FUNDAMENTAL THEOREM OF CALCULUS FOR THE LEBESGUE INTEGRAL (UPDATED December 23, 2019.)

These notes are meant to be a companion to the Lecture Notes from the lectures of Prof. Simcha Horowitz[1], as an alternative to Chapter III. We change some of the presentation and the ordering of the material, and add proofs of the Vitali Lemma and Vitali Covering Theorem, and Lebesgue Density Theorem, among other things.

Reminder: Fundamental Theorem of Calculus. We recall the Fundamental Theorem of Calculus: If f is continuous on $[a,b]$, then the function defined by

$$F(x) = \int_a^x f(t)dt$$

is differentiable, and $F'(x) = f(x)$ for all $x \in [a,b]$. (In particular, $F \in C^1([a,b])$; that is, F is continuously differentiable on $[a,b]$.) Conversely, given any $F \in C^1([a,b])$, we can form $\tilde{F} \in C^1([a,b])$ by $\tilde{F}(x) = \int_a^x F'(x)dx$, and since $\tilde{F}'(x) = F'(x)$ for every x, we have that $\tilde{F} - F$ is constant on $[a,b]$, and therefore

$$\int_a^b F'(x)dx = F(b) - F(a) \tag{1}$$

The question arises as to whether such a result holds if $F \notin C^1$. What if F is simply differentiable, or differentiable almost-everywhere, with F' Lebesgue-integrable?

In general, the answer is no— the Cantor function $C(x)$ on $[0,1]$ is continuous and monotone increasing, differentiable with derivative 0 almost-everywhere (more precisely, differentiable with $C'(x) = 0$ at every point x not belonging to the Cantor set). However $C(0) = 0$ and $C(1) = 1$, so we have

$$\int_0^1 C'(x)dx = 0 \neq 1 = C(1) - C(0)$$

So we must place further conditions on F, in addition to F' integrable, in order for (1) to hold.

[1]Can be found on math-wiki.com page.

A Diagram and a Spoiler. We consider the following classes of functions, on a closed interval $[a,b]$:

- $C^1([a,b])$: functions continuously differentiable on $[a,b]$
- $Lip([a,b])$: functions that are Lipschitz-continuous on $[a,b]$; i.e. such that there exists a constant $M>0$ such that $|F(x)-F(y)|\le M|x-y|$ for all $x,y\in[a,b]$
- $AC([a,b])$: functions that are absolutely continuous on $[a,b]$ (to be defined below)
- $BV([a,b])$: functions of bounded variation on $[a,b]$ (to be defined below)
- $DAE([a,b])$: functions differentiable almost-everywhere on $[a,b]$.

For any closed interval $[a,b]$ we will show the following inclusions, and that each is a strict inclusion:

$$C^1([a,b])\subset Lip([a,b])\subset AC([a,b])\subset BV([a,b])\subset DAE([a,b]) \tag{2}$$

We know that (1) holds for $F\in C^1([a,b])$, this is the classical Fundamental Theorem of Calculus. We also saw in the example of the Cantor function, that (1) cannot hold for all $F\in DAE([a,b])$.

The main result of this chapter is that (1) holds for all $F\in AC([a,b])$; and conversely, if f is Lebesgue integrable, then the function $F(x)=\int_a^x f dm$ is absolutely continuous. In other words, the fundamental theorem can be extended by replacing $f\in C([a,b])$ with $f\in L^1([a,b])$, and $F\in C^1([a,b])$ with $F\in AC([a,b])$.

1. Definitions and Inclusions

In this section we give the definitions of $AC([a,b])$ and $BV([a,b])$, and discuss the inclusions in the diagram above (2).

1.1. Absolutely Continuous Functions.

Definition 1. *A function $F:[a,b]\to\mathbb{R}$ is called* **absolutely continuous** *on $[a,b]$ (denoted $F\in AC([a,b])$) $\iff$ for every $\epsilon>0$ there exists $\delta>0$ such that, for any finite collection of disjoint intervals $[a_k,b_k]\subset[a,b]$, $k=1,2,\ldots,n$, we have*

$$\sum_{k=1}^{n}(b_k-a_k)<\delta \qquad \Longrightarrow \qquad \sum_{k=1}^{n}|F(b_k)-F(a_k)|<\epsilon$$

We first remark that the case $n=1$ is equivalent to uniform continuity (which is equivalent to continuity on a closed interval). The difference between continuity and absolute continuity, is that "the δ can be divided". Whereas uniform continuity says that the function cannot vary too much over *one* small interval, absolute continuity is the stronger property that the function cannot vary too much over a union of intervals, provided their total length is less than δ.

It is easy to see that any Lipschitz-continuous function is absolutely continuous (see below). To understand what it looks like for a function to be continuous but not absolutely continuous, consider the Cantor function (we will show later that it is not absolutely continuous). Although the Cantor function is monotone and continuous, its increases are "squeezed" into the Cantor set— it is constant almost-everywhere. This leads to sharp increases in small neighborhoods of points belonging to the Cantor set (indeed, the Cantor function is not differentiable at these points). The relation between the δ-ϵ of continuity is not linear; the smaller we take ϵ, we must take δ much smaller (δ goes to 0 much faster than ϵ), because of these sharp increases in small neighborhoods of points belonging to the Cantor set. Thus if we divide δ over many smaller neighborhoods of points in the Cantor set, we can get the absolute-continuity condition to fail.

Note: This intuitive idea is NOT how we will show that $C(x) \notin AC([0,1])$; it is rather difficult to prove directly in this way. We will instead show that $C(x) \notin AC([0,1])$ through the failure of (1)... we bring this explanation for intuition only.

1.2. **Bounded Variation.** Let $F : [a,b] \to \mathbb{R}$. We consider a partition of the domain $a = x_0 < x_1 < x_2 < \cdots < x_n < x_{n+1} = b$, and call $\sum_{k=0}^{n} |F(x_{k+1}) - F(x_k)|$ the variation of F with respect to the partition.

Definition 2. *Let* $F : [a,b] \to \mathbb{R}$*. The* **total variation** *of* F *over* $[a,b]$ *is*

$$T_a^b(F) = \sup_{a<x_1<\cdots<x_n<b} \sum_{k=0}^{n} |F(x_{k+1}) - F(x_k)|$$

We say F *is of* **bounded variation** *on* $[a,b]$ *iff* $T_a^b(F) < \infty$.

For example, if F is monotone increasing, then $F(x_{k+1}) \geq F(x_k)$, and so for *any* partition the absolute-value in the sum can be removed and we get a telescoping sum $\sum_{k=0}^{n}(F(x_{k+1}) - F(x_k)) = F(x_{n+1}) - F(x_0) = F(b) - F(a)$. Similarly if F is monotone decreasing, we would get $T_a^b(F) = F(a) - F(b)$. If F is piecewise monotone, then we would get the sum of the monotone increases and decreases, since it is easy to check that $T_a^b(F) + T_b^c(F) = T_a^c(F)$ for any $a < b < c$.

We have the following useful characterization of functions of bounded variation:

Lemma 1. *A function* F *on* $[a,b]$ *is of bounded variation* $\iff$ *there exist monotone non-decreasing functions* $g, h : [a,b] \to \mathbb{R}$ *such that* $F(x) = g(x) - h(x)$ *for all* $x \in [a,b]$.

Proof. Let $g(x) = T_a^x(F)$, the total variation of F up to x. (We set $g(a) = 0$.) It is clear that g is monotone non-decreasing.

Now set $h(x) = g(x) - f(x)$. By definition, it is clear that $f = g - h$. It remains to show that h is monotone non-decreasing. But since, by definition of total variation we have for $y > x$ that $g(y) - g(x) = T_y^x(F) \geq |F(y) - F(x)| \geq F(y) - F(x)$, we have

$$y > x \Longrightarrow h(y) = g(y) - F(y) \geq g(x) - F(x) = h(x)$$

so h is monotone non-decreasing. □

Intuitively, if F is increasing, then g increases with F and h remains unchanged, while if F is decreasing, then the difference between F and g grows, so h increases.

1.3. $C^1([a,b]) \subset Lip([a,b])$**.** We now turn to the diagram (2), and wish to prove each inclusion, and show it is a proper inclusion. We begin with the left-most, that every continuously-differentiable function on a closed interval is Lipschitz-continuous.

If $F \in C^1([a,b])$, then F' is continuous on $[a,b]$, and thus bounded. Bounded derivative is already enough for Lipschitz-continuity: let $M = \sup_{x\in[a,b]} |F'(x)|$, and consider the Lagrange Mean Value Theorem for $x, y \in [a,b]$: there exists c between x and y such that

$$|F(x) - F(y)| = |F'(c)||x - y| \leq M|x - y|$$

and the Lipschitz condition is satisfied.

The inclusion is proper since eg. the function $F(x) = |x|$ on $[-1, 1]$ satisfies the Lipschitz condition (with constant $M = 1$) by the triangle inequality

$$\Big||x| - |y|\Big| \leq |x - y|$$

but F is not differentiable at 0 (and the derivative jumps there from -1 to 1).

1.4. $Lip([a,b]) \subset AC([a,b])$**.** Suppose now that $F \in Lip([a,b])$; that is, there exists a constant $M > 0$ such that $|F(x) - F(y)| \leq M|x - y|$ for every $x, y \in [a,b]$. Suppose further we have a finite union of disjoint intervals $[a_k, b_k] \subset [a,b]$, for $k = 1, 2, \ldots, n$. Then by the Lipschitz condition we have

$$\sum_{k=1}^{n} |F(b_k) - F(a_k)| \leq \sum_{k=1}^{n} M|b_k - a_k| = M \sum_{k=1}^{n} |b_k - a_k| < M\delta$$

and so for any $\epsilon > 0$, setting $\delta = \epsilon/M$ (with M the Lipschitz constant) yields the condition for absolute continuity.

The function $F(x) = \sqrt{x}$ is not Lipschitz-continuous on $[0,1]$, since the derivative $F'(x) = \frac{1}{2\sqrt{x}}$ is continuous on $(0,1]$ and tends to $\lim_{x\to 0} F'(x) = \infty$, so again by Lagrange for any $x \in (0,\epsilon)$ we have some $c \in (0,\epsilon)$ such that

$$|\sqrt{x} - \sqrt{0}| = |F'(c)| \cdot |x - 0| > \frac{1}{2\sqrt{\epsilon}}|x-0|$$

and there is no constant M that can satisfy the Lipschitz condition, because it would require $M > 1/\epsilon$ for all $\epsilon > 0$.

1.5. $AC([a,b]) \subset BV([a,b])$**.** Let $F \in AC([a,b])$, and set $\epsilon = 1$. Then there exists $\delta > 0$ such that for any subinterval $[\alpha,\beta] \subset [a,b]$ of length $\beta - \alpha < \delta$, and *any* partition $\alpha = x_0 < x_1 < \cdots < x_n < x_{n+1} = \beta$, we have

$$\sum_{k=0}^{n} |F(x_{k+1}) - F(x_k)| < 1$$

and, in particular, this implies that $T_\alpha^\beta(F) \leq 1$. Since this is true for any subinterval of length less than δ, we can divide $[a,b]$ into finitely many such intervals $a = \alpha_0 < \alpha_1 < \cdots < \alpha_n < \alpha_{n+1} = b$ such that each $\alpha_{j+1} - \alpha_j < \delta$, and

$$T_a^b = \sum_{k=0}^{n} T_{\alpha_k}^{\alpha_{k+1}} \leq \sum_{k=0}^{n} 1 = n+1 < \infty$$

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

The Cantor function $C(x)$ on $[0,1]$ discussed above is monotone, and therefore of bounded variation. (In fact, $T_0^1(C) = 1$.) As mentioned above, we will show later that the Cantor function is not absolutely continuous, and thus the inclusion is proper.

1.6. $BV([a,b]) \subset DAE([a,b])$**.** The Lebesgue Differentiation Theorem (Theorem 2 below) shows that every monotone function is differentiable almost everywhere, and since any function of bounded variation can be expressed as a difference of two monotone functions (Lemma 1 above), a function of bounded variation is thus also differentiable almost everywhere.

The function $F(x) = \sin(\frac{1}{x})$ is differentiable at all points except $x = 0$, thus is differentiable almost everywhere. But it is not of bounded variation on $[0,1]$, since the partition $x_j = \frac{2}{\pi(n-j)}$ has $|\sin(1/x_{j+1}) - \sin(1/x_j)| = 1$ for all j, and thus the sum over $j = 1, 2, \ldots n$ shows that $T_0^1(F) \geq n$. But this holds for all $n \in \mathbb{N}$, so $T_0^1(F) = \infty$.

Remark: A similar argument shows that $F(x) = x\sin(\frac{1}{x})$ is not of bounded variation, by using the fact that the harmonic series diverges (in place of $T_0^1(F) \geq \sum_{k=1}^{n} 1 = n$ one estimates $T_0^1(F) \geq \sum_{k=1}^{n} \frac{1}{n}$).

This gives an example of a *continuous* function that is differentiable almost-everywhere, but not of bounded variation.

2. Outline of the Main Theorems

Here we list the four steps that we will prove, in order to extend the Fundamental Theorem of Calculus to Lebesgue-integrable functions and absolutely continuous functions.

(1) If F is absolutely contunuous on $[a, b]$, then F is differentiable almost-everywhere on $[a, b]$, and F' is Lebesgue-integrable on $[a, b]$. (This will follow from the **Lebesgue Differentiation Theorem**, see section 4.)
(2) If f is Lebesgue-integrable on $[a, b]$, then the function $F(x) = \int_a^x f dm$ is absolutely continuous on $[a, b]$.
(3) If f is Lebesgue integrable on $[a, b]$ and $F(x) = \int_a^x f dm$, then $F'(x) = f(x)$ almost-everywhere on $[a, b]$.
(4) If $F \in AC([a, b])$, then $\int_a^b F' dm = F(b) - F(a)$.

3. Vitali's Lemma and the Vitali Covering Theorem

A key tool in the proofs of the Theorems listed above[2] is the Vitali Covering Theorem, which we prove in this section. We begin with the covering lemma:

Lemma 2 (Vitali's Lemma)**.** *Let $\mathcal{C}$ be a collection of balls in $\mathbb{R}^d$, of bounded radius. Then there exists a countable subcollection $\mathcal{D} \subset \mathcal{C}$ of* disjoint *balls, such that for every $C \in \mathcal{C}$ there exists $D \in \mathcal{D}$ from the subcollection such that $C \cap D \neq \emptyset$ and $r(D) \geq \frac{1}{2} r(C)$.*

In particular, setting $5B(x, r) := B(x, 5r)$ to be the concentric ball with radius multiplied by 5, we have

$$\bigcup_{C \in \mathcal{C}} C \subset \bigcup_{D \in \mathcal{D}} 5D$$

Proof. We first divide $\mathcal{C}$ into families according to radii:

$$\mathcal{C}_n = \{C \in \mathcal{C} : 2^{-n-1} R < r(C) \leq 2^{-n} R\}$$

We will construct our subcollection $\mathcal{D}$ by taking from each family a subcollection $\mathcal{D}_n \subset \mathcal{C}_n$ of balls, in the following way:

We first set $\mathcal{D}_0 \subset \mathcal{C}_0$ to be a maximal disjoint subcollection— i.e., a disjoint subcollection $\mathcal{D}_0$ such that any $C \in \mathcal{C}_0$ intersects some ball from $\mathcal{D}_0$, and hence $\mathcal{D}_0$ is not properly contained in any disjoint subcollection in $\mathcal{C}_0$. One can construct such a maximal disjoint subcollection by

[2]In particular, of the Lebesgue Differentiation Theorem; but also in the last step

first considering only those balls in $\mathcal{C}_0$ whose center is within 1 of the origin; by finite measure considerations, there can be only finitely many disjoint balls at distance at most 1 from the origin, with radii bounded below by $r(D) > 2^{-1}R$, so there is a maximal such subcollection. If we then consider balls in $\mathcal{C}_0$ whose centers are within $k \in \mathbb{Z}$ of the origin, then we can add (finitely many) disjoint balls until we again reach a maximal such subcollection. Inductively as $k \to \infty$ we construct a countable subcollection $\mathcal{D}_0$ of disjoint balls, and any $C \in \mathcal{C}_0 \backslash \mathcal{D}_0$ has a center at distance $d \in \mathbb{R}$ from the origin, and since $d < k$ for some $k \in \mathbb{Z}$, C must intersect a ball from $\mathcal{D}_0$ constructed at step k, since otherwise it would contradict the maximality. So we have a maximal disjoint subcollection $\mathcal{D}_0 \subset \mathcal{C}_0$ that is countable.

Now consider the collection

$$\mathcal{H}_1 = \{C \in \mathcal{C}_1 : C \cap D = \emptyset \quad \forall D \in \mathcal{D}_0\}$$

of balls in $\mathcal{C}_1$ that are disjoint from all balls we collected so far in $\mathcal{D}_0$. In the same way as above, we select a maximal disjoint subcollection $\mathcal{D}_1 \subset \mathcal{H}_1$, which is again countable. We continue inductively to define

$$\mathcal{H}_n = \left\{ C \in \mathcal{C}_n : C \cap D = \emptyset \quad \forall D \in \bigcup_{j=0}^{n-1} \mathcal{D}_0 \right\}$$

and finally $\mathcal{D} = \bigcup_{n=0}^{\infty} \mathcal{D}_n$.

Since each $\mathcal{D}_n$ is countable, the full collection $\mathcal{D}$ is countable. By construction, all balls in $\mathcal{D}$ are pairwise disjoint. It remains to show that for every $C \in \mathcal{C}$ there exists $D \in \mathcal{D}$ from the subcollection such that $C \cap D \neq \emptyset$ and $r(D) \geq \frac{1}{2} r(C)$.

If $C \in \mathcal{C}_n$ were disjoint from all $D \in \bigcup_{j=0}^{n} \mathcal{D}_n$, then in particular $C \in \mathcal{H}_n$ and is disjoint from all $D \in \mathcal{D}_n$, which would contradict the maximality of $\mathcal{D}_n$. Therefore there exists $D \in \bigcup_{j=0}^{n} \mathcal{D}_n$ such that $C \cap D \neq \emptyset$.

But $C \in \mathcal{C}_n$ implies that $r(C) \leq 2^{-n}R$, and $D \in \bigcup_{j=0}^{n} \mathcal{D}_n \subset \bigcup_{j=0}^{n} \mathcal{C}_n$ implies that $r(D) > 2^{-n-1}R$, and so

$$r(D) > 2^{-n-1}R = \frac{1}{2} \cdot 2^{-n}R \geq \frac{1}{2} r(C)$$

as required.

It remains to show that $\bigcup_{C \in \mathcal{C}} C \subset \bigcup_{D \in \mathcal{D}} 5D$. But since any $C \in \mathcal{C}$ intersects $D \in \mathcal{D}$ with radius $r(D) \geq \frac{1}{2} r(C)$, any point in C is at distance at most

$$diam(C) + r(D) = 2r(C) + r(D) \leq 5r(D)$$

from the center of D, and so $C \subset 5D$.

□

We use the Lemma to prove the covering theorem. We first define a **Vitali covering**:

Definition 3. *Let $E \subset \mathbb{R}^d$. A* **Vitali covering** *of E is a collection of balls $\mathcal{V}$, such that for every $x \in E$ and every $\epsilon > 0$, there exists a ball $C \in \mathcal{V}$ such that $x \in C$ and $r(C) < \epsilon$.*

Intuitively, this means a covering of E by balls as small as we choose. Naturally, there is tremendous overlap between the balls— each point must be covered by infinitely many balls from $\mathcal{C}$— but this gives us the freedom to collect many small balls that will cover E "efficiently".

Theorem 1 (Vitali's Covering Theorem)**.** *Let $E \in \mathbb{R}^d$ be bounded*[3]*, and $\mathcal{V}$ a Vitali covering of E. Then there exists a countable, disjoint subcollection of balls $\mathcal{D} = \{D_n\}_{n=1}^{\infty} \subset \mathcal{V}$ such that*

$$m^*\left(E \backslash \bigcup_{n=1}^{\infty} D_n\right) = 0$$

and moreover, for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$m^*\left(E \backslash \bigcup_{n=1}^{N} D_n\right) < \epsilon$$

Proof. Without loss of generality we may assume that all balls in $\mathcal{V}$ are of radius at most 1, since this collection is still a Vitali covering of E. We may also discard from $\mathcal{V}$ any balls that do not intersect E, and thus we may assume that $\bigcup_{C\in\mathcal{V}} C$ is a bounded set.

We construct the countable, disjoint collection $\mathcal{D}$ from the Vitali Lemma, and note that in this case each $\mathcal{D}_n$ is finite, since E is bounded. Thus we can order $\mathcal{D} = \{D_n\}_{n=1}^{\infty}$ by decreasing size $r(D_j) \geq r(D_{j+1})$.

We consider the sets

$$Z_N = \left\{ z \in E : z \notin \bigcup_{n=1}^{N} \overline{D_n} \right\}$$

for each n, and their intersection

$$Z = \bigcap_{N=1}^{\infty} Z_N = \left\{ z \in E : z \notin \bigcup_{n=1}^{\infty} \overline{D_n} \right\}$$

of points not contained in any $\overline{D_n}$. Since the boundary of each D_n has measure 0, their union does as well, and so we need only show

[3]The boundedness condition can be removed, but since we will only apply the Vitali Covering Theorem to bounded sets, we omit this extension.

that for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $m^*(Z_N) < \epsilon$; since $Z_{N+1} \subset Z_N$, this will show that $m^*(Z) \leq \lim_{N\to\infty} m^*(Z_N) = 0$.

Thus, let $\epsilon > 0$. Since the balls $\{D_n\}$ are disjoint and measurable, and their union is contained in a bounded set, we have

$$\sum_{n=1}^{\infty} m(D_n) = m\left(\bigcup_{n=1}^{\infty} D_n\right) < \infty$$

and therefore there exists $N \in \mathbb{N}$ such that the tail

$$\sum_{n>N} m(D_n) < \epsilon$$

Now take $z \in Z_N$. Since z is not contained in the closed set $\bigcup_{n=1}^{N} \overline{D_n}$ (a finite union of closed balls), there exists $\delta > 0$ such that the ball $B(z,\delta)$ does not intersect $\bigcup_{n=1}^{N} \overline{D_n}$. Since $\mathcal{V}$ is a Vitali covering, there exists a ball $C \in \mathcal{V}$ such that $z \in C$ and $C \subset B(z,\delta)$, and thus $C \cap D_n = \emptyset$ for all $n \leq N$. But by the Vitali Lemma, C must intersect some ball from $\mathcal{D}$, and so there exists $n > N$ such that $C \cap D_n \neq \emptyset$, and $C \subset 5D_n$.

We conclude that for every $z \in Z_N$, there exists $n > N$ and a $C \in \mathcal{V}$ such that $z \in C \subset 5D_n$. Therefore,

$$Z_N \subset \bigcup_{n>N} 5D_n$$

But since $m(5D_n) = 5^d \cdot m(D_n)$, we deduce that

$$m^*(Z_N) \leq m\left(\bigcup_{n>N} 5D_n\right) = \sum_{n>N} 5^d \cdot m(D_n) = 5^d \sum_{n>N} m(D_n) < 5^d \epsilon$$

as required. □

Corollary 1 (Vitali Covering Theorem, Second Version). *Let $E \subset \mathbb{R}^d$ be bounded, and suppose $\mathcal{V}$ is a Vitali covering of E. Then for every $\epsilon > 0$, there exists a finite disjoint subcollection $D_1, \ldots, D_n \in \mathcal{V}$ such that*

$$\begin{aligned} m\left(\bigcup_{j=1}^{n} D_n\right) &< m^*(E) + \epsilon \\ m^*\left(E \cap \bigcup_{j=1}^{n} D_n\right) &> m^*(E) - \epsilon \end{aligned}$$

The second line follows from the Vitali Covering Theorem (Theorem 1); the point of the Corollary is the first line, that we may take our subcollection $\{D_n\}$ to have measure not more than $m^*(E) + \epsilon$.

Proof. Let $\epsilon > 0$. By definition of outer measure, there exists an open set $U \supset E$ of measure $\mu(U) < m^*(E) + \epsilon$. Consider the collection

$$\mathcal{V}' = \{C \in \mathcal{V} : C \subset U\}$$

of those balls in $\mathcal{V}$ that are contained in U; we claim this is also Vitali covering of E. Since for every $z \in E$, the fact that U is open guarantees that there exists $\delta > 0$ such that the ball $B(z, \delta) \subset U$, and thus any ball of radius less than $\delta/2$ containing z will be contained in $B(z, \delta) \subset U$. Since $\mathcal{V}$ is a Vitali covering, for every $z \in E$ and $\epsilon > 0$ there exists a ball $C \in \mathcal{V}$ of radius less than ϵ, and so taking $\min\{\epsilon, \delta/2\}$ we can guarantee such a ball that is also conatined in U, and thus belongs to $\mathcal{V}'$. Thus $\mathcal{V}'$ is also a Vitali covering of E.

Now apply the Vitali Covering Theorem to $\mathcal{V}'$. The Theorem gives

$$m^* \left(E \cap \bigcup_{j=1}^{n} D_n \right) > m^*(E) - \epsilon$$

and since each $D_j \in \mathcal{V}'$ is contained in U, we have $\bigcup_{j=1}^{n} D_j \subset U$ and therefore also

$$m \left(\bigcup_{j=1}^{n} D_n \right) \leq m(U) < m^*(E) + \epsilon$$

□

4. Lebesgue Differentiation Theorem

Theorem 2 (Lebesgue Differentiation Theorem)**.** *Let F be monotone non-decreasing on $[a, b]$. Then F is differentiable almost-everywhere on $[a, b]$, the derivative $F'(x) \geq 0$ almost everywhere on $[a, b]$, and we have*

$$\int_a^b F' dm \leq F(b) - F(a)$$

Remarks:

(1) Note that the inequality in the conclusion can be strict; eg. for the Cantor function $C(x)$ on $[0, 1]$ we have

$$\int_0^1 C'(x) dx = 0 \neq 1 = C(1) - C(0)$$

We will show later that equality holds if F is absolutely continuous.

(2) Since F' is almost-everywhere non-negative and its integral is bounded by the finite quantity $F(b) - F(a)$, we conclude that F' is Lebesgue-integrable. In fact, since any function of bounded

variation can be written as the difference of two monotone functions $F(x) = g(x) - h(x)$, we get that $F \in BV([a,b])$ implies F is differentiable almost everywhere and $F' = g' - h'$ is Lebesgue integrable on $[a,b]$, since g' and h' are Lebesgue integrable.

Since absolutely continuous functions have bounded variation, we conclude that if $F \in AC([a,b])$ then F is differentiable almost everywhere and F' is Lebesgue-integrable on $[a,b]$.

Proof. The main step in the proof is to prove that the set

$$E = \left\{ x \in [a,b] : \liminf_{h\to 0} \frac{F(x+h) - F(x)}{h} < \limsup_{h\to 0} \frac{F(x+h) - F(x)}{h} \right\}$$

has measure zero; these are points where the lim inf of the ratio defining the derivative is bounded but the limit does not exist. Points where the ratio tends to ∞ will be handled later in the proof.

For this it is sufficient to show that for any $\alpha < \beta$ the set

$$E_{\alpha,\beta} = \left\{ x \in [a,b] : \liminf_{h\to 0} \frac{F(x+h) - F(x)}{h} < \alpha < \beta < \limsup_{h\to 0} \frac{F(x+h) - F(x)}{h} \right\}$$

has measure zero, since then E is a countable union of these sets and

$$m^*(E) = m^*\left(\bigcup_{\alpha,\beta\in\mathbb{Q}} E_{\alpha,\beta} \right) \le \sum_{\alpha,\beta\in\mathbb{Q}} m^*(E_{\alpha,\beta}) = 0$$

So consider one of the sets $E_{\alpha,\beta}$, and let $\epsilon > 0$. For every $x \in E_{\alpha,\beta}$, there exists a sequence $h_n \to 0$ such that

$$\frac{F(x+h_n) - F(x)}{h_n} < \alpha$$

These intervals $[x, x+h_n]$ (if $h_n > 0$) or $[x+h_n, x]$ (if $h_n < 0$) form a Vitali covering of $E_{\alpha,\beta}$, so by the second version of the Vitali Covering Theorem (Lemma 1) there is a finite collection of disjoint intervals $\{[x_n, y_n]\}_{n=1}^N$ — with $F(y_n) - F(x_n) < \alpha(y_n - x_n)$ for every $n = 1, 2, \ldots, N$ — such that

$$m^*(E_{\alpha,\beta}) - \epsilon < \sum_{n=1}^{N} (y_n - x_n) < m^*(E_{\alpha,\beta}) + \epsilon \tag{3}$$

and in particular

$$\sum_{n=1}^{N} (F(y_n) - F(x_n)) < \alpha \sum_{n=1}^{N} (y_n - x_n) < \alpha(m^*(E_{\alpha,\beta}) + \epsilon) \tag{4}$$

Intuitively, over intervals approximating $E_{\alpha,\beta}$ we have mild rate-of-change essentially bounded by α.

We now wish to use the $\beta < \limsup$ condition to get a contradiction to $m^*(E_{\alpha,\beta}) > 0$. For each of the intervals $[x_n, y_n]$ considered above, we discard the endpoints (they are a negligible set anyway) and again consider for each $z \in E_{\alpha,\beta} \cap (x_n, y_n)$ a sequence $h_m \to 0$ satisfying $\frac{F(z+h_m)-F(z)}{h_m} > \beta$, which again gives a Vitali covering of $E_{\alpha,\beta} \cap (x_n, y_n)$; since (x_n, y_n) is open we may assume all intervals in the Vitali covering are contained in (x_n, y_n). Again apply the second version of the Vitali Covering Theorem (Lemma 1) to extract a finite subcollection of intervals $\{[w_{l,n}, z_{l,n}]\}_{l=1}^{L_n}$, each contained $[w_{l,n}, z_{l,n}] \subset (x_n, y_n)$ and each satisfying

$$F(z_{l,n}) - F(w_{l,n}) > \beta \cdot (z_{l,n} - w_{l,n}) \tag{5}$$

and such that

$$m^*(E_{\alpha,\beta} \cap (x_n, y_n)) - \frac{\epsilon}{N} < \sum_{l=1}^{L_n} (z_{l,n} - w_{l,n}) < m^*(E_{\alpha,\beta} \cap (x_n, y_n)) + \frac{\epsilon}{N} \tag{6}$$

Now we can put together the inequalities, starting with (4) to get

$$\begin{aligned}
\alpha(m^*(E_{\alpha,\beta}) + \epsilon) &> \sum_{n=1}^{N} (F(y_n) - F(x_n)) \\
&\geq \sum_{n=1}^{N} \sum_{[w_{l,n}, z_{l,n}] \subset (x_n, y_n)} (F(z_{l,n}) - F(w_{l,n})) \\
&> \sum_{n=1}^{N} \left(\beta \cdot \sum_{[w_{l,n}, z_{l,n}] \subset (x_n, y_n)} (z_{l,n} - w_{l,n}) \right) \\
&\geq \beta \left(\sum_{n=1}^{N} \sum_{[w_{l,n}, z_{l,n}] \subset (x_n, y_n)} (z_{l,n} - w_{l,n}) \right) \\
&> \beta \sum_{n=1}^{N} \left(m^*(E_{\alpha,\beta} \cap (x_n, y_n)) - \frac{\epsilon}{N} \right) \\
&> \beta \left(m^*(E_{\alpha,\beta} \cap \bigcup_{n=1}^{N} (x_n, y_n)) - \epsilon \right) \\
&> \beta(m^*(E_{\alpha,\beta}) - 2\epsilon)
\end{aligned}$$

Letting $\epsilon \to 0$ we have

$$\alpha \cdot m^*(E_{\alpha,\beta}) \geq \beta \cdot m^*(E_{\alpha,\beta})$$

but since $\alpha < \beta$, we get a contradiction unless $m^*(E_{\alpha,\beta}) = 0$.

Therefore $m^*(E_{\alpha,\beta}) = 0$ for any α, β, and so

$$m^*(E) = m^*\left(\bigcup_{\alpha,\beta\in\mathbb{Q}} E_{\alpha,\beta}\right) \leq \sum_{\alpha,\beta\in\mathbb{Q}} m^*(E_{\alpha,\beta}) = 0$$

For technical reasons, we extend F to be $F(x) = F(b)$ for all $x > b$. We then set

$$DF_n(x) = \frac{F(x+\frac{1}{n}) - F(x)}{1/n}$$

We have shown above that the limit $\lim_{n\to\infty} DF_n(x) = DF(x) \in [0,\infty]$ exists almost-everywhere; it is clear by monotonicity of F that $DF_n(x) \geq 0$ for every x and n, and so by Fatou's Lemma

$$\begin{aligned}
\int_a^b DF dm &\leq \liminf_{n\to\infty} \int_a^b DF_n dm \\
&\leq \liminf_{n\to\infty} n \cdot \int_a^b [F(x+\frac{1}{n}) - F(x)]dm \\
&\leq \liminf_{n\to\infty} n \cdot \left(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} F dm - \int_a^b F dm\right) \\
&\leq \liminf_{n\to\infty} n \cdot \left(\int_b^{b+\frac{1}{n}} F dm - \int_a^{a+\frac{1}{n}} F dm\right)
\end{aligned}$$

Now, since $F(x) = F(b)$ for all $x \geq b$, and in particular on the interval $[b, b+\frac{1}{n}]$, the first integral on the right-hand side is

$$n\int_b^{b+\frac{1}{n}} F dm = n\int_b^{b+\frac{1}{n}} F(b) dm = n \cdot F(b) \cdot \frac{1}{n} = F(b)$$

Similarly, since F is monotone, we have $F(x) \geq F(a)$ for all $x \geq a$ and so

$$n\int_a^{a+\frac{1}{n}} F dm \geq n\int_a^{a+\frac{1}{n}} F(a) dm = n \cdot F(a) \cdot \frac{1}{n} = F(a)$$

Putting these together gives

$$\int_a^b DF dm \leq \liminf_{n\to\infty} n \cdot \left(\int_b^{b+\frac{1}{n}} F dm - \int_a^{a+\frac{1}{n}} F dm\right) \leq F(b) - F(a)$$

Now since $DF \geq 0$ almost-everywhere and $\int_a^b DF dm \leq F(b) - F(a) < \infty$, we have $DF(x) < \infty$ almost everywhere, and therefore F

is differentiable and $F'(x) = DF(x)$ almost everywhere. It follows that

$$\int_a^b F' dm \leq F(b) - F(a)$$

as well. ☐

Corollary 2. *If $F \in BV([a,b])$, then F is differentiable almost everywhere and F' is Lebesgue-integrable on $[a,b]$.*

Proof. $F \in BV([a,b])$ implies that $F(x) = g(x) - h(x)$ where g, h are monotone non-decreasing functions. By the Lebesgue Differentiation Theorem, g and h are differentiable almost-everywhere, and at such points F is differentiable as well and $F'(x) = g'(x) - h'(x)$. Since g' and h' are integrable on $[a,b]$, so is F'. ☐

Note that since any absolutely continuous function is of bounded variation, any $F \in AC([a,b])$ is differentiable almost-everywhere and F' is Lebesgue-integrable.

5. Integral of $f \in L^1([a,b])$ is absolutely continuous

The key step in showing absolute continuity is the following

Lemma 3. *Let $f \in L^1([a,b])$. Then for every $\epsilon > 0$ there exists $\delta > 0$, such that if $E \subset [a,b]$ is measurable with $m(E) < \delta$, then $\int_E |f| dm < \epsilon$*

Proof. This property is clear if f is bounded, since if $|f(x)| \leq M$ for all x, then

$$\int_E |f(x)| dm \leq \int_E M dm = M \cdot m(E)$$

and we may take $\delta = \epsilon / M$.

The challenge is to deal with f that is not bounded; however, if f is integrable, then it may be approximated by bounded functions.

For each $n \in \mathbb{N}$ define

$$f_n(x) = \min\{|f(x)|, n\}$$

This is an approximation to f that cuts off all values above n, so that f_n is bounded. Clearly the sequence $\{f_n\}_{n=1}^\infty$ is monotone, and if $f(x)$ is bounded then the sequence $f_n(x)$ is eventually constant, so that $\lim_{n\to\infty} f_n(x) = |f(x)|$; if $|f(x)| = \infty$, then clearly $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} n = \infty$. Thus in either case we have $\lim_{n\to\infty} f_n(x) = |f(x)|$ monotonically.

By Monotone Convergence Theorem, we have

$$\lim_{n\to\infty} \int_a^b f_n dm = \int_a^b f dm$$

so given $\epsilon > 0$, we may find n such that

$$\int_a^b (|f(x)| - f_n(x))dm < \epsilon/2$$

Now $|f_n(x)| \leq n$ is bounded, so choosing $\delta < \epsilon/2n$ we have for any $m(E) < \epsilon/2n$

$$\begin{aligned} \int_E |f|dm &= \int_E (|f| - f_n)dm + \int_E f_n dm < \int_X (|f| - f_n)dm + \int_E ndm \\ &< \frac{\epsilon}{2} + n \cdot m(E) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

□

Theorem 3. *If $f \in L^1([a,b])$ and $F(x) = \int_a^x f dm$, then $F \in AC([a,b])$.*

Proof. Since $f \in L^1([a,b])$, by the Lemma for every ϵ there exists δ such that for every $E \subset [a,b]$ of measure $m(E) < \delta$, we have $\int_E |f|dm < \epsilon$. Let $[a_k, b_k] \subset [a,b]$ be disjoint intervals with $\sum_{k=1}^n (b_k - a_k) < \delta$; this means that $E = \bigcup_{n=1}^k [a_k, b_k]$ satisfies $m(E) < \delta$.

Now for any k,

$$F(b_k) - F(a_k) = \int_a^{b_k} f dm - \int_a^{a_k} f dm = \int_{a_k}^{b_k} f dm$$

so that

$$\begin{aligned} \sum_{k=1}^n |F(b_k) - F(a_k)| &= \sum_{k=1}^n \left| \int_{a_k}^{b_k} f dm \right| \\ &\leq \sum_{k=1}^n \int_{a_k}^{b_k} |f|dm \\ &\leq \int_E |f|dm < \epsilon \end{aligned}$$

Thus F satisfies the definition of absolute continuity. □

6. If $f \in L^1([a,b])$, then $\frac{d}{dx}(\int_a^x f dm) = f(x)$ almost everywhere

For this section we will need the following fact, which states that an integrable function can be characterized by its definite integrals:

Lemma 4. *Let $f \in L^1([a,b])$, and suppose that for every $c \in [a,b]$, we have $\int_a^c f dm = 0$. Then $f(x) = 0$ for almost every $x \in [a,b]$.*

Proof. Let $E = \{x \in [a,b] : f(x) \neq 0\}$, and suppose by contradiction that $m(E) > 0$. This set E is the union of two sets, one where $f(x) > 0$ and one where $f(x) < 0$; at least one of them must have positive measure, so without loss of generality assume $f(x) > 0$ on E (otherwise replace f with $-f$).

The first step is to find a *closed* set $F \subset E$ such that $m(F) > 0$ also has positive measure. For this consider the (measurable) set $[a,b] \cap E^c$, whose measure is equal to $b-a-m(E) < b-a$, and for $\epsilon > 0$ sufficiently small select an open covering $U \supset ([a,b] \cap E^c)$ of measure $b-a-\epsilon$. Then $F = [a,b] \backslash U \subset E$ is a closed set contained in E of positive measure $m(F) = \epsilon > 0$.

Since $F \subset E = \{x \in [a,b] : f(x) > 0\}$ and $m(F) > 0$, we must have $\int_F f dm > 0$. On the other hand F is closed, so $(a,b) \backslash F$ is open, and can be decomposed into a countable union of disjoint open intervals (ordered eg. by size) $\{(\alpha_n, \beta_n)\}_{n=1}^{\infty}$, so that

$$(a,b) = F \cup \bigcup_{n=1}^{\infty} (\alpha_n, \beta_n)$$

and moreover

$$\int_a^b f dm = \int_F f dm + \sum_{n=1}^{\infty} \int_{\alpha_n}^{\beta_n} f dm$$

But the hypothesis means that for any interval (α_n, β_n) we have

$$\int_{\alpha_n}^{\beta_n} f dm = \int_a^{\beta_n} f dm - \int_a^{\alpha_n} f dm = 0 - 0 = 0$$

so that $\int_F f dm > 0$ gives

$$\int_a^b f dm = \int_F f dm + \sum_{n=1}^{\infty} \int_{\alpha_n}^{\beta_n} f dm = \int_F f dm + 0 > 0$$

in contradiction to the hypothesis. □

Corollary 3. *If $f, g \in L^1([a,b])$ and $\int_a^c f dm = \int_a^c g dm$ for all $c \in [a,b]$, then $f(x) = g(x)$ for almost-every $x \in [a,b]$.*

Proof. Apply the Theorem to $\int_a^c (f-g) dm = 0$ for every $c \in [a,b]$. Thus $f(x) - g(x) = 0$ for almost every $x \in [a,b]$. □

We can now use this to prove the main Theorem of this section:

Theorem 4. *Let $f \in L^1([a,b])$, and define for each $x \in [a,b]$*

$$F(x) = \int_a^x f dm$$

Then $F'(x) = f(x)$ for almost-every $x \in [a,b]$.

Proof. We know from a previous section that F is differentiable almost-everywhere in $[a, b]$ and $F' \in L^1([a, b])$. The goal will be to show that $\int_a^c F' dm = \int_a^c f dm$ for every $c \in [a, b]$, and then deduce the conclusion of the Theorem from Corollary 3.

We will divide into three cases; consider first the case where f is bounded, that is there exists $M > 0$ such that $|f(x)| \leq M$ for all $x \in [a, b]$. Extend f by defining $f(x) = 0$ for all $x > b$, and define

$$DF_n(x) = \frac{F(x + \frac{1}{n}) - F(x)}{1/n}$$

We know that F is differentiable almost-everywhere, and thus $F'(x) = \lim_{n\to\infty} DF_n(x)$ almost-everywhere in $[a.b]$. By definition we have $DF_n(x) = n \int_x^{x+\frac{1}{n}} f dm$, and since $|f(x)| \leq M$ for every $x \in [a, b]$, we see

$$|DF_n(x)| = n \left| \int_x^{x+\frac{1}{n}} f dm \right| \leq n \int_x^{x+\frac{1}{n}} M = M$$

is uniformly bounded for all $x \in [a, b]$ and all $n \in \mathbb{N}$. Thus we may use the Bounded Convergence Theorem to deduce for any $c \in [a, b]$

$$\begin{aligned}
\int_a^c F' dm &= \lim_{n\to\infty} \int_a^c DF_n dm = \lim_{n\to\infty} n \int_a^c \left(F(x + \frac{1}{n}) - F(x) \right) dm \\
&= \lim_{n\to\infty} n \left(\int_{a+\frac{1}{n}}^{c+\frac{1}{n}} F dm - \int_a^c F dm \right) \\
&= \lim_{n\to\infty} n \int_c^{c+\frac{1}{n}} F dm - \lim_{n\to\infty} n \int_a^{a+\frac{1}{n}} F dm \\
&= F(c) - F(a) = F(c) = \int_a^c f dm
\end{aligned}$$

since F is (absolutely) continuous we have $\lim_{n\to\infty} n \int_x^{x+\frac{1}{n}} F dm = F(x)$ for any x, and the definition of F means that $F(a) = 0$.

Thus we have $\int_a^c F' dm = \int_a^c f dm$ for every $c \in [a, b]$, and we deduce from Corollary 3 that $F'(x) = f(x)$ for almost every $x \in [a, b]$.

Now we consider the case where $f(x) \geq 0$ for all $x \in [a, b]$, but not necessarily bounded. In this case we cutoff f by defining

$$f_n(x) = \min\{f(x), n\}$$

we further define $g_n(x) = f(x) - f_n(x) \geq 0$ and also

$$\begin{aligned} F_n(x) &= \int_a^x f_n dm \\ G_n(x) &= \int_a^x g_n dm \end{aligned}$$

We note that since $g_n(x) \geq 0$ for all $x \in [a,b]$, we have for $x < y$

$$G_n(y) - G_n(x) = \int_a^y g_n dm - \int_a^x g_n dm = \int_x^y g_n dm \geq 0$$

so for each fixed n, the function G_n is a monotone non-decreasing function on $[a,b]$, which in particular means that G_n is differentiable with $G_n' \geq 0$ almost-everywhere on $[a,b]$. By definition of $g_n = f - f_n$, we have $F(x) = F_n(x) + G_n(x)$, all monotone and therefore differentiable almost-everywhere, whereby for almost-every $x \in [a,b]$ we have

$$F'(x) = F_n'(x) + G_n'(x)$$

Now, by the previous case we know $F_n'(x) = f_n(x)$ almost-everywhere, since $f_n(x) \leq n$ is bounded. Moreover $G_n'(x) \geq 0$, so we have

$$F'(x) = f_n(x) + G_n'(x) \geq f_n(x)$$

But since this inequality holds for every $n \in \mathbb{N}$ we may send $n \to \infty$ and deduce $F'(x) \geq \lim_{n\to\infty} f_n(x) = f(x)$ almost everywhere, which by monotinicity of the integral implies that for every $c \in [a,b]$

$$\int_a^c F' dm \geq \int_a^c f dm = F(c) - F(a)$$

But since F is monotone non-decreasing, we have by the Lebesgue Differentiation Theorem that $\int_a^c F' dm \leq F(c) - F(a)$, so putting these together we conclude

$$\int_a^c F' dm = F(c) - F(a) = \int_a^c f dm$$

for every $c \in [a,b]$, and so again by Corollary 3 we get $F'(x) = f(x)$ for almost every $x \in [a,b]$.

To conclude, we consider general Lebesgue-integrable f on $[a,b]$, and decompose $f = f^+ - f^-$ into positive and negative parts, with

$$\begin{aligned} G(x) &= \int_a^x f^+ dm \\ H(x) &= \int_a^x f^- dm \end{aligned}$$

The second case above shows that $G'(x) = f^+(x)$ and $H'(x) = f^-(x)$ almost everywhere, and so almost-everywhere

$$F'(x) = G'(x) - H'(x) = f^+(x) - f^-(x) = f(x)$$

and we are done. □

7. $F \in AC([a,b])$ IMPLIES THAT $\int_a^b F' dm = F(b) - F(a)$

As mentioned above, this property does not hold for the Cantor function, despite being continuous and monotone (so of bounded variation)—the proof will have to use the absolute continuity in a crucial way. The key point is that an absolutely continuous function cannot have vanishing derivative almost-everywhere without being constant... this shows that an absolutely continuous function is truly determined by its derivative almost-everywhere.

Theorem 5. *Let $F \in AC([a,b])$, and assume that $F'(x) = 0$ for almost-every $x \in [a,b]$. Then F is constant on $[a,b]$.*

Proof. It is sufficient to show that $F(b) = F(a)$, since b may be replaced in the hypotheses by any $c \in [a,b]$.

Define $E = \{x \in [a,b] : F'(x) = 0\}$, and let $\epsilon > 0$, and the corresponding $\delta > 0$ in the definition of absolute continuity of F. Since $F'(x) = 0$ for every $x \in E$, there intervals $[x, x+h]$ or $[x+h, x]$ with h as small as we wish, such that

$$\left| \frac{F(x+h) - F(x)}{h} \right| < \epsilon$$

These intervals form a Vitali covering of E, and so we may select a finite collection of disjoint intervals $\{(x_k, y_k)\}_{k=1}^n$ such that

$$\sum_{k=1}^{n} (y_k - x_k) > b - a - \delta \tag{7}$$

and for each k we have

$$|F(y_k) - F(x_k)| < \epsilon(y_k - x_k) \tag{8}$$

The idea is that, on the one hand F cannot vary much over the intervals $\{(x_k, y_k)\}$, because of the condition (8) coming from the vanishing derivative. On the other hand, the remaining intervals have total length less than δ, and so absolute continuity says F cannot vary much over this residual set either; and so $F(b)$ is close to $F(a)$.

More precisely, let's define $a = y_0$ and $b = x_{n+1}$ for convenience, and decompose $[a,b]$ into the "Vitali" intervals $\{[x_k, y_k]\}_{k=1}^n$ and the residual intervals $\{(y_k, x_{k+1})\}_{k=0}^{n}$. Write

$$\begin{aligned}
&|F(b) - F(a)| \\
&= \left|F(x_{n+1}) - F(y_n) + F(y_n) - F(x_n) + F(x_n) + \cdots + F(y_1) - F(x_1) + F(x_1) - F(y_0)\right| \\
&\leq \sum_{k=1}^{n} |F(y_k) - F(x_k)| + \sum_{k=0}^{n} |F(x_{k+1}) - F(y_k)| \\
&< \epsilon(b - a - \delta) + \sum_{k=0}^{n} |F(x_{k+1}) - F(y_k)|
\end{aligned}$$

But since the residual intervals have total length

$$\sum_{k=0}^{n} (x_{k+1} - y_k) < \delta$$

the absolute continuity condition for F yields

$$\sum_{k=0}^{n} |F(x_{k+1}) - F(y_k)| < \epsilon$$

and so finally we see

$$\begin{aligned}
|F(b) - F(a)| &\leq \epsilon(b - a - \delta) + \epsilon \\
&< \epsilon(b - a + 1)
\end{aligned}$$

Let $\epsilon \to 0$ to get $|F(b) - F(a)| = 0$. □

Corollary 4. *If $F \in AC([a, b])$, then $\int_a^b F' dm = F(b) - F(a)$.*

Proof. We know that F is differentiable almost-everywhere and F' is Lebesgue-integrable, so we may define $G(x) = \int_a^x F' dm$; we wish to show that $G(x) - F(x) = -F(a)$ is a constant. We moreover know that G is absolutely continuous and $G'(x) = F'(x)$ for almost every $x \in [a, b]$, by Theorem 4. It follows from Theorem 5 that since $G(x) - F(x)$ is an absolutely continuous function with almost-everywhere-vanishing derivative, it is constant. Since $G(a) = 0$, we see that $G(x) - F(x) = -F(a)$ for all x, whereby

$$\int_a^x F' dm = G(x) = F(x) - F(a)$$

for all $x \in [a, b]$. □

8. Lebesgue Density Theorem

Let $E \subset \mathbb{R}$ be measurable, and define the **relative density** of E at $x \in \mathbb{R}$ to be

$$\lim_{\epsilon \to 0} \frac{m\left(E \cap [x-\epsilon, x+\epsilon]\right)}{m([x-\epsilon, x+\epsilon])}$$

Theorem 6 (Lebesgue Density Theorem for $\mathbb{R}$)**.** *Let $m(E) > 0$. Then for almost-every $x \in E$, we have*

$$\lim_{\epsilon \to 0} \frac{m\left(E \cap [x-\epsilon, x+\epsilon]\right)}{m([x-\epsilon, x+\epsilon])} = 1$$

Remark: It follows that the relative density of E for almost-every $x \in E^c$ is equal to 0, by applying to E^c in place of E. Thus in some versions, the Theorem is stated as "almost-every $x \in \mathbb{R}$ has relative-density either 0 or 1".

Proof. Since the result is local, we may assume E is contained in an interval $[a, b] \subset \mathbb{R}$.

Now define

$$F(x) = \int_a^x 1_E dm$$

Since 1_E is Lebesgue-integrable, F is absolutely continuous and we have $F'(x) = 1_E(x)$ for almost-every x. But this means $F'(x) = 1$ for almost every $x \in E$, which means

$$\begin{aligned} 1 = \lim_{\epsilon \to 0^+} \frac{F(x+\epsilon) - F(x)}{\epsilon} &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} m(E \cap [a, x+\epsilon]) - m(E \cap [a, x]) \\ &= \lim_{\epsilon \to 0} \frac{m(E \cap [x, x+\epsilon])}{\epsilon} \end{aligned}$$

Doing the same for $\epsilon < 0$ gives the similar limit

$$\lim_{\epsilon \to 0} \frac{m(E \cap [x-\epsilon, x])}{\epsilon} = 1$$

and so putting the two together we have

$$\begin{aligned} \lim_{\epsilon \to 0} \frac{m(E \cap [x-\epsilon, x+\epsilon])}{m([x-\epsilon, x+\epsilon])} &= \lim_{\epsilon \to 0} \frac{m(E \cap [x-\epsilon, x]) + m(E \cap [x, x+\epsilon])}{2\epsilon} \\ &= \lim_{\epsilon \to 0} \left(\frac{m(E \cap [x-\epsilon, x])}{2\epsilon} + \frac{m(E \cap [x, x+\epsilon])}{2\epsilon} \right) \\ &= \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

□

Exercise: Let $E \subset \mathbb{R}$ be a measurable set invariant under rational translations; i.e., for every $q \in \mathbb{Q}$ and $x \in E$ we have $x + q \in E \iff x \in E$.

Prove that either $m(E) = 0$ or $m(E^c) = 0$.

Exercise: For sets $A, B \subset \mathbb{R}$, define the **sumset**

$$A + B = \{a + b : a \in A, b \in B\} \subset \mathbb{R}$$

Prove that if A is measurable of positive measure $m(A) > 0$, then the set $A + A \subset \mathbb{R}$ contains an interval.