $f \in L^1(\mathbb{R}^n)$ and $g \in C_c(\mathbb{R}^n)$, we have

$$\begin{aligned} (f - g)^* &\leq f^* + g^* = f^*, \\ f^* = (f - g + g)^* &\leq (f - g)^* + g^* = (f - g)^*, \end{aligned}$$

which shows that $(f - g)^* = f^*$.

If $f \in L^1(\mathbb{R}^n)$, then we claim that there is a constant C, depending only on n, such that for every $0 < t < \infty$

$$\mu\left(\{x \in \mathbb{R}^n : f^*(x) > t\}\right) \leq \frac{C}{t} \|f\|_{L^1}. \tag{6.7}$$

To show this, we estimate

$$\begin{aligned} f^*(x) &\leq \sup_{r>0} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy\right] \\ &\leq \sup_{r>0} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, dy\right] + |f(x)| \\ &\leq Mf(x) + |f(x)|. \end{aligned}$$

It follows that

$$\{f^* > t\} \subset \{Mf + |f| > t\} \subset \{Mf > t/2\} \cup \{|f| > t/2\}.$$

By the Hardy-Littlewood theorem,

$$\mu\left(\{x \in \mathbb{R}^n : Mf(x) > t/2\}\right) \leq \frac{2 \cdot 3^n}{t} \|f\|_{L^1},$$

and by the Chebyshev inequality

$$\mu\left(\{x \in \mathbb{R}^n : |f(x)| > t/2\}\right) \leq \frac{2}{t} \|f\|_{L^1}.$$

Combining these estimates, we conclude that (6.7) holds with $C = 2\,(3^n + 1)$.

Finally suppose that $f \in L^1(\mathbb{R}^n)$ and $0 < t < \infty$. From Theorem 4.27, for any $\epsilon > 0$, there exists $g \in C_c(\mathbb{R}^n)$ such that $\|f - g\|_{L^1} < \epsilon$. Then

$$\begin{aligned} \mu\left(\{x \in \mathbb{R}^n : f^*(x) > t\}\right) &= \mu\left(\{x \in \mathbb{R}^n : (f - g)^*(x) > t\}\right) \\ &\leq \frac{C}{t} \|f - g\|_{L^1} \\ &\leq \frac{C\epsilon}{t}. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that

$$\mu\left(\{x \in \mathbb{R}^n : f^*(x) > t\}\right) = 0,$$

and hence since

$$\{x \in \mathbb{R}^n : f^*(x) > 0\} = \bigcup_{k=1}^{\infty} \{x \in \mathbb{R}^n : f^*(x) > 1/k\}$$

that

$$\mu\left(\{x \in \mathbb{R}^n : f^*(x) > 0\}\right) = 0.$$

This proves the result. $\square$

The set of points x for which the limits in Theorem 6.10 exist for a suitable definition of $f(x)$ is called the Lebesgue set of f.

Definition 6.11. If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then a point $x \in \mathbb{R}^n$ belongs to the Lebesgue set of f if there exists a constant $c \in \mathbb{R}$ such that

$$\lim_{r \to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - c| \, dy\right] = 0.$$

If such a constant c exists, then it is unique. Moreover, its value depends only on the equivalence class of f with respect to pointwise a.e. equality. Thus, we can use this definition to give a canonical pointwise a.e. representative of a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ that is defined on its Lebesgue set.

Example 6.12. The Lebesgue set of the step function f in Example 6.3 is $\mathbb{R} \setminus \{0\}$. The point 0 does not belong to the Lebesgue set, since

$$\lim_{r \to 0^+} \left[\frac{1}{2r} \int_{-r}^{r} |f(y) - c| \, dy\right] = \frac{1}{2}\left(|c| + |1 - c|\right)$$

is nonzero for every $c \in \mathbb{R}$. Note that the existence of the limit

$$\lim_{r \to 0^+} \left[\frac{1}{2r} \int_{-r}^{r} f(y) \, dy\right] = \frac{1}{2}$$

is not sufficient to imply that 0 belongs to the Lebesgue set of f.

6.6. Signed measures

A signed measure is a countably additive, extended real-valued set function whose values are not required to be positive. Measures may be thought of as a generalization of volume or mass, and signed measures may be thought of as a generalization of charge, or a similar quantity. We allow a signed measure to take infinite values, but to avoid undefined expressions of the form $\infty - \infty$, it should not take both positive and negative infinite values.

Definition 6.13. Let $(X, \mathcal{A})$ be a measurable space. A signed measure ν on X is a function $\nu : \mathcal{A} \to \overline{\mathbb{R}}$ such that:

(a) $\nu(\varnothing) = 0$;
(b) ν attains at most one of the values ∞, $-\infty$;
(c) if $\{A_i \in \mathcal{A} : i \in \mathbb{N}\}$ is a disjoint collection of measurable sets, then

$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i).$$

We say that a signed measure is finite if it takes only finite values. Note that since $\nu\left(\bigcup_{i=1}^{\infty} A_i\right)$ does not depend on the order of the A_i, the sum $\sum_{i=1}^{\infty} \nu(A_i)$ converges unconditionally if it is finite, and therefore it is absolutely convergent. Signed measures have the same monotonicity property (1.1) as measures, with essentially the same proof. We will always refer to signed measures explicitly, and 'measure' will always refer to a positive measure.

Example 6.14. If $(X, \mathcal{A}, \mu)$ is a measure space and $\nu^+, \nu^- : \mathcal{A} \to [0, \infty]$ are measures, one of which is finite, then $\nu = \nu^+ - \nu^-$ is a signed measure.

Example 6.15. If $(X, \mathcal{A}, \mu)$ is a measure space and $f : X \to \overline{\mathbb{R}}$ is an $\mathcal{A}$-measurable function whose integral with respect to μ is defined as an extended real number, then $\nu : \mathcal{A} \to \overline{\mathbb{R}}$ defined by

$$\nu(A) = \int_A f \, d\mu \tag{6.8}$$

is a signed measure on X. As we describe below, we interpret f as the derivative $d\nu/d\mu$ of ν with respect to μ. If $f = f^+ - f^-$ is the decomposition of f into positive and negative parts then $\nu = \nu^+ - \nu^-$, where the measures $\nu^+, \nu^- : \mathcal{A} \to [0, \infty]$ are defined by

$$\nu^+(A) = \int_A f^+ \, d\mu, \qquad \nu^-(A) = \int_A f^- \, d\mu.$$

We will show that any signed measure can be decomposed into a difference of singular measures, called its Jordan decomposition. Thus, Example 6.14 includes all signed measures. Not all signed measures have the form given in Example 6.15. As we discuss this further in connection with the Radon-Nikodym theorem, a signed measure ν of the form (6.8) must be absolutely continuous with respect to the measure μ.

6.7. Hahn and Jordan decompositions

To prove the Jordan decomposition of a signed measure, we first show that a measure space can be decomposed into disjoint subsets on which a signed measure is positive or negative, respectively. This is called the Hahn decomposition.

Definition 6.16. Suppose that ν is a signed measure on a measurable space X. A set $A \subset X$ is positive for ν if it is measurable and $\nu(B) \geq 0$ for every measurable subset $B \subset A$. Similarly, A is negative for ν if it is measurable and $\nu(B) \leq 0$ for every measurable subset $B \subset A$, and null for ν if it is measurable and $\nu(B) = 0$ for every measurable subset $B \subset A$.

Because of the possible cancelation between the positive and negative signed measure of subsets, $\nu(A) > 0$ does not imply that A is positive for ν, nor does $\nu(A) = 0$ imply that A is null for ν. Nevertheless, as we show in the next result, if $\nu(A) > 0$, then A contains a subset that is positive for ν. The idea of the (slightly tricky) proof is to remove subsets of A with negative signed measure until only a positive subset is left.

Lemma 6.17. *Suppose that ν is a signed measure on a measurable space $(X, \mathcal{A})$. If $A \in \mathcal{A}$ and $0 < \nu(A) < \infty$, then there exists a positive subset $P \subset A$ such that $\nu(P) > 0$.*

Proof. First, we show that if $A \in \mathcal{A}$ is a measurable set with $|\nu(A)| < \infty$, then $|\nu(B)| < \infty$ for every measurable subset $B \subset A$. This is because ν takes at most one infinite value, so there is no possibility of canceling an infinite signed measure to give a finite measure. In more detail, we may suppose without loss of generality that $\nu : \mathcal{A} \to [-\infty, \infty)$ does not take the value ∞. (Otherwise, consider $-\nu$.) Then $\nu(B) \neq \infty$; and if $B \subset A$, then the additivity of ν implies that

$$\nu(B) = \nu(A) - \nu(A \setminus B) \neq -\infty$$

since $\nu(A)$ is finite and $\nu(A \setminus B) \neq \infty$.

Now suppose that $0 < \nu(A) < \infty$. Let

$$\delta_1 = \inf \{\nu(E) : E \in \mathcal{A} \text{ and } E \subset A\}.$$

Then $-\infty \leq \delta_1 \leq 0$, since $\varnothing \subset A$. Choose $A_1 \subset A$ such that $\delta_1 \leq \nu(A_1) \leq \delta_1/2$ if δ_1 is finite, or $\mu(A_1) \leq -1$ if $\delta_1 = -\infty$. Define a disjoint sequence of subsets $\{A_i \subset A : i \in \mathbb{N}\}$ inductively by setting

$$\delta_i = \inf \left\{\nu(E) : E \in \mathcal{A} \text{ and } E \subset A \setminus \left(\bigcup_{j=1}^{i-1} A_j\right)\right\}$$

and choosing $A_i \subset A \setminus \left(\bigcup_{j=1}^{i-1} A_j\right)$ such that

$$\delta_i \leq \nu(A_i) \leq \frac{1}{2}\delta_i$$

if $-\infty < \delta_i \leq 0$, or $\nu(A_i) \leq -1$ if $\delta_i = -\infty$.

Let

$$B = \bigcup_{i=1}^{\infty} A_i, \qquad P = A \setminus B.$$

Then, since the A_i are disjoint, we have

$$\nu(B) = \sum_{i=1}^{\infty} \nu(A_i).$$

As proved above, $\nu(B)$ is finite, so this negative sum must converge. It follows that $\nu(A_i) \leq -1$ for only finitely many i, and therefore δ_i is infinite for at most finitely many i. For the remaining i, we have

$$\sum \nu(A_i) \leq \frac{1}{2} \sum \delta_i \leq 0,$$

so $\sum \delta_i$ converges and therefore $\delta_i \to 0$ as $i \to \infty$.

If $E \subset P$, then by construction $\nu(E) \geq \delta_i$ for every sufficiently large $i \in \mathbb{N}$. Hence, taking the limit as $i \to \infty$, we see that $\nu(E) \geq 0$, which implies that P is positive. The proof also shows that, since $\nu(B) \leq 0$, we have

$$\nu(P) = \nu(A) - \nu(B) \geq \nu(A) > 0,$$

which proves that P has strictly positive signed measure. □

The Hahn decomposition follows from this result in a straightforward way.

Theorem 6.18 (Hahn decomposition)**.** *If ν is a signed measure on a measurable space $(X, \mathcal{A})$, then there is a positive set P and a negative set N for ν such that $P \cup N = X$ and $P \cap N = \varnothing$. These sets are unique up to ν-null sets.*

Proof. Suppose, without loss of generality, that $\nu(A) < \infty$ for every $A \in \mathcal{A}$. (Otherwise, consider $-\nu$.) Let

$$m = \sup\{\nu(A) : A \in \mathcal{A} \text{ such that } A \text{ is positive for } \nu\},$$

and choose a sequence $\{A_i : i \in \mathbb{N}\}$ of positive sets such that $\nu(A_i) \to m$ as $i \to \infty$. Then, since the union of positive sets is positive,

$$P = \bigcup_{i=1}^{\infty} A_i$$

is a positive set. Moreover, by the monotonicity of of ν, we have $\nu(P) = m$. Since $\nu(P) \neq \infty$, it follows that $m \geq 0$ is finite.

Let $N = X \setminus P$. Then we claim that N is negative for ν. If not, there is a subset $A' \subset N$ such that $\nu(A') > 0$, so by Lemma 6.17 there is a positive set $P' \subset A'$ with $\nu(P') > 0$. But then $P \cup P'$ is a positive set with $\nu(P \cup P') > m$, which contradicts the definition of m.

Finally, if P', N' is another such pair of positive and negative sets, then

$$P \setminus P' \subset P \cap N',$$

so $P \setminus P'$ is both positive and negative for ν and therefore null, and similarly for $P' \setminus P$. Thus, the decomposition is unique up to ν-null sets. □

To describe the corresponding decomposition of the signed measure ν into the difference of measures, we introduce the notion of singular measures, which are measures that are supported on disjoint sets.

Definition 6.19. Two measures μ, ν on a measurable space $(X, \mathcal{A})$ are singular, written $\mu \perp \nu$, if there exist sets $M, N \in \mathcal{A}$ such that $M \cap N = \varnothing$, $M \cup N = X$ and $\mu(M) = 0$, $\nu(N) = 0$.

Example 6.20. The δ-measure in Example 2.36 and the Cantor measure in Example 2.37 are singular with respect to Lebesgue measure on $\mathbb{R}$ (and conversely, since the relation is symmetric).

Theorem 6.21 (Jordan decomposition)**.** *If ν is a signed measure on a measurable space $(X, \mathcal{A})$, then there exist unique measures $\nu^+, \nu^- : \mathcal{A} \to [0, \infty]$, one of which is finite, such that*

$$\nu = \nu^+ - \nu^- \qquad \text{and } \nu^+ \perp \nu^-.$$

Proof. Let $X = P \cup N$ where P, N are positive, negative sets for ν. Then

$$\nu^+(A) = \nu(A \cap P), \qquad \nu^-(A) = -\nu(A \cap N)$$

is the required decomposition. The values of $\nu^\pm$ are independent of the choice of P, N up to a ν-null set, so the decomposition is unique. □

We call ν^+ and ν^- the positive and negative parts of ν, respectively. The total variation $|\nu|$ of ν is the measure

$$|\nu| = \nu^+ + \nu^-.$$

We say that the signed measure ν is σ-finite if $|\nu|$ is σ-finite.

6.8. Radon-Nikodym theorem

The absolute continuity of measures is in some sense the opposite relationship to the singularity of measures. If a measure ν singular with respect to a measure μ, then it is supported on different sets from μ, while if ν is absolutely continuous with respect to μ, then it supported on on the same sets as μ.

Definition 6.22. Let ν be a signed measure and μ a measure on a measurable space $(X, \mathcal{A})$. Then ν is absolutely continuous with respect to μ, written $\nu \ll \mu$, if $\nu(A) = 0$ for every set $A \in \mathcal{A}$ such that $\mu(A) = 0$.

Equivalently, $\nu \ll \mu$ if every μ-null set is a ν-null set. Unlike singularity, absolute continuity is not symmetric.

Example 6.23. If μ is Lebesgue measure and ν is counting measure on $\mathcal{B}(\mathbb{R})$, then $\mu \ll \nu$, but $\nu \not\ll \mu$.

Example 6.24. If $f : X \to \overline{\mathbb{R}}$ is a measurable function on a measure space $(X, \mathcal{A}, \mu)$ whose integral with respect μ is well-defined as an extended real number and the signed measure $\nu : \mathcal{A} \to \overline{\mathbb{R}}$ is defined by

$$\nu(A) = \int_A f \, d\mu,$$

then (4.4) shows that ν is absolutely continuous with respect to μ.

The next result clarifies the relation between Definition 6.22 and the absolute continuity property of integrable functions proved in Proposition 4.16.

Proposition 6.25. *If ν is a finite signed measure and μ is a measure, then $\nu \ll \mu$ if and only if for every $\epsilon > 0$, there exists $\delta > 0$ such that $|\nu(A)| < \epsilon$ whenever $\mu(A) < \delta$.*

Proof. Suppose that the given condition holds. If $\mu(A) = 0$, then $|\nu(A)| < \epsilon$ for every $\epsilon > 0$, so $\nu(A) = 0$, which shows that $\nu \ll \mu$.

Conversely, suppose that the given condition does not hold. Then there exists $\epsilon > 0$ such that for every $k \in \mathbb{N}$ there exists a measurable set A_k with $|\nu|(A_k) \geq \epsilon$ and $\mu(A_k) < 1/2^k$. Defining

$$B = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} A_j,$$

we see that $\mu(B) = 0$ but $|\nu|(B) \geq \epsilon$, so ν is not absolutely continuous with respect to μ. □

The Radon-Nikodym theorem provides a converse to Example 6.24 for absolutely continuous, σ-finite measures. As part of the proof, from [**4**], we also show that any signed measure ν can be decomposed into an absolutely continuous and singular part with respect to a measure μ (the Lebesgue decomposition of ν). In the proof of the theorem, we will use the following lemma.

Lemma 6.26. *Suppose that μ, ν are finite measures on a measurable space $(X, \mathcal{A})$. Then either $\mu \perp \nu$, or there exists $\epsilon > 0$ and a set P such that $\mu(P) > 0$ and P is a positive set for the signed measure $\nu - \epsilon\mu$.*

Proof. For each $n \in \mathbb{N}$, let $X = P_n \cup N_n$ be a Hahn decomposition of X for the signed measure $\nu - \frac{1}{n}\mu$. If

$$P = \bigcup_{n=1}^{\infty} P_n \qquad N = \bigcap_{n=1}^{\infty} N_n,$$

then $X = P \cup N$ is a disjoint union, and

$$0 \leq \nu(N) \leq \frac{1}{n}\mu(N)$$

for every $n \in \mathbb{N}$, so $\nu(N) = 0$. Thus, either $\mu(P) = 0$, when $\nu \perp \mu$, or $\mu(P_n) > 0$ for some $n \in \mathbb{N}$, which proves the result with $\epsilon = 1/n$. □

Theorem 6.27 (Lebesgue-Radon-Nikodym theorem)**.** *Let ν be a σ-finite signed measure and μ a σ-finite measure on a measurable space $(X, \mathcal{A})$. Then there exist unique σ-finite signed measures ν_a, ν_s such that*

$$\nu = \nu_a + \nu_s \qquad \text{where } \nu_a \ll \mu \text{ and } \nu_s \perp \mu.$$

Moreover, there exists a measurable function $f : X \to \overline{\mathbb{R}}$, uniquely defined up to μ-a.e. equivalence, such that

$$\nu_a(A) = \int_A f \, d\mu$$

for every $A \in \mathcal{A}$, where the integral is well-defined as an extended real number.

Proof. It is enough to prove the result when ν is a measure, since we may decompose a signed measure into its positive and negative parts and apply the result to each part.

First, we assume that μ, ν are finite. We will construct a function f and an absolutely continuous signed measure $\nu_a \ll \mu$ such that

$$\nu_a(A) = \int_A f \, d\mu \qquad \text{for all } A \in \mathcal{A}.$$

We write this equation as $d\nu_a = f \, d\mu$ for short. The remainder $\nu_s = \nu - \nu_a$ is the singular part of ν.

Let $\mathcal{F}$ be the set of all $\mathcal{A}$-measurable functions $g : X \to [0, \infty]$ such that

$$\int_A g \, d\mu \leq \nu(A) \qquad \text{for every } A \in \mathcal{A}.$$

We obtain f by taking a supremum of functions from $\mathcal{F}$. If $g, h \in \mathcal{F}$, then $\max\{g, h\} \in \mathcal{F}$. To see this, note that if $A \in \mathcal{A}$, then we may write $A = B \cup C$ where

$$B = A \cap \{x \in X : g(x) > h(x)\}, \qquad C = A \cap \{x \in X : g(x) \leq h(x)\},$$

and therefore

$$\int_A \max\{g, h\} \, d\mu = \int_B g \, d\mu + \int_C h \, d\mu \leq \nu(B) + \nu(C) = \nu(A).$$

Let

$$m = \sup\left\{\int_X g \, d\mu : g \in \mathcal{F}\right\} \leq \nu(X).$$

Choose a sequence $\{g_n \in \mathcal{F} : n \in \mathbb{N}\}$ such that

$$\lim_{n\to\infty} \int_X g_n \, d\mu = m.$$

By replacing g_n with $\max\{g_1, g_2, \dots, g_n\}$, we may assume that $\{g_n\}$ is an increasing sequence of functions in $\mathcal{F}$. Let

$$f = \lim_{n\to\infty} g_n.$$

Then, by the monotone convergence theorem, for every $A \in \mathcal{A}$ we have

$$\int_A f \, d\mu = \lim_{n\to\infty} \int_A g_n \, d\mu \leq \nu(A),$$

so $f \in \mathcal{F}$ and

$$\int_X f \, d\mu = m.$$

Define $\nu_s : \mathcal{A} \to [0, \infty)$ by

$$\nu_s(A) = \nu(A) - \int_A f \, d\mu.$$

Then ν_s is a positive measure on X. We claim that $\nu_s \perp \mu$, which proves the result in this case. Suppose not. Then, by Lemma 6.26, there exists $\epsilon > 0$ and a set P with $\mu(P) > 0$ such that $\nu_s \geq \epsilon\mu$ on P. It follows that for any $A \in \mathcal{A}$

$$\begin{aligned}
\nu(A) &= \int_A f \, d\mu + \nu_s(A) \\
&\geq \int_A f \, d\mu + \nu_s(A \cap P) \\
&\geq \int_A f \, d\mu + \epsilon\mu(A \cap P) \\
&\geq \int_A (f + \epsilon\chi_P) \, d\mu.
\end{aligned}$$

It follows that $f + \epsilon\chi_P \in \mathcal{F}$ but

$$\int_X (f + \epsilon\chi_P) \, d\mu = m + \epsilon\mu(P) > m,$$

which contradicts the definition of m. Hence $\nu_s \perp \mu$.

If $\nu = \nu_a + \nu_s$ and $\nu = \nu_a' + \nu_s'$ are two such decompositions, then $\nu_a - \nu_a' = \nu_s' - \nu_s$ is both absolutely continuous and singular with respect to μ which implies that it is zero. Moreover, f is determined uniquely by ν_a up to pointwise a.e. equivalence.

Finally, if μ, ν are σ-finite measures, then we may decompose

$$X = \bigcup_{i=1}^{\infty} A_i$$

into a countable disjoint union of sets with $\mu(A_i) < \infty$ and $\nu(A_i) < \infty$. We decompose the finite measure $\nu_i = \nu|_{A_i}$ as

$$\nu_i = \nu_{ia} + \nu_{is} \qquad \text{where } \nu_{ia} \ll \mu_i \text{ and } \nu_{is} \perp \mu_i.$$

Then $\nu = \nu_a + \nu_s$ is the required decomposition with

$$\nu_a = \sum_{i=1}^{\infty} \nu_{ia}, \qquad \nu_s = \sum_{i=1}^{\infty} \nu_{ia}$$

is the required decomposition. □

The decomposition $\nu = \nu_a + \nu_s$ is called the Lebesgue decomposition of ν, and the representation of an absolutely continuous signed measure $\nu \ll \mu$ as $d\nu = f \, d\mu$ is the Radon-Nikodym theorem. We call the function f here the Radon-Nikodym derivative of ν with respect to μ, and denote it by

$$f = \frac{d\nu}{d\mu}.$$

Some hypothesis of σ-finiteness is essential in the theorem, as the following example shows.

Example 6.28. Let $\mathcal{B}$ be the Borel σ-algebra on $[0, 1]$, μ Lebesgue measure, and ν counting measure on $\mathcal{B}$. Then μ is finite and $\mu \ll \nu$, but ν is not σ-finite. There is no function $f : [0, 1] \to [0, \infty]$ such that

$$\mu(A) = \int_A f \, d\nu = \sum_{x \in A} f(x).$$

There are generalizations of the Radon-Nikodym theorem which apply to measures that are not σ-finite, but we will not consider them here.

6.9. Complex measures

Complex measures are defined analogously to signed measures, except that they are only permitted to take finite complex values.

Definition 6.29. Let $(X, \mathcal{A})$ be a measurable space. A complex measure ν on X is a function $\nu : \mathcal{A} \to \mathbb{C}$ such that:

(a) $\nu(\varnothing) = 0$;
(b) if $\{A_i \in \mathcal{A} : i \in \mathbb{N}\}$ is a disjoint collection of measurable sets, then

$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i).$$

There is an analogous Radon-Nikodym theorems for complex measures. The Radon-Nikodym derivative of a complex measure is necessarily integrable, since the measure is finite.

Theorem 6.30 (Lebesgue-Radon-Nikodym theorem)**.** *Let ν be a complex measure and μ a σ-finite measure on a measurable space $(X, \mathcal{A})$. Then there exist unique complex measures ν_a, ν_s such that*

$$\nu = \nu_a + \nu_s \qquad \text{where } \nu_a \ll \mu \text{ and } \nu_s \perp \mu.$$

Moreover, there exists an integrable function $f : X \to \mathbb{C}$, uniquely defined up to μ-a.e. equivalence, such that

$$\nu_a(A) = \int_A f \, d\mu$$

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

To prove the result, we decompose a complex measure into its real and imaginary parts, which are finite signed measures, and apply the corresponding theorem for signed measures.