Chapter 2

Integration

2.1 Measurable functions

Intuition. Continuous functions are the functions that behave well with respect to the open sets (topology). It is natural to define measurable functions as those that behave well with respect to measurable sets.

Definition 2.1. *(i) Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be two measurable spaces. A function $f : X \to Y$ is called $(\mathcal{M}, \mathcal{N})$-measurable) (or just* ***measurable*** *if σ-algebras are clear from the context) if $f^{-1}(E) \in \mathcal{M}$ for every $E \in \mathcal{N}$.*

(ii) We say that $f : X \to Y$ is ***Borel measurable*** *if it is $(\mathcal{B}(X), \mathcal{B}(Y))$-measurable.*

Remarks 2.2. (a) To check $(\mathcal{M}, \mathcal{N})$-measurability, it is sufficient to check if $f^{-1}(E) \in \mathcal{M}$ for every E in a family that generate $\mathcal{N}$ (e.g., Y-open sets if $\mathcal{N} = \mathcal{B}(Y)$). Indeed, f^{-1} behaves well under (countable) unions, intersections, and complements.

Proposition 2.3. *(i) Any continuous function between two topological spaces is Borel measurable.*

(ii) If f, g are $(\mathcal{M}, \mathcal{N})$-measurable, then so are $f + g$, fg, f/g (if well-defined).

(iii) If f is $(\mathcal{M}, \mathcal{N})$-measurable and g is $(\mathcal{N}, \mathcal{S})$-measurable, then $g \circ f$ is $(\mathcal{M}, \mathcal{S})$-measurable.

(iv) If $\{f_j\}_{j=1}^{\infty}$ is a family of $(\mathcal{M}, \mathcal{B}(\mathbb{R}))$-measurable functions, then

$$\sup_j f_j(x), \quad \inf_j f_j(x), \quad \limsup_{j\to\infty} f_j(x), \quad \liminf_{j\to\infty} f_j(x)$$

are all $(\mathcal{M}, \mathcal{B}(\mathbb{R} \cup \{\pm\infty\}))$-measurable.

(v) If $\{f_j\}_{j=1}^{\infty}$ is a family of $(\mathcal{M}, \mathcal{B}(\mathbb{R}))$-measurable functions that is pointwise convergent, then its limit is measurable.

Remarks 2.4. (a) In particular, $f^+(x) := \max(f(x), 0)$ and $f^-(x) := -\min(f(x), 0)$ are both measurable.

(b) If $f_j \to f(x)$ pointwise μ-almost everywhere, then on a μ-null set $f(x)$ can behave arbitrarily, so doesn't have to be measurable in general. If μ is complete however, then f is measurable in that case too.

Proof. (iv) Note that $\liminf_{j\to\infty} y_j = \sup_{n\geq 1} \inf_{j\geq n} y_j$, and similarly $\limsup_{j\to\infty} y_j = \inf_{n\geq 1} \sup_{j\geq n} y_j$, so just need inf and sup. Let $g = \sup_j f_j$, then $g^{-1}((a, \infty]) = \bigcup_{j=1}^{\infty} f_j^{-1}((a, \infty])$, and these intervals generate $\mathcal{B}(\mathbb{R} \cup \{\pm\infty\})$, so we're done.

(v) follows from $\lim = \limsup = \liminf$ whenever the limit exists. ☐

2.2 Integration of simple functions: definition of the Lebesgue integral

Intuition. Now we want to define the notion of the integral of a measurable function with respect to a given measure. We will do this in steps: first for simple (i.e., "piece-wise constant") functions, then for positive functions, then for real-valued functions, and finally for complex functions.

In what follows, let χ_E be the characteristic function of a set E: equal to 1 if $x \in E$ and 0 otherwise.

Definition 2.5. *Let $(X, \mathcal{M})$ be a measurable space. A* ***simple function*** *$\phi(x)$ is a finite linear combination of characteristic functions of disjoint measurable sets:*

$$\phi(x) = \sum_{j=1}^{n} a_j \chi_{E_j},$$

where $E_j \in \mathcal{M}$, $E_j \cap E_k = \varnothing$ if $j \neq k$.

Definition 2.6. *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

If $\phi(x) = \sum_{j=1}^{n} a_j \chi_{E_j}$ is a simple function $X \to [0, +\infty)$, then for any $E \in \mathcal{M}$, we define the ***(Lebesgue) integral*** *of ϕ with respect to μ over E by*

$$\int_E \phi \, d\mu = \sum_{j=1}^{n} a_j \mu(E_j \cap E)$$

(with the convention $0 \cdot \infty = 0$).

Remark 2.7. Note that $\int_X \chi_E \, d\mu = \mu(E)$.

2.3 Integration of positive functions: definition of the Lebesgue integral

Let us denote L^+ to be the space of all measurable functions $X \to [0, \infty]$.

Proposition 2.8. *Let $f \in L^+$. Then there exists a sequence of simple functions $\{\phi_n\}$ such that $0 \leq \phi_1(x) \leq \phi_2(x) \leq \ldots \leq f(x)$ such that $\phi_n(x) \to f(x)$ pointwise.*

Remark 2.9. An analogue with $0 \leq |\phi_1(x)| \leq |\phi_2(x)| \leq \ldots \leq |f(x)|$ holds for complex-valued measurable functions.

Proof. Divide the range into subintervals $I_{k,n} := (\frac{k}{2^n}, \frac{k+1}{2^n}]$ and $I_{0,n} := (2^n, \infty]$ and define $\phi_n := \sum \frac{k}{2^n} \chi_{f^{-1}(I_{k,n})} + 2^n \chi_{f^{-1}(I_{0,n})}$ (draw a pic!) ☐

Definition 2.10. *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

If $f(x) \in L^+$, we define the ***(Lebesgue) integral*** *of f with respect to μ over E by*

$$\int_E f \, d\mu = \sup \left\{ \int_E \phi \, d\mu : 0 \leq \phi \leq f, \phi \text{ simple} \right\}$$

Remarks 2.11. (a) The two definitions agree when f is a simple function.

(b) As we saw in Proposition 2.8, for any $f \in L^+$ there exists a sequence of simple functions such that $0 \leq f_n(x) \uparrow f(x)$ ("converges monotonically from below") for every x, in which case $\int_E f_n \, d\mu \uparrow \int_E f \, d\mu$. This will follow from the Monotone Convergence Theorem.

2.4 Integration of positive functions: basic properties

Proposition 2.12. *Assume all the sets and functions below are measurable, and all the functions are in L^+.*

(i) $\int_E (f+g)\, d\mu = \int_E f\, d\mu + \int_E g\, d\mu$.

(ii) If $0 \le c < \infty$ *is a constant then* $\int_E cf\, d\mu = c \int_E f\, d\mu$.

(iii) If $0 \le f \le g$ *on* E*, then* $0 \le \int_E f\, d\mu \le \int_E g\, d\mu$.

(iv) If $A \subseteq B$ *and* $f \in L^+$ *then* $\int_A f\, d\mu \le \int_B f\, d\mu$.

(v) If $\mu(E) = 0$ *then* $\int_E f\, d\mu = 0$*, even if* $f(x) = \infty$ *for every* $x \in E$.

(vi) $\int_E f\, d\mu = \int_X \chi_E f\, d\mu$.

(vii) If $f \in L^+$ *then* $\int_E f\, d\mu = 0$ *iff* $f = 0$ *a.e. on* E.

Remark 2.13. As we are about to see, property (i) for functions in L^+ can be extended to infinite sums!

Proof. For simple functions, all these properties follow directly from the definition.

For arbitrary positive functions: (ii), (iii), (iv), (v), (vi) are also immediate from the definition. For (i) and (vii), one can first prove Lebesgue's Monotone Convergence Theorem (see Theorem 2.14), which establishes Remark 2.11(b), and then the rest becomes easy. □

2.5 Integration of positive functions: main theorems

Intuition. L^+ functions are very easy to deal with since there is no issue of convergence: their Lebesgue integral always exist in $[0, \infty]$.

Theorem 2.14 (Lebesgue's Monotone Convergence Theorem)**.** *Suppose:*

(i) $f_j \in L^+$ *for every* j*;*

(ii) $f_n(x) \uparrow f(x)$ *("monotonically converges from below") as* $n \to \infty$ *for every* $x \in X$.

Then

$$\int_X f_n\, d\mu \to \int_X f\, d\mu$$

Remarks 2.15. (a) As we discuss in the next section, in all of these results we can replace "for every x" with "for μ-almost every x".

Proof. Let $\lim_{n\to\infty} \int f_n = I \in [0, +\infty]$ (non-decreasing numbers must have a limit). Need to show I is equal to $\int f$.

Since $f_n \le f$, we get $\int f_n \le \int f$, so that $I \le \int f$.

To show $I \ge \int f$, need to use the definition of $\int f$. So choose any simple function ϕ with $0 \le \phi \le f$. Ideally we would like to have $f_n \ge \phi$ since then $\int f_n \ge \int \phi$, implying $I \ge \int f$. But $f_n \ge \phi$ can of course fail, so we choose an arbitrary $0 < c < 1$, and let $E_n = \{x : f_n \ge c\phi(x)\}$, Then E_n is increasing sequence of sets whose union is X. So $\int f_n \ge \int_{E_n} f_n \ge c \int_{E_n} \phi$. Now take $n \to \infty$ to get $I \ge c \int \phi$ (this is believable, but technically we are using Proposition 2.22(iv) here). Then take $c \to 1$ and sup over all ϕ to get $I \ge \int f$. □

Theorem 2.16 (Interchanging integral and sum for L^+)**.**

$$\int_X \left(\sum_{j=1}^{\infty} f_j(x) \right) d\mu = \sum_{j=1}^{\infty} \int_X f_j(x)\, d\mu$$

for any $f_j \in L^+$.

Proof. For finite sums: approximate by simple functions from below monotonically, and use monotone convergence theorem: $\int \sum_{j=1}^N f_j = \sum_{j=1}^N \int f_j$.

For infinite sums: take $N \to \infty$ using monotone convergence theorem. ☐

Theorem 2.17 (Fatou's Lemma)**.** *For any $f_n \in L^+$,*

$$\int_X \left(\liminf_{n\to\infty} f_n \right) d\mu \le \liminf_{n\to\infty} \int_X f_n\, d\mu.$$

Remarks 2.18. (a) In particular, if all the limits exist, then $\int \lim f_n \le \lim \int f_n$ for $f_n \in L^+$.

(b) The strict inequality may occur: e.g., $f_n = \chi_{[n,n+1]}$ or $f_n = n\chi_{(0,1/n)}$: in both cases "mass is escaping to infinity". So to get $\lim \int f_n = \int \lim f_n$ one needs to forbid these kind of behaviours. See the Dominated Convergence Theorem later on.

Proof. $\liminf_{k\to\infty} = \lim_{k\to\infty} \inf_{n\ge k}$, so first we notice that $\inf_{n\ge k} f_n \le f_j$ for any $j \ge k$, so $\int \inf_{n\ge k} f_n \le \int f_j$ for every $j \ge k$, so that $\int \inf_{n\ge k} f_n \le \inf_{j\ge k} \int f_j$. Now take $k \to \infty$ and use the monotone convergence theorem. ☐

2.6 $L^1(\mu)$: integration of real and complex functions

Definition 2.19. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Define $L^1(\mu)$ (**space of Lebesgue integrable functions**) to be the collection of all measurable functions $f : X \to \mathbb{C}$ for which*

$$\int_X |f|\, d\mu < \infty$$

Definition 2.20. *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

*(i) If $f(x) \in L^1(\mu)$ is real-valued, then we decompose $f(x) = f^+(x) - f^-(x)$, where $f^+(x) := \max(f(x), 0)$ and $f^-(x) := -\min(f(x), 0)$, and we define the **(Lebesgue) integral** of f with respect to μ over E by*

$$\int_E f\, d\mu = \int_E f^+\, d\mu - \int_E f^-\, d\mu,$$

(which is finite since $\int_X |f|\, d\mu < \infty$).

*(ii) If $f(x) \in L^1(\mu)$ is complex-valued, then we decompose $f(x) = \operatorname{Re} f + i \operatorname{Im} f$, and we define the **(Lebesgue) integral** of f with respect to μ over E by*

$$\int_E f\, d\mu = \int_E \operatorname{Re} f\, d\mu + i \int_E \operatorname{Im} f\, d\mu,$$

(which is finite since $\int_X |f|\, d\mu < \infty$).

Remarks 2.21. (a) If $f(x) = f^+(x) - f^-(x)$ (real-valued) but not in $L^1(\mu)$, we may still write $\int_E f\,d\mu$ to mean $\int_E f^+\,d\mu - \int_E f^-\,d\mu$, provided that one of the integrals on the right-hand side is finite (so that $\int_E f\,d\mu$ is either $+\infty$ or $-\infty$).

(b) If two functions f and g are equal to each other μ-almost everywhere (that is, $\mu(\{x : f(x) \neq g(x)\}) = 0$, then it is easy to see that $\int_E f\,d\mu = \int_E g\,d\mu$ for any set E. Therefore from now on, as members of $L^1(\mu)$, f and g will not be distinguished from each other. More formally, we are redefining $L^1(\mu)$ as the quotient space, where each equivalence class consists of functions equal μ-almost everywhere. We can (and will) therefore even include functions that are undefined or infinite on a μ-null set.

(c) Another reason for the discussion in (b) is that $L^1(\mu)$ will be viewed as a metric space with distance $\rho(f,g) = \int_X |f-g|\,d\mu$, but then $\rho(f,g) = 0$ iff $f = g$ holds only if we stop distinguishing between functions in the same equivalence class.

(d) In fact, with the convention in (b), $L^1(\mu)$ becomes a normed space with the norm $||f||_1 := \int_X |f|\,d\mu$, and it can be shown it's complete (i.e., $L^1(\mu)$ is a Banach space).

2.7 $L^1(\mu)$: basic properties

Proposition 2.22. *Let $f, g \in L^1(\mu)$.*

(i) (Linearity) For any constants $\alpha, \beta \in \mathbb{C}$, we have $\alpha f + \beta g \in L^1(\mu)$ and

$$\int_X (\alpha f + \beta g)\,d\,\mu = \alpha \int_X f\,d\mu + \beta \int_X g\,d\mu.$$

(ii) (Monotonicity) If $f \leq g$ (a.e.) on E then $\int_E f\,d\mu \leq \int_E g\,d\mu$.

(iii) (Triangle inequality) $\left|\int_X f\,d\mu\right| \leq \int_X |f|\,d\mu$.

(iv) For any $f \in L^1$, $E \mapsto \int_E f\,d\mu$ is a complex measure, in particular, it is σ-additive:

$$\int_E f\,d\mu = \sum_{j=1}^{\infty} \int_{E_j} f\,d\mu$$

if E is the union of disjoint μ-measurable sets E_j.

Remark 2.23. The measure $E \mapsto \int_E f\,d\mu$ in (iv) can sometimes be denoted by $f\,d\mu$.

Proof. (i) $\alpha f + \beta g \in L^1(\mu)$ follows from $|\alpha f + \beta g| \leq |\alpha|\,|f| + |\beta|\,|g|$. To show linearity for the real case, we let $h = f + g$, which can be rewritten as $h^+ + f^- + g^- = f^+ + g^+ + h^-$, which means $\int h^+ + \int f^- + \int g^- = \int f^+ + \int g^+ + \int h^-$ by linearity for positive functions. Then for complex case, one just splits each function into real and imaginary part. For constant multiplication arguments are similar.

(ii) is easy by the similar arguments.

(iii) For real functions, this is immediate from $-|f| \leq f \leq |f|$ and (ii). For complex f, let $\alpha = \overline{\mathrm{sgn}(\int f)}$. Then $|\int f| = \alpha \int f = \int \alpha f$. So $\int \alpha f$ is real and $|\int f| = \mathrm{Re} \int \alpha f = \int \mathrm{Re}(\alpha f) \leq \int |\alpha f| = \int |f|$.

(iv) For simple functions in L^+, this follows from the definition. For general L^+ functions, it follows from $\chi_E f = \sum_{j=1}^{\infty} \chi_{E_j} f$ and then Theorem 2.16. For general complex $f \in L^1(\mu)$, one uses Theorem 2.26 instead (proved below). ☐

2.8 $L^1(\mu)$: Dominated Convergence Theorem

Theorem 2.24 (Lebesgue's Dominated Convergence Theorem)**.** *Suppose:*

(i) $\lim_{n\to\infty} f_n(x) = f(x)$ *for (almost) every* $x \in X$*;*

(ii) $|f_n(x)| \leq g(x)$ *for some* $g \in L^1(\mu)$*.*

Then $f \in L^1(\mu)$ *and*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu$$

Moreover, $\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0$.

Remarks 2.25. (a) $\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0$ can also be written as $||f_n - f||_1 \to 0$ and is called L^1-convergence or convergence in L^1-norm.

Proof. [F55]

Intuition for the proof: Fatou's lemma says that for positive functions some mass can escape to infinity in the limit, but nothing extra can appear. Our f_j is not positive, but $f_j + g \geq 0$ is, so we can get the same conclusion. Now, we can apply the same argument to $-f_j$ (or rather, to $-f_j + g \geq 0$), so Fatou's lemma will say that no extra mass can appear for $-f_j$, that is, no mass can escape for f_j.

Rigorously: we assume f_n are real-valued, as otherwise we just take real and imaginary parts. Then $g + f_n \geq 0$ and $g - f_n \geq 0$, so by Fatou's lemma, we get $\int(g+f) \leq \liminf \int(g+f_n)$ and $\int(g-f) \leq \liminf \int(g-f_n)$. These two inequalities give us: $\liminf \int f_n \geq \int f \geq \limsup \int f_n$, which implies $\lim \int f_n = \int f$. Finally, to show $\int |f_n - f| \to 0$, just apply the previous case to the sequence $|f_n - f|$ which is dominated by $2g \in L^1(\mu)$. ☐

Theorem 2.26 (Interchanging integral and sum for L^1)**.** *Suppose* $\sum_{j=1}^\infty \int |f_j| \, d\mu < \infty$*. Then* $\sum_{j=1}^\infty f_j(x)$ *converges a.e. to a function* $f(x) \in L^1(\mu)$ *and*

$$\int_X \left(\sum_{j=1}^\infty f_j(x) \right) d\mu = \sum_{j=1}^\infty \int_X f_j(x) \, d\mu.$$

Moreover, $||f - \sum_{j=1}^n f_j||_1 \to 0$.

Proof. [F55]

First of all, $g = \sum_{j=1}^\infty |f_j(x)|$ is in L^+ and by Theorem 2.16, $\int \sum = \sum \int$, so $g \in L^1(\mu)$. $L^1(\mu)$ functions are μ-almost everywhere finite, so g is a.e. convergent. Since $|f| \leq g$, we also get $f \in L^1$ and f a.e. finite.

Now note that $|\sum_{j=1}^n f_j| \leq g$ for any n, so by the Dominated Convergence Theorem we can interchange $\int$ and $\sum$.

$||f - \sum_{j=1}^n f_j||_1 \to 0$ follows by applying the Dominated Convergence Theorem to $f - \sum_{j=1}^n f_j$. ☐

2.9 Lebesgue vs Riemann integrals

Intuition. Riemann integral is constructed by dividing the domain of the input function f into subintervals. One can think of the approximation by simple functions in the Lebesgue integral (see Section 2.3) as the division of the *range* of the function f into subintervals. Somehow Lebesgue's approach works better because it takes into account properties of f, whereas Riemann's division completely ignores it.

Theorem 2.27. *Let* $f : [a,b] \to \mathbb{R}$.

(a) If f is Riemann integrable on $[a,b]$ then f is Lebesgue integrable and $\int_a^b f(x)\,dx = \int_{[a,b]} f\,dm$.

(b) f is Riemann integrable iff f is bounded and the set of discontinuity points has Lebesgue measure zero.

Proof. [F57], [F,Ex.2.3.23]

(i) Take a sequence of meshes that include the preceding ones. For each take the piecewise constant upper and lower functions for the Riemann integral $\overline{f}_n$ and $\underline{f}_n$. Since these are monotone and bounded, the limits $\overline{f}(x) = \lim \overline{f}_n(x)$ and $\overline{f}(x) = \lim \underline{f}_n(x)$ exist for all x. The corresponding lower and upper Darboux sums for $\overline{f}_n$ and $\underline{f}_n$ agree with their Lebesgue integrals. By Dominated Convergence Theorem, Lebesgue integral and limit can be interchanged, so we get $\int \overline{f}\,dm = \lim \int \overline{f}_n\,dm$. But the latter is equal to the limit of the upper Darboux sums, that is equal to the Riemann integral $\int_a^b f\,dx$. Similarly for $\int \underline{f}\,dm$. This means $\overline{f} = \underline{f}$ a.e., and therefore $\overline{f} = \underline{f} = f$ a.e. since $\overline{f} \geq f \geq \underline{f}$. This shows that f is Lebesgue measurable and Lebesgue integral of f is equal to its Riemann integral.

(ii) is delegated to exercises. ☐

2.10 Product measures

Recall the definition of the product σ-algebra $\mathcal{M} \otimes \mathcal{N}$ in Section 1.3.

Now, for any set $E \in X \times Y$, let

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

Theorem 2.28. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measure spaces. For any set $E \in \mathcal{M} \otimes \mathcal{N}$, the function $x \mapsto \nu(E_x)$ is $\mathcal{M}$-measurable and the function $y \mapsto \mu(E^y)$ is $\mathcal{N}$-measurable, and*

$$\int \nu(E_x)\,d\mu(x) = \int \mu(E^y)\,d\nu(y). \tag{2.10.1}$$

Remarks 2.29. (a) Because of Remark 2.7, the equality (2.10.1) can equivalently be rewritten as

$$\int \left(\int \chi_E(x,y)\,d\nu(y) \right) d\mu(x) = \int \left(\int \chi_E(x,y)\,d\mu(x) \right) d\nu(y),$$

so can be regarded as a baby Fubini theorem.

(b) The condition of σ-finiteness is important (see [F, Ex. 2.5.45]).

Proof. [F64–67]

For simplicity, we will suppose both μ and ν are finite and assume everything is measurable (which is not obvious, but hopefully quite believable).

Denote $\mathcal{C}$ to be the set of all $E \in \mathcal{M} \otimes \mathcal{N}$ for which (2.10.1) holds. For $A \times B$ with measurable A, B, this is clear. If we show that $\mathcal{C}$ is a σ-algebra, then $\mathcal{C}$ contains the minimal σ-algebra containing all of $A \times B$, and therefore $\mathcal{C} \supseteq \mathcal{M} \otimes \mathcal{N}$, so we are done.

To show that $\mathcal{C}$ is a σ-algebra:

Step 1: $\mathcal{C}$ is a "monotone class" (collection of sets closed under countable increasing unions and countable decreasing intersections): indeed, let $\{E_n\}$ in $\mathcal{C}$ increases and E is their union. Then $\nu((E_n)_x)$ increase pointwise to $\nu(E_x)$. So Monotone Convergence Theorem can be applied. For decreasing $\{E_n\}$, one applies Dominated Convergence Theorem.

Step 2: Then show that the minimally generated monotone class is the minimally generated σ-algebra. Details are omitted [F66]. ☐

Definition 2.30. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measure spaces. On the measurable space $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ we define the* ***product measure*** *$\mu \times \nu$ to be the measure given by*

$$(\mu \times \nu)(E) := \int \nu(E_x)\, d\mu(x) = \int \mu(E^y)\, d\nu(y) \tag{2.10.2}$$

for $E \in \mathcal{M} \otimes \mathcal{N}$.

Remark 2.31. That the two quantities in (2.10.2) are equal is established in Theorem 2.28. That $\mu \times \nu$ is really a measure (that is, that it's countably additive) can be proved by using Theorem 2.16.

2.11 Tonelli and Fubini theorems

Theorem 2.32. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measure spaces. Then*

(i) (Tonelli Theorem = Fubini Theorem for positive functions) If $f \in L^+(X \times Y)$, then

$$\int f\, d(\mu \times \nu) = \int \left(\int f(x, y)\, d\nu(y) \right) d\mu(x) = \int \left(\int f(x, y)\, d\mu(x) \right) d\nu(y) \tag{2.11.1}$$

(ii) (Fubini Theorem) If $f \in L^1(X \times Y)$, then (2.11.1) *holds.*

Remarks 2.33. (a) The theorems are also implicitly stating that all the integrand functions in (2.11.1) are a.e. integrable.

(b) Conditions are important: σ-finiteness, as well as of course $f \in L^1(X \times Y)$.

(c) Typically, one first uses Tonelli theorem to check the condition $f \in L^1(X \times Y)$, at which point one is allowed to use Fubini.

Proof. [F67–68]

(i) If f is a characteristic function, then we proved this in Theorem 2.28. For general $f \in L^+$, we approximate f with simple $\phi_n \nearrow f$, and then use Monotone Convergence Theorem.

(ii) If f is real, then split $f = f^+ - f^-$ and then apply Tonelli to each f^+ and f^-. Similarly if f is complex, then split into real and imaginary parts. ☐

2.12 Absolutely continuous and singular measures

Definition 2.34. *Let μ be a positive measure on $(X, \mathcal{M})$. A (positive or signed) measure ν on $(X, \mathcal{M})$ is called* ***absolutely continuous with respect to*** *μ if $\mu(A) = 0$ implies $\nu(A) = 0$. We write $\nu \ll \mu$.*

Remark 2.35. Often one says that a measure is "absolutely continuous" if it is a measure on $\mathbb{R}$ (or $\mathbb{R}^n$) that is absolutely continuous with respect to the Lebesgue measure m (or m^n, respectively).

Example 2.36. If $f \in L^1(\mu)$, then as we saw in Proposition 2.22(iv), $E \mapsto \int_E f\, d\mu$ (or $f\, d\mu$ for short) is a complex measure. It is clear (see Proposition 2.12(v)) that this measure is absolutely continuous with respect to μ. As we show in Section 2.14 the converse is also true!

Proposition 2.37. *Let μ be a positive measure on $(X, \mathcal{M})$. Let $\nu : \mathcal{M} \to (-\infty, +\infty)$ be a signed measure. Then $\nu \ll \mu$ iff for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu(E) < \delta$ implies $|\nu(E)| < \varepsilon$.*

Proof. For signed measure, the proof will follow once we have the Jordan Decomposition Theorem below. So let us just prove the case for positive measures ν.

$\Leftarrow$ is clear, so let us prove $\Rightarrow$. If the ε-δ condition fails then there exists an $\varepsilon > 0$ such that for some sets E_n: $\nu(E_n) \geq \varepsilon$ and $\mu(E_n) < 2^{-n}$. Let $F_k = \cup_{n=k}^{\infty} E_n$ and $F = \cap_{j=1}^{\infty} F_j$. Then $\mu(F_k) < \sum_k^{\infty} 2^{-n} = 2^{1-k}$, so $\mu(F) = 0$. However $\nu(F_k) \geq \varepsilon$, so by continuity from above (Theorem 1.19(iv)), we get $\nu(F) \geq \varepsilon$ since ν is finite. ☐

Corollary 2.38. *If $f \in L^1(\mu)$ then for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu(E) < \delta$ implies $|\int_E f\, d\mu| < \varepsilon$.*

Definition 2.39. *Let μ be a (positive or signed) measure on $(X, \mathcal{M})$. A (positive or signed) measure ν on $(X, \mathcal{M})$ is called* ***singular with respect to*** *μ if there exists a measurable set $A \in \mathcal{M}$ such that $\mu(A) = 0$ and $\nu(A^c) = 0$. We write $\mu \perp \nu$.*

Remark 2.40. Since being singular is a symmetric relation, we also say that μ and ν are mutually singular.

Proposition 2.41. *Let all the measures below be defined on the same measurable space.*

(i) If $\nu \ll \mu$ and $\mu \ll \lambda$ then $\nu \ll \lambda$.

(ii) If $\nu_1 \ll \mu$ and $\nu_2 \ll \mu$ then $\nu_1 + \nu_2 \ll \mu$.

(iii) If $\nu_1 \perp \mu$ and $\nu_2 \perp \mu$ then $\nu_1 + \nu_2 \perp \mu$.

(iv) If $\nu \ll \mu$ and $\nu \perp \mu$ then $\nu = 0$.

2.13 Decompositions for signed measures

Intuition. Basically every signed measure can be decomposed into difference of two positive measures.

We will say that a set E is positive for a signed measure ν if $\nu(F) \geq 0$ for all measurable $F \subseteq E$. Similarly for negative.

Theorem 2.42 (Hahn Decomposition Theorem)**.** *If ν is a signed measure that there exist two disjoint sets P and N whose union is the whole space X such that P is positive for ν and N is negative for ν.*

Remark 2.43. The decomposition is unique up to the ν-null sets.

Proof. [F86–87]

Assume $\nu : \mathcal{M} \to [-\infty, +\infty)$. Let $m \in \mathbb{R}$ be $\sup \nu(E)$ over all positive sets E. Then $\nu(P_j) \to m$ for some positive $\{P_j\}$. Let $P = \cup P_j$. By countable additivity, it is easy to see that P is positive for ν, and therefore by continuity from below for measures (or MCT), $\nu(P) = m$. We only need to prove that P^c is negative for ν.

Suppose not, that is, $\nu(A) \in (0, \infty)$ for some $A \subseteq P^c$. Then A cannot be positive as otherwise $P \cup A$ would be positive set of larger measure than m. Then we can find $A_1 \subseteq A$ with $\nu(A_1) > \nu(A)$ (because A is not positive, so $\nu(C) < 0$ for some $C \subseteq A$, so we can take $A_1 = A \setminus C$). In fact, let us choose A_1 and n_1 with $\nu(A_1) > \nu(A) + \frac{1}{n_1}$ with minimal $n_1 \in \mathbb{N}$. Repeat the construction to find $A_j \subseteq A_{j-1}$ with $\nu(A_j) > \nu(A_{j-1}) + \frac{1}{n_j}$ and $n_j \in \mathbb{N}$ minimal. Let $A_\infty = \cap_j A_j$. Since $\infty > \nu(A_\infty) = \lim \nu(A_j) > \sum n_j^{-1}$, we get $n_j \to \infty$. But we can apply the construction above again to find $B \subseteq A_\infty$ such that $\nu(B) > \nu(A_\infty) + \frac{1}{n}$ for some integer $n \in \mathbb{N}$. Since $n_j \to \infty$, we get $n_j > n$ for some j, which contradicts the construction of n_j and A_j. ☐

Theorem 2.44 (Jordan Decomposition Theorem). *If ν is a signed measure, then there exists positive measures ν^+ and ν_- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.*

Remarks 2.45. (a) From the "almost" uniqueness of the Hahn Decomposition follows that Jordan Decomposition if in fact unique.

(b) At least one of ν^+ and ν_- is a finite measure.

Proof. [F87]

Define $\nu^+(E) = \nu(E \cap P)$ and $\nu^-(E) = |\nu(E \cap N)|$, where P and N are as in the proof of the last theorem. □

2.14 Radon–Nikodym theorem

Intuition. This theorem is the inverse of the observation we made in Example 2.36.

Theorem 2.46 (Radon–Nikodym Theorem). *Let μ and ν be two σ-finite positive measures on $(X, \mathcal{M})$. $\nu \ll \mu$ iff there exists $f \in L^+$, such that*

$$d\nu = f\, d\mu$$

(in the sense $\int_E d\nu = \int_E f\, d\mu$ for any measurable E). Moreover, ν is finite iff $f \in L^1(\mu)$.

Remarks 2.47. (a) Same arguments lead to similar results for complex measures.

(b) It can be shown that f is unique (μ-a.e.).

(c) The function f is referred to as the Radon–Nikodym derivative and is denoted by $\frac{d\nu}{d\mu}$.

Proof. [F90]

Suppose that both ν and μ are finite, as for the σ-finite case we just divide X into countable union of disjoint sets.

Let

$$\mathcal{F} = \left\{ f \in L^+ : \int_E f\, d\mu \leq \int_E d\nu \text{ for all } E \in \mathcal{M} \right\}.$$

Note that $f, g \in \mathcal{F}$ implies $h = \max(f, g) \in \mathcal{F}$. Indeed, if $A = \{x : f(x) > g(x)\}$, then $\int_E h d\mu = \int_{E\cap A} f\, d\mu + \int_{E\setminus A} g\, d\mu \leq \nu(E \cap A) + \nu(E \setminus A) = \nu(E)$.

We want to choose "maximal" f. To this end, let $a = \sup\{\int_X f\, d\mu : f \in \mathcal{F}\}$ ($a \leq \nu(X) < \infty$), and choose $\{f_n\} \subset \mathcal{F}$ such that $\int_X f_n\, d\mu \to a$. By redefining $f_n := \max\{f_1, \ldots, f_n\}$, we can assume f_n is non-decreasing. By Monotone Convergence Theorem, $\lim f_n(x) =: f(x)$ belongs to $\mathcal{F}$ and has $\int f\, d\mu = a$.

Define a measure $d\lambda := d\nu - f\, d\mu$ which is a positive measure. It must in fact be zero as otherwise we could add a bit more to f. Rigorously: let $P_n \cup N_n$ be the Hahn decompositon for $d\nu - f\, d\mu - \frac{1}{n} d\mu$. If $\mu(P_n) > 0$, then $f + \frac{1}{n}\chi_E \in \mathcal{F}$ which contradicts to the maximality of f. Therefore $\mu(P_n) = 0$, and so $\mu(P) = 0$ (where $P = \cup P_n$) by continuity from below. By absolute continuity of ν, we get $d\nu - f\, d\mu$ is 0 on P. For the negative sets, let $N = \cap N_j = P^c$. Then $N \subseteq N_j$ for every j, so $0 \leq \int_N (d\nu - f\, d\mu) \leq \frac{1}{n}\mu(N)$ which implies $d\nu - f\, d\mu$ is 0 on N. □

2.15 Lebesgue Decomposition theorem

Theorem 2.48 (Lebesgue Decomposition Theorem)**.** *Let μ and ν be two σ-finite positive measures on $(X, \mathcal{M})$. Then there exist unique measures ν_{ac} and ν_s such that*

$$\nu = \nu_{ac} + \nu_s,$$

such that $\nu_{ac} \ll \mu$ and $\nu_s \perp \mu$.

Remarks 2.49. (a) Same arguments lead to similar results for complex measures.

(b) Combining with the Radon–Nikodym theorem, we obtain that for any two σ-finite measures, one has $d\nu = f\, d\mu + d\nu_s$.

Proof. [F90]

Again, arguing as in the previous theorem, we can suppose that both ν and μ are finite.

Now observe that $\nu \ll \nu + \mu$. By the Radon–Nikodym theorem there exists a measurable function $f \in L^+$ such that $\nu(E) = \int_E f\, d\nu + \int_E f\, d\mu$. Rearranging, we get $\int_E (1-f)\, d\nu = \int_E f\, d\mu \geq 0$, so $1 - f \geq 0$ ν-a.e. Let $A := \{x : f(x) = 1\}$ and $B := \{x : 0 \leq f(x) < 1\}$. Then $\nu(A) = \int_A 1\, d\mu + \int_A 1\, d\nu = \mu(A) + \nu(A)$, so $\mu(A) = 0$.

Now define $\nu_s(E) := \nu(E \cap A)$ and $\nu_{ac} := \nu(E \cap A^c)$. Then $\nu_s \perp \mu$ because $\mu(A) = 0$. We just need to show that $\nu_{ac} \ll \mu$. If $\mu(E) = 0$, then $\nu_{ac}(E) = \int_{E\cap B} d\nu = \nu(E \cap B) = \int_{E\cap B} f\, d\nu + 0$. This means that $\int_{E\cap B}(1-f)d\nu = 0$. Since $1 - f > 0$ ν-a.e., we get $\nu(E \cap B) = 0$, that is $\nu_{ac}(E) = 0$.

To prove uniqueness: if $\nu_s + \nu_{ac} = \widetilde{\nu}_s + \widetilde{\nu}_{ac}$, then $\nu_s - \widetilde{\nu}_s = \widetilde{\nu}_{ac} - \nu_{ac}$ is both singular and a.c. with respect to μ. This implies it is 0 by Proposition 2.41(iv). ☐

2.16 Lebesgue–Stieltjes measures on $\mathbb{R}$: decomposition with respect to the Lebesgue measure

Let μ be a Borel measure on $\mathbb{R}$ that is finite on compacts. As we know from Section 1.11, μ is a Lebesgue–Stieltjes measure μ_F for some non-decreasing right-continuous function $F : \mathbb{R} \to \mathbb{R}$. The Lebesgue integral $\int_E f d\mu_F$ (for $f \in L^1(\mu_F)$) is called the Lebesgue–Stieltjes integral, often denoted by $\int_E f(x)\, dF(x)$.

Remarks 2.50. (a) For a non-decreasing right-continuous function $F : \mathbb{R} \to \mathbb{R}$, one can use the construction analogous to the Riemann integration but instead of Δx_j one uses $F(x_{j+1}) - F(x_j)$. This construction leads to the so-called Riemann–Stieltjes integral. The relation between the Riemann–Stieltjes integral and the Lebesgue–Stieltjes integral is similar to the relation between the Riemann and the Lebesgue integral.

(b) One can relax the requirement that F is non-decreasing (or even real-valued) but instead require that F has bounded total variation. The corresponding measure μ_F is then signed (or complex, respectively).

Definition 2.51. *(i) We say that $x \in \mathbb{R}$ is a **pure point** (or mass point, or point mass) of μ if $\mu(\{x\}) > 0$.*

*(ii) We call μ a **continuous measure** if $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}$ (that is, if μ has no pure points).*

*(iii) We call μ is a **pure point measure** if $\mu(A) \neq 0$ implies that there exists $x \in A$ such that $\mu(\{x\}) > 0$.*

Remark 2.52. Since μ is finite on compacts, one can easily show that μ is a pure point measure iff $\mu(A) = \sum_{x\in A} \mu(\{x\})$ for every A. Put it another way, iff μ is a countable (or finite) linear combination of Dirac measures.

Theorem 2.53. *Let μ be a Borel measure on $\mathbb{R}$ that is finite on compacts. Then there exist unique measures μ_{ac}, μ_{sc}, μ_{pp} such that*

$$\mu = \mu_{ac} + \mu_{sc} + \mu_{pp},$$

where $\mu_{ac} \ll m$, μ_{sc} is $\perp m$ and continuous, and μ_{pp} is $\perp m$ and pure point.

Remark 2.54. We call μ_{ac}, μ_{sc}, μ_{pp} the "absolutely continuous", "singular continuous", and "pure point" parts of μ, respectively.

Proof. Let P be the set of pure points of μ. Then define $\mu_{pp}(A) := \sum_{x\in A\cap P} \mu(\{x\})$ and $\mu_{cont} := \mu - \mu_{pp}$. Clearly μ_{pp} and μ_{cont} are pure point and continuous measures, respectively. Then apply the Lebesgue decomposition to μ_{cont} with respect to m. □

Examples 2.55. (a) Let F be continuously-differentiable. Then one can show that $\int_E f(x)dF(x) = \int_E f(x)F'(x)dm$. In other words, $d\mu_F = F'(x)dm$ is absolutely-continuous with singular-continuous and pure point parts being zero. This can be generalized to the situations where F' exists only m-a.e. with $F' \in L^1(m)$, plus an extra condition on F (that is, fittingly, called "absolute continuity"). As example (c) below shows, this extra condition is in fact required.

(b) Let $F(x)$ be 0 for $x < 0$ and 1 for $x \geq 0$. Then it is clear that $\mu_F((a, b]) = 1$ if $0 \in (a, b]$ and 0 otherwise. In other words μ_F is δ_0, the Dirac measure at $x = 0$. From the definition of the integral, one can easily check that $\int_{\mathbb{R}} f(x)dF = f(0)$ (functions in the space $L^1(\mu_F)$ in this case are completely determined by $f(0)$, i.e., $L^1(\mu_F)$ is just a one-dimensional vector space).

More generally, if F (right-continuous and non-decreasing) is piecewise constant, then each jump discontinuity $F(x_j + 0) - F(x_j - 0) = \alpha_j$ of F will contribute $\alpha_j\delta_{x_j}$ to μ_F.

(c) Let $F(x)$ be the Cantor function (recall that it's continuous on $[0, 1]$) continued by $F(x) = 0$ for $x \leq 0$, $F(x) = 1$ for $x \geq 1$. We call the corresponding measure μ_F the Cantor measure. By continuity of F, we get $\mu_F(\{x\}) = 0$ for every $x \in \mathbb{R}$, so μ_F is continuous. Notice that on $[0, 1] \setminus C$, F is constant, which implies that $\mu_F([0, 1] \setminus C) = 0$. Since $m(C) = 0$, we obtain that the Cantor measure μ_F is singular with respect to the Lebesgue measure: $\mu_F \perp m$. In other words, μ_F is singular-continuous: $\mu_F = (\mu_F)_{sc}$.

2.17 Lebesgue integrals in $\mathbb{R}^n$: Hardy–Littlewood maximal function

Intuition. In the next few sections we will restrict ourselves to $\mathbb{R}^n$ and head towards the analogue of the Fundamental Theorem of Calculus for the Lebesgue integrals. Note that $\frac{d}{dx}\int_a^x f(x)dx$ can be rewritten as $\lim_{r\to 0} \frac{1}{m(I_r)} \int_{I_r} f(x)dm$, where I_r is the interval with endpoints x and $x+r$. Instead of I_r we can also use the interval with endpoints $x - r$ and $x + r$, of course. This motivates the objects we work with in this section.

In Sections 2.17–2.23, we will start writing dx for the (n-dimensional) Lebesgue measure instead of $dm(x)$ or $d^n m(x)$. Also, "measurable", "integrable", and "a.e." will be meant with respect to the Lebesgue measure.

Definition 2.56. *A function $f : \mathbb{R}^n \to \mathbb{C}$ is called locally integrable (denoted by $f \in L^1_{loc}$) if f is Lebesgue measurable and $\int_K |f(x)|\, dx < \infty$ for every bounded measurable set $K \subset \mathbb{R}^n$.*

Definition 2.57. *For $f \in L^1_{loc}$:*

(i) Define $A_r f(x)$ be the average value of f over $B(x, r) := \{y \in \mathbb{R}^n : |x - y| < r\}$:

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

(ii) Define the ***Hardy–Littlewood maximal function*** $Hf(x)$ *by*

$$Hf(x) = \sup_{r>0} A_r|f|(x) = \sup_{r>0} \frac{1}{m(B(r,x))} \int_{B(r,x)} |f(y)|\, dy$$

Remarks 2.58. (a) It can be shown that $A_r f$ (for any $r > 0$) and Hf are measurable [F96].

(b) Note that A_r is linear but H_r is not!

The next theorem says that for $f \in L^1$, Hf cannot be too large on a large set.

Theorem 2.59 (Maximal inequality)**.** *For any $f \in L^1$ and any $\alpha > 0$,*

$$m(\{x : Hf(x) > \alpha\}) \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f(x)| dx$$

Proof. [F96]

First we prove the following lemma.

Lemma 2.60 (Covering lemma)**.** *Let U be the union of a collection of open balls. For any $c < m(U)$, we can find a disjoint finite subcollection $B_1, \ldots, B_k$ of balls such that $c < 3^n \sum_{j=1}^k m(B_j)$.*

Proof. Since m is regular, in particular inner regular, we can find a compact $K \subseteq U$ such that $c < m(K)$. By compactness we can find a finite subcover with the balls $B(x_j, r_j)$, $j = 1, \ldots, N$. Without loss of generality, assume $B(x_j, r_j)$ are ordered so that $r_1 \geq r_2 \geq \ldots \geq r_N$. Choose B_1 to be $B(x_1, r_1)$ and discard all the rest $B(x_j, r_j)$ that intersect B_1. Choose B_2 the next remaining $B(x_j, r_j)$ and so on until we run out of the balls. The resulting collection $B_1, \ldots, B_k$ is then disjoint. If B_j^* is B_j but with tripled radius, then from the construction, it is easy to see that $\bigcup_{j=1}^k B_j^*$ contains all of the $\{B(x_j, r_j)\}_{j=1}^N$, and so K. Therefore

$$c < m(K) \leq \sum_{j=1}^k m(B_j^*) = 3^n \sum_{j=1}^k m(B_j).$$

□

Let $E_\alpha = \{x : Hf(x) > \alpha\}$. For each $x \in E_\alpha$ there is some radius $r_x > 0$ such that $A_{r_x}|f|(x) > \alpha$, that is, $m(B(r_x, x)) < \frac{1}{\alpha} \int_{B(r_x,x)} |f| dx$. The balls $B(r_x, x)$ cover E_α. By the covering lemma, for any $c < m(E_\alpha)$, we get a finite disjoint subcover such that $c < 3^n \sum_{j=1}^k m(B_j) \leq \frac{3^n}{\alpha} \sum_{j=1}^k \int_{B_j} |f(y)| dy \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f(y)| dy$. □

2.18 Lebesgue integrals in $\mathbb{R}^n$: Lebesgue differentiation theorem, weak version

Theorem 2.61 (Lebesgue Differentiation Theorem, weak version)**.** *If $f \in L^1_{loc}$, then $\lim_{r \to 0} A_r f(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$, that is*

$$\lim_{r \to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} f(y)\, dy = f(x) \text{ a.e.} \tag{2.18.1}$$

Remark 2.62. Equivalently,

$$\lim_{r \to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} (f(y) - f(x))\, dy = 0 \text{ a.e.} \tag{2.18.2}$$

Proof. [F97]

Let us prove the statement of the theorem for a continuous function g first. By continuity for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|y - x| < \delta$ implies $|g(x) - g(y)| < \varepsilon$. Therefore for $r < \delta$:

$$|A_r g(x) - g(x)| = \frac{1}{m(B(r,x))} \left| \int_{B(r,x)} (g(y) - g(x)) dy \right| < \varepsilon.$$

This shows $\lim_{r\to 0} A_r g(x) = g(x)$ for any x.

Now for the general f, first note that the statement of the theorem depends on the local properties only, so we may assume $f \in L^1$ instead of L^1_{loc}.

Now, we claim that for any $\varepsilon > 0$ we can find $g \in C_c(\mathbb{R}^n)$ (continuous function with compact support) such that $||f - g||_1 < \varepsilon$. This will be proved later in larger generality (for regular measures and for any of the L^p spaces) in Section 3.24.

Thus $\limsup_{r\to 0} |A_r f(x) - f(x)| = \limsup_{r\to 0} |A_r(f - g) + (A_r g - g) + (g - f)| \le H(f - g) + 0 + |f - g|$.

Thus, $F_\alpha := \{x : \limsup |A_r f - f| > \alpha\} \subseteq \{x : |f - g| > \alpha/2\} \cup \{x : H(f - g) > \alpha/2\}$. But $m(\{x : H(f - g) > \alpha/2\}) \le \frac{2\cdot 3^n}{\alpha} ||f - g||_1$ and $m(\{x : |f - g| > \alpha/2\}) \le \frac{2}{\alpha} ||f - g||_1$ (Chebyshev's inequality), so that $m(F_\alpha) \le C\varepsilon$ for some constant $C > 0$. Since $\varepsilon > 0$ is arbitrary, we get $m(F_\alpha) = 0$ for any $\alpha > 0$. Therefore

$$m(\{x : \limsup_{r\to 0} |A_r f - f| > 0\}) = m(\cup_{n=1}^{\infty} F_{1/n}) = 0.$$

□

2.19 Lebesgue integrals in $\mathbb{R}^n$: Lebesgue points and the Lebesgue set

Definition 2.63. *(i) For $f \in L^1_{loc}$, if*

$$\lim_{r\to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} |f(y) - f(x)| \, dy = 0 \tag{2.19.1}$$

*holds at a point $x \in \mathbb{R}^n$, then we call x a **Lebesgue point** of f.*

*(ii) The **Lebesgue set** L_f of $f \in L^1_{loc}$ is defined to be the set of all Lebesgue points of f.*

Remark 2.64. The condition (2.19.1) is strictly stronger then the condition (2.18.2), so at this point we don't know if any Lebesgue point exists.

2.20 Lebesgue integrals in $\mathbb{R}^n$: Lebesgue differentiation theorem, strong version

Theorem 2.65 (Lebesgue Differentiation Theorem, strong version)**.** *If $f \in L^1_{loc}$, then*

$$\lim_{r\to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} |f(y) - f(x)| \, dy = 0 \tag{2.20.1}$$

holds for a.e. x, that is, $m((L_f)^c) = 0$.

Proof. [F98]

Let $\mathbb{Q}^2$ be the collection of all points in $\mathbb{C}$ with rational real and imaginary parts. For each $c \in \mathbb{Q}^2$, apply Theorem 2.61 to the function $|f(x) - c|$: except for x in a m-null set E_c, we have

$$\lim_{r \to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} |f(y) - c| \, dy = |f(x) - c|. \tag{2.20.2}$$

Let $E = \bigcup_{c \in \mathbb{Q}^2} E_c$. Then $m(E) = 0$, and we claim that E^c contains only Lebesgue points of f. Indeed, let $x \notin E$, so that (2.20.2) holds for every $c \in \mathbb{Q}^2$. We find $c \in \mathbb{Q}^2$ within $\varepsilon > 0$ distance from $f(x)$. Then $|f(x) - c| \leq \varepsilon$ and $|f(y) - f(x)| \leq |f(y) - c| + \varepsilon$, so that $\limsup_{r \to 0} \frac{1}{m(B(r,x))} \int_{m(B(r,x))} |f(y) - f(x)| dy \leq |f(x) - c| + \frac{1}{m(B(r,x))} \varepsilon m(B(r,x)) \leq 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we get that $x \notin E$ is a Lebesgue point. □

2.21 Lebesgue integrals in $\mathbb{R}^n$: Lebesgue differentiation theorem, generalized strong version

Intuition. Recall that in the Fundamental Theorem of Calculus (see *Intuition* in Section 2.17), we could have chosen an interval centered at x, or an interval that touches x. So it seems that there should be nothing special about the centered ball in (2.18.2) or (2.20.1). This in indeed the case, but there'll be some minor requirement on the domain of integration.

Definition 2.66. *We say that a family $\{E_r\}_{r>0}$ of Borel sets in $\mathbb{R}^n$* ***shrinks to x nicely*** *if $E_r \subseteq B(r,x)$ for each $r > 0$ and there is a constant $\alpha > 0$ such that $m(E_r) > \alpha m(B(r,x))$.*

Remarks 2.67. (a) E_r doesn't need to contain x.

(b) E.g., we can take $E_r = \{x + ry : y \in U\}$, where U is some chosen Borel subset of $B(1,0)$ with $m(U) > 0$.

Theorem 2.68 (Lebesgue Differentiation Theorem, generalized strong version)**.** *If $f \in L^1_{loc}$, then for any Lebesgue point of f,*

$$\lim_{r \to 0} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)| \, dy = 0,$$
$$\lim_{r \to 0} \frac{1}{m(E_r)} \int_{E_r} f(y) \, dy = f(x),$$

for any family $\{E_r\}_{r>0}$ that nicely shrinks to x.

Proof. [F99–100]

This follows from the previous theorem and the trivial bounds $m(E_r) \geq \alpha m(B(r,x))$ and $\int_{E_r} \leq \int_{B(r,x)}$. □

2.22 Lebesgue integrals in $\mathbb{R}^n$: differentiation of measures

Intuition. Note that the result of the (weak) Lebesgue differentiation theorem can be viewed as $\lim_{r \to 0} \frac{\mu(B(r,x))}{m(B(r,x))} = f(x)$ where $d\mu = f(x) dm$ (note that $f(x) = \frac{d\mu}{dx}$). It is natural to ask what happens when μ is not necessarily absolutely continuous. As we show in the next theorem, the singular component disappears in this limit.

Theorem 2.69. *Let μ be a Borel measure on $\mathbb{R}^n$ that is finite on compacts. Let $d\mu = f dm + d\mu_s$ be its Lebesgue–Radon–Nikodym decomposition. Then*

$$\lim_{r\to 0} \frac{\mu(E_r)}{m(E_r)} = f(x) \quad \text{for a.e. } x \in \mathbb{R}^n$$

and for any family $\{E_r\}_{r>0}$ that shrinks nicely to x.

Remarks 2.70. (a) In particular, if $\mu \perp m$, then $\lim_{r\to 0} \frac{\mu(B(r,x))}{m(B(r,x))} = 0$ for m-a.e. $x \in \mathbb{R}^n$.

(b) It is curious (but not that surprising if one thinks about it!) that if $\mu \perp m$, then $\lim_{r\to 0} \frac{\mu(B(r,x))}{m(B(r,x))} = \infty$ for $\underline{\mu\text{-a.e.}}$ $x \in \mathbb{R}^n$ (see [Rudin, 7.15] for a proof).

Proof. Since $d\mu = f dm + d\mu_s$, and we know this result for the a.c. measures, we just need to show that the limit above is zero a.e. for singular measures. Moreover, we can pass from E_r to $B(r,x)$ by the same trick as in the last proof.

Assuming $\mu \perp m$, let $\mu(A) = 0$, $m(A^c) = 0$. As mentioned in Remark 1.38, any μ on $\mathbb{R}^n$ that is finite on compacts is regular, therefore outer regular. Therefore given $\varepsilon > 0$, we can approximate A from above by an open set U such that $\mu(U) < \varepsilon$. Let $F_k = \{x \in A : \limsup \frac{\mu(B(r,x))}{m(B(r,x))} > \frac{1}{k}\}$. Then each $x \in F_k$ has a ball $B_x \subseteq U$ such that $\mu(B_x) > \frac{1}{k} m(B_x)$. By the covering lemma (Lemma 2.60), for any $c < m(F_k)$ we can choose a finite subcover of F_k so that $c \le 3^n \sum_{j=1}^N m(B_j) \le 3^n k \sum_{j=1}^N \mu(B_j) \le 3^n k \mu(U) \le 3^n k\varepsilon$. Taking $c \to m(F_k)$ and $\varepsilon \to 0$, we get $m(F_k) = 0$ for each k. □

2.23 Lebesgue–Stieltjes measures on $\mathbb{R}$: Fundamental Theorem of Calculus

Theorem 2.71 (Fundamental Theorem of Calculus for Lebesgue Integrals)**.** *If $f : \mathbb{R} \to \mathbb{C}$ is in $L^1(m)$ and*

$$F(x) = \int_{(-\infty,x]} f\, dx,$$

then $F'(x) = f(x)$ at every Lebesgue point of f (in particular F is a.e. differentiable).

Remark 2.72. However, one should not expect $\int_{[a,x]} F'(x)dx = F(x) - F(a)$ under the mere assumption that F' exists a.e. and $F' \in L^1(m)$ (for a counterexample, choose F to be the Cantor function). Indeed, one needs the extra condition that F is an absolutely continuous function.

Proof. Use Theorem 2.68 with $E_r = (x, x+r]$ and then again with $E_r = (x-r, x]$ to get that both the right-hand and left-hand derivatives, respectively, are equal to $f(x)$ at every Lebesgue point of f. □

2.24 Riesz(–Markov) Representation Theorem

Intuition. Note that if μ is a Borel measure on $(X, \mathcal{M})$ that is finite on compacts, then $C_c(X)$ (compactly supported continuous functions) are integrable, and the $C_c(X) \mapsto \mathbb{C}$ map $f \mapsto \int f\, d\mu$ is a positive linear functional. Riesz/Riesz–Markov theorem investigates the converse of this result.

Throughout this section, let X be a locally compact (every point has a neighbourhood with compact closure) Hausdorff (any two points can be separated with disjoint neighbourhoods) space. This is required in order for Urysohn's lemma to be valid, which in turn is required in order for $C_c(X)$ to be rich enough.

Definition 2.73. *(i) We say that a map $I : C_c(X) \to \mathbb{C}$ is a **linear functional** if $I(f+g) = I(f) + I(g)$ and $I(\alpha f) = \alpha I(f)$ for $\alpha \in \mathbb{C}$.*

*(ii) We say that a functional is **bounded** if there is $C > 0$ such that $|I(f)| \le C\,||f||_\infty$.*

*(iii) We say that a functional is **positive** if $f(x) \ge 0$ (for all x) implies $I(f) \ge 0$.*

Remark 2.74. We discuss linear functionals in a more general setting in Section 3.7 later.

Theorem 2.75 (Riesz(–Markov) representation theorem)**.** *Let I be a positive linear functional on $C_c(X)$. Then there exists a unique Borel measure μ that is finite on compacts, outer regular, and inner regular on open sets, such that $I(f) = \int_X f\,d\mu$ for all $f \in C_c(X)$.*

Remarks 2.76. (a) Uniqueness fails if we don't require regularity properties!

(b) In "reasonable" spaces (namely, locally compact, Hausdorff, such that every open set is σ-compact (countable union of compacts)), any measure that is finite on compacts is in fact regular (see [Rudin, 2.18]).

(c) I is a bounded functional iff μ is finite.

(d) There is also an analogue of this result for complex linear functionals.

Proof. [F212–215]

Full proof is in the book. We will only sketch the rough idea. Let us write $f \prec U$ if $0 \le f \le 1$ and $\mathrm{supp} f \subseteq U$.

To get μ from I, we define

$$\mu(U) = \sup\{I(f) : f \in C_c(X), f \prec U\}$$

for open sets U. To extend this to a measure we use outer measure and Carathéodory theorem: let

$$\mu^*(E) = \inf\{\mu(U) : U \supseteq E, U \text{ open}\}$$

for any set E. One shows that μ^* is an outer measure and that every Borel set is μ^*-measurable. This proves that $\mu|_{\mathcal{B}(X)}$ is a Borel measure, and we are just left with showing that μ satisfies the regularity properties and that $I(f) = \int f\,d\mu$ for $f \in C_c(X)$. □
