Appendix

In this appendix, we describe without proof some results from real analysis which help to understand weak and distributional derivatives in the simplest context of functions of a single variable. Proofs are given in [**10**] or [**12**], for example. These results are, in fact, easier to understand from the perspective of weak and distributional derivatives of functions, rather than pointwise derivatives.

3.A. Functions

For definiteness, we consider functions $f : [a,b] \to \mathbb{R}$ defined on a compact interval $[a,b]$. When we say that a property holds almost everywhere (*a.e.*), we mean *a.e.* with respect to Lebesgue measure unless we specify otherwise.

3.A.1. Lipschitz functions. Lipschitz continuity is a weaker condition than continuous differentiability. A Lipschitz continuous function is pointwise differentiable almost everwhere and weakly differentiable. The derivative is essentially bounded, but not necessarily continuous.

Definition 3.51. A function $f : [a,b] \to \mathbb{R}$ is uniformly Lipschitz continuous on $[a,b]$ (or Lipschitz, for short) if there is a constant C such that

$$|f(x) - f(y)| \leq C\,|x - y| \qquad \text{for all } x, y \in [a,b].$$

The Lipschitz constant of f is the infimum of constants C with this property.

We denote the space of Lipschitz functions on $[a,b]$ by $\mathrm{Lip}[a,b]$. We also define the space of locally Lipschitz functions on $\mathbb{R}$ by

$$\mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \in \mathrm{Lip}[a,b] \text{ for all } a < b\}.$$

By the mean-value theorem, any function that is continuous on $[a,b]$ and pointwise differentiable in (a,b) with bounded derivative is Lipschitz. In particular, every function $f \in C^1([a,b])$ is Lipschitz, and every function $f \in C^1(\mathbb{R})$ is locally Lipschitz. On the other hand, the function $x \mapsto |x|$ is Lipschitz but not C^1 on $[-1,1]$. The following result, called Rademacher's theorem, is true for functions of several variables, but we state it here only for the one-dimensional case.

Theorem 3.52. *If $f \in \mathrm{Lip}[a,b]$, then the pointwise derivative f' exists almost everywhere in (a,b) and is essentially bounded.*

It follows from the discussion in the next section that the pointwise derivative of a Lipschitz function is also its weak derivative (since a Lipschitz function is absolutely continuous). In fact, we have the following characterization of Lipschitz functions.

Theorem 3.53. *Suppose that $f \in L^1_{\mathrm{loc}}(a,b)$. Then $f \in \mathrm{Lip}[a,b]$ if and only if f is weakly differentiable in (a,b) and $f' \in L^\infty(a,b)$. Moreover, the Lipschitz constant of f is equal to the sup-norm of f'.*

Here, we say that $f \in L^1_{\mathrm{loc}}(a,b)$ is Lipschitz on $[a,b]$ if is equal almost everywhere to a (uniformly) Lipschitz function on (a,b), in which case f extends by uniform continuity to a Lipschitz function on $[a,b]$.

Example 3.54. The function $f(x) = x_+$ in Example 3.3 is Lipschitz continuous on $[-1, 1]$ with Lipschitz constant 1. The pointwise derivative of f exists everywhere except at $x = 0$, and is equal to the weak derivative. The sup-norm of the weak derivative $f' = \chi_{[0,1]}$ is equal to 1.

Example 3.55. Consider the function $f : (0, 1) \to \mathbb{R}$ defined by

$$f(x) = x^2 \sin\left(\frac{1}{x}\right).$$

Since f is C^1 on compactly contained intervals in $(0, 1)$, an integration by parts implies that

$$\int_0^1 f\phi'\,dx = -\int_0^1 f'\phi\,dx \qquad \text{for all } \phi \in C_c^\infty(0,1).$$

Thus, the weak derivative of f in $(0, 1)$ is

$$f'(x) = -\cos\left(\frac{1}{x}\right) + 2x\sin\left(\frac{1}{x}\right).$$

Since $f' \in L^\infty(0, 1)$, f is Lipschitz on $[0, 1]$,

Similarly, if $f \in L^1_{\text{loc}}(\mathbb{R})$, then $f \in \text{Lip}_{\text{loc}}(\mathbb{R})$, if and only if f is weakly differentiable in $\mathbb{R}$ and $f' \in L^\infty_{\text{loc}}(\mathbb{R})$.

3.A.2. Absolutely continuous functions. Absolute continuity is a strengthening of uniform continuity that provides a necessary and sufficient condition for the fundamental theorem of calculus to hold. A function is absolutely continuous if and only if its weak derivative is integrable.

Definition 3.56. A function $f : [a, b] \to \mathbb{R}$ is absolutely continuous on $[a, b]$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sum_{i=1}^{N} |f(b_i) - f(a_i)| < \epsilon$$

for any finite collection $\{[a_i, b_i] : 1 \leq i \leq N\}$ of non-overlapping subintervals $[a_i, b_i]$ of $[a, b]$ with

$$\sum_{i=1}^{N} |b_i - a_i| < \delta$$

Here, we say that intervals are non-overlapping if their interiors are disjoint. We denote the space of absolutely continuous functions on $[a, b]$ by AC$[a, b]$. We also define the space of locally absolutely continuous functions on $\mathbb{R}$ by

$$\text{AC}_{\text{loc}}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \in \text{AC}[a, b] \text{ for all } a < b\}.$$

Restricting attention to the case $N = 1$ in Definition 3.56, we see that an absolutely continuous function is uniformly continuous, but the converse is not true (see Example 3.58).

Example 3.57. A Lipschitz function is absolutely continuous. If the function has Lipschitz constant C, we may take $\delta = \epsilon/C$ in the definition of absolute continuity.

Example 3.58. The Cantor function f in Example 3.5 is uniformly continuous on $[0,1]$, as is any continuous function on a compact interval, but it is not absolutely continuous. We may enclose the Cantor set in a union of disjoint intervals the sum of whose lengths is as small as we please, but the jumps in f across those intervals add up to 1. Thus for any $0 < \epsilon \leq 1$, there is no $\delta > 0$ with the property required in the definition of absolute continuity. In fact, absolutely continuous functions map sets of measure zero to sets of measure zero; by contrast, the Cantor function maps the Cantor set with measure zero onto the interval $[0,1]$ with measure one.

Example 3.59. If $g \in L^1(a,b)$ and

$$f(x) = \int_a^x g(t)\, dt$$

then $f \in \mathrm{AC}[a,b]$ and $f' = g$ pointwise *a.e.* (at every Lebesgue point of g). This is one direction of the fundamental theorem of calculus.

According to the following result, the absolutely continuous functions are precisely the ones for which the fundamental theorem of calculus holds. This result may be regarded as giving an explicit characterization of weakly differentiable functions of a single variable.

Theorem 3.60. *A function $f : [a,b] \to \mathbb{R}$ is absolutely continuous if and only if: (a) the pointwise derivative f' exists almost everywhere in (a,b); (b) the derivative $f' \in L^1(a,b)$ is integrable; and (c) for every $x \in [a,b]$,*

$$f(x) = f(a) + \int_a^x f'(t)\, dt.$$

To prove this result, one shows from the definition of absolute continuity that if $f \in \mathrm{AC}[a,b]$, then f' exists pointwise *a.e.* and is integrable, and if $f' = 0$, then f is constant. Then the function

$$f(x) - \int_a^x f'(t)\, dt$$

is absolutely continuous with pointwise *a.e.* derivative equal to zero, so the result follows.

Example 3.61. We recover the function $f(x) = x_+$ in Example 3.3 by integrating its derivative $\chi_{[0,\infty)}$. On the other hand, the pointwise *a.e.* derivative of the Cantor function in Example 3.5 is zero, so integration of its pointwise derivative (which exists *a.e.* and is integrable) gives zero instead of the original function.

Integration by parts holds for absolutely continuous functions.

Theorem 3.62. *If $f, g : [a,b] \to \mathbb{R}$ are absolutely continuous, then*

$$\int_a^b f g'\, dx = f(b)g(b) - f(a)g(a) - \int_a^b f' g\, dx \tag{3.18}$$

where f', g' denote the pointwise a.e. derivatives of f, g.

This result is not true under the assumption that f, g that are continuous and differentiable pointwise *a.e.*, as can be seen by taking f, g to be Cantor functions on $[0,1]$.

In particular, taking $g \in C_c^\infty(a,b)$ in (3.18), we see that an absolutely continuous function f is weakly differentiable on (a,b) with integrable derivative, and the weak derivative is equal to the pointwise *a.e.* derivative. Thus, we have the following characterization of absolutely continuous functions in terms of weak derivatives.

Theorem 3.63. *Suppose that* $f \in L^1_{\mathrm{loc}}(a,b)$. *Then* $f \in \mathrm{AC}[a,b]$ *if and only if* f *is weakly differentiable in* (a,b) *and* $f' \in L^1(a,b)$.

It follows that a function $f \in L^1_{\mathrm{loc}}(\mathbb{R})$ is weakly differentiable if and only if $f \in \mathrm{AC}_{\mathrm{loc}}(\mathbb{R})$, in which case $f' \in L^1_{\mathrm{loc}}(\mathbb{R})$.

3.A.3. Functions of bounded variation. Functions of bounded variation are functions with finite oscillation or variation. A function of bounded variation need not be weakly differentiable, but its distributional derivative is a Radon measure.

Definition 3.64. The total variation $\mathrm{V}_f([a,b])$ of a function $f : [a,b] \to \mathbb{R}$ on the interval $[a,b]$ is

$$\mathrm{V}_f([a,b]) = \sup\left\{\sum_{i=1}^{N} |f(x_i) - f(x_{i-1})|\right\}$$

where the supremum is taken over all partitions

$$a = x_0 < x_1 < x_2 < \cdots < x_N = b$$

of the interval $[a,b]$. A function f has bounded variation on $[a,b]$ if $\mathrm{V}_f([a,b])$ is finite.

We denote the space of functions of bounded variation on $[a,b]$ by $\mathrm{BV}[a,b]$, and refer to a function of bounded variation as a BV-function. We also define the space of locally BV-functions on $\mathbb{R}$ by

$$\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \in \mathrm{BV}[a,b] \text{ for all } a < b\}.$$

Example 3.65. Every Lipschitz continuous function $f : [a,b] \to \mathbb{R}$ has bounded variation, and

$$\mathrm{V}_f([a,b]) \le C(b-a)$$

where C is the Lipschitz constant of f.

A BV-function is bounded, and an absolutely continuous function is BV; but a BV-function need not be continuous, and a continuous function need not be BV.

Example 3.66. The discontinuous step function in Example 3.4 has bounded variation on the interval $[-1,1]$, and the continuous Cantor function in Example 3.5 has bounded variation on $[0,1]$. The total variation of both functions is equal to one. More generally, any monotone function $f : [a,b] \to \mathbb{R}$ has bounded variation, and its total variation on $[a,b]$ is equal to $|f(b) - f(a)|$.

Example 3.67. The function

$$f(x) = \begin{cases} \sin(1/x) & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases}$$

is bounded $[0,1]$, but it is not of bounded variation on $[0,1]$.

Example 3.68. The function

$$f(x) = \begin{cases} x\sin(1/x) & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases}$$

is continuous on $[0,1]$, but it is not of bounded variation on $[0,1]$ since its total variation is proportional to the divergent harmonic series $\sum 1/n$.

The following result states that any BV-functions is a difference of monotone increasing functions. We say that a function f is monotone increasing if $f(x) \leq f(y)$ for $x \leq y$; we do not require that the function is strictly increasing.

Theorem 3.69. *A function $f : [a,b] \to \mathbb{R}$ has bounded variation on $[a,b]$ if and only if $f = f_+ - f_-$, where $f_+, f_- : [a,b] \to \mathbb{R}$ are bounded monotone increasing functions.*

To prove the theorem, we define an increasing variation function $v : [a,b] \to \mathbb{R}$ by $v(a) = 0$ and

$$v(x) = \mathrm{V}_f\left([a,x]\right) \qquad \text{for } x > a.$$

We then choose f_+, f_- so that

$$f = f_+ - f_-, \qquad v = f_+ + f_-, \tag{3.19}$$

and show that f_+, f_- are increasing functions.

The decomposition in Theorem 3.69 is not unique, since we may add an arbitrary increasing function to both f_+ and f_-, but it is unique if we add the condition that $f_+ + f_- = \mathrm{V}_f$.

A monotone function is differentiable pointwise *a.e.*, and thus so is a BV-function. In general, a BV-function contains a singular component that is not weakly differentiable in addition to an absolutely continuous component that is weakly differentiable

Definition 3.70. A function $f \in \mathrm{BV}[a,b]$ is singular on $[a,b]$ if the pointwise derivative f' is equal to zero *a.e.* in $[a,b]$.

The step function and the Cantor function are examples of non-constant singular functions.[4]

Theorem 3.71. *If $f \in \mathrm{BV}[a,b]$, then $f = f_{ac} + f_s$ where $f_{ac} \in \mathrm{AC}[a,b]$ and f_s is singular. The functions f_{ac}, f_s are unique up to an additive constant.*

The absolutely continuous part f_{ac} of f is given by

$$f_{ac}(x) = \int_a^x f'(x)\,dx$$

and the remainder $f_s = f - f_{ac}$ is the singular part. We may further decompose the singular part into a jump-function (such as the step function) and a singular continuous part (such as the Cantor function).

For $f \in \mathrm{BV}[a,b]$, let $D \subset [a,b]$ denote the set of points of discontinuity of f. Since f is the difference of monotone functions, it can only contain jump discontinuities at which its left and right limits exist (excluding the left limit at a and the right limit at b), and D is necessarily countable.

If $c \in D$, let

$$[f](c) = f(c^+) - f(c^-)$$

denote the jump of f at c (with $f(a^-) = f(a)$, $f(b^+) = f(b)$ if $a, b \in D$). Define

$$f_p(x) = \sum_{c \in D \cap [a,x]} [f](c) \qquad \text{if } x \notin D.$$

[4]Sometimes a singular function is required to be continuous, but our definition allows jump discontinuities.

Then f_p has the same jump discontinuities as f and, with an appropriate choice of $f_p(c)$ for $c \in D$, the function $f - f_p$ is continuous on $[a, b]$. Decomposing this continuous part into and absolutely continuous and a singular continuous part, we get the following result.

Theorem 3.72. *If $f \in \mathrm{BV}[a, b]$, then $f = f_{ac} + f_p + f_{sc}$ where $f_{ac} \in \mathrm{AC}[a, b]$, f_p is a jump function, and f_{sc} is a singular continuous function. The functions f_{ac}, f_p, f_{sc} are unique up to an additive constant.*

Example 3.73. Let $Q = \{q_n : n \in \mathbb{N}\}$ be an enumeration of the rational numbers in $[0, 1]$ and $\{p_n : n \in \mathbb{N}\}$ any sequence of real numbers such that $\sum p_n$ is absolutely convergent. Define $f : [a, b] \to \mathbb{R}$ by $f(0) = 0$ and

$$f(x) = \sum_{a \le q_n \le x} p_n \qquad \text{for } x > 0.$$

Then $f \in \mathrm{BV}[a, b]$, with

$$\mathrm{V}_f[a, b] = \sum_{n \in \mathbb{N}} |p_n|.$$

This function is a singular jump function with zero pointwise derivative at every irrational number in $[0, 1]$.

3.B. Measures

We denote the extended real numbers by $\overline{\mathbb{R}} = [-\infty, \infty]$ and the extended nonnegative real numbers by $\overline{\mathbb{R}}_+ = [0, \infty]$. We make the natural conventions for algebraic operations and limits that involve extended real numbers.

3.B.1. Borel measures.

The Borel σ-algebra of a topological space X is the smallest collection of subsets of X that contains the open and closed sets, and is closed under complements, countable unions, and countable intersections. Let $\mathcal{B}$ denote the Borel σ-algebra of $\mathbb{R}$, and $\overline{\mathcal{B}}$ the Borel σ-algebra of $\overline{\mathbb{R}}$.

Definition 3.74. A Borel measure on $\mathbb{R}$ is a function $\mu : \mathcal{B} \to \overline{\mathbb{R}}_+$, such that $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{n \in \mathbb{N}} E_n\right) = \sum_{n \in \mathbb{N}} \mu(E_n)$$

for any countable collection of disjoint sets $\{E_n \in \mathcal{B} : n \in \mathbb{N}\}$.

The measure μ is finite if $\mu(\mathbb{R}) < \infty$, in which case $\mu : \mathcal{B} \to [0, \infty)$. The measure is σ-finite if $\mathbb{R}$ is a countable union of Borel sets with finite measure.

Example 3.75. Lebesgue measure $\lambda : \mathcal{B} \to \overline{\mathbb{R}}_+$ is a Borel measure that assigns to each interval its length. Lebesgue measure on $\mathcal{B}$ may be extended to a complete measure on a larger σ-algebra of Lebesgue measurable sets by the inclusion of all subsets of sets with Lebesgue measure zero. Here we consider it as a Borel measure.

Example 3.76. For $c \in \mathbb{R}$, the unit point measure $\delta_c : \mathcal{B} \to [0, \infty)$ supported on c is defined by

$$\delta_c(E) = \begin{cases} 1 & \text{if } c \in E, \\ 0 & \text{if } c \notin E. \end{cases}$$

This measure is a finite Borel measure. More generally, if $\{c_n : n \in \mathbb{N}\}$ is a countable set of points in $\mathbb{R}$ and $\{p_n \geq 0 : n \in \mathbb{N}\}$, we define a point measure

$$\mu = \sum_{n \in \mathbb{N}} p_n \delta_{c_n}, \qquad \mu(E) = \sum_{c_n \in E} p_n.$$

This measure is σ-finite, and finite if $\sum p_n < \infty$.

Example 3.77. Counting measure $\nu : \mathcal{B} \to \overline{\mathbb{R}}_+$ is defined by $\nu(E) = \#E$ where $\#E$ denotes the number of points in E. Thus, $\nu(\emptyset) = 0$ and $\nu(E) = \infty$ if E contains infinitely many points. This measure is not σ-finite.

In order to describe the decomposition of measures, we introduce the idea of singular measures that 'live' on different sets.

Definition 3.78. Two measures $\mu, \nu : \mathcal{B} \to \overline{\mathbb{R}}_+$ are mutually singular, written $\mu \perp \nu$, if there is a set $E \in \mathcal{B}$ such that $\mu(E) = 0$ and $\nu(E^c) = 0$.

We also say that μ is singular with respect to ν, or ν is singular with respect to μ. In particular, a measure is singular with respect to Lebesgue measure if it assigns full measure to a set of Lebesgue measure zero.

Example 3.79. The point measures in Example 3.76 are singular with respect to Lebesgue measure.

Next we consider signed measures which can take negative as well as positive values.

Definition 3.80. A signed Borel measure is a map $\mu : \mathcal{B} \to \overline{\mathbb{R}}$ of the form

$$\mu = \mu_+ - \mu_-$$

where $\mu_+, \mu_- : \mathcal{B} \to \overline{\mathbb{R}}_+$ are Borel measures, at least one of which is finite.

The condition that at least one of μ_+, μ_- is finite is needed to avoid meaningless expressions such as $\mu(\mathbb{R}) = \infty - \infty$. Thus, μ takes at most one of the values ∞, $-\infty$.

According to the Jordan decomposition theorem, we may choose μ_+, μ_- in Definition 3.80 so that $\mu_+ \perp \mu_-$, in which case the decomposition is unique. The total variation of μ is then measure $|\mu| : \mathcal{B} \to \overline{\mathbb{R}}_+$ defined by

$$|\mu| = \mu_+ + \mu_-.$$

Definition 3.81. Let $\mu : \mathcal{B} \to \overline{\mathbb{R}}_+$ be a measure. A signed measure $\nu : \mathcal{B} \to \overline{\mathbb{R}}$ is absolutely continuous with respect to μ, written $\nu \ll \mu$, if $\mu(E) = 0$ implies that $\nu(E) = 0$ for any $E \in \mathcal{B}$.

The condition $\nu \ll \mu$ is equivalent to $|\nu| \ll \mu$. In that case ν 'lives' on the same sets as μ; thus absolute continuity is at the opposite extreme to singularity. In particular, a signed measure ν is absolutely continuous with respect to Lebesgue measure if it assigns zero measure to any set with zero Lebesgue measure,

If $g \in L^1(\mathbb{R})$, then

$$\nu(E) = \int_E g \, dx \tag{3.20}$$

defines a finite signed Borel measure $\nu : \mathcal{B} \to \mathbb{R}$. This measure is absolutely continuous with respect to Lebesgue measure, since $\int_E g \, dx = 0$ for any set E with Lebesgue measure zero.

If $g \geq 0$, then ν is a measure. If the set $\{x : g(x) = 0\}$ has non-zero Lebesgue measure, then Lebesgue measure is not absolutely continuous with respect to ν. Thus $\nu \ll \mu$ does not imply that $\mu \ll \nu$.

The Radon-Nikodym theorem (which holds in greater generality) implies that every absolutely continuous measure is given by the above example.

Theorem 3.82. *If ν is a Borel measure on $\mathbb{R}$ that is absolutely continuous with respect to Lebesgue measure λ then there exists a function $g \in L^1(\mathbb{R})$ such that ν is given by (3.20).*

The function g in this theorem is called the Radon-Nikodym derivative of ν with respect to λ, and is denoted by

$$g = \frac{d\nu}{d\lambda}.$$

The following result gives an alternative characterization of absolute continuity of measures, which has a direct connection with the absolute continuity of functions.

Theorem 3.83. *A signed measure $\nu : \mathcal{B} \to \overline{\mathbb{R}}$ is absolutely continuous with respect to a measure $\mu : \mathcal{B} \to \overline{\mathbb{R}}_+$ if and only if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\mu(E) < \delta$ implies that $|\nu(E)| \leq \epsilon$ for all $E \in \mathcal{B}$.*

3.B.2. Radon measures. The most important Borel measures for distribution theory are the Radon measures. The essential property of a Radon measure μ is that integration against μ defines a positive linear functional on the space of continuous functions ϕ with compact support,

$$\phi \mapsto \int \phi \, d\mu.$$

(See Theorem 3.96 below.) This link is the fundamental connection between measures and distributions. The condition in the following definition characterizes all such measures on $\mathbb{R}$ (and $\mathbb{R}^n$).

Definition 3.84. A Radon measure on $\mathbb{R}$ is a Borel measure that is finite on compact sets.

We note in passing that a Radon measure μ has the following regularity property: For any $E \in \mathcal{B}$,

$$\mu(E) = \inf \{\mu(G) : G \supset E \text{ open}\}, \quad \mu(E) = \sup \{\mu(K) : K \subset E \text{ compact}\}.$$

Thus, any Borel set may be approximated in a measure-theoretic sense by open sets from the outside and compact sets from the inside.

Example 3.85. Lebesgue measure λ in Example 3.75 and the point measure δ_c in Example 3.76 are Radon measures on $\mathbb{R}$.

Example 3.86. The counting measure ν in Example 3.77 is not a Radon measure since, for example, $\nu[0, 1] = \infty$. This measure is not outer regular: If $\{c\}$ is a singleton set, then $\nu(\{c\}) = 1$ but

$$\inf \{\nu(G) : c \in G, \, G \text{ open}\} = \infty.$$

The following is the Lebesgue decomposition of a Radon measure.

Theorem 3.87. *Let μ, ν be Radon measures on $\mathbb{R}$. There are unique measures ν_{ac}, ν_s such that*

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

3.B.3. Lebesgue-Stieltjes measures. Given a Radon measure μ on $\mathbb{R}$, we may define a monotone increasing, right-continuous distribution function $f : \mathbb{R} \to \mathbb{R}$, which is unique up to an arbitrary additive constant, such that

$$\mu(a, b] = f(b) - f(a).$$

The function f is right-continuous since

$$\lim_{x \to b^+} f(b) - f(a) = \lim_{x \to b^+} \mu(a, x] = \mu(a, b] = f(b) - f(a).$$

Conversely, every such function f defines a Radon measure μ_f, called the Lebesgue-Stieltjes measure associated with f. Thus, Radon measures on $\mathbb{R}$ may be characterized explicitly as Lebesgue-Stieltjes measures.

Theorem 3.88. *If $f : \mathbb{R} \to \mathbb{R}$ is a monotone increasing, right-continuous function, there is a unique Radon measure μ_f such that*

$$\mu_f(a, b] = f(b) - f(a)$$

for any half-open interval $(a, b] \subset \mathbb{R}$.

The standard proof is due to Carathéodory. One uses f to define a countably sub-additive outer measure μ_f^* on all subsets of $\mathbb{R}$, then restricts μ_f^* to a measure on the σ-algebra of μ_f^*-measurable sets, which includes all of the Borel sets [**10**].

The Lebesgue-Stieltjes measure of a compact interval $[a, b]$ is given by

$$\mu_f[a, b] = \lim_{x \to a^-} \mu_f(x, b] = f(b) - \lim_{x \to a^-} f(a).$$

Thus, the measure of the set consisting of a single point is equal to the jump in f at the point,

$$\mu_f\{a\} = f(a) - \lim_{x \to a^-} f(a),$$

and $\mu_f\{a\} = 0$ if and only if f is continuous at a.

Example 3.89. If $f(x) = x$, then μ_f is Lebesgue measure (restricted to the Borel sets) in $\mathbb{R}$.

Example 3.90. If $c \in \mathbb{R}$ and

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

then μ_f is the point measure δ_c in Example 3.76.

Example 3.91. If f is the Cantor function defined in Example 3.5, then μ_f assigns measure one to the Cantor set C and measure zero to $\mathbb{R} \setminus C$. Thus, μ_f is singular with respect to Lebesgue measure. Nevertheless, since f is continuous, the measure of any set consisting of a single point, and therefore any countable set, is zero.

If $f : \mathbb{R} \to \mathbb{R}$ is the difference $f = f_+ - f_-$ of two right-continuous monotone increasing functions $f_+, f_- : \mathbb{R} \to \mathbb{R}$, at least one of which is bounded, we may define a signed Radon measure $\mu_f : \mathcal{B} \to \overline{\mathbb{R}}$ by

$$\mu_f = \mu_{f_+} - \mu_{f_-}.$$

If we add the condition that $\mu_{f_+} \perp \mu_{f_-}$, then this decomposition is unique, and corresponds to the decomposition of f in (3.19).

3.C. Integration

A function $\phi : \mathbb{R} \to \overline{\mathbb{R}}$ is Borel measurable if $\phi^{-1}(E) \in \mathcal{B}$ for every $E \in \overline{\mathcal{B}}$. In particular, every continuous function $\phi : \mathbb{R} \to \mathbb{R}$ is Borel measurable.

Given a Borel measure μ, and a non-negative, Borel measurable function ϕ, we define the integral of ϕ with respect to μ as follows. If

$$\psi = \sum_{i \in \mathbb{N}} c_i \chi_{E_i}$$

is a simple function, where $c_i \in \overline{\mathbb{R}}_+$ and χ_{E_i} is the characteristic function of a set $E_i \in \mathcal{B}$, then

$$\int \psi \, d\mu = \sum_{i \in \mathbb{N}} c_i \mu(E_i).$$

Here, we define $0 \cdot \infty = 0$ for the integral of a zero value on a set of infinite measure, or an infinite value on a set of measure zero. If $\phi : \mathbb{R} \to \overline{\mathbb{R}}_+$ is a non-negative Borel-measurable function, we define

$$\int \phi \, d\mu = \sup \left\{ \int \psi \, d\mu : 0 \leq \psi \leq \phi \right\}$$

where the supremum is taken over all non-negative simple functions ψ that are bounded from above by ϕ.

If $\phi : \mathbb{R} \to \overline{\mathbb{R}}$ is a general Borel function, we split ϕ into its positive and negative parts,

$$\phi = \phi_+ - \phi_-, \qquad \phi_+ = \max(\phi, 0), \quad \phi_- = \max(-\phi, 0),$$

and define

$$\int \phi \, d\mu = \int \phi_+ \, d\mu - \int \phi_- \, d\mu$$

provided that at least one of these integrals is finite.

The continual annoyance of excluding $\infty - \infty$ as meaningless is often viewed as a defect of the Lebesgue integral, which cannot cope directly with the cancelation between infinite positive and negative components. For example, the improper integral

$$\int_0^\infty \frac{\sin x}{x} \, dx = \frac{\pi}{2}$$

does not hold as a Lebesgue integral since $|\sin(x)/x|$ is not integrable. Nevertheless, other definitions of the integral — such as the Henstock-Kurzweil integral — have not proved to be as useful.

Example 3.92. The integral of ϕ with respect to Lebesgue measure λ in Example 3.75 is the usual Lebesgue integral

$$\int \phi \, d\lambda = \int \phi \, dx.$$

Example 3.93. The integral of ϕ with respect to the point measure δ_c in Example 3.76 is

$$\int \phi \, d\delta_c = \phi(c).$$

Note that $\phi = \psi$ pointwise *a.e.* with respect to δ_c if and only if $\phi(c) = \psi(c)$.

Example 3.94. If f is absolutely continuous, the associated Lebesgue-Stieltjes measure μ_f is absolutely continuous with respect to Lebesgue measure, and

$$\int \phi \, d\mu_f = \int \phi f' \, dx.$$

Next, we consider linear functionals on the space $C_c(\mathbb{R})$ of linear functions with compact support.

Definition 3.95. A linear functional $I : C_c(\mathbb{R}) \to \mathbb{R}$ is positive if $I(\phi) \geq 0$ whenever $\phi \geq 0$, and locally bounded if for every compact set K in $\mathbb{R}$ there is a constant C_K such that

$$|I(\phi)| \leq C_K \|\phi\|_\infty \qquad \text{for all } \phi \in C_c(\mathbb{R}) \text{ with } \operatorname{spt} \phi \subset K.$$

A positive functional is locally bounded, and a locally bounded functional I defines a distribution $I \in \mathcal{D}'(\mathbb{R})$ by restriction to $C_c^\infty(\mathbb{R})$. We also write $I(\phi) = \langle I, \phi \rangle$. If μ is a Radon measure, then

$$\langle I_\mu, \phi \rangle = \int \phi \, d\mu$$

defines a positive linear functional $I_\mu : C_c(\mathbb{R}) \to \mathbb{R}$, and if μ_+, μ_- are Radon measures, then $I_{\mu_+} - I_{\mu_-}$ is a locally bounded functional.

Conversely, according to the following Riesz representation theorem, all locally bounded linear functionals on $C_c(\mathbb{R})$ are of this form

Theorem 3.96. *If $I : C_c(\mathbb{R}) \to \mathbb{R}_+$ is a positive linear functional on the space of continuous functions $\phi : \mathbb{R} \to \mathbb{R}$ with compact support, then there is a unique Radon measure μ such that*

$$I(\phi) = \int \phi \, d\mu.$$

If $I : C_c(\mathbb{R}) \to \overline{\mathbb{R}}_+$ is locally bounded linear functional, then there are unique Radon measures μ_+, μ_- such that

$$I(\phi) = \int \phi \, d\mu_+ - \int \phi \, d\mu_-.$$

Note that the functional $\mu = \mu_+ - \mu_-$ is not well-defined as a signed Radon measure if both μ_+ and μ_- are infinite.

Every distribution $T \in \mathcal{D}'(\mathbb{R})$ such that

$$\langle T, \phi \rangle \leq C_K \|\phi\|_\infty \qquad \text{for all } \phi \in C_c^\infty(\mathbb{R}) \text{ with } \operatorname{spt} \phi \subset K$$

may be extended by continuity to a locally bounded linear functional on $C_c(\mathbb{R})$, and therefore is given by $T = I_{\mu_+} - I_{\mu_-}$ for Radon measures μ_+, μ_- . We typically identify a Radon measure μ with the corresponding distribution I_μ. If μ is absolutely continuous with respect to Lebesgue measure, then $\mu = \mu_f$ for some $f \in \mathrm{AC}_{\mathrm{loc}}(\mathbb{R})$, meaning that

$$\mu_f(E) = \int_E f' \, dx,$$

and I_μ is the same as the regular distribution $T_{f'}$. Thus, with these identifications, and denoting the Radon measures by $\mathcal{M}$, we have the following local inclusions:

$$\mathrm{AC} \subset \mathrm{BV} \subset L^1 \subset \mathcal{M} \subset \mathcal{D}'.$$

The distributional derivative of an AC function is an integrable function, and the following integration by parts formula shows that the distributional derivative of a BV function is a Radon measure.

Theorem 3.97. *Suppose that* $f \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ *and* $g \in \mathrm{AC}_c(\mathbb{R})$ *is absolutely continuous with compact support. Then*

$$\int g\, d\mu_f = -\int f g'\, dx.$$

Thus, the distributional derivative of $f \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ is the functional I_{μ_f} associated with the corresponding Radon measure μ_f. If

$$f = f_{ac} + f_p + f_{sc}$$

is the decomposition of f into a locally absolutely continuous part, a jump function, and a singular continuous function, then

$$\mu_f = \mu_{ac} + \mu_p + \mu_{sc},$$

where μ_{ac} is absolutely continuous with respect to Lebesgue measure with density f'_{ac}, μ_p is a point measure of the form

$$\mu_p = \sum_{n\in\mathbb{N}} p_n \delta_{c_n}$$

where the c_n are the points of discontinuity of f and the p_n are the jumps, and μ_{sc} is a measure with continuous distribution function that is singular with respect to Lebesgue measure. The function is weakly differentiable if and only if it is locally absolutely continuous.

Thus, to return to our original one-dimensional examples, the function x_+ in Example 3.3 is absolutely continuous and its weak derivative is the step function. The weak derivative is bounded since the function is Lipschitz. The step function in Example 3.4 is not weakly differentiable; its distributional derivative is the δ-measure. The Cantor function f in Example 3.5 is not weakly differentiable; its distributional derivative is the singular continuous Lebesgue-Stieltjes measure μ_f associated with f.

We summarize the above discussion in a table.

Function	*Weak Derivative*
Smooth (C^1)	Continuous (C^0)
Lipschitz	Bounded (L^∞)
Absolutely Continuous	Integrable (L^1)
Bounded Variation	Distributional derivative is Radon measure

The correspondences shown in this table continue to hold for functions of several variables, although the study of fine structure of weakly differentiable functions and functions of bounded variation is more involved than in the one-dimensional case.