1 Differentiation of Lebesgue Integrals

Now that we have examined the Lebesgue measure theory and the integration theory that follows it, specifically L^1 and L^p spaces, we turn our attention to differentiation. Differentiation is understood as an inverse operation compared to integration and we seek to formalize the fundamental theorem of calculus in the Lebesgue setting. There are two fundamental theorems of calculus (FTC) questions that we would like to answer:

1. FTC1: If f is integrable on [a,b] and F is its indefinite integral $F(x) = \int_a^x f(y)dy$, then does this imply that F is differentiable (at least for a.e. x) and that specifically $F' = f$?

2. FTC2: What conditinos on a function F on [a,b] guarantee that the derivative exists for a.e. x, that this function F is integrable and that moreover the fundamental theorem of calculus holds:

$$F(b) - F(a) = \int_a^b F'(x)dx$$

In order to extend the FTC to general Lebesgue-measurable functions, we define the notion of a locally integrable function. That is:

A function $f : \mathbb{R}^d \to \mathbb{R}$ is locally integrable if it is Lebesgue measurable and

$$\int_K |f(x)|dx < \infty$$

That is it is integrable over a compact subset of its domain.

To write the FTC1, we usually say:

$$f(x) = \frac{d}{dx}\int_a^x f(y)dy$$

whichcan be written in terms of symmetric differences based on the definition of the derivative. Say we define $F(x) = \int_a^x f(y)dy$, then:

$$f(x) = \lim_{h\to 0}\frac{F(x+h) - F(x)}{h} = \lim_{r\to 0}\int_x^{x+r} f(y)dy$$

which can also be written as:

$$f(x) = \lim_{r\to 0^+}\frac{1}{2r}\int_{x-r}^{x+r} f(y)dy$$

with the nth dimension version just using balls with radius r and centered at $x \in \mathbb{R}^d$. This form of the FTC1 is useful because we can then make use of our notion of Lebesgue measure and the maximal functions.

1.1 Mean Value Theorem

Before diving into Lebesgue theory for differentiation, we should remind ourselves of the important Mean-Value theorem (MVT), which states:

Let $f : [a, b] \to \mathbb{R}$ be a continuous function on a closed interval [a, b] and differentiable on the open interval (a, b) with a < b. Then there exists some $c \in (a, b)$ such that:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

1.2 Differentiation of the Lebesgue Integral

To define the differentiation of the integral, we define a function f on [a,b] that is integrable. We define:

$$F(x) = \int_a^x f(y)dy, \quad a \leq x \leq b$$

We call F the indefinite integral of f. To define the differentiation of F, we recall the definition of the derivative as the limit of a quotient:

$$F'(x) = \lim_{h \to 0} \frac{F(x + h) - F(x)}{h}$$

The averaging problem in differentiation of integrals We are interested in the following problem: If f is integrable, then is it true that if you take an average of functions over a ball, as you take the limit of the measure of the ball going to zero, that the limit is achieved for a.e. x?

The answer is yes when f is continuous at x. If f is integrable on $\mathbb{R}^d$, and is continuous at x, then:

$$f(x) = \lim_{m(B) \to 0} \frac{1}{m(B)} \int_B f(y)dy = f(x)$$

Proof. (*) For every $\epsilon > 0$, there exists a $\delta > 0$ such that if $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.

If say radius of ball B is $< \delta/2$ (worst case any two points have at most distance δ), and $x \in B$, then (*) holds whenever $y \in B$ as well.

We want f(x) to be close to this "averaging" function that we are interested in. Note that:

$$|f(x) - \frac{1}{m(B)} \int_B f(y)dy| = |\frac{1}{m(B)} \int_B (f(x) - f(y)dy|$$

Then we have that:

$$|\frac{1}{m(B)}\int_B (f(x)-f(y)dy| \tag{1}$$

$$\leq \frac{1}{m(B)}\int_B |f(x)-f(y)|dy \quad \text{(By tri-ineq)} \tag{2}$$

$$\leq \frac{1}{m(B)}\epsilon m(B) = \epsilon \tag{3}$$

Since ϵ was arbitrary, then we have our equality. □

Now, we would like to answer this question for not only an x where f(x) is continous, but absolutely at **every x**. This leads us to the Hardy-Littlewood Maximal function, where we will see that the answer is yes.

Note that the definition above can be rewritten as:

$$\frac{1}{h}\int_x^{x+h} f(y)dy = \frac{1}{|I|}\int_I f(y)dy$$

where $I = (x, x+h)$ interval, and $|I|$ is the length of that interval. Essentially the above is saying that we take an **average** value of f over the interval I. Then as you take the limit as $|I|$ goes to zero, then we expect these averages to tend to f(x). Now, we want to know when does th following equation hold? (What constraints are needed on the functions and the points x)

$$\lim_{|I|\to 0,\ x\in I} \frac{1}{|I|}\int_I f(y)dy = f(x)$$

or for higher dimensions in more generality:

$$\lim_{m(B)\to 0,\ x\in B} \frac{1}{m(B)}\int_B f(y)dy = f(x)$$

for a.e. x. The limit is taken as the volume of open balls containing x (in high dimensions) goes towards zero.

1.3 Covering Lemma To Demonstrate Weak-L1 Property

In order to prove the Hardy-littlewood theorem, which will tell us that the maximal functions of an integrable function are weak-L^1, then we need the notion of a covering lemma, which states that any **finite** collection of open balls in $\mathbb{R}^d$ has a **disjoint subcollection** that can cover a fraction of the original collection. The fraction is a exponent based on the dimension of the space.

[Vitali Covering Argument] Suppose $B = \{B_1, B_2, ..., B_N\}$ is a finite collection of open balls in $\mathbb{R}^d$. Then there exists a disjoint sub-collection of $B_{i_1}, B_{i_2}, ..., B_{i_k}$ of B such that:

$$m(\bigcup_{l=1}^{N} B_l) \leq 3^d \sum_{j=1}^{k} m(B_{i_j})$$

That is we may find a disjoint sub-collection of balls that covers a fraction of the region covered by the original collection. This fraction scales exponentially as a function of the dimensionality of the space.

Proof. Suppose B and B' are two balls that intersect, with radius of B' (call it r') being less than or equal to B (call it r). Then $B' \subset \tilde{B} = \{y \in \mathbb{R}^d : |x - y| \leq 3r\}$, where $\tilde{B}$ is the ball centered at x with three-times the radius of B.

Note that we need three times the radius of B to fully cover B'. If B and B' have the same radius and are touching then any radius less then 3r would not fully cover B'.

The measure of $\tilde{B}$ by the dilation property of measure is $m(\tilde{B}) = 3^d m(B)$.

Now, we can state an iterative procedure for choosing balls based on the above: We begin by choosing B'_1 that is the ball with the largest radius from our collection. Then we remove all balls in our collection that intersect with B'_1. Then choose the next-largest ball from our remaining collection. In this way, we obtain a subcollection of balls, $\{B'_1, ..., B'_k\}$ that are disjoint by definition. Now if we expand each of these balls as done above by a factor of 3, then we get that they cover the original collection plus more.

$$\bigcup_{i=1}^{N} B_i \subset \bigcup_{i=1}^{k} \tilde{B}'_i$$

Now by monotonicity, disjoint finite additivity and dilation-properties of measure, we have:

$$m(\bigcup_{i=1}^{N} B_i) \leq m(\bigcup_{i=1}^{k} \tilde{B}'_i) = 3^d \sum_{i=1}^{k} m(B'_i)$$

□

1.4 Hardy-Littlewood Maximal Function and Weak-L1 Functions

First, we define the Weak-L1 space.

A measurable function, $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is in weak-$L^1(\mathbb{R}^d)$ if there exists a constant C_f that depends on the function f, but not on t, such that:

$$m(\{x \in \mathbb{R}^d : |f(x)| > t\}) \leq \frac{C}{t}$$

By the Chebyshev inequality, we have that for integrable functions:

$$m(\{x \in \mathbb{R}^d : |f(x)| > t\}) \leq \frac{1}{t}||f||_{L^1}$$

where C is exactly the L1 norm of f. So, weak-L1 functions are a generalization of L1 functions, where the constant term depends on the function.

If f is integrable on $\mathbb{R}^d$, then we define its maximal function f^* as:

$$f^*(x) = \sup_{x \in B} \frac{1}{m(B)} \int_B |f(y)| dy, \quad x \in \mathbb{R}^d$$

where the supremum is taken over all balls containing the point x. Note, we replace the limit in the averaging problem with a supremum and f by its absolute value. Now, one might ask what properties this f^* function has? The following theorem states that it is measurable, finite for a.e. x, and satisfies a "weak-type" inequality, where it is not "that much larger" then f(x).

We shall also see that f^* is not necessarily integrable, even if f is integrable. The inequality in the theorem part iii) is sort of the "best-substitute" for integrability.

[Properties of the Maximal Function] Suppose f is integrable on $\mathbb{R}^d$, then the following are true:

i) f^* is measurable

ii) $f^*(x) < \infty$ for a.e. x

iii) f is a weak-L1 function: $m(\{x \in \mathbb{R}^d : \; f^*(x) > \alpha\}) \leq \frac{A}{\alpha} ||f||_{L^1(\mathbb{R}^d)}$ for all $\alpha > 0$ where $A = 3^d$ and $||f|| = \int |f(x)| dx$.

Proof. f^* is measurable. We will just show that the set $E_\alpha = \{x \in \mathbb{R}^d : f^*(x) > \alpha\}$ is open. Note that if $x \in E_\alpha$, then there exists a ball centered at x with radius ϵ, such that $x \in B$ and

$$\frac{1}{m(B)} \int_B |f(y)| dy > \alpha$$

Then if you take any point $y \in \mathbb{R}^d$ such that $|y - x| < \epsilon/2$, then y will also belong to B as well and hence belongs to E_α. Therefore, E_α is open, and hence measurable. □

Proof. If we take set $E_\infty = \{x : f^*(x) = \infty\}$, then it is a subset of E_α. If we take the limit as $\alpha \to \infty$, then by the third property we have that:

$$m(E_\infty) \leq 0$$

and so must have measure zero. Therfore $f^* < \infty$ for a.e. x. □

Proof. We fix $\alpha > 0$ and let:

$$E_\alpha = \{x \in \mathbb{R}^d : f^* > \alpha\}$$

as the set of interest where f^* is bigger then our fixed α.

The measure of that set can be written as the supremum over compact subsets.

$$m(E_\alpha) = \sup\{m(K) : K \subset E_\alpha \text{ compact}\}$$

Since this is a supremum, it suffices to show that:

$$m(K) \leq \frac{C}{t} \int |f(y)| dy$$

for every compact subset K of E_α. Now that we are dealing with a compact subset, by the definition of compactness, there exists a finite subcover. Now that there is a finite subcover, we know by the Vitali Covering lemma that there exists a subcollection of balls that cover a fraction of the original subset up to a factor of 3^d.

We call our finite subcover: $\{B_1, ..., B_N\}$. Then if $x \in K$, then there is an open ball B_x centered at x such that:

$$\frac{1}{m(B_x)} \int_{B_x} |f(y)|dy > \alpha$$

by definition of the E_α set, or in other words:

$$m(B_x) < \frac{1}{\alpha} \int_{B_x} |f(y)|dy$$

By taking a finite sucollection of our subcover that are all disjoint, $\{B_1', ..., B_k'\}$, we get the series of inequalities:

$$m(K) \le \sum_{i=1}^{N} m(B_i) \tag{4}$$

$$\le 3^d \sum_{i=1}^{k} m(B_i') \tag{5}$$

$$\le \frac{3^d}{\alpha} \sum_{i=1}^{k} \int_{B_i'} |f(y)|dy \tag{6}$$

$$= \frac{3^d}{\alpha} \int_{\cup_{j=1}^{k}} B_j' |f(y)|dy \quad \text{(by def of integral)} \tag{7}$$

$$\le \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(y)|dy \quad \text{(By monotonicity)} \tag{8}$$

$\square$

1.5 Lebesgue Differentiation Theorem

Now that we have an understanding of the maximal function, we are interested in a solution to the averaging problem. The Lebesgue differentiation theorem tells us that if a function is locally integrable, then we can perform the FTC1, and take the derivative of the integral to get the function inside the integrand.

If $f \in L^1_{loc}(\mathbb{R}^d)$, then:

$$\lim_{m(B(x)) \to 0} \frac{1}{m(B)} \int_B f(y)dy = f(x)$$

for a.e. x.

Proof. □

Now that we have a theorem stating when we can differentiate an integral (i.e. when it is locally integrable), then we can define a set of points x, for which the limits exists for a suitable function f. This set is called the Lebesgue set.

If $f \in L^1_{loc}$, then a point $x \in \mathbb{R}^d$ belongs to the Lebesgue set) of f, if there exists a constant $c \in \mathbb{R}$ such that:

$$\lim_{r \to 0} \frac{1}{m(B_r(x))} \int_B |f(y) - c| dy = 0$$

That is if you have points in the Lebesgue set, then you have the results of the HWL maximal theorem. The Lebesgue set of f, does depend on which f we choose. It has some nice properties that allow us to take a wide variety of averages and recover points of the functions.

1.6 Good Kernels and Approximating the Identity

Now that we have explored the averaging problem and shown various properties of the maximal function in relation to differentiation of an integral, we want to now look at averages of functions given as convolutions. Namely:

$$(f * K_\delta)(x) = \int_{\mathbb{R}^d} f(x-y) K_\delta(y) dy$$

where f is an integrable function, which is **fixed**, while the K_δ varies over a class of fucntions, referred to as **kernels**. We are interested in convolutions and their integrals because they come up in the Fourier transform.

[Good Kernels] We specifically call K_δ **good kernels** if they are integrable and satisfy the following conditions for $\delta > 0$:

1. Normalization to 1: $\int_{\mathbb{R}^d} K_\delta(x) dx = 1$
2. Bounded L1 norm: $\int_{\mathbb{R}^d} |K_\delta(x)| dx \leq A$
3. For every $\eta > 0$, we have: $\int_{|x| \geq \eta} |K_\delta(x)| dx \to 0$ as $\delta \to 0$

[Approximations to the Identity Kernels] Restricting the class of kernels to be more narrow then "good kernels", we define approximations to the identity kernels as K_δ that are integrable and satisfy the following conditions:

1. Normalization to 1: $\int_{\mathbb{R}^d} K_\delta(x) dx = 1$
2. Bounded at every δ: $|K_\delta(x)| \leq A\delta^{-d}$ for all $\delta > 0$
3. $|K_\delta(x)| \leq A\delta / |x|^{d+1}$ for all $\delta > 0$ and $x \in \mathbb{R}^d$

Note that condition i) is the same as a good kernel, but conditions ii and iii are more restrictive then what is mentioned in good kernels. Therefore, all "approximations to the identity kernels" are also "good kernels".

1.7 Differentiation of Functions

Now, we explore the second Fundamental theorem of calculus in the view of Lebesgue measure. That is, under what general conditions does the following hold?

$$F(b) - F(a) = \int_a^b F'(x)dx$$

There are two problems related to existence and measurability.

First, off the RHS is complicated if $F'(x)$ does not exist. The exemplar is the Weierstrass function that is continuous, but nowhere differentiable. So F being continuous is not enough of a condition to guarantee the FTC2.

Second, if $F'(x)$ exists, it might not be Lebesgue measurable.

1.7.1 Bounded Variation - To Enable Differentiability Almost Everywhere

We study a class of functions that have **bounded variation** that will give us existence of the derivative. We will see that functions of bounded variation are differentiable a.e. But we will see by usage of the Cantor-lebesgue function, that it does not guarantee the validity of the FTC2. After bounded variation, we will explore this next concept of **absolute continuity** that will give us the necessary conditions for the FTC2. First we define a few keywords.

Let γ be a parametrized curve in the plane: $z(t) = (x(t), y(t))$ with $a \leq t \leq b$. x and y are continuous real valued functions supported on [a, b].

[Rectifiability of a curve] We say that the curve is rectifiable if there exists M $< \infty$ such that for any partition $P = \{t_0, t_1, ..., t_N\}$ such that $a = t_0 < t_1 < ... < t_N = b$ of [a,b]:

$$\sum_{j=1}^{N} |z(t_j) - z(t_{j-1})| \leq M$$

[Length of a curve] The length $L(\gamma)$ of curve, γ is the supremum over all partition of the sum:

$$L(\gamma) = \sup_P \sum_{j=1}^{N} |z(t_j) - z(t_{j-1})|$$

or note that $L(\gamma) = \inf_M \{\sum_{j=1}^{N} |z(t_j) - z(t_{j-1})| \leq M\}$. That is the infimum over all M's that satisfy the rectifiability of the curve.

With these two definitions, one is interested in determining when rectifiability occurs. What conditions on the functions x and y are needed?

Suppose F(t) is a complex-valued function defined on [a,b] with a partition P on the interval. We define the **variation** of F on this **particular partition** as:

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|$$

[Bounded Variation of a Function] The function F is of bounded variation if the variations of F over **all partitions** are bounded.

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| \leq M$$

for all partitions $a = t_0 < t_1 < ... < t_N = b$.

By the triangle inequality, if we take refinements of a partition P, then the variation of F on the refinement is greater then or equal to the variation of F on P. Bounded variation is a useful property as we will see in this next theorem because it guarantees curve rectifiability (a necessary **and** sufficient condition).

[Functions of Bounded Variation Form Rectifiable Curves] A curve parametrized by (x(t), y(t)) with $a \leq t \leq b$ is rectifiable if and only if x and y are functions of bounded variation.

What does it mean to b e a function of bounded variation? It means that a function cannot oscillate too often with very high amplitudes. Examples of bounded variation functions are:

1. real-valued, monotonic and bounded functions
2. differentiable with F' bounded functions

In addition to bounded variation, we also can define the total variation of a function f on [a,x] by:

$$T_F(a, x) = \sup_{P} \sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|$$

where the sup is taken over partitions of [a,x]. There is also the notion of positive and negative variation when F is real-valued. Note in total and bounded variation, F may be complex-valued, since we take the absolute value.

The positive and negative variation are:

$$P_F(a, x) = \sup \sum_{+} F(t_j) - F(t_{j-1})$$

where the sum is over all j such that $F(t_j) \geq F(t_{j-1})$.

$$N_F(a, x) = \sum \sum_{-} -(F(t_j) - F(t_{j-1}))$$

where the sum takes over all $F(t_j) \leq F(t_{j-1})$. That is positive/negative variation occurs over the "monotonic" regions of [a,x]. Positive and negative variation are useful since they require you to only look at a simplified expression. They are related to the function and the total variation as follows:

[BV Functions can be written in form of positive and negative variation (or as total variation)] Suppose F is real-valued and of bounded variation on [a,b]. Then for all $a \leq x \leq b$:

$$F(x) - F(a) = P_F(a, x) - N_F(a, x)$$

and

$$T_F(a, x) = P_F(a, x) + N_F(a, x)$$

Proof. $\square$

Next, we will see that a function of BV is equivalent to saying that the function is the difference of two bounded monotonic functions. We make use of the previous lemma specifically to create those two monotonic functions using the positive and negative variation of the function.

[BV functions are the difference of two increasing bounded functions]

Real-valued F on [a,b] is of BV if and only if F is the difference of two increasing bounded functions.

A consequence of this theorem is to see that the total variation of a function and the curve parametrized by a continuous function are closely tied. Namely that the length of the curve between two points, A and B: L(A,B) is equivalent to the $T_F(A, B)$, where F is the curve. Thus $L(A, B)$ is a continuous function of B (and of A). Thus we observe that if a function of BV is continous, then so is its total variation.

Now, we come to a result that guarantees the existence of a derivative. Namely, if a function is of BV, then it is differentiable a.e.

[F being BV means F is differentiable a.e.] If F is of BV on [a,b], then F is differentiable on [a,b] a.e.

A corollary gets us close to the FTC, but with an inequality instead of equality. Here, F is increasing and continous, then we get a derivative a.e. In addition, F will be measurable, non-negative.

If F is increasing and continous, then F' exists a.e. F' is also measurable, non-negative and obeys the following inequality:

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

If F is bounded, then F' is integrable.

In order to prove the theorem, there is a useful Rising-sun Lemma by Riesz.

[Rising-sun lemma] Suppose G is real-valued and continous on $\mathbb{R}$.

E is the set of points x such that:

$$G(x + h) > G(x)$$

for some $h = h_x > 0$. That is E is the set of points, where the function G translated to the right by a small amount is greater.

If E is non-empty, then it must be open and hence can be written as a countable disjoint union of open intervals: $E = \bigcup(a_k, b_k)$. If (a_k, b_k) is a finite interval in this union, then we have that $G(b_k) = G(a_k)$

Proof. We assume E is non-empty.

First, we show that E is open.

Since E is open, then it can be written as a countable disjoint union of open intervals.

Next, we show the equality of G at the endpoints of the disjoint intervals. ☐

Now, if we consider the case when $a = a_k$, then we get the following corollary:

G is as defined in the Rising-sun lemma 1.7.1. If $a = a_k$, then we have $G(a_k) \leq G(b_k)$.

1.7.2 Why is Bounded Variation Not Enough for Fundamental Theorem of Calculus?

From the results above, we cannot go any further. We will need a stronger notion then Bounded Variation to guarantee the integration of a derivative.

The Cantor-lebesgue function is a continuous function that is increasing and of bounded variation, but:

$$\int_a^b F'(x)dx \neq F(b) - F(a)$$

The Cantor-Lebesgue function is a sequence of continuous increasing functions $\{F_n\}$ such that:

$$|F_{n+1}(x) - F_n(x)| \leq 2^{-n-1}$$

F_n converges uniformly to a continuous limit F. Since m(C) = 0 (the measure of the Cantor set is zero), then $F'(x) = 0$ a.e.

Therefore, a function of bounded variation **guarantees the existence of a derivative a.e., but not the Fundamental theorem of calculus**. By taking a stronger notion, we will guarantee existence of derivatives a.e., and ALSO the FTC.

1.7.3 Absolute Continuity - To Enable Second Fundamental Theorem of Calculus

Absolute continuity will allow us to write:

$$F(b) - F(a) = \int_a^x F'(x)dx$$

We define it as:

[Absolute continuity] A function F on [a,b] is absolutely continuous if for any $\epsilon > 0$, there exists a $\delta > 0$ such that:

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| < \epsilon$$

whenever $\sum_{k=1}^{N} (b_k - a_k) < \delta$.

1. Absolute continuity implies continuity and uniform continuity
2. If F is absolutely continuous on a bounded interval, then it is of BV. In addition its total variation is absolutely continuous
3. If F is the antiderivative of an integrable function, then F is absolutely continuous.

Note, that absolute continuity is in essence a necessary condition for the FTC.

Finally, we obtain the FTC using absolute continuity as a necessary condition in the following theorem.

[FTC - integrating a derivative] If F is absolutely continuous on [a,b], then F' exists a.e. and is integrable. Moreover:

$$F(x) - F(a) = \int_a^x F'(y)dy$$

for all $a \leq x \leq b$

1.8 Jump Functions and Their Differentiability

We next analyze a class of monotonic functions that are not continuous. This will allow us to show that there exists derivatives a.e. even for non-continuous functions (analagous to Thm 1.7.1)

1.9 Isoperimetric Inequality and Rectifiable Curves