1.4. CONSTRUCTION OF LEBESGUE-STIELTJES MEASURES

In this section we shall put to use the Carathéodory-Hahn theory, in order to construct measures with certain desirable properties – first on the real line (§§ 1.4.A, 1.4.B), then on Euclidean spaces (§ 1.4.C), and finally on infinite-dimensional spaces (§ 1.6.A).

A: MEASURES ON THE REAL LINE

To motivate the developments that follow, imagine trying to "invert" the process of Exercise 3.3: we *start* with an increasing, right-continuous function F, and try to *find* a measure μ_F on the Borel sets of the real line, such that (1.7) holds.

4.1 Definition: Distribution Function. An increasing, right-continuous function $F : \mathbf{R} \to \mathbf{R}$ is called *distribution function on* $\mathbf{R}$. A distribution function with $F(-\infty) \equiv \lim_{x\downarrow -\infty} F(x) = 0$, $F(\infty) \equiv \lim_{x\to\infty} F(x) = 1$ is called *probability distribution function.*

Suppose now that we are given a distribution function F. The Lebesgue-Stieltjes measure μ_F is the measure μ of Theorem 3.3, constructed on the Borel σ-algebra $\mathcal{B}(\mathbf{R})$ in the following way:

We define $\ell_F((a,b]) := F(b) - F(a)$ for intervals of the type $(a,b]$, $-\infty < a \le b \le \infty$; recall that the collection $\mathcal{G}$ of such intervals is an elementary family, and denote by $\mathcal{E}$ the algebra of all finite disjoint unions of such intervals (Definition 3.3, Exercise 3.1). We have $\mathcal{B}(\mathbf{R}) = \sigma(\mathcal{E}) = \sigma(\mathcal{G})$, and set

$$\nu(\cup_{j=1}^n (a_j, b_j]) := \sum_{j=1}^n \left[F(b_j) - F(a_j) \right], \qquad \cup_{j=1}^n (a_j, b_j] \in \mathcal{E} \tag{4.1}$$

whenever the sets $\{(a_j, b_j]\}_{1\le j\le n}$ are disjoint. Since F is increasing, this set-function ν is non-negative, finitely-additive, and agrees with ℓ_F on $\mathcal{G}$. It is also easily seen that ν is well-defined, in the sense that *(4.1) is independent of how a given set in $\mathcal{E}$ is decomposed into a finite union of disjoint intervals.*

Indeed, if $\{(a_j, b_j]\}_{1\le j\le n}$ are disjoint and $\cup_{j=1}^n (a_j, b_j] = (a,b]$, then we have $b = b_1 > a_1 = \cdots = b_{n-1} > a_{n-1} = b_n > a_n = a$ (by relabelling some indices if necessary) and $\sum_{j=1}^n [F(b_n) - F(a_n)] = F(b) - F(a)$.

More generally, let $\{I_i\}_{1\le i\le m}$ and $\{J_j\}_{1\le j\le n}$ be finite collections of disjoint intervals in $\mathcal{G}$ such that $\cup_{i=1}^m I_i = \cup_{j=1}^n J_j$; then this same reasoning gives $\sum_{i=1}^m \nu(I_i) = \sum_{i=1}^m \sum_{j=1}^n \nu(I_i \cap J_j) = \sum_{j=1}^n \nu(J_j)$, so ν is indeed well defined.

• Let us verify now the conditions in Theorem 3.3(i), in particular that *ν is a pre-measure on $\mathcal{E}$*; in other words, that for any collection of disjoint intervals $\{(a_n, b_n]\}_{n\in\mathbf{N}} \subseteq \mathcal{G}$ with $\cup_{n\in\mathbf{N}} (a_n, b_n] \in \mathcal{E}$, we have the countable additivity property

$$\nu(\cup_{n\in\mathbf{N}}(a_n,b_n]) = \sum_{n\in\mathbf{N}} \ell_F((a_n,b_n]) \,.$$

Since $\cup_{n\in\mathbf{N}}(a_n,b_n]$ is the union of finitely-many intervals in $\mathcal{G}$ (recall that we are assuming this union to be in the algebra $\mathcal{E}$), we may partition the collection $\{(a_n,b_n]\}_{n\in\mathbf{N}}$ into finitely many sub-collections so that the union of the intervals in each of these sub-collections is a *single* interval in $\mathcal{G}$. Thus, it is enough to establish the following.

4.1 Proposition: *Whenever an interval $(a,b]$ can be written as a countable disjoint union $(a,b] = \cup_{n\in\mathbf{N}}(a_n,b_n]$, we have $\sum_{n\in\mathbf{N}} \ell_F((a_n,b_n]) = \ell_F((a,b])$.*

Proof: It is straightforward to see that, whenever $(a,b] = \cup_{n=1}^N(a_n,b_n]$ is a *finite* disjoint union, we have the additivity property $\ell_F((a,b] = \sum_{n=1}^N \ell_F((a_n,b_n])$. Indeed, one establishes separately the implications

$$\cup_{n=1}^N(a_n,b_n] \subseteq (a,b] \quad\Longrightarrow\quad \sum_{n=1}^N \ell_F((a_n,b_n]) \,\le\, \ell_F((a,b])\,,$$

$$(a,b] \subseteq \cup_{n=1}^N(a_n,b_n] \quad\Longrightarrow\quad \ell_F((a,b]) \,\le\, \sum_{n=1}^N \ell_F((a_n,b_n])$$

(the second does not even need the intervals to be disjoint), and the finite additivity follows.

For an infinite collection $\{(a_n,b_n]\}_{n\in\mathbf{N}}$ of disjoint intervals in $\mathcal{G}$ with $(a,b] = \cup_{n\in\mathbf{N}}(a_n,b_n]$, we let $N\to\infty$ in the first of these implications and obtain the inequality $\sum_{n\in\mathbf{N}} \ell_F((a_n,b_n]) \le \ell_F((a,b])$.

• To obtain the reverse inequality and finish the proof, it seems like a very obvious step to pass from a finite to an infinite collection. Yet, in order to be able to make this step, one needs (some version of) the Heine-Borel theorem.

One way to do this is to select, for any given $\varepsilon \in (0, b-a)$ and $n\in\mathbf{N}$, a real number $b_n' > b_n$ with $F(b_n') - F(b_n) < \varepsilon 2^{-n}$ (the right-continuity of $F(\cdot)$). Then $[a+\varepsilon, b] \subseteq \cup_{n\in\mathbf{N}}(a_n,b_n')$, so by the Heine-Borel theorem there exists a *finite* collection $(a_{n_1},b_{n_1}'),\cdots,(a_{n_K},b_{n_K}')$ of open intervals that covers $[a+\varepsilon,b]$; and from our previous discussion

$$F(b)-F(a+\varepsilon) \,=\, \ell_F\big((a+\varepsilon,b]\big) \,\le\, \sum_{k=1}^K \ell_F\big((a_{n_k},b_{n_k}']\big)$$

$$\le\, \sum_{k=1}^K \big[\, \ell_F\big((a_{n_k},b_{n_k}]\big) + \varepsilon 2^{-n_k} \,\big] \,\le\, \sum_{n\in\mathbf{N}} \ell_F\big((a_n,b_n]\big) + \varepsilon \,.$$

Letting $\varepsilon \downarrow 0$ and using again the right-continuity of $F(\cdot)$, we arrive at the desired inequality $F(b)-F(a) \,=\, \ell_F\big((a,b]\big) \,\le\, \sum_{n\in\mathbf{N}} \ell_F\big((a_n,b_n]\big)$. ⋄

- Having done all this, we proceed in three steps:

(i) First, Theorem 3.2 provides an outer measure μ^* defined on all subsets of $\mathbf{R}$ via the recipe (3.7).

(ii) Then the Carathéodory Theorem 3.1 guarantees that the restriction $\overline{\mu}_F$ of this outer measure to the σ-algebra $\mathcal{M} \equiv \mathcal{M}_F$ of its measurable sets, is a *σ-finite measure.*

(iii) Finally, the Hahn Extension Theorem 3.3 asserts that this construction leads to a unique measure $\mu \equiv \mu_F$ on $\mathcal{B}(\mathbf{R}) = \sigma(\mathcal{E})$ with the property (1.7):

$$\mu_F((a,b]) = F(b) - F(a) = \ell_F((a,b]) \quad \text{for every} \quad (a,b] \in \mathcal{G}\,.$$

This construction gives immediately the property

$$\begin{aligned} \overline{\mu}_F(E) \;&=\; \inf\left\{ \sum_{n\in\mathbf{N}} [F(b_n) - F(a_n)] \,\Big|\, E \subseteq \bigcup_{n\in\mathbf{N}} (a_n, b_n] \right\} \\ &=\; \inf\left\{ \sum_{n\in\mathbf{N}} \overline{\mu}_F\big((a_n, b_n]\big) \,\Big|\, E \subseteq \bigcup_{n\in\mathbf{N}} (a_n, b_n] \right\}; \end{aligned} \tag{4.1$'$}$$

see Exercise 4.1 for further elaboration.

Integrals with respect to this measure are denoted as

$$\int_{(a,b]} f\, d\mu_F \;\equiv\; \int_a^b f(x)\, dF(x)\,, \qquad \text{for} \quad -\infty \le a < b < \infty. \tag{4.2}$$

According to Exercise 3.5, the completion of the measure space $(\mathbf{R}, \mathcal{B}(\mathbf{R}), \mu_F)$ is the σ-finite measure space $(\mathbf{R}, \mathcal{M}_F, \overline{\mu}_F)$. The measure $\overline{\mu}_F$ is regular, in the sense of Exercise 1.5; it is called the **Lebesgue-Stieltjes measure** induced by the distribution function F. It is defined on the σ-algebra $\mathcal{M}_F \equiv \mathcal{M}$ of so-called *$F-$Lebesgue-Stieltjes measurable sets,* which is constructed as in Theorems 3.1-3.3 and satisfies

$$\mathcal{B}(\mathbf{R}) \subset \mathcal{M}_F \subset \mathcal{P}(\mathbf{R})\,. \tag{4.3}$$

Here $\mathcal{P}(\mathbf{R})$ is the collection of all subsets of $\mathbf{R}$. We shall see in Appendix A, that both inclusions in (4.3) are typically strict.

For the choice $F(x) \equiv x$, the resulting $\overline{\lambda} \equiv \overline{\mu}_F$ is the **Lebesgue measure on the real line**, and the class $\mathcal{L} \equiv \mathcal{M}_F$ is the *σ-algebra of Lebesgue-measurable sets.* This class is invariant under translations and dilations, and so is Lebesgue measure, in the sense that $E + s \in \mathcal{L}$, $rE \in \mathcal{L}$ and

$$\overline{\lambda}(E+s) = \overline{\lambda}(E)\,, \qquad \overline{\lambda}(r\,E) = r\,\overline{\lambda}(E)$$

for every $E \in \mathcal{L}$, $s \in \mathbf{R}$, $r > 0$. The restriction $\lambda = \overline{\lambda}\big|_{\mathcal{B}(\mathbf{R})}$ to the σ-algebra of Borel sets, is also called "Lebesgue measure"; it has the property $\lambda((a,b]) = b - a$ of (1.6).

Discsussion: Every singleton $\{x\}$ with $x \in \mathbf{R}$ has Lebesgue measure zero; thus the same is true for every countable set. Enumerate as $\mathbf{Q}_1 = \{\varrho_n\}_{n \in \mathbf{N}} \subset [0,1]$ all the rational numbers of the unit interval, observe that $\lambda(\mathbf{Q}_1) = 0$ from the above discussion, and for any given $\varepsilon > 0$ set $I_n = (\varrho_n - \varepsilon 2^{-(n+1)}, \varrho_n + \varepsilon 2^{-(n+1)})$. Then the set $G := (0,1) \cap (\cup_{n \in \mathbf{N}} I_n)$ is dense (topologically "large") in the unit interval; its closure $\overline{G}$ is [0,1]. But is measure-theoretically "minuscule": its Lebesgue measure is $\lambda(G) \le \sum_{n \in \mathbf{N}} \varepsilon 2^{-n} \le \varepsilon$.

We shall see in Appendix A that *there exist sets with the cardinality of the continuum and with Lebesgue measure zero.*

4.1 Exercise : For any $E \in \mathcal{M}_F$ we have

$$\overline{\mu}_F(E) = \inf \left\{ \sum_{n \in \mathbf{N}} \overline{\mu}_F\big((a_n, b_n)\big) \,\Big|\, E \subseteq \bigcup_{n \in \mathbf{N}} (a_n, b_n) \right\},$$

as well as

$$\overline{\mu}_F(E) = \inf_{\substack{U \in \mathcal{O} \\ E \subseteq U}} \overline{\mu}_F(U) = \sup_{\substack{K \in \mathcal{K} \\ K \subseteq E}} \overline{\mu}_F(K),$$

where $\mathcal{O}$ (respectively, $\mathcal{K}$) are all the open (resp., compact) subsets of the real line.

4.2 Exercise : For any $E \in \mathcal{M}_F$ with $\bar{\mu}_F(E) < \infty$, and any $\varepsilon > 0$, there exists a finite union of open intervals $\cup_{n=1}^N I_n =: U$, such that $\overline{\mu}_F(E \triangle U) < \varepsilon$.

4.3 Exercise : For any real numbers $a < b$, we have $\mu_F(a) = F(a) - F(a-)$, as well as

$$\mu_F((a,b]) = F(b) - F(a), \qquad \mu_F([a,b)) = F(b-) - F(a-),$$

$$\mu_F([a,b]) = F(b) - F(a-), \qquad \mu_F((a,b)) = F(b-) - F(a).$$

4.4 Exercise : For any function $f : \mathbf{R} \to \mathbf{R}$ which is integrable with respect to Lebesgue measure λ, show that $F(t) = \int_{-\infty}^t f(x)\,dx$, $t \in \mathbf{R}$ (in the notation of (4.2)) defines a continuous function.

B: CONNECTIONS WITH THE RIEMANN INTEGRAL

Let us discuss briefly the relation of the Lebesgue integral that we have constructed, to the more familiar from classical calculus *Riemann integral.* Consider a compact interval $[a,b]$ on the real line, and a bounded function $f : [a,b] \to \mathbf{R}$ on it. Corresponding to every partition $\Pi = \{t_j\}_{j=0}^n$ of the interval, with $a = t_0 < t_1 < \cdots < t_n = b$, we define the upper- and lower- *Darboux sums*

$$\overline{S}(f;\Pi) := \sum_{j=1}^{n} \overline{M}_j\,(t_j - t_{j-1})\,, \qquad \underline{S}(f;\Pi) := \sum_{j=1}^{n} \underline{M}_j\,(t_j - t_{j-1})$$

respectively, where $\overline{M}_j = \sup_{t_{j-1}<t\le t_j} f(t)$, $\underline{M}_j = \inf_{t_{j-1}<t\le t_j} f(t)$. In terms of these sums we introduce the *upper-* and *lower-Riemann integrals* of f on $[a,b]$ as

$$\overline{R}(f) \equiv \overline{R}(f;[a,b]) := \inf_{\Pi\in\mathcal{P}} \overline{S}(f;\Pi)\,, \qquad \underline{R}(f) \equiv \underline{R}(f;[a,b]) := \sup_{\Pi\mathcal{P}} \underline{S}(f;\Pi)\,,$$

where $\mathcal{P}$ is the set of all possible partitions of the interval $[a,b]$. Whenever the upper- and lower-Riemann integrals coincide, we say that f is *Riemann-integrable* on $[a,b]$, and call the common value $R(f) \equiv \int_a^b f(x)\,dx$ the **Riemann integral** of f on $[a,b]$. The first exercise below asserts that the Lebesgue integral subsumes the Riemann integral, and strictly. The third exercise, though, points to the fact that a so-called *improper Riemann integral* may exist even in situations where the Lebesgue integral fails to exist.

4.5 Exercise: A bounded function $f : [a,b] \to \mathbf{R}$ is Riemann-integrable,

(i) if and only if the set $\mathcal{D}$ of its discontinuities has $\lambda(\mathcal{D}) = 0$;

(ii) only if it is Lebesgue-measurable (thus also Lebesgue-integrable) on $[a,b]$, in which case $R(f) = I(f)$.

4.6 Exercise: If $\mathbf{Q}$ denotes the set of rational numbers in $[a,b]$, show that the function $f = \chi_{\mathbf{Q}}$ is Lebesgue-integrable but *not* Riemann-integrable.

4.7 Exercise: For a function $f : [a,\infty)$ we can define the *improper Riemann integral* as

$$R(f;[a,\infty)) := \lim_{b\to\infty} \int_a^b f(x)\,dx = \lim_{b\to\infty} R(f;[a,b])$$

whenever this limit exists and is finite. Show that the improper Riemann integral for f may exist, while that for $|f|$ does not; and that not all improper-Riemann-integrable functions are Lebesgue-integrable. (*Hint:* Take $f(x) = \sin(x)\,/\,x$ and $a = 1$.)

4.8 Exercise: Show that $\lim_{n\to\infty} \int_0^\infty \frac{e^{-nx}}{\sqrt{x}}\,dx = 0$.

4.9 Exercise: Show that the function $F(\cdot)$ defined by the improper Riemann integral $F(t) = \int_0^\infty x^2\,e^{-tx}\,dx$ for $0 < t < \infty$ is continuous, and that its derivative exists and is given $F'(t) = -\int_0^\infty x^3\,e^{-tx}\,dx$.

C: MEASURES ON EUCLIDEAN SPACES

Let us try to examine how the theory of §1.4.A can be extended to higher-dimensional Euclidean spaces $(\Omega, \mathcal{F}) \equiv (\mathbf{R}^d, \mathcal{B}(\mathbf{R}^d))$ with $d \geq 2$. Suppose that μ is a $\sigma-$finite measure on this space, and consider the function $F(x) := \mu((-\infty, x]),\ x \in \mathbf{R}^d$ with the notation

$$a \leq b \ \Leftrightarrow_{\text{def}} \ a_i \leq b_i\,,\ \forall\, i = 1, \cdots, d \quad \text{and} \quad (a, b] := \{\omega \in \mathbf{R}^d \,|\, a_i < \omega_i \leq b_i,\ \ \forall\, i = 1, \cdots, d\}$$

for vectors $\omega = (\omega_1, \cdots, \omega_d) \in \mathbf{R}^d$. The so-defined function $F : \mathbf{R}^d \to [0, \infty)$ is called *cumulative distribution function of* μ. Setting $\Delta_{b_i - a_i}\, g(a_1, \cdots, a_{i-1}, a_i, a_{i+1}, \cdots, a_d) := g(a_1, \cdots, a_{i-1}, b_i, a_{i+1}, \cdots, a_d) - g(a_1, \cdots, a_{i-1}, a_i, a_{i+1}, \cdots, a_d)$ and $\Delta_{b-a}\, g(a) := \Delta_{b_1 - a_1} \cdots \Delta_{b_d - a_d}\, g(a_1, \cdots, a_d)$, we observe that this function satisfies

$$\lim_{\substack{b \to a \\ a \leq b}} F(b) \ = \ F(a)\,, \quad \forall \ \ a \in \mathbf{R}^d\,, \tag{4.4}$$

as well as the analogue of the property (1.7), namely $\Delta_{b-a}\, F(a) = \mu((a, b]) \geq 0\,,\ \ \forall\ a \leq b$ in $\mathbf{R}^d$. For instance, in the case $d = 2$ the left-hand side of this last expression is $\Delta_{b_1 - a_1}\, (\Delta_{b_2 - a_2}\, F(a_1, b_1)) = \Delta_{b_2 - a_2}\, F(b_1, a_2) - \Delta_{b_2 - a_2}\, F(a_1, a_2) = [F(b_1, b_2) - F(b_1, a_2)] - [F(a_1, b_2) - F(a_1, a_2)] = \mu((a, b])$. We have the following generalization of the notions in Definition 4.1.

4.2 Definition: Distribution Function on $\mathbf{R}^d$. A Borel-measurable function $F : \mathbf{R}^d \to [0, \infty)$ that satisfies the right-continuity property (4.4), as well as

$$\Delta_{b-a}\, F(a) \geq 0\,, \qquad \forall\ a \leq b \quad \text{in} \ \ \mathbf{R}^d\,, \tag{4.5}$$

is called a *distribution function on* $\mathbf{R}^d$. A distribution function that satisfies also

$$\lim_{x_1 \to \infty,\, \cdots,\, x_d \to \infty} F(x_1, \cdots, x_d) = 1 \qquad \text{and} \qquad \lim_{x_i \downarrow -\infty} F(x_1, \cdots, x_d) = 0\,,\ \forall\ i = 1, \cdots, d$$

is called a *probability distribution function* $\mathbf{R}^d$.

For $d = 1$, the distribution functions are the right-continuous, increasing functions. For $d \geq 2$, a function $F : \mathbf{R}^d \to [0, \infty)$ can be increasing and right-continuous in each of its variables separately, but fail to be a distribution function. Consider, for instance, the function $F(x_1, x_2) = \chi_{\{x_1 + x_2 \geq 0\}}$ on $\mathbf{R}^2$; with $a = (-1, -1)$ and $b = (2, 2)$ we have $\Delta_{b-a}\, F(a) \ = \ F(2, 2) - F(2, -1) - F(-1, 2) + F(-1, -1) = -1$.

4.3 Definition: Product Distribution Function. If $F_1, \cdots, F_d$ are distribution functions on $\mathbf{R}$, then

$$F(x) \equiv F(x_1, \cdots, x_d) := F_1(x_1) \cdots F_d(x_d), \qquad x \in \mathbf{R}^d \tag{4.6}$$

is a distribution function on $\mathbf{R}^d$, called the *product distribution function of* $F_1, \cdots, F_d$. If each $F_1, \cdots, F_d$ is a probability distribution function, then so is the product $F : \mathbf{R}^d \to [0,1]$ of (4.6).

We can follow now the procedure of §1.4.A. For a given distribution function $F : \mathbf{R}^d \to \mathbf{R}$, we define $\nu((a,b]) := \Delta_{b-a} F(a)$ on the class $\mathcal{G}$ of rectangles of the form $(a,b]$. We allow (some of) the coördinates of b to become $+\infty$, and (some of) the coördinates of a to become $-\infty$; in such cases, we replace $(a,b]$ by $(a,b] \cap \mathbf{R}^d$, whenever the former appears; we also allow $a = b$, that is, for the rectangle to become the empty set. With these conventions, $\mathcal{G}$ is an elementary family.

Consider then the algebra $\mathcal{E}$ consisting of finite disjoint unions of such rectangles, and extend ν to $\mathcal{E}$ by the recipe of (4.1). It can be seen, thanks to the conditions of Definition 4.2, that ν is a pre-measure on $\mathcal{E}$, so that the Hahn extension Theorem 3.3 can be invoked to ensure that ν has a unique extension μ_F to the $\sigma-$algebra of Borel sets $\mathcal{B}(\mathbf{R}^d) = \sigma(\mathcal{E})$ of $\mathbf{R}^d$. The completion $\bar{\mu}_F$ of μ_F is the **Lebesgue-Stieltjes measure induced by F on $\mathcal{B}(\mathbf{R}^d)$**; both $\bar{\mu}_F$ and μ_F are regular (Exercise 1.5). Furthermore, if F is a probability distribution function, then μ_F is a probability measure. By analogy with (4.2), integrals with respect to this measure are denoted as

$$\int_{(a,b]} f \, d\mu_F \;\equiv\; \int_{a_1}^{b_1} \cdots \int_{a_d}^{b_d} f(x_1, \cdots, x_d)\, dF(x_1, \cdots, x_d), \tag{4.2$'$}$$

for $-\infty \le a_i < b_i < \infty$, $i = 1, \cdots, d$.

4.1 Example : Lebesgue Measure on $\mathcal{B}(\mathbf{R}^d)$. The measure λ, induced by the product distribution function $F(x) = x_1 \cdots x_d$ on $\mathcal{B}(\mathbf{R}^d)$, assigns to each rectangle its volume

$$\lambda((a,b]) = \prod_{i=1}^{d} (b_i - a_i), \quad (a,b] \in \mathcal{G}.$$

The completion $\bar{\lambda}$ of λ is called **Lebesgue measure on $\mathcal{B}(\mathbf{R}^d)$**. Just as in Remark 4.1, it is translation-, reflection- and rotation-invariant.

D: FUNCTIONS OF A REAL VARIABLE

Suppose $h : [a,b] \to \mathbf{R}$ is a continuous function on a finite interval of the real line. Then the Fundamental Theorem of Calculus asserts that

$$\frac{d}{dx}\left(\int_a^x h(u)\,du\right) = h(x) \tag{4.7}$$

holds for every $x \in [a,b]$: *the derivative of the indefinite integral of a continuous function coincides with the function.* Without the continuity assumption, it is reasonable to expect that (4.7) will typically fail; but "not by much" if the function h is integrable, as the following result shows.

4.1 Theorem: Lebesgue's Differentiation. *If $h : [a,b] \to \mathbf{R}$ is integrable with respect to Lebesgue measure λ on $[a,b]$, then (4.7) holds for $\lambda-$a.e. $x \in [a,b]$.*

Proof: Writing $h = h^+ - h^-$, we see that it is enough to consider $h \geq 0$. With $H(x) := \int_a^x h(u)\,du$, we need then to establish

$$\lim_{\delta \downarrow 0} \frac{H(x+\delta) - H(x)}{\delta} = \lim_{\delta \downarrow 0} \frac{H(x) - H(x-\delta)}{\delta} = h(x)\,, \qquad \forall\; x \in [a,b] \setminus B \tag{4.8}$$

for some Borel subset B of $[a,b]$ with $\lambda(B) = 0$. Now for every rational number $q \in \mathbf{Q}$, let us define

$$h_q(x) := \big(h(x) - q\big)^+\,, \qquad H_q(x) := \int_a^x h_q(u)\,du \tag{4.9}$$

and note

$$\frac{d}{dx} H_q(x) = 0\,, \quad \text{for } \lambda - \text{a.e. } x \in [a,b] \quad \text{with } \; h(x) \leq q$$

from Exercise 4.11 below applied to the measure $\mu_q(A) = \int_A h_q(u)\,du$. Consequently, we shall take

$$B := \bigcup_{q \in \mathbf{Q}} \left\{ x \in [a,b] \,\middle|\, h(x) \leq q \;\text{ and }\; \frac{d}{dx} H_q(x) \neq 0 \right\} \in \mathcal{B}([a,b])$$

in (4.8), since $\lambda(B) = 0$ for this set. Now from the definition (4.9) we have the inequality

$$\frac{1}{\delta} \int_x^{x+\delta} h(u)\,du \leq q + \frac{1}{\delta}\big[H_q(x+\delta) - H_q(x)\big]$$

for $a \leq x < x+\delta \leq b$, as well as

$$\frac{1}{\delta} \int_{x-\delta}^{x} h(u)\,du \leq q + \frac{1}{\delta}\big[H_q(x) - H_q(x-\delta)\big]$$

for $a \le x-\delta < x \le b$. For $x \notin B$ and $h(x) < q$, we have $\frac{d}{dx} H_q(x) = 0$ and thus

$$D^+ H(x) \,:=\, \overline{\lim_{\delta\downarrow 0}}\, \frac{1}{\delta} \int_x^{x+\delta} h(u)\, du \,\le\, q\,;$$

similarly, we obtain

$$D^- H(x) \,:=\, \overline{\lim_{\delta\downarrow 0}}\, \frac{1}{\delta} \int_{x-\delta}^{x} h(u)\, du \,\le\, q\,.$$

Letting $q \downarrow h(x)$, this leads to $D^+ H(x) \le h(x)\,,\ \ D^- H(x) \le h(x)\,.$ Considering the function $-h$ instead of h, we obtain likewise:

$$D_+ H(x) \,:=\, \underline{\lim}_{\delta\downarrow 0}\, \frac{1}{\delta} \int_x^{x+\delta} h(u)\, du \,\ge\, h(x)\,,$$

$$D_- H(x) \,:=\, \underline{\lim}_{\delta\downarrow 0}\, \frac{1}{\delta} \int_{x-\delta}^{x} h(u)\, du \,\ge\, h(x)\,.$$

Thus, all four numbers $D^\pm H(x)\,,\ D_\pm H(x)$ are equal to $h(x)$, and (4.8) follows. $\diamond$

4.10 Exercise: Let $\mathcal{U} = \{I_\alpha\}_{\alpha\in A}$ be a collection of open intervals of the real line, with $I = \cup_{\alpha\in A} I_\alpha$ a bounded set. Then, for any $t < \lambda(I)$, there exists a finite sub-collection $\{I_1, \cdots, I_r\} \subset \mathcal{U}$ of disjoint intervals, such that $t \le 3 \cdot \sum_{j=1}^r \lambda(I_j)\,.$

4.11 Exercise: Using Exercise 4.10, show that if μ is a finite measure on $\mathcal{B}([a,b])$, and if $A \in \mathcal{B}([a,b])$ has $\mu(A) = 0$, then

$$\frac{d}{dx}\, \mu\big([a,x]\big) \,=\, 0\,, \qquad \text{for } \ \lambda-\text{a.e. } \ x \in A\,.$$

4.1 Definition: Functions of Finite Variation. For any function $f : [a,b] \to \mathbf{R}$ and any partition $\Pi = \{x_0, x_1, \cdots, x_n\}$ of the form $a = x_0 < x_1 < \cdots < x_n = b$ for the bounded interval $[a,b]$, let

$$V^f(\Pi) \,:=\, \sum_{j=1}^n \big| f(x_j) - f(x_{j-1}) \big|\,, \qquad V^f \equiv V^f(a,b) \,:=\, \sup_{\Pi\in\mathcal{P}} V_f(\Pi)$$

denote the *variation of f on the partition* Π, and the *total variation of f on* $[a,b]$, respectively. If $\widetilde{\Pi}$ is a refinement of Π, that is, every point of Π is also a point of $\widetilde{\Pi}$, then clearly $V^f(\Pi) \le V^f(\widetilde{\Pi})$. We say that f **is of finite variation on** $[a,b]$, if $V^f(a,b) < \infty$.

It is checked relatively easily that $V^f(a,b) = V^f(a,c) + V^f(c,b)$, for $a < c < b$. Every increasing function $F : [a,b] \to \mathbf{R}$ is of finite variation $V^F(a,b) = F(b) - F(a)$, because then $V^F(\Pi) = F(b) - F(a)$ for any partition Π of the interval $[a,b]$. Likewise, every function $f : [a,b] \to \mathbf{R}$ of the form

$$f \;=\; F - G\,, \quad \text{where both } \; F\,, G \; \text{ are increasing}\,, \tag{4.10}$$

is of finite variation. As it turns out, all functions of finite variation are of this type.

4.2 Theorem: *A function $f : [a,b] \to \mathbf{R}$ is of finite variation, if and only if it is of the form (4.10).*

Proof: If f is of finite variation, then the functions

$$F(x) \;:=\; V^f(a,x)\,, \quad G(x) \;:=\; F(x) - f(x)\,, \qquad x \in [a,b] \tag{4.11}$$

are increasing; indeed, we have $V^f(a,x) \;\leq\; V^f(a,y)$ and

$$G(y) - G(x) \;=\; F(y) - F(x) - \big[f(y) - f(x)\big] \;\geq\; 0 \quad \text{for} \;\; a \leq x < y \leq b\,,$$

the latter because $f(y) - f(x) \;\leq\; V^f(a,y) - V^f(a,x) = V^f(x,y)\,.$ $\diamond$

4.2 Proposition: *The discontinuities of every function of finite variation are exhausted by a set which is at most countable.*

Proof: It suffices to consider an increasing function $F : [a,b] \to \mathbf{R}$; then the limits $F(x+) \;:=\; \lim_{u \downarrow x} F(u) \;\geq\; \lim_{v \uparrow x} F(v) \;=:\; F(x-)$ exist at every $x \in (a,b)$, and with $a \leq y < x < z \leq b$ we have $[F(x) - F(y)] + [F(z) - F(x)] \leq [F(b) - F(a)]$, thus

$$[F(x) - F(x-)] + [F(x+) - F(x)] \;\leq\; [F(b) - F(a)]\,.$$

Similarly, for any finite set $E \subset (a,b)$:

$$\sum_{x \in E} \Big([F(x) - F(x-)] + [F(x+) - F(x)] \Big) \;\leq\; [F(b) - F(a)]\,.$$

Therefore, the set $\{\, x \in [a,b] \,:\, F(x) - F(x-) > 1/n \;\text{ or }\; F(x+) - F(x) > 1/n \,\}$ is finite for any $n \in \mathbf{N}$, which means that the set $\{\, x \in [a,b] \,:\, F(x) > F(x-) \;\text{ or }\; F(x+) > F(x) \,\}$ is at most countable. $\diamond$

4.1 Remark: Note that *any countable set can be the set of discontinuities of some increasing function.* Just consider the rational numbers $\mathbf{Q} = \{q_n\}_{n \in \mathbf{N}}$, and define the probability distribution function

$$F(x) \;:=\; \sum_{\substack{n \in \mathbf{N} \\ q_n \leq x}} 2^{-n}\,, \quad x \in \mathbf{R}\,.$$

Indeed, it is straightforward to see that this function is right-continuous and increasing with $F(-\infty) = 0$, $F(+\infty) = 1$. It increases by jumps only, at *each* rational number; thus, the set of points of discontinuity is dense in the real line.

4.12 Exercise: The function f defined by $f(x) = x^\alpha \cdot \sin\big(1/x^\beta\big)$, $0 < x \le 1$ and $f(0) = 0$, is of finite variation on the interval $[0,1]$ if $\alpha > \beta > 0$, but not if $\beta > \alpha > 0$.

4.13 Exercise: Consider the *van der Waerden function*

$$f(x) := \sum_{n \in \mathbf{N}_0} 4^{-n}\, h(4^n x)\,, \quad x \in \mathbf{R}$$

where h is defined by $h(x) = x$ for $0 \le x \le 1/2$, by $h(x) = 1 - x$ for $1/2 \le x \le 1$, and then by periodicity (period 1) on the entire real line. Show that f is continuous, but fails to be differentiable at any $x_0 \in \mathbf{R}$. Argue also that f cannot be of finite variation on *any* interval. $\big($*Hint:* Consider increments of the type $\pm 4^n\,[f(x_0 \pm 4^{-n}) - f(x_0)]\,\big)$.

We shall continue this study of functions of a real variable in §1.7.A.

D: ELEMENTS OF CALCULUS FOR FUNCTIONS OF FINITE VARIATION

Let us consider now a right-continuous function $A : [0,\infty) \to \mathbf{R}$ with $A(0) = 0$ which is of finite variation $V(t) \equiv V^A(0,t)$ on every compact interval $[0,t]$. We know that we can decompose this function as $A = A^+ - A^-$, where $A^\pm$ are right-continuous and increasing functions with $A^\pm(0) = 0$ and

$$V(t) = A^+(t) + A^-(t)\,, \qquad A^\pm(t) = \mu^\pm\big((0,t]\big) \qquad \text{for all } \ t \in [0,\infty)\,,$$

for two measures μ^+, μ^- on $\mathcal{B}([0,\infty))$. In particular, $\Delta A(t) := A(t) - A(t-) = \mu^+(\{t\}) - \mu^-(\{t\})$. As a consequence, Lebesgue-Stieltjes integrals

$$\int_0^t h(s)\, dA(s) \;:=\; \int_{(0,t]} h\, d\mu^+ - \int_{(0,t]} h\, d\mu^- \;, \qquad 0 \le t < \infty$$

can be defined for any Borel-measurable $h : [0,\infty) \to \mathbf{R}$ which is locally bounded (that is, bounded on compact intervals); the resulting function $t \mapsto \int_0^t h(s)\, dA(s)$ is then right-continuous and of finite variation on compact intervals. From Proposition 4.2, the function A can have at most countably-many discontinuities, so

$$A^c(t) \;:=\; A(t) - \sum_{0 \le s \le t} \Delta A(s)\,, \qquad 0 \le t < \infty$$

defines well a *continuous* function, of finite variation on compact intervals.

4.14 Exercise: Integration by Parts. With A as above, and with B another function with the same properties, we have

$$A(t)B(t) \;=\; \int_0^t A(s)\,dB(s) \;+\; \int_0^t B(s-)\,dA(s)$$

for every $0 \le t < \infty$ and, symmetrically,

$$A(t)B(t) \;=\; \int_0^t A(s-)\,dB(s) \;+\; \int_0^t B(s-)\,dA(s) \;+\; \sum_{0 \le s \le t} \Delta A(s)\,\Delta B(s)\,.$$

4.15 Exercise: Chain Rule. With A as above, and with $\Phi : \mathbf{R} \to \mathbf{R}$ a function of class $\mathcal{C}^1(\mathbf{R})$, the composite function $t \mapsto \Phi\big(A(t)\big)$ is also of finite variation on compact intervals, and we have the chain rule

$$\Phi\big(A(t)\big) = \Phi(0) + \int_0^t \Phi'\big(A(s-)\big)\,dA(s) + \sum_{0 \le s \le t} \Big(\Phi\big(A(s)\big) - \Phi\big(A(s-)\big) - \Phi'\big(A(s-)\big)\Delta A(s)\Big)\,.$$

4.16 Exercise: An Integral Equation. With A as above,

$$Z(t) \;=\; \prod_{0 \le s \le t} \big(1 + \Delta A(s)\big) \cdot e^{\,A^c(t)}\;, \qquad 0 \le t < \infty$$

provides the unique locally bounded solution of the integral equation

$$Z(t) \;=\; 1 + \int_0^t Z(s-)\,dA(s)\;, \qquad 0 \le t < \infty\,.$$

4.17 Exercise: Change of Variables. With A as above and *increasing*, consider its inverse

$$\Gamma(u) \;:=\; \inf\{t \ge 0 \,|\, A(t) > u\,\}\,, \qquad 0 \le u < \infty\,,$$

with the conventions that the infimum of an empty set is $+\infty$, and $\Gamma(0-) := 0$.

Show that this function is increasing, right-continuous and that we have for it

$$\Gamma(u-) \;=\; \inf\{t \ge 0 \,|\, A(t) \ge u\,\} \quad \text{and} \quad A\big(\Gamma(u)\big) \ge u\,, \qquad 0 \le u < \infty$$

as well as

$$A(t) \;=\; \inf\{\,u \ge 0 \,|\, \Gamma(u) > t\,\} \quad \text{and} \quad \Gamma\big(A(t)\big) \ge t\,, \qquad 0 \le t < \infty$$

(imperative: draw a picture!). Show also that for any Borel-measurable function $h : [0,\infty) \to [0,\infty)$ we have the *change-of-variable* formula

$$\int_{[0,\infty)} h(s)\,dA(s) \;=\; \int_0^\infty h\big(\Gamma(u)\big)\,\chi_{\{\Gamma(u)<\infty\}}\,du\,.$$