Real Analysis II: Study Note

Chapter 6: Differentiation Theory

Tianqi Zhang

Department of Economics

University of Southern California

Spring 2026

Contents

1 Re-proving the Fundamental Theorem of Calculus

The classical FTC has two parts:

$$\text{(I)} \quad \frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f = f(x), \qquad\qquad \text{(II)} \quad \int_a^b f' = f(b) - f(a).$$

Lebesgue's theory allows both to be proved with minimal regularity. Part I holds for any locally integrable f, almost everywhere. Part II holds whenever f is absolutely continuous. The central tools are the Lebesgue Differentiation Theorem and the theory of bounded variation.

Roadmap: LDT (Part I) → Hardy–Littlewood maximal function → Lebesgue set → integral kernels & approximations to identity → BV and AC (Part II).

2 The Lebesgue Differentiation Theorem (Part I)

Setup. Differentiation is an averaging result: $\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f = \frac{1}{h}\int_x^{x+h} f$ as $h \to 0$. Generalizing to $\mathbb{R}^d$: let B_r denote a ball of radius r, and write $B \searrow x$ to mean B is a ball containing x with $\operatorname{diam}(B) \to 0$.

Theorem 2.1: Lebesgue Differentiation Theorem

If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ (locally integrable), then for a.e. $x \in \mathbb{R}^d$:

$$\lim_{B \searrow x} \frac{1}{|B|}\int_B f(y)\,\mathrm{d}y = f(x).$$

Here the limit is over all balls B containing x with $\operatorname{diam}(B) \to 0$, and $|B|$ denotes the Lebesgue measure of B.

Remark 2.1. *When f is continuous at x, the conclusion holds trivially by continuity. The theorem is non-trivial at points where f may be discontinuous, and the exceptional set (where convergence fails) has measure zero.*

Theorem 2.2: Lebesgue Differentiation Theorem — Stronger Form

If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, then for a.e. $x \in \mathbb{R}^d$:

$$\lim_{B \searrow x} \frac{1}{|B|}\int_B |f(y) - f(x)|\,\mathrm{d}y = 0.$$

This is strictly stronger than the signed average version above.

Definition 2.3: Hardy–Littlewood Maximal Function

For $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, the **Hardy–Littlewood maximal function** is

$$f^*(x) \;:=\; \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y)| \, \mathrm{d}y,$$

where the supremum is over all balls B containing x.

Theorem 2.4: Hardy–Littlewood Maximal Inequality

There exists a constant $A_d > 0$ (depending only on dimension d) such that for all $f \in L^1(\mathbb{R}^d)$ and $\alpha > 0$:

$$m(\{f^* > \alpha\}) \;\le\; \frac{A_d}{\alpha} \|f\|_1.$$

This is a **weak-type** $(1,1)$ bound. For $p > 1$, f^* is also bounded on L^p: $\|f^*\|_p \le C_{d,p} \|f\|_p$.

Remark 2.2. *f^* is measurable. The maximal inequality is the key tool in the proof of the LDT: it controls the set where ball averages can be large, enabling the a.e. convergence argument.*

Definition 2.5: Lebesgue Set

The **Lebesgue set** of f is

$$\operatorname{Leb}(f) := \left\{ x : \lim_{B \searrow x} \frac{1}{|B|} \int_B |f(y) - f(x)| \, \mathrm{d}y = 0 \right\}.$$

Points in $\operatorname{Leb}(f)$ are called **Lebesgue points** of f. The condition is stronger than the signed average version of the LDT: not only do ball averages converge to $f(x)$, but the L^1 local oscillation around $f(x)$ vanishes.

Corollary 2.6

For $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, almost every $x \in \mathbb{R}^d$ is a Lebesgue point: $m(\mathbb{R}^d \setminus \operatorname{Leb}(f)) = 0$.

Example.

$$f(x) = \begin{cases} -1 & x < 0 \\ 1 & x \ge 0. \end{cases}$$

At $x = 0$: $\frac{1}{2h} \int_{-h}^{h} f = 0 \neq f(0) = 1$, so $0 \notin \operatorname{Leb}(f)$. At all other points f is locally constant, so $x \in \operatorname{Leb}(f)$ for $x \neq 0$.

3 Integral Kernels and Approximations to the Identity

Many classical PDEs have solutions given by **convolution** with an integral kernel $K \in L^1(\mathbb{R}^d)$:

Definition 3.1: Convolution

$$(K * f)(x) \ := \int_{\mathbb{R}^d} K(x-y)\, f(y)\, \mathrm{d}y = \int_{\mathbb{R}^d} K(y)\, f(x-y)\, \mathrm{d}y.$$

Motivation: two PDE kernels.

1) Poisson Kernel. Solve $\Delta u = 0$ on the upper half-plane $\{(x,y) \in \mathbb{R}^2 : y > 0\}$ with $u|_{y=0} = f$ ($f : \mathbb{R} \to \mathbb{R}$). The solution kernel is:

$$K_y(x) = \frac{1}{\pi} \cdot \frac{y}{x^2+y^2}, \qquad u(x,y) = (K_y * f)(x) = \int_{\mathbb{R}} \frac{y}{\pi((x-t)^2+y^2)}\, f(t)\, \mathrm{d}t.$$

Claim: As $y \to 0^+$, $K_y * f \to f$ at every Lebesgue point of f.

2) Heat Kernel. Solve $\partial_t u = \Delta u$ on $\mathbb{R}^d$ with $u(x,0) = f(x)$:

$$H_t(x) = (4\pi t)^{-d/2} \exp\left(-\frac{|x|^2}{4t}\right), \quad t > 0.$$

This is a Gaussian with variance $2t$ in each coordinate (connection to Brownian motion: H_t is the transition density of $\sqrt{2}$-scaled Brownian motion). Solution: $u(\,\cdot\,,t) = H_t * f$. The rescaling parameter is $\delta = \sqrt{4t}$, so $H_t(x) = \delta^{-d} H_1(x/\delta)$ with $\delta \to 0$ as $t \to 0$.

For fixed K, rescale by $\delta > 0$: $K_\delta(x) := \delta^{-d} K(x/\delta)$. In the Poisson kernel $\delta = y$; in the heat kernel $\delta = \sqrt{t}$.

Lemma 3.2

For f continuous and bounded, $K \in L^1(\mathbb{R}^d)$ with $K_\delta(x) = \delta^{-d} K(x/\delta)$ and $\int K = 1$, we have $K_\delta * f \to f$ pointwise as $\delta \to 0$.

Definition 3.3: Approximation to the Identity

$\{K_\delta\}_{\delta>0} \subset L^1(\mathbb{R}^d)$ is an **approximation to the identity** if:

(i) $\int K_\delta = 1$ (unit mass).

(ii) $|K_\delta(x)| \le A\delta^{-d}$ (bounded peak).

(iii) $|K_\delta(x)| \le A\delta/|x|^{d+1}$ (tail envelope decay).

When $\delta \to 0$, convolution with K_δ is just f itself (a delta function in the limit).

Theorem 3.4: Approx. to Identity Convergence

If $\{K_\delta\}$ is an approximation to the identity on $\mathbb{R}^d$ and $f \in L^1(\mathbb{R}^d)$, then

$$K_\delta * f(x) \ \longrightarrow \ f(x) \quad \text{for all } x \in \mathrm{Leb}(f), \ \text{a.e.}$$

4 The FTC Part II: Differentiation of Functions

Part II of the FTC asks: when is f recoverable from f' via $f(b) - f(a) = \int_a^b f'$? Since f' is in an integral, it only needs to exist a.e. The right condition is absolute continuity.

4.1 Bounded Variation

Definition 4.1: Bounded Variation

$f : [a, b] \to \mathbb{R}$ is of **bounded variation** (BV) if the **total variation** is finite:

$$T_f(a, b) := \sup_{\mathcal{P}} \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| < \infty,$$

where the supremum is over all Riemann partitions $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$.

Examples.

(i) If f is bounded and monotone: $T_f(a, b) = |f(b) - f(a)|$.

(ii) $f(x) = x^a \sin(1/x)$ for $x \in (0, 1]$, $f(0) = 0$: f is continuous at 0 for $a > 0$. However, f is BV on $[0, 1]$ if and only if $a > 1$. At $a = 1$ the total variation is comparable to the harmonic series $\sum 1/n = \infty$, so f is *not* BV when $a \leq 1$.

Lemma 4.2: BV and Monotone Decomposition

If $f : [a, b] \to \mathbb{R}$ is increasing, then $T_f(a, b) = f(b) - f(a)$. For a general BV function, the arc length of the graph satisfies length $= \int_a^b \sqrt{1 + (f')^2}$ whenever f is absolutely continuous, but not in general for merely BV functions (which may have jump discontinuities).

Theorem 4.3: BV Decomposition

f is BV if and only if f can be written as the difference of two bounded increasing functions.

4.2 Absolute Continuity

Definition 4.4: Absolute Continuity

$f : [a, b] \to \mathbb{R}$ is **absolutely continuous** (AC) if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any finite collection of disjoint subintervals (a_i, b_i) with $\sum_i (b_i - a_i) < \delta$:

$$\sum_i |f(b_i) - f(a_i)| \; < \; \varepsilon.$$

Remark 4.1. *AC ⇒ UC (uniform continuity). AC is strictly stronger: the Cantor function is UC but* not *AC.*

The Cantor function. Every $x \in \mathcal{C}$ has a base-3 (ternary) expansion using only digits $\{0, 2\}$. Define f on $\mathcal{C}$ by replacing each digit 2 with 1 in the ternary expansion of x, then reading the result as a base-2 (binary) number: $f(x) = \sum_j a_j/2^j$ where $a_j \in \{0, 1\}$. Extend f to $[0, 1]$ by making it constant on each deleted open interval (middle-thirds). Then $f : [0, 1] \to [0, 1]$ is continuous, non-decreasing, and surjective, but $f' = 0$ a.e. (since $m(\mathcal{C}) = 0$ and f is constant off $\mathcal{C}$), so $\int_0^1 f' = 0 \neq 1 = f(1) - f(0)$. The FTC fails: f is not AC.

Connection to integration. If $h \in L^1(a, b)$ and $f(x) = \int_a^x h(t)\,\mathrm{d}t$, then f is AC. This follows from the absolute continuity of the integral (Chapter 2): for any $\varepsilon > 0$, take δ as given by that result, and use that $\sum_i |f(b_i) - f(a_i)| = \sum_i \left|\int_{a_i}^{b_i} h\right| \leq \int_{\bigcup_i (a_i, b_i)} |h| < \varepsilon$.

Theorem 4.5: FTC Part II

$f : [a, b] \to \mathbb{R}$ is AC if and only if f is differentiable a.e., $f' \in L^1(a, b)$, and

$$f(x) - f(a) \;=\; \int_a^x f'(t)\,\mathrm{d}t \quad \text{for all } x \in [a, b].$$

In particular $\int_a^b f' = f(b) - f(a)$.

Key steps in the proof.

- **Step 1:** Every monotone increasing $f : [a, b] \to \mathbb{R}$ is differentiable a.e. (proved using Vitali covering lemma and the rising sun lemma).
- **Step 2:** Every BV function is differentiable a.e. (follows from Step 1 via the Jordan decomposition $f = f_1 - f_2$ into increasing functions).
- **Step 3:** If f is AC and $f' = 0$ a.e., then f is constant (Theorem 4.6).
- **Step 4:** For AC f, set $G(x) := f(x) - \int_a^x f'$; then G is AC and $G' = 0$ a.e., so G is constant by Step 3, giving $f(x) - f(a) = \int_a^x f'$.

Theorem 4.6

For f AC on $[a, b]$: if $f' = 0$ a.e., then f is constant.