Suppose that μ, ν are two measures on the *same* measurable space $(\Omega, \mathcal{F})$. We say that ν *is absolutely continuous with respect to* μ, and write $\nu < \mu$, if $A \in \mathcal{F}$ and $\mu(A) = 0$ imply $\nu(A) = 0$. For example, Exercise 2.3 (i) shows that this is the case when

$$\nu(A) = \int_A h \, d\mu \quad \forall \ A \in \mathcal{F} \quad \text{for some} \quad h \in \mathbf{L}^1(\mu) \cap \mathbf{L}^0_+ \, . \tag{7.1}$$

We shall see below that this example is not so special: to wit, if μ is $\sigma-$finite and ν is finite, then $\nu < \mu$ implies (7.1). This is the content of the celebrated Radon-Nikodým Theorem 7.2.

We say that μ, ν are *equivalent* (and write $\mu \sim \nu$), if they are mutually absolutely continuous, that is $\nu < \mu$ and $\mu < \nu$. Finally, we say that μ, ν are *singular* (and write $\mu \perp \nu$), if there exists a measurable set $A \in \mathcal{F}$ such that $\mu(A) = \nu(A^c) = 0$.

7.1 THEOREM : LEBESGUE DECOMPOSITION. *Let* $(\Omega, \mathcal{F})$ *be a measurable space, and that* μ, ν *are two* σ*-finite measures on it. Then there exist measures* ν_{ac} , ν_s *with*

$$\nu = \nu_{ac} + \nu_s \ ; \qquad \nu_{ac} < \mu \, , \quad \nu_s \perp \mu \, ,$$

and this decomposition is unique.

For instance, let $\lambda\big|_{[a,b]}$ denote Lebesgue measure on the interval $[a, b]$, and take $\mu = \lambda\big|_{[0,2]}$, $\nu = \lambda\big|_{[1,3]}$. Then $\nu_{ac} = \lambda\big|_{[1,2]}$ and $\nu_s = \lambda\big|_{[2,3]}$.

7.2 THEOREM: RADON-NIKODÝM. *Let the two measures* μ, ν *on the measurable space* $(\Omega, \mathcal{F})$ *be* σ*-finite and finite, respectively, with* $\nu < \mu$. *Then there exists a unique (up to* $\mu-$*a.e. equivalence) function* $h \in \mathbf{L}^1(\mu) \cap \mathbf{L}^0_+$ *such that* $\nu(A) = \int_A h \, d\mu \, , \ \forall \ A \in \mathcal{F}$ *as in (7.1).*

The function $h : \Omega \to [0, \infty)$ of (7.1) is called the **Radon-Nikodým derivative** of ν with respect to μ, and is denoted $h = \frac{d\nu}{d\mu}$. This notation suggests correct conclusions. For instance, if $\nu < \mu$ and $f \in \mathbf{L}^1(\nu)$, then $\int_\Omega f \left(\frac{d\nu}{d\mu} \right) d\mu = \int_\Omega f \, d\nu$; and if $\rho < \nu < \mu$ are finite measures, then $\frac{d\rho}{d\mu} = \frac{d\rho}{d\nu} \cdot \frac{d\nu}{d\mu}$, $\mu -$ a.e.

We shall provide a unified proof of Theorems 7.1 and 7.2, in several steps.

Step 1: Reduction to the finite-measure case $\mu(\Omega) + \nu(\Omega) < \infty$. By the assumptions of Theorem 7.1, we can write $\Omega = \cup_{n \in \mathbf{N}} E_n$, where $\{E\}_{n \in \mathbf{N}}$ are pairwiswe disjoint and

$\mu(E_n) + \nu(E_n) < \infty$. For any measure ρ on $(\Omega, \mathcal{F})$, let $\rho_n(C) := \rho(C \cap E_n)$, $n \in \mathbf{N}$ so that $\rho(C) = \sum_{n \in \mathbf{N}} \rho_n(C)$. Then $\rho(C) = 0$ is equivalent to $\rho_n(C) = 0$, $\forall\, n \in \mathbf{N}$; and we have $\nu < \mu \Leftrightarrow \nu_n < \mu, \forall\, n \in \mathbf{N} \Leftrightarrow \nu_n < \mu_n, \forall\, n \in \mathbf{N}$ as well as $\nu \perp \mu \Leftrightarrow \nu_n \perp \mu, \forall\, n \in \mathbf{N} \Leftrightarrow \nu_n \perp \mu_n, \forall\, n \in \mathbf{N}$.

On the other hand, for any sequence $\{\rho_n\}_{n \in \mathbf{N}}$ of measures on $(\Omega, \mathcal{F})$ with $\rho_n(E_n^c) = 0$, $\forall\, n \in \mathbf{N}$, the recipe $\rho(C) := \sum_{n \in \mathbf{N}} \rho_n(C)$ defines a measure for which $\rho < \mu \Leftrightarrow \rho_n < \mu_n, \forall\, n \in \mathbf{N}$ as well as $\rho \perp \mu \Leftrightarrow \rho_n \perp \mu_n, \forall\, n \in \mathbf{N}$. Thus, in proving Theorem 7.1, it is enough to assume that both measures μ, ν are finite.

Now suppose we have proved Theorem 7.2 on each of the sets E_n, with some measurable $h_n : E_n \to [0, \infty)$, $\forall\, n \in \mathbf{N}$. Then $h(\omega) \equiv h_n(\omega)$, $\omega \in E_n$ defines a measurable function $h : \Omega \to [0, \infty)$ for which $\int_A h \, d\mu = \sum_{n \in \mathbf{N}} \int_{A \cap E_n} h_n \, d\mu_n = \sum_{n \in \mathbf{N}} \nu(A \cap E_n) = \nu(A) < \infty$, $\forall\, A \in \mathcal{F}$, by assumption of Theorem 7.2; in other words, $h \in \mathbf{L}^1(\mu) \cap \mathbf{L}^0_+$, and Theorem 7.2 then holds in all generality.

Step 2: Hilbert-space argument (von Neumann (1940)). Consider the Hilbert space $\mathcal{H} := \mathbf{L}^2 \equiv \mathbf{L}^2(\Omega, \mathcal{F}, \mu + \nu)$, and observe that $\nu \le \mu + \nu$ and that the identity map from $\mathbf{L}^2$ into $\mathbf{L}^1 \equiv \mathbf{L}^1(\Omega, \mathcal{F}, \mu + \nu)$ is continuous. This is because the linear operator $\mathcal{H} \ni f \mapsto T(f) := \int_\Omega f \, d\nu \in \mathbf{R}$ is bounded, thus also continuous:

$$|T(f)| \le \int_\Omega |f| \, d\nu \le \int_\Omega |f| \, d(\mu + \nu) = ||f||_1 \le ||f||_2 \cdot \sqrt{\mu(\Omega) + \nu(\Omega)} \,,$$

from Exercise 5.7. Thus, from the Riesz Representation Theorem $B.2$ (Appendix B), there exists a function $g \in \mathcal{H}$ such that for every $f \in \mathcal{H}$ we have

$$\int_\Omega f \, d\nu = \int_\Omega f g \, d(\mu + \nu) \,. \tag{7.2}$$

Remark: Note that, if indeed $\nu = \int h \, d\mu$ as postulated in the Radon-Nikodým theorem, then (7.2) becomes

$$\int_\Omega f h \, d\mu = \int_\Omega f g \, d\mu + \int_\Omega f g h \, d\mu \,, \quad \forall \;\; f \in \mathcal{H} \,,$$

suggesting rather strongly that $h(1 - g) = g$ should hold $\mu-$a.e. and that we should choose $h = g/(1 - g)$ on $\{g \ne 1\}$. This observation will be very valuable below.

The equation (7.2) can be written equivalently, for every $f \in \mathcal{H}$, as

$$\int_\Omega f(1 - g) \, d\nu = \int_\Omega f g \, d\mu \,, \tag{7.2$'$}$$

$$\int_\Omega f(1 - g) \, d(\mu + \nu) = \int_\Omega f \, d\mu \,. \tag{7.2$''$}$$

It is easy to see that $0 \le g \le 1$, $(\mu+\nu)$–a.e. For if $(\mu+\nu)(\{g<0\}) > 0$, then (7.2) with $f = \chi_{\{g<0\}}$ gives $\nu(\{g<0\}) = \int_{\{g<0\}} g\, d(\mu+\nu) < 0$; and if $(\mu+\nu)(\{g>1\}) > 0$, then $(7.2)''$ with $f = \chi_{\{g>1\}}$ gives $\mu(\{g>1\}) = \int_{\{g>1\}} (1-g)\, d(\mu+\nu) < 0$. Both these conclusions are absurd.

In conclusion, we may assume that $0 \le g(\omega) \le 1$, $\forall\, \omega \in \Omega$; *then (7.2)-$(7.2)''$ hold for any non-negative, measurable* $f \in \mathbf{L}^0_+$, by the Monotone Convergence Theorem.

Step 3: Existence in Theorem 7.1. Let us start by observing that the set $A := \{\omega \in \Omega \,|\, g(\omega) = 1\}$ has measure $\mu(A) = 0$ (just read $(7.2)'$ with $f = \chi_A$). In other words, we have $g < 1$, μ–a.e. Thus, for the two measures

$$\nu_s(E) := \nu(E \cap A), \quad \nu_{ac}(E) := \nu(E \cap A^c); \qquad E \in \mathcal{F} \tag{7.3}$$

we have $\nu = \nu_{ac} + \nu_s$, $\nu_s \perp \mu$ (because $\mu(A) = 0$, $\nu_s(A^c) = \nu(\emptyset) = 0$), as well as $\nu_{ac} < \mu$. This latter conclusion can be seen from the fact that, if $E \subseteq A^c = \{g < 1\}$ has $\mu(E) = 0$, then reading $(7.2)'$ with $f = \chi_E$ gives $\int_E (1-g)\, d\nu = \int_E g\, d\mu = 0$, whence $\nu(E) = 0$, $\nu_{ac}(E) = \nu(E) = 0$.

Step 4: Uniqueness in Theorem 7.1 . Suppose that $\nu = \rho + \sigma$, where $\rho < \mu$, $\sigma \perp \mu$. Then $\rho(A \cap E) = 0$ (because $\mu(A) = 0$), whence

$$\nu_s(E) = \nu(E \cap A) = \rho(A \cap E) + \sigma(A \cap E) = \sigma(A \cap E) \le \sigma(E)$$

for any measurable set $E \in \mathcal{F}$; consequently $\nu_s \le \sigma$ and $\nu_{ac} - \rho = \sigma - \nu_s$ is a measure, both absolutely *and* singular with respect to μ. Thus, this measure is identically equal to zero: $\nu_{ac} \equiv \rho$, $\nu_s \equiv \sigma$.

Step 5: Proof of Theorem 7.2 . Now suppose $\nu \equiv \nu_{ac}$ ($\nu_s \equiv 0$) and define a non-negative, measurable function as $h := g/(1-g)$ on A^c, $h := 0$ on A. Then $f := h\chi_E$, $E \in \mathcal{F}$ is in $\mathbf{L}^0_+$ and we get $\int_E h(1-g)\, d(\mu+\nu) = \int_E h\, d\mu$, thus

$$\begin{aligned}
\int_E h\, d\mu &= \int_{E\cap A^c} g\, d(\mu+\nu) && \text{(from } (7.2)''\text{)} \\
&= \int_\Omega g\chi_{E\cap A^c}\, d(\mu+\nu) \;=\; \nu(E \cap A^c) && \text{(from (7.2))} \\
&= \nu_{ac}(E) = \nu(E). && \text{(by assumption)}
\end{aligned}$$

The uniqueness follows easily.

7.1 Example: If C is the Cantor set and F the Cantor function of Appendix A, then $\mu_F([0,1] \setminus C) = 0$ and $\lambda(C) = 0$; in other words, the Lebesgue-Stieltjes measure

μ_F induced on $\mathcal{B}([0,1])$ by the Cantor function F, is singular with respect to Lebesgue measure λ.

7.1 Remark: Signed Measures. On a measurable space $(\Omega, \mathcal{F})$ let us consider a $\sigma-$finite measure μ as well as the collection $\mathcal{SM}(\mu)$ of its **signed measures**: that is, all set functions $\nu : \mathcal{F} \to \mathbf{R}$ that can be written as a difference

$$\nu = \nu_1 - \nu_2 \quad \text{of two finite measures } \nu_1, \nu_2 \text{ with } \nu_1 < \mu, \ \nu_1 < \mu .$$

In particular, an element $\nu \in \mathcal{SM}(\mu)$ is a *countably additive*, real-valued set function, namely

$$\nu\left(\bigcup_{n\in\mathbf{N}} E_n\right) = \sum_{n\in\mathbf{N}} \nu(E_n) \quad \text{for any sequence } \{E_n\}_{n\in\mathbf{N}} \subseteq \mathcal{F} \text{ of pairwise disjoint sets,}$$

and is absolutely continuous with respect to μ in the sense that $\nu(E) = 0$ holds for any $E \in \mathcal{F}$ with $\mu(E) = 0$. The prototypical example is

$$\nu(E) = \int_E f\, d\mu\,, \quad E \in \mathcal{F} \qquad \text{for some } f \in \mathbf{L}^1(\mu)\,;$$

for then we can write $\nu = \nu_+ - \nu_-$ with $\nu_\pm(E) = \int_E f^\pm\, d\mu\,,\, E \in \mathcal{F}$.

On the other hand, for every signed measure $\nu = \nu_1 - \nu_2 \in \mathcal{SM}(\mu)$ as above, the Radon-Nikodým Theorem 7.2 gives $\nu_i(E) = \int_E f_i\, d\mu\,,\, E \in \mathcal{F}$ with $f_i \in \mathbf{L}^1(\mu) \cap \mathbf{L}^0_+$ for $i = 1, 2$, thus $\nu(E) = \int_E f\, d\mu$ for $f := f_1 - f_2 \in \mathbf{L}^1(\mu)$.

In other words, every element (function) of $\mathbf{L}^1(\mu)$ is seen to correspond to an element (signed measure) of $\mathcal{SM}(\mu)$, and vice-versa.

7.1 Definition: If μ, ν are probability measures on $(\Omega, \mathcal{F})$, we define the *total variation distance*

$$||\mu - \nu|| := \sup_{A\in\mathcal{F}} |\mu(A) - \nu(A)| \tag{7.4}$$

between them, as well as the *relative entropy of μ with respect to ν:*

$$H(\mu|\nu) := \int_\Omega \log\left(\frac{d\mu}{d\nu}\right) d\mu\,, \quad \text{if } \mu < \nu \tag{7.5}$$

and $H(\mu|\nu) = \infty$ otherwise. It can be shown that we have *Csiszár's inequality*

$$2\,||\mu - \nu||^2 \leq H(\mu|\nu)\,. \tag{7.6}$$

7.1 Exercise: Scheffé's Theorem. Let μ, $\{\mu_n\}$ be probability measures on $(\Omega, \mathcal{F})$ which are absolutely continuous with respect to the 'reference' measure λ (not necessarily finite). Set $X = d\mu/d\lambda$, $X_n = d\mu_n/d\lambda$ for $n \in \mathbf{N}$, and assume that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ holds for $\lambda-$a.e. $\omega \in \Omega$. Show that

$$||\mu_n - \mu|| = \frac{1}{2} \int |X_n - X| \, d\lambda \longrightarrow 0\,, \quad \text{as } n \to \infty\,. \tag{7.7}$$

7.2 Exercise: Every distribution function $F : \mathbf{R} \to \mathbf{R}$ with $F(\infty) - F(-\infty) < \infty$ has a *unique* decomposition of the form $F = F_{ac} + F_{sc} + F_d$, where

(i) $F_{ac}(x) = \int_{-\infty}^{x} f(u)\, du$, $x \in \mathbf{R}$, for some $f \in \mathbf{L}^+ \cap \mathbf{L}^1(\lambda)$, is the part which is *absolutely continuous* with respect to Lebesgue measure λ;

(ii) $F_{sc}(\cdot)$ is continuous but *singular,* in the sense that the corresponding Lebesgue-Stieltjes measure $\mu_{F_{sc}}$ on $\mathcal{B}(\mathbf{R})$ is singular with respect to λ; and

(iii) $F_d(\cdot)$ is *purely discontinuous:* there exist an (at most) countable set $\mathcal{K} \subset \mathbf{R}$ and a collection $\{p(k)\}_{k\in\mathcal{K}}$ of nonnegative numbers, such that $F(x) = \sum_{\substack{k\in\mathcal{K}\\ k\le x}} p(k)$, $x \in \mathbf{R}$.

7.3 Exercise: Csiszár's inequality. Prove the inequality (7.6). (*Hint:* Use the identity in (7.7).)

7.4 Exercise: Let $\{\mu_\alpha\}_{\alpha\in I}$ and ν be probability measures on $(\Omega, \mathcal{F})$ with $\mu_\alpha < \nu$ for every $\alpha \in I$. If we have $\sup_{\alpha\in I} H(\mu_\alpha|\nu) < \infty$, then the family of random variables $\{d\mu_\alpha \,/\, d\nu\}_{\alpha\in I}$ is $\nu-$uniformly integrable.

7.5 Exercise: Projective distance between measures. Let $\mathcal{V}$ be the space of all finite measures on $(\Omega, \mathcal{F})$. For any two elements λ and μ of this space, we shall write $\lambda \le \mu$ if $\lambda(E) \le \mu(E)$ holds for every $E \in \mathcal{F}$; and we shall say that λ *and* μ *are comparable,* if $\alpha\lambda \le \mu \le \beta\lambda$ for suitable real constants $0 < \alpha \le \beta < \infty$.

The *projective distance*, also called "Hölder metric", is defined on $\mathcal{V}$ as

$$h(\lambda,\mu) = \log \left(\frac{\sup_{E\in\mathcal{F},\, \mu(E)>0} \left(\frac{\lambda(E)}{\mu(E)}\right)}{\inf_{E\in\mathcal{F},\, \mu(E)>0} \left(\frac{\lambda(E)}{\mu(E)}\right)} \right)$$

if λ and μ are comparable, and by $h(\lambda,\mu) = \infty$ otherwise.

(i) Show that $h(\cdot,\cdot)$ is a pseudo-metric on $\mathcal{V}$, thus a true metric on the space of all probability measures on $(\Omega, \mathcal{F})$ that are comparable with a given, fixed measure $\lambda_0 \in \mathcal{V}$.

(ii) For a linear operator $T : \mathcal{V} \to \mathcal{V}$ we define the *Birkhoff contraction coëfficient*

$$\varrho(T) := \sup_{0<h(\lambda,\mu)<\infty} \left(\frac{h(T\lambda, T\mu)}{h(\lambda,\mu)} \right)$$

as well as the *$h-$diameter of* $T\mathcal{V}$, the transform of $\mathcal{V}$ by T, namely

$$D(T) := \sup_{\substack{\lambda,\mu \text{ comparable} \\ \lambda\in\mathcal{V},\, \mu\in\mathcal{V}}} h(T\lambda, T\mu)\,.$$

Show that

$$\varrho(T) = \tanh\left(\frac{1}{4}\, D(T)\right) .$$

(iii) Show that in the notation of (7.4) for the total variation distance we have

$$||\lambda - \mu|| \;\leq\; 1 \wedge \left(\frac{e^{\,h(\lambda,\mu)} \;-1}{2}\right) \;\leq\; \frac{h(\lambda,\mu)}{\log 3} \;.$$

A: FUNCTIONS OF A REAL VARIABLE

We shall resume in this subsection the study of functions of a real variable that we started in §1.4.D. Let us recall the concept and properties of finite variation from Definition 4.4, Theorem 4.2 and Proposition 4.1.

7.1 Proposition: *Every function* $f : [a,b] \to \mathbf{R}$ *of finite variation is differentiable* $\lambda-$*a.e. on* $[a,b]$.

Proof: From Theorem 4.2 and Proposition 4.1, is suffices to assume that f is increasing and right-continuous. For the finite measure μ on $\mathcal{B}([a,b])$ defined by the recipe

$$\mu\big((a,x]\big) \;:=\; f(x) - f(a) \qquad \text{for} \quad x \in [a,b]\,, \tag{7.8}$$

the Lebesgue Decomposition Theorem 7.1 gives $\mu = \mu_{ac} + \mu_s$, where μ_{ac} is absolutely continuous and μ_s is singular with respect to Lebesgue measure λ on $[a,b]$. In particular,

$$\mu_{ac}(A) \;=\; \int_A h\, d\lambda\,, \qquad A \in \mathcal{B}([a,b]) \tag{7.9}$$

by the Radon-Nikodým Theorem 7.2, for some $h : [a,b] \to [0,\infty)$ in $\mathbf{L}^1(\lambda)$. The equations (7.8), (7.9) imply

$$f(x) \;=\; f(a) + \int_a^x h(u)\, du \;+\; \mu_s\big((a,x]\big)\,, \quad a \leq x \leq b\,. \tag{7.10}$$

Now there exists a Borel set $A \subset [a,b]$ with $\lambda([a,b] \setminus A) = \mu_s(A) = 0$. From Exercise 4.11 we have $\frac{d}{dx}\,\mu_s\big((a,x]\big) \;=\; 0$ for $\lambda-$a.e. $x \in [a,b]$; whereas from this and Lebesgue's Differentiation Theorem 4.1 we conclude that $f'(x)$ exists and equals $h(x)$, for $\lambda-$a.e. $x \in [a,b]$. ◇

7.2 Definition: Absolute Continuity of Functions. A function $f : [a,b] \to \mathbf{R}$ is called *absolutely continuous*, if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

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

holds for any collection $\big\{(a_i, b_i)\big\}_{i=1}^{N}$ (finite, or countably infinite) of disjoint open intervals for which $\sum_{i=1}^{N}(b_i - a_i) < \delta$.

Every absolutely continuous function is uniformly continuous (just take $N = 1$ in the above definition), and of finite variation; see Exercise 7.7. In particular, every such function f can be written in the form $f = F - G$ of (4.10), where the increasing functions F and G in (4.11) are now also absolutely continuous.

If $f : [a,b] \to \mathbf{R}$ is is everywhere differentiable with derivative f' which is bounded, then f is absolutely continuous, because we have $|f(b_i) - f(a_i)| \le \max |f'| \cdot (b_i - a_i)$ by the mean-value theorem.

The following result asserts that a function $f : [a,b] \to \mathbf{R}$ is absolutely continuous, if and only if it is of the form (7.10) with $\mu_s \equiv 0$, i.e. $f(x) = f(a) + \mu\big((a,x]\big)$, $a \le x \le b$ for some signed measure μ which is absolutely continuous with respect to Lebesgue measure.

7.2 Proposition: *A function $f : [a,b] \to \mathbf{R}$ is absolutely continuous, if and only if it is the indefinite integral*

$$f(x) \, = \, f(a) + \int_a^x h(u)\, du \, , \quad a \le x \le b \tag{7.11}$$

of some $h \in \mathbf{L}^1([a,b])$; in which case $f'(x)$ exists and equals $h(x)$ for $\lambda-$a.e. $x \in [a,b]$.

Proof: (i) Suppose that (7.11) holds and, without loss of generality, that $h \ge 0$. Then $\mu(A) = \int_A h \, d\lambda$, $A \in \mathcal{B}([a,b])$ defines a finite measure with $\mu < \lambda$: for any given $\varepsilon > 0$, there exists $\delta > 0$ such that $\mu(A) < \varepsilon$ holds for every $A \in \mathcal{B}([a,b])$ with $\lambda(A) < \delta$. Take $A = \cup_i (a_i, b_i)$ as in the definition of absolute continuity $\big($in particular, with $\lambda(A) = \sum_i (b_i - a_i) < \delta \big)$ to conclude from (7.11) that

$$\sum_i \big[f(b_i) - f(a_i) \big] \, = \, \sum_i \int_{(a_i, b_i)} h \, d\lambda \, = \, \int_A h \, d\lambda \, = \, \mu(A) \, < \, \varepsilon \, .$$

Thus, f is absolutely continuous. The last claim follows from Theorem 4.1.

(ii) Now assume that f is absolutely continuous (and, without loss of generality, also increasing). If we can show $\mu < \lambda$ for the measure μ of (7.8), then the Radon-Nikodým

Theorem 7.2 will give $\mu\big((a,x]\big) = \int_a^x h(u)\,du\,,\ x \in [a,b]$ for some $h \geq 0$ in $\mathbf{L}^1(\lambda)$, which will lead to (7.11).

To argue this, suppose $\lambda(A) = 0$ for some $A \in \mathcal{B}([a,b])$. From the regularity property (1.9), consider a decreasing sequence $\{G_n\}_{n\in\mathbf{N}}$ of open sets $G_n = \cup_i \big(a_i^{(n)}, b_i^{(n)}\big)$, each of them a union of disjoint open intervals and such that $\lambda(G_n) \downarrow \lambda(A) = 0\,,\ \mu(G_n) \downarrow \mu(A)$. Now the absolute continuity of f implies that for arbitrary $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\sum_i \left(b_i^{(n)} - a_i^{(n)}\right) \;=\; \lambda(G_n) \;<\; \delta \tag{7.12}$$

gives $\sum_i [f(b_i^{(n)}) - f(a_i^{(n)})] < \varepsilon$. Select n large enough so that (7.12) holds, and note

$$\mu(G_n) \;=\; \sum_i \mu\Big(\big(a_i^{(n)}, b_i^{(n)}\big)\Big) \;=\; \sum_i \left[f\left(b_i^{(n)}\right) - f\left(a_i^{(n)}\right)\right] \;<\; \varepsilon\,.$$

Letting $n \to \infty$ we obtain $0 \leq \mu(A) \leq \varepsilon$, thus $\mu(A) = 0$ from the arbitrariness of $\varepsilon > 0$. $\diamond$

7.6 Exercise: Show that a function $f : [a,b] \to \mathbf{R}$

- can be continuous without being of finite variation;
- can be continuous and increasing (thus, of finite variation) without being absolutely continuous;
- that satisfies a Lipschitz condition $|f(x) - f(y)| \leq K\,|x-y|\,,\ \forall\ a \leq x, y \leq b$ for some $K \in (0,\infty)$, is absolutely continuous;
- can be continuous and everywhere differentiable, without being absolutely continuous (e.g., $f(0) = 0$ and $f(x) = x^2 \cdot \sin(1/x^2)$ for $0 < x \leq 1$).

7.7 Exercise: If a function $f : [a,b] \to \mathbf{R}$ is absolutely continuous on the bounded interval $[a,b]$, then it is of finite variation.