Notes on the Lebesgue-Radon-Nikodym Theorem

Eric A. Carlen[1]
Rutgers University

December 9, 2014

1 Introduction

1.1 DEFINITION (Mutually Singular)**.** Two positive measures μ_1 and μ_2 on a measurable space $(X, \mathcal{M})$ are *mutually singular* in case there is a measurable set A so that

$$\mu_1(A^c) = 0 \qquad \text{and} \qquad \mu_2(A) = 0 \ . \tag{1.1}$$

We denote the mutual singularity of μ_1 and μ_2 by writing $\mu_1 \perp \mu_2$.

Note that when (1.1) is satisfied, for any $E \in \mathcal{M}$,

$$\mu_1(E) = \mu_1(A \cap E) \qquad \text{and} \qquad \mu_2(E) = \mu_2(A^c \cap E) \ .$$

in this sense, "μ_1 lives on A, and μ_2 lives on the complement of A".

1.2 EXAMPLE. Let $(X, \mathcal{M}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Let μ_1 be Lebesgue measure on $\mathbb{R}$, and let μ_2 be the point mass at the origin, often called the *Dirac mass*. That is, for all $E \in \mathcal{B}_{\mathbb{R}}$, $\mu_2(E) = 1$ if $0 \in E$ and $\mu_2(E) = 0$ otherwise. Then with $A = \mathbb{R}\backslash\{0\}$, (1.1) is satisfied, and so μ_1 and μ_2 are mutually singular.

The measure μ_2 is the Lebesgue-Stieltjes measure associated to the right continuous function F where

$$F(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

through $\mu_2((a, b]) = F(b) - F(a)$.

For a second – more interesting – example, let $(X, \mathcal{M}) = ([0, 1], \mathcal{B}_{[0,1]})$, and let $F : [0, 1] \to [0, 1]$ be the Cantor function, which is continuous and monotone non-decreasing. Hence there is a unique Lebesgue-Stieltjes measure μ_2 such that $\mu_2((a, b]) = F(b) - F(a)$ for all $a < b$ in $[0, 1]$. Let C be the Cantor set. Then $\mu_2(C^c) = 0$ while the Lebesgue measure of C is zero. Thus, taking μ_1 to be Lebesgue measure, and $A = C^c$, (1.1) is again satisfied, and μ_1 and μ_2 are mutually singular.

1.3 DEFINITION (Absolutely continuous)**.** Let μ_1 and μ_2 be two measures on a measurable space $(X, \mathcal{M})$. Then μ_1 is *absolutely continuous* with respect to μ_2 in case for all measurable sets A,

$$\mu_2(A) = 0 \quad \Rightarrow \quad \mu_1(A) = 0 \ . \tag{1.2}$$

1.4 EXAMPLE. Let $(X, \mathcal{M})$ be a measure space, and let μ_2 be a measure on $(X, \mathcal{M})$. Let $h \geq 0$ be an integrable function on $(X\mathcal{M}, \mu_2)$. Define a mesure μ_1 on $(X, \mathcal{M})$ by

$$\mu_1(E) = \int_E h \mathrm{d}\mu_2$$

for all $E \in \mathcal{M}$. Then, as we have seen, μ_1 is a finite measure on $(X, \mathcal{M})$ with $\mu_1(X) = \|h\|_1$.

If $\mu_2(E) = 0$, then $1_E h = 0$ a.e. with respect to μ_2, and so

$$\mu_1(E) = \int_E h \mathrm{d}\mu_2 = \int_X 1_E h \mathrm{d}\mu_2 = 0$$

since the integral of a measurable integrand that equals zero almost everywhere is zero.

The *Radon-Nikodym Theorem*, proved below, says that when μ_1 and μ_2 are finite, all examples of absolute continuity are of this type.

The Lebesgue Decomposition Theorem provides conditions under which, given two measure μ and ν on $(X, \mathcal{M})$, there exists two measures: $\mu^{(\mathrm{s})}$ and $\mu^{(\mathrm{ac})}$ such that $\mu^{(\mathrm{s})} \perp \nu$, $\mu^{(\mathrm{ac})} << \nu$, and $\mu = \mu^{(\mathrm{s})} + \mu^{(\mathrm{ac})}$. The next lemma says that whenever such a decomposition of μ exists, the components $\mu^{(\mathrm{s})}$ and $\mu^{(\mathrm{ac})}$ are uniquely determined. Thus, it makes sense to refer to them, respectivley, as the singular and absolutely contininuous parts of μ with respect to ν.

1.5 LEMMA (Uniquenes of the Lebesgue Decomposition). *Let μ and ν be two measures on $(X, \mathcal{M})$. Suppose there exist measures λ_j and ρ_j, $j = 1, 2$, such that $\lambda_j \perp \nu$, $\rho_j << \nu$ and $\mu = \lambda_j + \rho_j$ for $j = 1, 2$. Then $\lambda_1 = \lambda_2$ and $\rho_1 = \rho_2$.*

Proof. By definition, there exists sets $A_j\ in \mathcal{M}$ such that

$$\nu(A_j) = 0 \qquad \text{and} \lambda_j(A_j^c) = 0 \quad \text{for } j = 1, 2\ .$$

Let $b - A_1 \cup A_2$. Then $\nu(B) \leq \nu(A_1) + \nu(A_2) = 0$. Hence $\nu(B) = 0$. Consequently, $\rho_j(B) = 0$ for $j = 1, 2$. Next, $B^c = A_1^c \cap A_2^c \subset A_j^c$ for $j = 1, 2$. Hence $\lambda_j(B^c) = 0$ for $j = 1, 2$.

Now, for any $E \in \mathcal{M}$, and each $j = 1, 2$.

$$\rho_j(E) = \rho_j(E \cap B) + \rho_j(E \cap B^c) = \rho_j(E \cap B^c) = \rho_j(E \cap B^c) + \lambda_j(E \cap B^c) = \mu(E \cap B)$$

where we have used, successively, the fact that $\rho_j(E \cap B) = 0$, $\lambda_j(E \cap B^c) = 0$, and $\mu = \lambda_j + \rho_j$. Thus, $\rho_j(E) = \mu(E \cap B)$ for $j = 1, 2$, which shows that $\rho_1 = \rho_2$. Finaly $\lambda_j = \mu - \rho_j$, so $\lambda_1 = \lambda_2$. □

2 The Main Theorems

2.1 THEOREM (Lebesgue Decomposition Theorem). *Let μ_1 and μ_2 be two finite measures on a measurable space $(X, \mathcal{M})$ Then there are measures $\mu_1^{(\mathrm{s})}$ and $\mu_1^{(\mathrm{ac})}$ so that*

$$\mu_1 = \mu_1^{(\mathrm{s})} + \mu_1^{(\mathrm{ac})}$$

where $\mu_1^{(\mathrm{s})}$ and μ_2 are mutually singular, and where $\mu_1^{(\mathrm{ac})}$ is absolutely continuous with respect to μ_2. Moreover, this decomposition into a singular and absolutely continuous parts is unique.

2.2 THEOREM (Radon–Nikodym Theorem). *Let μ_1 and μ_2 be two finite measures on a on a measurable space $(X, \mathcal{M})$. If μ_1 is absolutely continuous with respect to μ_2, there is a function h that is integrable with respect to μ_2 such that for all $E \in \mathcal{M}$,*

$$\mu_1(E) = \int_E h \mathrm{d}\mu_1 \ , \tag{2.1}$$

and moreover, h is unique up to a.e. equivalence.

The following proof of these theorems is due to Von Neumann.

Proof. Let μ_1 and μ_2 be two finite measures on $\mathcal{M}$. Define the positive finite Borel measure ν by

$$\nu = \mu_1 + \mu_2 \ .$$

Let $\mathcal{H}$ denote $L^2(X, \mathcal{M}, \nu)$. For all $f \in \mathcal{H}$, by the fact that $\nu \geq \mu_2$, and then the Cauchy-Schwarz inequality,

$$\int_X |f| \mathrm{d}\mu_2 \leq \int_X 1|f| \mathrm{d}\nu \leq \left(\int_X 1 \mathrm{d}\nu\right)^{1/2} \left(\int_X |f|^2 \mathrm{d}\nu\right)^{1/2} = (\nu(X))^{1/2} \left(\int_X |f|^2 \mathrm{d}\nu\right)^{1/2} \ . \tag{2.2}$$

Thus, for all $f \in \mathcal{H}$, $f \in L^1(X, \mathcal{M}, \mu_2)$, and we may define a linear functional L on $\mathcal{H}$ by

$$L(f) = \int_X f \mathrm{d}\mu_2 \ .$$

It follows from (2.2) that for all $f \in \mathcal{H}$,

$$|L(f)| \leq \int_X |f| \mathrm{d}\mu_2 \leq (\nu(X))^{1/2} \|f\|_{\mathcal{H}} \ .$$

Therefore, L is bounded, and by the Riesz Representation Theorem, there exists a unique function $g \in \mathcal{H}$ such that

$$\int_X f \mathrm{d}\mu_2 = \int_X f g \mathrm{d}\nu \tag{2.3}$$

for all $f \in \mathcal{H}$. Since $\nu = \mu_1 + \mu_2 \geq \mu_2$, it follows immediately that for all $f \geq 0$,

$$\int_X f \mathrm{d}\nu \geq \int_X f g \mathrm{d}\nu \geq 0 \ . \tag{2.4}$$

Hence, for any $E \in \mathcal{M}$, $\nu(E) \geq \int_E g \mathrm{d}\nu \geq 0$, and this means that

$$0 \leq g(x) \leq 1$$

almost everywhere with respect to ν.

Now let $A = \{ \ x \ : \ g(x) > 0 \ \}$, or, what is the same, $A^c = \{x \ : \ g(x) = 0\}$. Taking $f = 1_{A^c}$ in (2.3), we see that

$$\mu_2(A^c) = 0 \ .$$

Therefore, if we define a measure $\mu_1^{(\mathrm{s})}$ by

$$\mu_1^{(\mathrm{s})}(E) = \mu_1(A^c \cap E) \quad \text{for all} \quad E \in \mathcal{M} \ , \tag{2.5}$$

$$\mu_1^{(s)}(A) = 0 \ .$$

Since $\mu_1^{(s)}(A) = 0$ and $\mu_2(A^c) = 0$, $\mu_1^{(s)}$ and μ_2 are mutually singular.

Next define $\mu_1^{(ac)}$ by

$$\mu_1^{(ac)} = \mu_1 - \mu_1^{(s)} \ ,$$

or, what is the same,

$$\mu_1^{(ac)}(E) = \mu_1(E \cap A)$$

for all $E \in \mathcal{M}$. It remains to find h, which we shall show is given by $h = (1-g)/g$ on A. To see this, use $\nu = \mu_1 + \mu_2$ to rewrite (2.3) as

$$\int_X f(1-g)\mathrm{d}\mu_2 = \int_X fg\mathrm{d}\mu_1 \tag{2.6}$$

for all $f \in \mathcal{H}$.

Now let E be any measurable subset of A, and for each positive integer N define

$$f_N = 1_E \min\{g^{-1}, N\} \ .$$

Since $g > 0$ on E, g^{-1} is defined and finite and

$$1_E g^{-1} = \lim_{N\to\infty} f_N \tag{2.7}$$

almost everywhere. Moreover, since f_N is bounded, it belongs to $\mathcal{H}$. Hence from (2.6),

$$\int_X f_N(1-g)\mathrm{d}\mu_2 = \int_X f_N g\mathrm{d}\mu_1 \ .$$

By (2.7) and the Lebesgue Monotone Convergence Theorem,

$$\begin{aligned}\int_E \frac{1-g}{g}\mathrm{d}\mu_2 &= \lim_{N\to\infty}\int_X f_N(1-g)\mathrm{d}\mu_2 \\ &= \lim_{N\to\infty}\int_X f_N g\mathrm{d}\mu_1 \\ &= \mu_1(E) \ .\end{aligned}$$

Taking $E = A$,

$$\int_A \frac{1-g}{g}\mathrm{d}\mu_2 = \mu_1(A) \le \mu_1(X) < \infty \ .$$

Hence the non–negative measurable function h defined by

$$h(x) = \begin{cases} 0 & \text{if } x \in A^c \\ (1-g(x))/g(x) & \text{if } x \in A \end{cases}$$

is integrable with respect to μ_2 and for all measurable sets E,

$$\mu_1^{(ac)}(E) = \mu_1(E \cap A) = \int_E h\mathrm{d}\mu_2 \ . \tag{2.8}$$

It follows immediately that if $\mu_2(E) = 0$, then $\mu_1^{(ac)}(E) = 0$, so that $\mu_1^{(ac)}$ is indeed absolutely continuous with respect to μ_2.

This proves the existence of the Lebesgue decomposition. The uniqueness is provided by Lemma 1.5. Finally, since for $h, \tilde{h} \in Ł^1(X, \mathcal{M}, \mu_2)$,

$$\int_E h \mathrm{d}\mu_2 = \int_E \tilde{h} \mathrm{d}\mu_2$$

for all $E \in \mathcal{M}$ if and only if $h = \tilde{h}$ a.e. with respect to μ_2. Thus, the function h in the Radon-Nikodym Theorem is unique. □

2.3 Remark. We have stated and proved the Lebesgue Dcomposition Theorem and the Radon-Nikodym Theorem for finite measures. However, its application in the σ-finite case is immediate from this since then there exists a countable partion of X into measurable sets of finie measure for both μ_1 and μ_2. (Consider countable partitions for each measure separately, and then take intersections.) The theorems may be applied on each set in the partition separately. The resulting Radon-Nikodym derivative is then integrable on each of the sets in the partition, but may not be integrable on the whole space. All the same, it is uniquely defined and is integrable on any measurable set on which both μ_1 and μ_2 are finite.

3 Transformations of Lebesgue measure under homeomorphisms with a Lipschitz inverse

Let $K \subset \mathbb{R}^n$ be compact, and let μ denote the restriction of Lebesgue measure m to K. That is, for all Borel sets E, $\mu(E) = m(E \cap K)$.

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a homeomorphism such that T^{-1} is a Lipschitz transformation on the compact set $T(K)$. That is, there exists a finite M such that

$$|T^{-1}(x) - T^{-1}(y)| \leq M|x - y|$$

for all $x, y \in T(K)$. Equivalently

$$|T(x) - T(y)| \geq \frac{1}{M}|x - y|$$

for all $x, y \in K$.

For example if T is defined on an open set U containing K , and is continuously differentiable on U, and the Jacobian determinant $\det(DT(x))$ is non-zero everywhere on K, these conditions conditions are readily verified.

Our main goal in this section is to show that for such a transformation T, $T\#\mu$ is absolutely continuous with respect to Lebesgue measure. In a later section we shall return to the computation of the Radon-Nikodym derivative and show that is equals $|\det(DT(x))|$.

To prove that $T\#\mu << m$, we recall that for all Borel sets E,

$$T\#\mu(E) = m(T^{-1}(E))$$

by the very definition of $T\#\mu$. Thus $T\#\mu << m$ if and only if for all Borel sets E with $m(E) = 0$, it is the case that $m(T^{-1}(E)) = 0$. Since our hypothesis is that T^{-1} is Lipschitz on $T(K)$, it suffices to prove the following, in which we reverse the roles of T and its inverse to keep the notation simple.

3.1 THEOREM. *Let $K \subset \mathbb{R}^n$ be compact and suppose that T is a Lipschitz function on K. Let μ^* denote Lebesgue outer measure on $\mathbb{R}^n$. If $E \subset K$ is such that $\mu^*(E) = 0$, then $\mu^*(T(E)) = 0$.*

Proof. Suppose that $\mu^*(E) = 0$. Then for every $\epsilon > 0$, there exist a countable covering of E by dyadic rational half open rectangles R_j such that

$$\sum_{j=1}^{\infty} m(R_j) \leq \epsilon \ . \tag{3.1}$$

Let $R = (a_1, b_1] \times \cdots \times (a_n, b_n]$ be any finite volume dyadic rational half open rectangle. Letting 2^m be the largest denominator among the numbers $\{a_1, \dots a_n, b_1, \dots . b_n\}$, Then R is a finite union of half open cubes of side length 2^{-m}. Thus, without loss of generality, we may suppose that each rectangle R_j in (3.1) is a cube.

Let L be the Lipschitz constant of T. If ℓ is the side length of a cube R, then the diameter of the set $T(R)$ is at most $L\sqrt{n}\ell$. Let B denote the unit ball in $\mathbb{R}^n$. Any set of diameter $\sqrt{n}\ell L$ is contained in a ball of radius $L\sqrt{n}\ell/2$, and hence has a Lebesgue out measure of at most

$$m_n(B) \left[\frac{L\sqrt{n}\ell}{2}\right]^n = m_n(B) \left[\frac{L\sqrt{n}}{2}\right]^n m_n(R)$$

since $m_n(R) = \ell^n$. Since $\{T(R_j\}_{j\in\mathbb{N}}$ is a cover of $T(E)$, it follows that

$$\mu^*(T(E)) \leq \sum_{j=1}^{\infty} m_n(B) \left[\frac{L\sqrt{n}}{2}\right]^n \sum_{j=1}^{\infty} m_n(R_j) \leq m_n(B) \left[\frac{L\sqrt{n}}{2}\right]^n \epsilon \ .$$

Since $\epsilon > 0$ is arbitrary, $\mu^*(T(E)) = 0$. □

4 The Lebesgue Differentiation Theorem

4.1 DEFINITION (Hardy-Littlewood maximal function)**.** A function f on $\mathbb{R}^n$ is *locally integrable* with respect to Lebesgue measure m on $\mathbb{R}^n$ in case $\int_K |f| \mathrm{d}m < \infty$ for all compact $K \subset \mathbb{R}^n$. For f locally integrable, the *Hardy-Littlewood maximal function Mf* is defined by

$$Mf(x) = \sup_{r>0} \frac{1}{m(B_r(x))} \int_{B_r(x)} f(y) \mathrm{d}m(y) \ . \tag{4.1}$$

Fix $r, r' > 0$ and $x, x' \in \mathbb{R}^n$. As $x' \to x$ in $\mathbb{R}^n$ and $r' \to r$ in $(0, \infty)$, $m(B_{r'}(x') \Delta B_r(x)) \to 0$. Since f is integrable on $B_{2r}(x)$, for all $\epsilon > 0$, there is a $\delta > 0$ so that whenever $E \subset B_{2r}(x)$ and $m(E) < \delta$, $\int_E |f| \mathrm{d}m < \epsilon$. It follows that there is an $\eta > 0$ so that whenever $|r - r'| + |x - x'| < \eta$,

$$\left| \int_{B_{r'}(x')} f \mathrm{d}m - \int_{B_r(x)} f \mathrm{d}m \right| \leq \int_{B_{r'}(x') \Delta B_r(x)} |f| \mathrm{d}m < \epsilon \ .$$

It follows that for each $r > 0$, the function $f_r(x)$ given by

$$f_r(x) = \frac{1}{m(B_r(x))} \int_{B_r(x)} f(y) \mathrm{d}m(y) \tag{4.2}$$

is jointly continuous in r and x.

As a result of the continuity in r, the supremum in (4.1) is the same as the supremum over *rational* $r > 0$, and then, since the supremum of a countable set of measurable functions is measurable, $Mf \in L^+(\mathbb{R}^n, \mathcal{L}, m)$.

4.2 LEMMA. *For $f \in L^1(\mathbb{R}^n, \mathcal{L}, m)$ and $r > 0$, let the function f_r be defined by (4.2). Then for all $r > 0$, $f_r \in L^1(\mathbb{R}^n, \mathcal{L}, m)$ and*

$$\lim_{r\to 0} \|f_r - f\|_1 = 0 \ . \tag{4.3}$$

Proof. Since continuous and compactly supported functions are dense in $L^1(\mathbb{R}^n, \mathcal{L}, m)$, for all $\epsilon > 0$, there exists a continuous, compactly supported function g such that $\|g - f\|_1 < \epsilon$. Since continuous, compactly supported functions are uniformly continuous, for all $\eta > 0$, there exists a $\delta > 0$ so that whenever $|y - x| < \delta$, $|g(y) - g(x)| < \eta$. But then with g_r defined in terms of g by (4.2) with g in place of f, $|g_r(x) - g(x)| < \eta$ for all x and all $r < \delta$. In other words, $\lim_{r\to 0} g_r = g$, uniformly in x. Now let $R > 0$ be such that the support of g is contained in $B_R(0)$. Then for all $r < 1$, the support of g_r is contained in $B_{R+1}(0)$. Since uniform convergence on compact sets implies L^1 convergence, $\lim_{r\to 0} \|g_r - g\|_1 = 0$. We now return to f itself.

The function $1_{B_r(x)}(y)$, as a function of x and y, is measurable on $\mathbb{R}^n \times \mathbb{R}^n$, and hence $1_{B_r(x)}(y)|f(y)|$ is measurable on $\mathbb{R}^n \times \mathbb{R}^n$. By Tonelli's Theorem,

$$\int_{R^n} |f_r| \mathrm{d}m \le \frac{1}{m(B_r(0))} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} 1_{B_r(x)}(y)|f(y)|\mathrm{d}m(x) \right) \mathrm{d}m(y) = \int_{\mathbb{R}^n} |f(y)|\mathrm{d}m(y) = \|f\|_1 \ .$$

Applying the same argument to $f - g$, with g as above, and noting that $(f - g)_r = f_r - g_r$, we see that

$$\|f_r - g_r\|_1 \le \|f - g\|_1 \le \epsilon \ . \tag{4.4}$$

Now we use Minkowski's inequality: $\|f_r - f\|_1 \le \|f_r - g_r\|_1 + \|g_r - g\|_1 + \|g - f\|_1$. Combining this with (4.4), we have

$$\|f_r - f\|_1 \le \|g_r - g\|_1 + 2\epsilon \ .$$

This together with the first part of the proof shows that $\limsup_{r\to 0} \|f_r - f\|_1 \le 2\epsilon$, and then since $\epsilon > 0$ is arbitrary, proves the lemma. □

Since a subsequence of every L^1 convergent sequence converges almost everywhere, there exists a sequence $\{r_k\}_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} r_k = 0$ such that

$$\lim_{k\to\infty} f_{r_k}(x) = f(x) \quad \text{for a.e. } x \ . \tag{4.5}$$

The Lebesgue Differentiation Theorem says much more: It says that

$$\lim_{r\to 0} f_r(x) = f(x) \quad \text{for a.e. } x \ . \tag{4.6}$$

In fact, from this central result, we shall extract several important corollaries.

4.3 THEOREM (The Lebesgue Differentiation Theorem)**.** *Let $f \in L^1(\mathbb{R}^n, \mathcal{L}, m)$. Then there is a set $E \in \mathcal{L}$ with $m(E) = 0$ such that for all $x \in E^c$,*

$$f(x) = \lim_{r\to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} f(y)\mathrm{d}m(y) \ .$$

We shall prove this theorem using two results with that each have many other applications: The Hardy-Littlewood maximal function bound and the Finite Vitali Covering Lemma. This line of argument is different from that of Lebesgue, and is due to Wiener.

4.4 LEMMA (Finite Vitali Covering Lemma)**.** *Let (X,d) be a metric space. Let $\{B_1,\dots,B_N\}$ be a finite collection of open balls in X. Let $U = \bigcup_{j=1}^N B_j$. Let $B_j = B_{r_j}(x_j)$. Define $\widehat{B}_j = B_{3r_j}(x_j)$. That is, $\widehat{B}_j$ is the open ball in X with the same center as B_j, but 3 times the radius.*

There exits a subset $J \subset \{1,\dots,N\}$ such that:

(1) The balls B_j with $j \in J$ are mutually disjoint.

(2) $U \subset \bigcup_{j\in J} \widehat{B}_j$.

Proof. Let r_j denote the radius of B_j. We define J inductively. Pick j_1 so that $r_{j_1} \geq r_k$ for all $k \in \{1,\dots,N\}$. Suppose that the set $\{j_1,\dots,j_m\}$ has been selected. Let L_m be given by

$$L_m = \{\ell \ : \ B_\ell \bigcap \left(B_{j_1} \bigcup \cdots \bigcup B_{j_m}\right) = \emptyset\} \ .$$

If $L_m = \emptyset$, define $J = \{j_1,\dots,j_m\}$. Otherwise, choose j_{m+1} such that $r_{m+1} \geq r_k$ for all $k \in L_m$. Since $\{B_1,\dots,B_N\}$ is finite, the selection process must terminate at some finite point, and then each ball in $\{B_1,\dots,B_N\}$ has non-empty intersection with some B_j such that $j \in J$.

Now suppose $k \notin J$. Let m be the least value of m such that

$$B_k \bigcap \left(B_{j_1} \bigcup \cdots \bigcup B_{j_m}\right) \neq \emptyset \ .$$

Then $r_k \leq r_m$. Thus, $B_k \cap B_j \neq \emptyset$ for some $j \in J$, and $r_j \geq r_k$. By the triangle inequality, $B_k \subset \widehat{B}_j$. Since $k \notin J$ is arbitrary, this shows that for all $k \notin J$, $B_k \subset \bigcup_{j\in J} \widehat{B}_j$. It is evident that for $k \in J$, $B_k \subset \bigcup_{j\in J} \widehat{B}_j$. This proves that $U \subset \bigcup_{j\in J} \widehat{B}_j$. □

4.5 Remark. There is wide variety of covering lemmas. The full Vitalli Covering Lemma applies to uncountable covers by balls or cubes in $\mathbb{R}^n$. A particularly deep and powerful covering lemma is that of Besicovich. For our present purpose, the simple Lemma 4.4 suffices.

4.6 THEOREM (Hardy-Littlewood Maximal Theorem)**.** *Let $f \in L^1(\mathbb{R}^n, \mathcal{L}, m_n)$, and let Mf denote the Hardy-Littlewood maximal function of f. Then for all $\alpha > 0$,*

$$m_n(\{\ x \ : \ Mf(x) > \alpha\}) \leq 3^n \frac{\|f\|_1}{\alpha} \ .$$

Proof. Let $K \subset \{\ x \ : \ Mf(x) > \alpha\}$ be compact. for each $x \in K$, choose r_x such that

$$\frac{1}{m(B_{r_x}(x))} \int_{B_{r_x}(x)} f(y)\mathrm{d}m(y) > \alpha \ . \tag{4.7}$$

Then $\{B_{r_x}(x) \ : \ x \in K\}$ is an open cover of K. Since K is compact, some finite subset

$$\{B_1,\dots,B_N\} \subset \{B_{r_x}(x) \ : \ x \in K\}$$

covers K. By the Finite Vitali Covering Lemma, there exists a set $J \subset \{1,\dots,N\}$ with the properties *(1)* and *(2)* described in the lemma. Then, using the notation of the lemma,

$$m_n(K) \leq m_n\left(\bigcup_{j\in J} \widehat{B}_j\right) \leq \sum_{j\in J} m_n(\widehat{B}_j) = 3^n \sum_{j\in J} m_n(B_j) \ .$$

However, by (4.7),

$$m_n(B_j) \leq \frac{1}{\alpha}\int_{B_j} |f|\mathrm{d}m_n \ .$$

Thus

$$m_n(K) \leq \frac{3^n}{\alpha}\sum_{j\in J}\int_{B_j} |f|\mathrm{d}m_n \ .$$

Finally since the B_j with $j \in J$ are mutually disjoint,

$$\sum_{j\in J}\int_{B_j} |f|\mathrm{d}m_n = \int_{\cup_{j\in J}B_j} |f|\mathrm{d}m_n \leq \int_{\mathbb{R}^n} |f|\mathrm{d}m_n \ .$$

Altogether $m_n(K) \leq \dfrac{3^n}{\alpha}\displaystyle\int_{\mathbb{R}^n} |f|\mathrm{d}m_n$.

Finally, by the inner regularity of Lebesgue measure,

$$m_n\left(\{\ x\ :\ Mf(x) > \alpha\}\right) = \max\{m(K)\ :\ K \text{ compact and } K \subset \{\ x\ :\ Mf(x) > \alpha\}\}\ .$$

□

To prepare for the proof of the Lebesgue Differentiation Theorem, we make the following definitions: For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathcal{L}, m_n)$, we define

$$\Omega_+ f(x) = \limsup_{r\to 0} f_f r\ , \quad \Omega_- f(x) = \liminf_{r\to 0} f_r\ , \quad \text{and}\Omega f(x) = \Omega_+ f(x) - \Omega_+ f(x)$$

where f_r is defined by (4.2). $\Omega f(x)$ is called the *oscillation of f at x*.

It is clear that if $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathcal{L}, m_n)$, then $\Omega_+(f+g) \leq \Omega_+ f + \Omega_+ g$ and $\Omega_-(f+g) \geq \Omega_- f + \Omega_- g$. Consequently,

$$\Omega(f+g) \leq \Omega f + \Omega g\ . \tag{4.8}$$

Proof of Theorem 4.3. Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathcal{L}, m_n)$. Pick $\epsilon > 0$. Since the restriction of f to any compact set is integrable, and since $f_r(x)$ only depends on the values of f on $B_r(x)$, we may assume without loss of generality that $f \in L^1$. Pick a continuous compactly supported function g such that $\|f - g\|_1 < \epsilon$, and let $h = f - g$ so that $f = h + g$. Since g is continuous, $\Omega g(x) = 0$ for all x. Then by (4.8),

$$\Omega f(x) = \Omega h(x) \tag{4.9}$$

for all x. By the definitions of the oscillation and the maximal functions, $\Omega h(x) \leq 2Mh(x)$ for all x. Then by Theorem 4.6, for all $\alpha > 0$,

$$m_n\left(\{x\ :\ \Omega h(x) > \alpha\ \}\right) \leq 2\frac{3^n}{\alpha}\|h\|_1\ .$$

Combining this with (4.9) and recalling that $\|h\|_1 < \epsilon$, we have

$$m_n\left(\{x\ :\ \Omega f(x) > \alpha\ \}\right) \leq 2\frac{3^n}{\alpha}\epsilon\ .$$

Since $\epsilon > 0$ is arbitrary, this means that $m_n\left(\{x\ :\ \Omega f(x) > \alpha\ \}\right) = 0$ for all $\alpha > 0$, and thus that $\Omega f = 0$ a.e. Since $\lim_{r\to 0} f_r(x)$ exists for all x such that $\Omega f(x) = 0$, it remains only to identify the limit as f. This, however, is a direct consequence of (4.5). □

Let $f \in Ł^1_{\rm loc}(\mathbb{R}^n, \mathcal{L}, m_n)$, and let $c \in \mathbb{C}$. Applying Theorem 4.3 to $g = |f - c|$, we see that

$$\lim_{r\to 0} \frac{1}{m(B_r(x))}\int_{B_r(x)} |f(y) - c|{\rm d}m(y) = |f(x) - c| \tag{4.10}$$

on the complement of a set E_c with $m_n(E_c) = 0$. Now let $\{c_k\}_{k\in\mathbb{N}}$ be a dense sequence in $\mathbb{C}$. For each k, let E_{c_k} be the exceptional set on which (4.10) do not hold with $c = c_k$. Let $E = \bigcup_{k\in\mathbb{N}} E_k$, and note that $m_n(E) = 0$.

Fix any $x \in E^c$, and any $\epsilon > 0$. Choose k such that $|c_k - f(x)| < \epsilon$. Then for all y, $|f(y) - f(x)| < |f(y) - c_k| + \epsilon$. and hence

$$\limsup_{r\to 0} \frac{1}{m(B_r(x))}\int_{B_r(x)} |f(y) - f(x)|{\rm d}m(y) \le \limsup_{r\to 0} \frac{1}{m(B_r(x))}\int_{B_r(x)} |f(y) - c_k|{\rm d}m(y) + \epsilon < 2\epsilon\ .$$

Thus, for all $x \in E^c$,

$$\lim_{r\to 0} \frac{1}{m(B_r(x))}\int_{B_r(x)} |f(y) - f(x)|{\rm d}m(y) = 0\ . \tag{4.11}$$

4.7 DEFINITION. Let $f \in L^1_{\rm loc}(\mathbb{R}^n, \mathcal{L}, m)$. The Lebesgue set of f is the set of all x such that (4.11) is valid.

The complement of the Lebesgue set is contained a set of measure zero, as we have seen, and hence belongs to $\mathcal{L}$. It follows that the Lebesgue set belongs to $\mathcal{L}$.

There is one further refinement that will be useful to us: We not limited to averaging over balls centered on x.

4.8 DEFINITION (Nicely shrinking to a point)**.** A set $\{E_r\}_{r>0}$ of Borel sets in R^n *shrinks nicely to* $x \in \mathbb{R}^n$ in case:

(1) $E_r \subset B_r(x)$ for each $r > 0$.

(2) There is a constant $\alpha > 0$ such that $m(E_r) \ge \alpha m(B_r(x))$ for each $r > 0$.

Note that is is not required that $x \in E_r$ for any r.

The point of the definition is that for all $r > 0$,

$$\frac{1}{m(E_r)}\int_{E_r} |f(y)-f(x)|{\rm d}m(y) \le \frac{1}{m(E_r)}\int_{B_r(x)} |f(y)-f(x)|{\rm d}m(y) \le \frac{1}{\alpha}\frac{1}{m(B_r(x))}\int_{B_r(x)} |f(y)-f(x)|{\rm d}m(y)\ .$$

Consequently we have the final form of the Lebesgue Differentiation Theorem:

4.9 THEOREM. *Let $f\ inL^1_{\rm loc}(\mathbb{R}^n, \mathcal{L}, m)$. Then for each x in the Lebesgue set of f, and each set $\{E_r\}_{r>0}$ of sets that shrink nicely to x,*

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

5 Lebesgue decomposition for locally finite Borel measures on $\mathbb{R}^n$

Let ν be a locally finite Borel measure on $\mathbb{R}^n$. Then ν is inner and outer regular, and $\nu(K) < \infty$ for all compact K. Let $\nu = \mu + f\mathrm{d}m$ be the Lebesgue decomposition of ν with respect to Lebesgue measure m, so that f is the Radon-Nikodym derivative of the absolutely continuous part. Since ν is locally finite, so are μ and $f\mathrm{d}m$. Hence they too are inner and outer regular and finite on compact sets.

The next theorem gives a more explicit form of the Lebesgue decomposition of ν with respect to Lebesgue measure.

5.1 THEOREM (Lebesgue decomposition for locally finite Borel measures on $\mathbb{R}^n$)**.** *Let ν be a locally finite Borel measure on $\mathbb{R}^n$. Then for m-almost every $x \in \mathbb{R}^n$, and each set $\{E_r\}_{r>0}$ of sets that shrink nicely to x,*

$$\lim_{r\to 0} \frac{\nu(E_r)}{m(E_r)} = f(x) \tag{5.1}$$

where $f\mathrm{d}m$ is the absolutely continuous part of ν.

Proof. $\nu = \mu + f\mathrm{d}m$ be the Lebesgue decomposition of ν with respect to Lebesgue measure m. As noted above, μ is outer regular. Since μ and m are mutually singular, there is a Borel set A such that $\mu(A) = 0$ and $m(A^c) = 0$. Fix $k \in \mathbb{N}$, and define the set F by

$$F = \left\{ \ x \in A \ : \ \limsup_{r\to 0} \frac{\mu(B_r(x))}{m(B_r(x))} > \frac{1}{k} \ \right\} \ .$$

Fix $\epsilon > 0$. Let K be compact and U open with $K \subset F \subset U$ be such that $\mu(U) \le \mu(F) + \epsilon$ and $m(K) > m(F) - \epsilon$. Note that we are considering the Lebesgue measure of K, and hence using the inner regularity of Lebesgue measure, and the μ-measure of U, and hence using the outer regularity of μ. Since by definition $F \subset A$ and $\mu(A) = 0$, $\mu(U) < \epsilon$.

Be the definition of F. for each x in F, and hence each $x \in K$, there is an $r > 0$ such that $R_r(x) \subset U$ and

$$\frac{\mu(B_r(x))}{m(B_r(x))} > \frac{1}{k}$$

Since K is compact, finitely many of the balls cover K. By the Finite Vitali Covering Lemma, there is a finite set $\{x_1, \dots, x_N\}$ of points in K and a corresponding set $\{r_1, \dots, r_N\}$ of positive numbers such that the balls $\{B_{r_1}(x_1), \dots, B_{r_N}(x_N)\}$ are mutually disjoint, but such that

$$K \subset \bigcup_{j=1}^N B_{3r_j}(x_j) \ .$$

It then follows that

$$m(K) \le \sum_{j=1}^N m\left(B_{3r_j}(x_j)\right) = 3^n \sum_{j=1}^N m\left(B_{r_j}(x_j)\right) \le \frac{3^n}{k} \sum_{j=1}^N \mu\left(B_{r_j}(x_j)\right) \le \frac{3^n}{k}\mu(U)$$

Then since $m(F) \le m(K) + \epsilon$ and $\mu(U) < \epsilon$,

$$m(F) \le \epsilon\left(1 + \frac{3^n}{k}\right) \ .$$

Since $\epsilon > 0$ is arbitrary, $m(F) = 0$ for all k. It now follows that on the complement of a set B of Lebesgue measure zero,

$$\limsup_{r\to 0} \frac{\mu(B_r(x))}{m(B_r(x))} = 0 \ .$$

Since whenever $\{E_r\}_{r>0}$ shrinks nicely to x, there is an $\alpha > 0$ so that

$$\frac{\mu(E_r)}{m(E_r)} \leq \frac{1}{\alpha}\frac{\mu(B_r(x))}{m(B_r(x))} \ ,$$

we have that on the same set B^c,

$$\limsup_{r\to 0} \frac{\mu(E_r)}{m(E_r)} = 0 \ .$$

Therefore, for $x \in B^c$ and the Lebesgue set of x,

$$\lim_{r\to 0} \frac{\nu(E_r)}{m(E_r)} = \lim_{r\to 0} \frac{f\mathrm{d}m(E_r)}{m(E_r)} = \lim_{r\to 0} \frac{1}{m(E_r)}\int_{E_r} f\mathrm{d}m \ ,$$

where the last equality is valid by the Lebesgue Differentiation Theorem. □

Our first application is a general change-of-variables formula. We shall make use of Rademacher's Theorem which asserts that Lipschitz functions are differentiable almost everywhere with respect to Lebesgue measure. Let T be a homeomorphism from $\mathbb{R}^n$ onto $\mathbb{R}^n$ that whose inverse S is Lipschitz on every compact $K \subset \mathbb{R}^n$. Let $\nu = T\#m$. We have proved that ν is absolutely continuous with respect to Lebesgue measure. We now compute the Radon-Nikodym derivative in terms of the derivative of S, DS. Recall that to says that S is differentiable at x means that there is an $n \times n$ matrix A such that for all $\epsilon > 0$, there is a $\delta_\epsilon > 0$ such that

$$|y - x| < \delta_\epsilon \qquad \Rightarrow \qquad |S(y) - [S(x) + A(y - x)| < \epsilon|y - x| \ .$$

The matrix A, or rather the linear transformation that it represents, is called the derivative of S at x, and is denoted $DS(x)$.

5.2 THEOREM. *Let T be a homeomorphism from $\mathbb{R}^n$ onto $\mathbb{R}^n$ that whose inverse S is Lipschitz on every compact $K \subset \mathbb{R}^n$. Let $\nu = T\#m$. Then*

$$\nu = |\det(DS(x))|m \ . \tag{5.2}$$

In particular, for all non-negative Borel function g on $\mathbb{R}^n$,

$$\int_{\mathbb{R}^n} g\mathrm{d}m = \int_{\mathbb{R}^n} (g \circ S)|\det(DS)|\mathrm{d}m \ .. \tag{5.3}$$

Proof. We have proved that ν is absolutely continuous with respect to Lebesgue measure. Let f denote the Radon-Nikodym derivative. By the previous theorem, and the definition of ν, at every point of the Lebesgue set of f,

$$f(x) = \lim_{r\to 0} \frac{\nu(B_r(x))}{m(B_r(x))} = \lim_{r\to 0} \frac{m(S((B_r(x)))}{m(B_r(x))} \ .$$

By Rademacher's Theorem. S is differentiable almost everywhere. Suppose that x is a point in the Lebesgue set of f at which S is differentiable. Then

Fix $\epsilon > 0$ and $r < \delta_\epsilon/2$. Define

$$F_r = \{S(y) \ : \ |y-x| < r\ \} \qquad \text{and} \qquad G_r = \{[S(x) + A(y-x) \ : |y-x| < r\ \}\ .$$

Since $F_r = S(B_r(x))$, $m(F_r) = \nu(B_r(x))$.

Suppose that $\det(A) = 0$. Then the image of $\mathbb{R}^n$ under A lies in a subspace of dimension at most $n-1$. In particular $G(r)$ is contained in a ball of radius $\|A\|r$ in an affine subspace of $\mathbb{R}^n$ of dimension at most $n-1$. Then, since every point in F_r is within a distance ϵr of a point of G_r, F_r is contained in a set of Lebesgue measure at most

$$\omega_{n-1}(1+\epsilon)^{n-1}\epsilon r^n$$

where ω_{n-1} denotes the $n-1$ dimensional Lebesgue measure of the unit ball in $\mathbb{R}^{n-1}$. In particular,

$$\frac{m(S((B_r(x)))}{m(B_r(x))} \leq \frac{\omega_{n-1}}{\omega_n}(1+\epsilon)^{n-1}\epsilon\ .$$

Since $\epsilon > 0$ is arbitrary, this shows

$$\lim_{r\to 0}\frac{\nu(B_r(x))}{m(B_r(x))} = 0 = |\det(A)|\ .$$

Next, suppose that $\det(A) \neq 0$. Define $\widetilde{S}(x) = A^{-1}S(x)$. Then the derivative of $\widetilde{S}$ at x is I, the $n\times n$ identity matrix, so that for all $\epsilon > 0$, there is a $\delta_\epsilon > 0$ such that

$$|y-x| < \delta_\epsilon \qquad \Rightarrow \qquad |\widetilde{S}(y) - [\widetilde{S}(x) + (y-x)| < \epsilon|y-x|\ .$$

Fix $1 > \epsilon > 0$ and $r < \delta_\epsilon/2$. Define

$$\widetilde{F}_r = \{\widetilde{S}(y) \ : \ |y-x| < r\ \} \qquad \text{and} \qquad \widetilde{G}_r = \{[\widetilde{S}(x) + (y-x) \ : |y-x| < r\ \}\ .$$

Then $G_r = B_r(\widetilde{S}(x))$, and $B_{(1-\epsilon)r}(\widetilde{S}(x)) \subset \widetilde{F}_r \subset B_{(1+\epsilon)r}(\widetilde{S}(x))$. Consequently

$$(1-\epsilon)^n < \frac{m(\widetilde{F}_r)}{m(B_r(x))} < (1+\epsilon)^n \tag{5.4}$$

By what we know about the transformation of Lebesgue measure under invertible linear transformations,

$$m(\widetilde{F}_r) = m(A^{-1}S(B_r(x)) = |\det(A^{-1})|m(S(B_r(x))) = |\det(A^{-1})|\nu(B_r(x))\ .$$

Combining this with (5.4), in which ϵ may be arbitrarily small, it follows that

$$\lim_{r\to 0}\frac{\nu(B_r(x))}{m(B_r(x))} = |\det(A)|\ .$$

Since $A = DS(x)$, the proof of (5.2) is complete.

Next, by the general change of variables formula, for any non-negaitve Borel functions g, since $\nu = T\sharp m$,

$$\int_{\mathbb{R}^n} g\circ T\mathrm{d}m = \int_{\mathbb{R}^n} g\mathrm{d}\nu\ .$$

Now replace g by $g\circ S$ and apply (5.2). □

Our next application of Theorem 5.1 concerns monotone functions of $\mathbb{R}$. Let F be a monotone non-decreasing function from $\mathbb{R}$ to $\mathbb{R}$. For $t \in \mathbb{R}$, define

$$F(t+) = \lim_{h \downarrow 0} F(t+h) \ ,$$

and define $G(t) = F(t+)$. Note that $F(t+) - F(t) \geq 0$, and for all $N \in \mathbb{N}$

$$F(N) - F(-N) \geq \sum_{t \in [-N,N)} (F(t+) - F(t)) \ . \tag{5.5}$$

Dince the left side is finite, $F(t+) - F(t) = 0$ for all except at most countably many values of t. The function $G(t)$ is monotone increasing and right continuous. Let μ_G be the corresponding Lebesgue-Stieltjes measure. Then

$$G(t+h) - G(t) = \begin{cases} \mu_G((t, t+h]) & h > 0 \\ \mu_G((t-h, t]) & h < 0 \end{cases} .$$

Let $\{E_h\}_{h>0}$ be given by $E_h = (t, t+h]$. Then $\{E_h\}_{h>0}$ shrinks nicely to t (with $\alpha = 1/2$). Since μ_G is a locally finite Borel measure, Thoerem 5.1 ensures that

$$\lim_{h \downarrow 0} \frac{G(t+h) - G(t)}{h} = \lim_{h \downarrow 0} \frac{\mu_G(E_h)}{m(E_h)} \tag{5.6}$$

exists almost everyhere. Replacing E_h by $(t-h, t]$, and repeating the argument, we see that G is differentiable almost everywhere.

Next, let $H = G - F$. This function is non-zero only at the discontinuity points of F, but these may be dense. We shall now show, again using Theorem 5.1, that for almost every t, H is differentiable at t, and $H'(t) = 0$.

Let $\{t_j\}_{j \in \mathbb{N}}$ be an enumeration of the points at which $H \neq 0$. Defined a Borel measure ν by

$$\nu = \sum_{j=1}^{\infty} H(t_j) \delta_{t_j}$$

where δ_{t_j} is the Dirac mass at t_j. By (5.5), ν is locally finite. We now estimate

$$|H(t+h) - H(t)| \leq H(t+h) + H(t) \leq \nu([t - |h|, t + |h|])$$

since no matter the sign of h, t and $t+h$ are both included in $[t - |h|, t + |h|]$. Define $E_r = [t - r/2, t + r/2]$. Then the sets $\{E_r\}_{r>0}$ shrink nicely to t, and so

$$\lim_{|h| \to 0} \frac{\nu(E_{2|h|})}{m(E_{2|h|})} \geq \frac{1}{2} \limsup_{|h| \to 0} \frac{|H(t+h) - H(t)|}{|h|} = 0$$

almost everywhere since ν is singular with respect to Lebesgue measure.

Thus excluding sets of measure zero where G is not diferentiable and where H is not differentiable with derivative $H' = 0$, we have a set of full measure on which F is differentiable, and $F' = G' + H' = G'$. We have proved:

5.3 THEOREM. *Let F be a monotone non-decreasing function from $\mathbb{R}$ to $\mathbb{R}$. Let $G(t) = F(t+)$. Then F and G are differentiable almost everywhere, and $G' = F'$ almost everywhere.*

Now suppose that G is a monotone nondecreasing and right continuous function from $\mathbb{R}$ to $\mathbb{R}$. Let μ_G be the associated Lebesgue-Stieljes measure. Let

$$\mu_G = \lambda + gm$$

be its Lebesgue decomposition with respect to Lebesgue measure m where g is the Radon-Nikodym derivative and λ is the singular component. By Theorem 5.1 and (5.6) once more, for almost every t,

$$g(t) = G'(t) \ .$$

Thefore, for any $a < b$,

$$G(b) - G(a) = \mu_G((a,b]) = \lambda((a,b]) + \int_{[a,b]} G'(t)\mathrm{d}m(t) \ .$$

We see that the Funadamental Theorem of Calculus is valid for G on all intervals $[a,b]$ if and only if μ_G is absolutely continuous with respect to Lebesgue measure.

The absolute continuity of G can be characterized directly in terms of G. Suppose that μ_G is absolutely continuous with respect to m. Let g be the Radon-Nikodym derivative. Then since $\{g\}$ is uniformly integrable, for all $\epsilon > 0$, there is a $\delta > 0$ so that when E is a Borel set with $m(E) < \delta$, then $\int_E g\mathrm{d}m < \epsilon$. This means that whenever $m(E) < \delta$, $\mu_G(E) < \epsilon$. In particular, if E is any finite union of disjoint intervals $E = \bigcup_{j=1}^N (a_j, b_j]$,

$$\sum_{j=1}^N (b_j - a_j) < \delta \quad \Rightarrow \quad \sum_{j=1}^N (G(b_j) - G(a_j)) < \epsilon \ . \tag{5.7}$$

On the other hand, suppose that for all $\epsilon > 0$, there exists a $\delta > 0$ such that (5.7) is valid for any finite union of disjoint intervals $E = \bigcup_{j=1}^N (a_j, b_j]$.

Then for any set F with $m(F) < \delta$, there is a countable cover of F be a sequence of disjoint intervals $\{(a_j, b_j]\}_{j\in\mathbb{N}}$ such that

$$\sum_{j=1}^\infty m((a_j, b_j]) < \delta \ ,$$

since the Lebesgue outer measure of F is the same as $m(F)$. It follows that for all N, $\sum_{j=1}^N (G(b_j) - G(a_j)) < \epsilon$, and then by continuity from below that

$$\mu_G\left(\bigcup_{j=1}^\infty (a_j, b_j]\right) < \epsilon \ .$$

Since $F \subset \bigcup_{j=1}^\infty (a_j, b_j]$, this shows $\mu_G(F) < \epsilon$.

This brings us to the following definition:

5.4 DEFINITION (Absolutely continuous functions)**.** A fuction $F : \mathbb{R} \to \mathbb{C}$ is *absolutely continuous* in case for every $\epsilon > 0$ there is a $\delta > 0$ such that whenever $\{(a_j, b_j)\}_{j=1,\dots,n}$ is any finite set of disjoint open intervals satisfying

$$\sum_{j=1}^{n}(b_j - a_j) < \delta \ . \tag{5.8}$$

it is the case that

$$\sum_{j=1}^{n}|F(b_j) - F(a_j)| < \epsilon \ . \tag{5.9}$$

If for some $-\infty < a < b < \infty$, and the condition is satisfied whenever each of the intervals lie in $[a,b]$, we say that F is absolutely continuous on $[a,b]$.

If F is absolutely continuous, and is a linear combination of real valued, and monotone non-decreasing functions, we then we have that F is differentiable almost everywhere, and for all $a < b$,

$$F(b) - F(a) = \int_{[a,b]} F'(t)\mathrm{d}m(t) \ .$$

We now introduce the class of cuntions for which such a decomposition is valid:

5.5 DEFINITION (Functions of bounded variation on $\mathbb{R}$)**.** Let $F : \mathbb{R} \to \mathbb{C}$. For each $x \in \mathbb{R}$ define

$$T_F(x) = \sup\left\{ \sum_{j=1}^{n}|F(x_j) - F(x_{j-1})| \ : \ n \in \mathbb{N} \ , \ -\infty < x_0 < \cdots < x_n = x \right\} : \ . \tag{5.10}$$

so that T_F is a function with values n $[0,\infty]$. We say that F is a *bounded variation function* in case

$$\sup_{x\in\mathbb{R}} T_F(x) < \infty \ .$$

We denote the set of all bounded variation functions by BV.

Note that for $-\infty < a < b < \infty$,

$$T_F(b) - T_F(a) = \sup\left\{ \sum_{j=1}^{n}|F(x_j) - F(x_{j-1})| \ : \ n \in \mathbb{N} \ , \ -\infty \ , \quad a = x_0 < \cdots < x_n = b \right\} : \ . \tag{5.11}$$

It follows that T_F is monotone non-decreasing. We say that $F \in BV[a,b]$ if $T_F(b) - T_F(a) < \infty$.

5.6 LEMMA (Decomposition Lemma)**.** *If $F \in BV$, then for all $x < y$,*

$$T_F(y) - T_F(x) \geq |F(y) - F(x)| \ . \tag{5.12}$$

If $F \in BV$ is real, $T_F + F$ and $T_F - F$ are monotone non-decreasing.

Proof. Pick $x < y$, $\epsilon > 0$. Choose $-\infty < x_0 < \cdots < x_n = x$ so that

$$\sum_{j=1}^{n}|F(x_j) - F(x_{j-1})| \geq T_F(x) - \epsilon \ . \tag{5.13}$$

Then $-\infty < x_0 < \cdots < x_{n-1} < x < y$ is is one set of endpoints of a partition of $(-\infty, y]$ into intervals, and so

$$T_F(y) \geq \sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| + |F((y) - F(x)| \ . \tag{5.14}$$

Combining (5.13) and (5.14) we have

$$T_F(y) \geq T_F(x) - \epsilon + |F(y) - F(x)| \ .$$

Since $\epsilon > 0$ is arbitrary, (5.12) is proved. When F is real this is equivalent to

$$T_F(y) \pm F(y) \geq T_F(x) \pm F(x) \ .$$

which is what we sought to prove. □

5.7 DEFINITION (The space NBV)**.** The set of normalized BV functions, NBV, is given by

$$NBV = \{ \ F \in BV \ : \ F \text{ is right continuous} \quad \text{and} \quad F(-\infty) = 0 \ \} \ .$$

By what we have shown above, for every BV function F, $F - F(-\infty)$ is almost everywhere equal to an NBV function $\widetilde{F}$, and moreover $F' = \widetilde{F}$ a.e. m.

Next, note that for $F_1, F_2 \in BV$, for all x,

$$T_{F_1+F_2}(x) \leq T_{F_1}(x) + T_{F_2}(x) \ .$$

From this is follows easily that BV is a vector space over $\mathbb{C}$.

We next introduce a natural metric on BV. Define d_{BV} on $NBV \times NBV$ by

$$d_{BV}(F_1, F_2) = T_{F_1 - F_2}(\infty) \ .$$

Since

$$|F_1(x) - F_2(x)| \leq T_{F_1-F_2}(x) \leq T_{F_1-F_2}(\infty) \ , \tag{5.15}$$

$d_{BV}(F_1, F_2) = 0$ if and only if $F_1(x) = F_2(x)$ for all x. Since symmetry and the triangle inequality hold, d_{BV} is a metric.

Next, by (5.15), if $\{F_k\}$ is a Cauchy sequence in NBV equipped with the BV metric, for each x, $\{F_k(x)\}$ is a Cauchy sequence in $\mathbb{C}$. Hence $F(x) = \lim_{k\to\infty} F_k(x)$ exists for each x. It is easy to show that $F \in NBV$ and that $\lim_{k\to\infty} d_{BV}(F_k, F) = 0$. Thus, NBV equipped with the BV metric is a complete metric space. In fact, the metric is given by a *norm*: For $F \in NBV$, we define

$$\|F\|_{BV} = T_F(\infty) \ .$$

and we have that $d_{BV}(F_1, F_2) = \|F_1 - F_2\|_{BV}$ and that for $z \in \mathbb{C}$ and $F \in NBV$, $\|zF\|_{BV} = |z|\|F\|_{BV}$.

5.8 LEMMA. *Fix $a, b \in \mathbb{R}$ with $-\infty < a < b < \infty$. Let $F : [a, b] \to \mathbb{R}$ be absolutely continuous on $[a, b]$. Then:*

(1) F and T_F are uniformly continuos on $[a, b]$.

(2) $F \in BV[a, b]$.

Proof. Fix $\epsilon > 0$ and let $\delta > 0$ be such that whenver $\{(a_j, b_j)\}_{j=1,\ldots,n}$ is a finite set of disjoint open intervals such that (5.8) is satisfied, (5.9) is satisfied. It follows that if $y > x$ and $y - x < \delta$, then for every $x = x_0 < \cdots < x_n = y$, $\sum_{j=1}^{n}(x_j - x_{j-1}) < \delta$, and hence,

$$T_F(y) - T_F(x) \leq \epsilon \ ,$$

which shows that T_F is uniformly continuous. By Lemma 5.6, $|F(y) - F(x)| \leq T_F(y) - T_F(x) \leq \epsilon$, and so F too is uniformly continuous.

Next, fix $\epsilon = 1$ and let $\delta > 0$ be such that

$$\sum_{j=1}^{n}(b_j - a_j) < \delta \Rightarrow \sum_{j=1}^{n}|F(b_j) - F(a_j)| < \epsilon \ .$$

If $\delta \geq b - a$, this together with the definition (5.11) imply that $T_F(b) - T_F(a) \leq 1$. Otherwise, $T_F(a + \delta) - T_F(a) \leq 1$, and then covering $[b - a]$ by at most $(b - a)/\delta$ intervales of length at most δ, we see that $T_F(b) - T_F(a) \leq (b - a)/\delta$. □

5.9 THEOREM. *Let $-\infty < a < b < \infty$, and suppose that $F : [a, b] \to \mathbb{R}$ is absolutely continuous on $[a, b]$. Then F is differentaible for m a.e. $x \in (a, b)$, and $F'(x)$ is integrable, and for every $a \leq x < y \leq b$,*

$$F(y) - F(x) = \int_{[x,y]} F'(t)\mathrm{d}m \ . \tag{5.16}$$

Conversely, if F is is differentaible for m a.e. $x \in (a, b)$, and $F'(x)$ is integrable on $[a.b]$, and (5.16) is valid for all $a \leq x < y \leq b$, then F is absolutely continuous on $[a, b]$.

Proof. Let F be absolutely continuous on $[a, b]$. Then by Lemma 5.8, $F \in BV[a, b]$, and both F and T_F are contnuous. Hence we have the decomposition $F = F_1 - F_2$ where $F_1 = (T_F + F)/2$ and $F_2 = (T_F - F)/2$. Both F_1 and F_2 are monotone non-decreasing continuous functions and is T_F. Let μ_{F_1}, μ_{F_2} and μ_{T_F} be the correspinding Lebesgue-Stieltjes measures on $[a, b]$. We claim that both F_1 and F_2 are absolutely continouous. From here, the Fundamenta Theorem of Calculus may be applied separately to each of F_1 and F_2.

To see that F_1 is absolutely continuous, let $\mathrm{d}\mu_{F_1} = \mathrm{d}\lambda_{F_1} + f_1\mathrm{d}m$ be the Lebesgue decomposition of μ_{F_1}. Let A be a Borel subset of $\mathbb{R}$ such that $\lambda_{F_1}(A^c) = 0$ and $m(A) = 0$.

Then for all $\delta > 0$, there is a covering of A by a ountable disjoint sequence $\{(a_j, b_j]\}_{j\in\mathbb{N}}$ of half open intervals such that $\sum_{j=1}^{\infty}(b_j - a_j) < \delta$. By continuiuty from below, there is a finite N such that

$$\frac{1}{2}\lambda_{F_1}(\mathbb{R}) \leq \sum_{j=1}^{N}\lambda_{F_1}((a_j, b_j])) \qquad \text{and} \qquad \sum_{j=1}^{N}(b_j - a_j) < \delta \ .$$

But for any interval (x, y)

$$\lambda_{F_1}((x, y]) \leq \mu_{F_1}((x, y]) \leq \mu_{T_F}((x, y]) = T_F(y) - T_F(x) \ .$$

Thus,

$$\frac{1}{2}\lambda_{F_1}(\mathbb{R}) \leq \sum_{j=1}^{n}(T_F(b_j) - T_F(a_j)) \ .$$

In turn, each $T_F(b_j) - T_F(a_j)$ may be arbitraily well approcximated by a sum of the form

$$\sum_{k=1}^{n_j} |F(x_{k-1} - F(x_k)| \qquad a_j = x_0 < \cdots < x_{n_j} = b_j \ ,$$

and thus by subdividing the intervals further we have that

$$\frac{1}{2}\lambda_{F_1}(\mathbb{R}) - \epsilon \leq \sum_{j=1}^{N}(F(b_j) - F(a_j)) \ .$$

where we are no summing over the refined sets of intervals, and still have $\sum_{j=1}^{N}(b_j - a_j) < \delta$. Since F is absolutely continuous, we conclude that for an appropriate choice of $\delta > 0$.

$$\lambda_{F_1}([a,b]) \leq 4\epsilon \ .$$

Since $\epsilon > 0$ is arbitrary, $\lambda_{F_1} = 0$ and the same conclusion holds for λ_{F_2}. Hence both μ_{F_1} and μ_{F_2} are absolutely contnuous. Thus, for any $a \leq x < y \leq b$,

$$F_j(y) - F_j(x) = \int_{[x,y]} f_j \mathrm{d}m$$

for $j = 1, 2$, which as we have shown at the begiining means that F is differentiable a.e. m and $F'(x) = f_1(x) - f_2(x)$ a.e. m. The rest is stright-forward. □

6 Exercises

1. Let $\phi : [0,1] \to [0,1]$ be strictly increasing and continuously differentiable with $\phi(0) = 0$ and $\phi(1) = 1$. Let $F : [0,1] \to [0,1]$ be a function of bounded variation. Let $G = \phi \circ F \circ \phi^{-1}$. Show that $G \in BV[0,1]$, and that there is a constant K depending only on ϕ, and not F, such that

$$\|G\|_{BV} \leq K\|F\|_{BV} \ .$$

Find the least value of K for which this is true.

2. Let F and G be absolutely continuous on $[0,1]$. Is it necessarily true that FG is absolutely continuous on $[0,1]$? Justify your answer.

3. Let m be Lebesgue measure on $\mathbb{R}$. Let $T : [0,1] \to \mathbb{R}$ be an absolutely continuous functions. Let A be Borel subset of $[0,1]$ such that $m(A) = 0$. Prove that $m(T(A)) = 0$.