Lecture 8: Differentiation

Tianpei Xie

Aug. 2nd., 2015

Contents

1 Differentiation theorems

- **Remark** In these notes we explore the question of the extent to which these theorems continue to hold when the differentiability or integrability conditions on the various functions F, F', f are relaxed. Among the results proven in these notes are

  1. ***The Lebesgue differentiation theorem***, which roughly speaking asserts that ***the Fundamental Theorem of Calculus*** continues to hold for almost every x if f is merely ***absolutely integrable***, rather than *continuous*;

  2. A number of *differentiation theorems*, which assert for instance that *monotone*, *Lipschitz*, or *bounded variation functions* in one dimension are ***almost everywhere differentiable***; and

  3. ***The Second Fundamental Theorem of Calculus*** for *absolutely continuous functions.*

1.1 The Lebesgue Differentiation Theorem in One Dimension

- **Theorem 1.1** *(**Lebesgue differentiation theorem, one-dimensional case**).*
  Let $f : \mathbb{R} \to \mathbb{C}$ be an ***absolutely integrable*** *function, and let $F : \mathbb{R} \to \mathbb{C}$ be the definite integral $F(x) := \int_{[-\infty,x]} f(t)dt$. Then F is* **continuous** *and* ***almost everywhere differentiable****, and $F'(x) = f(x)$ for* ***almost every*** *$x \in \mathbb{R}$.*

- **Theorem 1.2** *(**Lebesgue differentiation theorem, second formulation**).*
  Let $f : \mathbb{R} \to \mathbb{C}$ be an ***absolutely integrable*** *function. Then*
  $$\lim_{h \to 0+} \frac{1}{h} \int_{[x,x+h]} f(t)dt = f(x) \tag{1}$$
  for almost every $x \in \mathbb{R}$, and
  $$\lim_{h \to 0+} \frac{1}{h} \int_{[x-h,x]} f(t)dt = f(x) \tag{2}$$
  for almost every $x \in \mathbb{R}$.

- **Remark** (***Density Argument***) [Tao, 2011]
  The conclusion (1) we want to prove is a ***convergence theorem*** - an assertion that for all functions f in a given class (in this case, *the class of absolutely integrable functions $f : \mathbb{R} \to \mathbb{R}$*), a certain sequence of *linear expressions $T_h f$* (in this case, *the right averages $T_h f(x) = \frac{1}{h} \int_{[x,x+h]} f(t)dt$*) *converge in some sense* (in this case, pointwise almost everywhere) to a specified limit (in this case, f).

  There is a general and very useful argument to prove such convergence theorems, known as ***the density argument***. This argument requires ***two ingredients***, which we state informally as follows:

  1. A ***verification*** of the convergence result for some "***dense subclass***" of "***nice***" functions f, such as *continuous functions*, *smooth functions*, *simple functions*, etc.. By "*dense*", we mean that a *general function f* in the *original class* can be ***approximated to arbitrary accuracy*** in a suitable sense by a function *in the nice subclass.*

2. A ***quantitative estimate*** that ***upper bounds*** *the* ***maximal fluctuation*** of the *linear expressions* $T_h f$ in terms of the "***size***" of the function f (where *the precise definition of "size" depends on the nature of the approximation* in the first ingredient).

Once one has these two ingredients, it is usually not too hard to put them together to obtain the desired convergence theorem for general functions f (*not just those in the dense subclass*).

- **Proposition 1.3** *(**Translation is continuous** in L^1). [Tao, 2011]*
Let $f : \mathbb{R}^d \to \mathbb{C}$ be an absolutely integrable function, and for each $h \in \mathbb{R}^d$, let $f_h : \mathbb{R}^d \to \mathbb{C}$ be ***the shifted function*** *$f_h(x) := f(x-h)$. Then f_h converges in L^1 norm to f as $h \to 0$, thus*
$$\lim_{h \to 0} \int_{\mathbb{R}^d} |f_h(x) - f(x)|\, dx = 0.$$

- **Exercise 1.4** *Let $f, g : \mathbb{R}^d \to \mathbb{C}$ be Lebesgue measurable functions such that f is* ***absolutely integrable*** *and g is* ***essentially bounded*** *(i.e.* ***bounded*** *outside of a null set). Show that* ***the convolution*** *$f * g : \mathbb{R}^d \to \mathbb{C}$ defined by the formula*
$$f * g(x) = \int_{\mathbb{R}^d} f(y)g(x-y)dy$$
is well-defined (in the sense that the integrand on the right-hand side is absolutely integrable) and that $f * g$ is a ***bounded, continuous*** *function.*

- **Remark** One drawback with ***the density argument*** is it gives convergence results which are ***qualitative*** rather than ***quantitative*** - there is no explicit bound on the rate of convergence.

1.2 The Lebesgue Differentiation Theorem in $\mathbb{R}^d$

1.2.1 Absolute Integrable Version

- **Theorem 1.5** *(**Lebesgue Differentiation Theorem (Absolute Integrable version)**) [Tao, 2011]*
Suppose $f : \mathbb{R}^d \to \mathbb{C}$ is ***absolutly integrable****. Then for almost every x, we have*
$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(z) - f(x)|\, dz = 0 \tag{3}$$
$$\text{and} \quad \lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(z)dz = f(x),$$
where $B(x,r) := \{y \in \mathbb{R}^d : \|x - y\| < r\}$ is the open ball of radius r centred at x.

- **Definition** A point x for which (3) holds is called ***a Lebesgue point*** of f; thus, for an ***absolutely integrable function*** f, *almost every point in $\mathbb{R}^d$ will be a Lebesgue point for $\mathbb{R}^d$*.

- The ***quantitative estimate*** we will need is *the Hardy-Littlewood maximal inequality*. First, we need to introduce *the Hardy-Littlewood maximal function*:

  Definition [Folland, 2013]
  If $f \in L^1_{loc}(\mathbb{R}^d)$, the ***Hardy-Littlewood maximal function*** $Hf(x)$ is defined as
$$Hf(x) \equiv \sup_{r>0} \frac{1}{m\,(B(r,x))} \int_{B(r,x)} |f(z)|\, dz$$

where $B(r,x) = \{y : \|y - x\| < r\}$, and the ***average value*** of f on $B(r,x)$ is

$$A_r f(x) = \frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} f(z) dz.$$

- **Remark** A useful variant of $Hf(x)$ (see [Stein and Shakarchi, 2009]) as

$$H^* f(x) \equiv \sup \left\{ \frac{1}{m\left(B\right)} \int_{B} |f(z)|\, dz,\;\; B \text{ is a ball }, x \in B \right\}.$$

- **Remark** *The Hardy-Littlewood maximal function* is an important function in the field of *(real-variable) harmonic analysis.*
- **Remark** The Hardy-Littlewood maximal function has the following properties:
  1. $(Hf)^{-1}(a,\infty) = \bigcup_{r>0} (A_r f)^{-1}(a,\infty)$ is open for any $a \in \mathbb{R}$, so the Hardy-Littlewood maximal function is *measureable.*
  2. Moreover, $Hf(x) < \infty, a.e. x$ is ***essentially bounded***.
  3. Note that $Hf \le H^* f \le 2^d Hf$
- We need to prove the following theorem:

  Theorem 1.6 *(**The Hardy-Littlewood Maximal Theorem**) [Stein and Shakarchi, 2009, Folland, 2013]*
  Suppose f is integrable, then

  1. $$H^* f(x) \equiv \sup \left\{ \frac{1}{m\left(B\right)} \int_{B} |f(z)|\, dz,\;\; B \textit{ is a ball }, x \in B \right\}.$$

     is measurable.
  2. $H^* f(x) < \infty$ *for a.e.* x.
  3. $H^* f$ *satisfies **the Hardy-Littlewood maximal inequality**:*

     $$m\left(\{x :\, H^* f(x) > \alpha\}\right) \le \frac{A}{\alpha} \|f\|_{L^1(\mathbb{R}^d)}$$

     for $\alpha > 0$, where $A = 3^d$, and $\|f\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f(x)|\, dx$.

  *Note that $H^* f \ge |f|, a.e. x$, but the above expression indicates that $H^* f$ **is not much larger than** $|f|$. However, we may not be able to assume $H^* f$ integrable for any f.*
- **Remark** In order to prove this theorem, we need to introduce concept of ***Vitali covering***:
  - **Definition** (***Vitali Covering***) [Royden and Fitzpatrick, 1988, Stein and Shakarchi, 2009]
    A collection $\mathcal{B}$ of balls $\{B\}$ is said to be a ***Vitali covering*** of a set E, (***covers** E **in Vitali sense***,) if for every $x \in E$, any $\eta > 0$, there is a *ball* $B \in \mathcal{B}$, such that $x \in B$ and $m(B) < \eta$. Thus *every point is covered by* **balls** *of* ***arbitrary small measure***.

– **Lemma 1.7** ***(Lebesgue number lemma)***
*For any **open covering** $\mathcal{A}$ of the **metric** space (X, d). If X is **compact**, there exists a number $\delta > 0$ such that for **any subset** of X having **diameter** $< \delta$, there exists an element of $\mathcal{A}$ containing it.*

– **Lemma 1.8** ***(Vitali Covering Lemma in elementary form)*** *[Stein and Shakarchi, 2009]*
Suppose $\mathcal{B} \equiv \{B_1, \ldots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^d$. Then there exists a disjoint sub-collection $B_{i_1}, B_{i_2}, \ldots, B_{i_k}$ of $\mathcal{B}$ that satisfies

$$m\left(\bigcup_{s=1}^{N} B_s\right) \le 3^d \sum_{j=1}^{k} m(B_{i_j})$$

*Loosely speaking, we may always find **a disjoint sub-collection of balls** that covers **a fraction of the region** covered by the original collection of balls.*

Proof: Observe that B and B' is a pair of balls that intersects, with the radius of B' being not greater than that of B. Then B' is contained in a ball $\tilde{B}$ that is concentric with B but with 3 times its radius.

First, we pick a ball B_{i_1} in $\mathcal{B}$ with *maximal* (largest) radius, and then delete it from $\mathcal{B}$ as well as any ball that intersect with B_{i_1}. Thus all deleted balls are contained in the ball $\tilde{B}_{i_1}$ concentric with B_{i_1} with 3-times its radius.

Then the remaining balls yield a new collection $\mathcal{B}'$, from which we could repeat the above procedure. After at most N steps, we obtain a collection of disjoint balls $\tilde{B}_{i_1}, \tilde{B}_{i_2}, \ldots, \tilde{B}_{i_k}$.

Finally, we need to prove that the disjoint balls satisfies the above inequality. We use the observation made at the beginning of the proof. Let $\tilde{B}_{i_j}$ be the ball that is concentric to B_{i_j}, but with 3-times its radius. Since any ball B in $\mathcal{B}$ must a ball B_{i_j} and have equal or smaller radius than B_{i_j}, we must have $B \subset \tilde{B}_{i_j}$, thus

$$\begin{aligned}
m\left(\bigcup_{s=1}^{N} B_s\right) &\le m\left(\bigcup_{j=1}^{k} \tilde{B}_{i_j}\right) \\
&\le \sum_{j=1}^{k} m\left(\tilde{B}_{i_j}\right) \\
&= 3^d \sum_{j=1}^{k} m\left(B_{i_j}\right)
\end{aligned}$$

In last equality, we use the fact that in $\mathbb{R}^d$ a dilation of a set by $\delta > 0$ results in the multiplication of δ^d of the Lebesgue measure. ∎

– We will use the following *Vitali Covering Lemma* to prove *the Hardy-Littlewood maximal theorem*:

Lemma 1.9 ***(Vitali Covering Lemma in general)*** *[Stein and Shakarchi, 2009, Folland, 2013]*
Suppose E is a set of finite measure and $\mathcal{B}$ is a Vitali covering of E. For any $\delta > 0$, we

can find finitely many balls $B_1, \ldots, B_N$ in $\mathcal{B}$ that are disjoint and so that

$$\sum_{i=1}^{N} m(B_i) \geq m(E) - \delta$$

Proof: We can apply the elementary lemman above iteratively, with the aim of exhausting the set E. It suffice to take δ sufficiently small, say $\delta < m(E)$, and using the just cited covering lemma, we can find an initial collection of disjoint balls $B_1, \ldots, B_{N_1}$ in $\mathcal{B}$ such that $\sum_{i=1}^{N_1} m(B_i) \geq \gamma\delta$, where $\gamma = 3^{-d}$.

Indeed, first we have $m(E') \geq \delta$ for an appropriate compact subset E' of E. Because of the compactness, we can cover E' with finitely many balls from $\mathcal{B}$, and then the previous lemma allows us to select a disjoint sub-collection of these balls $B_1, B_2, \ldots, B_{N_1}$ such that $\sum_{i=1}^{N_1} m(B_i) \geq \gamma\, m(E') \geq \gamma\delta$.

With $B_1, \ldots, B_{N_1}$ as our initial sequence of balls, we consider two possibilities: either $\sum_{i=1}^{N_1} m(B_i) \geq m(E) - \delta$ and we are done with $N = N_1$; or, contrariwise, $\sum_{i=1}^{N_1} m(B_i) < m(E) - \delta$. In the second case, with $E_2 = E - \bigcup_{i=1}^{N_1} \overline{B}_i$ so that $m(E_2) > \delta$. We then repeat the above procedure, by choosing a compact subset E_2' of E_2 with $m(E_2') > \delta$, and by noting that the balls in $\mathcal{B}$ that are disjoint from $\bigcup_{i=1}^{N_1} \overline{B}_i$ still cover E_2 and in fact gives a Vitali covering of E_2, and hence for E_2'. Then we can choose $B_i, N_1 < i \leq N_2$ so that $\sum_{i=N_1}^{N_2} m(B_i) \geq \gamma\delta$. Therefore, now $\sum_{i=1}^{N_2} m(B_i) \geq 2\gamma\delta$ and balls $B_i, 1 \leq i \leq N_2$ are disjoint.

Again we consider whether or not $\sum_{i=1}^{N_2} m(B_i) \geq m(E) - \delta$: if it is $N = N_2$; otherwise repeat for $E_3 = E - \bigcup_{i=1}^{N_2} \overline{B}_i$. If $k \geq (m(E) - \delta)/(\gamma\delta)$, then after at most k-steps, we should have selected a subcollection of disjoint balls $B_i, 1 \leq i \leq N_k$ with its sum of measures $\geq k\gamma\delta$. Thus

$$\sum_{i=1}^{N_k} m(B_i) \geq m(E) - \delta,$$

which completes our proof. ∎

– **Corollary 1.10** *[Stein and Shakarchi, 2009, Royden and Fitzpatrick, 1988]*
Follwing the setting above, we can arrange the choice of balls so that

$$m\left(E - \bigcup_{i=1}^{N} B_i\right) < 2\delta$$

Proof: Let $O \supset E$ be an open set that contains E with $m(O - E) < \delta$. We then choose balls that are contained in O. Then $(E - \bigcup_{i=1}^{N} B_i) \cup \bigcup_{i=1}^{N} B_i \subset O$.

$$\begin{aligned} m\left(E - \bigcup_{i=1}^{N} B_i\right) &\leq m(O) - m\left(\bigcup_{i=1}^{N} B_i\right) \\ &\leq m(E) + \delta - (m(E) - \delta) = 2\delta. \end{aligned}$$ ∎

- Now we turn to the proof of Theorem 1.6

Proof: 1. Note that the set $E_\alpha = \{x : H^* f(x) > \alpha\}$ is open, because if $\overline{x} \in E_\alpha$, then there exists a ball B such that $\overline{x} \in B$ and

$$\frac{1}{m(B)} \int_B |f(z)|\, dz > \alpha.$$

Now any x close enough to $\overline{x}$ will also belong to B; hence $x \in E_\alpha$ as well.

2. Follows from the fact that $\{x : H^* f(x) = \infty\} = \bigcap_{\alpha=1}^{\infty} \{x : H^* f(x) > \alpha\}$. So the measure approaches to 0 as $\alpha \to \infty$.

3. Use ***the Vitali covering lemma***. Let $E_\alpha = \{x : H^* f(x) > \alpha\}$. For each $x \in E_\alpha$, there exists a ball B_x that contains x, and

$$\frac{1}{m(B_x)} \int_{B_x} |f(z)|\, dz > \alpha.$$

So for each ball

$$m(B_x) < \frac{1}{\alpha} \int_{B_x} |f(z)|\, dz.$$

Fix a compact subset K of E_α. Since K is covered by $\bigcup_{x \in E_\alpha} B_x$, we have a finite subcover $K \subset \bigcup_{m=1}^{N} B_m$. The covering lemma guarantees the existence of a subcollection $B_{i_1}, \ldots, B_{i_k}$ of disjoint balls with

$$m\left(\bigcup_{m=1}^{N} B_m\right) \le 3^d \sum_{j=1}^{k} m(B_{i_j}).$$

Then

$$\begin{aligned} m(K) \le m\left(\bigcup_{m=1}^{N} B_m\right) &\le 3^d \sum_{j=1}^{k} m(B_{i_j}) \\ &\le \frac{3^d}{\alpha} \sum_{j=1}^{k} \int_{B_{i_j}} |f(z)|\, dz \\ &= \frac{3^d}{\alpha} \int_{\bigcup_{j=1}^{k} B_{i_j}} |f(z)|\, dz \quad (B_{i_j} \text{ are disjoint}) \\ &\le \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(z)|\, dz \end{aligned}$$

Since the above inequality holds for all $K \subset E_\alpha$ compact, we use the inner regularity of Lebesgue measure,

$$m(E_\alpha) = \sup_{K \subset E_\alpha,\ \text{compact}} m(K) \le \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(z)|\, dz. \quad \blacksquare$$

1.2.2 Local Integrable Version

- **Definition** [Stein and Shakarchi, 2009]
  A *measurable function* f on $\mathbb{R}^d$ is ***locally integrable***, i.e. $f \in L^1_{loc}(\mathbb{R}^d)$, if for every ball B the function $f(x)\mathbb{1}_B$ is *integrable*.

- This theorem follows from *the Hardy-Littlewood maximal inequality*

  Theorem 1.11 *[Stein and Shakarchi, 2009]*
  If $f \in L^1_{loc}(\mathbb{R}^d)$ is ***locally integrable****, then for the* ***average*** *of f, i.e.*

  $$A_r f(x) = \frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} f(z)dz,$$

  we have

  $$A_r f(x) \stackrel{a.e.}{\to} f(x), \quad r \to 0.$$

  Proof: It suffices to show that for $N \in \mathbb{N}$, $A_r f(x) \to f(x)$ for $a.e.x$ with $|x| \le N$. But for $|x| \le N$ and $r \le 1$ the values $A_r f(x)$ depend only on values $f(z)$ for $|z| \le N+1$, so by replacing f with $f\mathbb{1}\left\{B(N+1,0)\right\}$ we may assume that that $f \in L^1$.

  Given $\epsilon > 0$, we know that there exists a continuous integrable function g such that $\int |g(z) - f(z)|\, dz < \epsilon$. Continuity of g implies that for every $x \in \mathbb{R}^d$ and $\delta > 0$, there exists $r > 0$ such that $|g(y) - g(z)| < \delta$ whenever $|y - x| < r$, and hence,

  $$|A_r g(x) - g(x)| = \frac{1}{m\left(B(r,x)\right)} \left| \int_{B(r,x)} \left[g(z) - g(x)\right] dz \right| < \delta.$$

  Therefore $A_r g(x) \to g(x)$ as $r \to 0$ for every x, so

  $$\begin{aligned}\limsup_{r\to 0} |A_r f(x) - f(x)| &= \limsup_{r\to 0} |A_r(f-g)(x) + A_r g(x) - g(x) + g(x) - f(x)| \\ &\le H(f-g)(x) + 0 + |f-g|(x).\end{aligned}$$

  Hence, if

  $$\begin{aligned}F_\alpha &= \left\{x : \limsup_{r\to 0} |A_r f(x) - f(x)| > \alpha\right\}, \\ G_\alpha &= \left\{x : |f-g|(x) > \alpha\right\},\end{aligned}$$

  we have

  $$F_\alpha \subset G_{\alpha/2} \cup \left\{x : H(f-g) > \alpha/2\right\}.$$

  But $\frac{\alpha}{2} m(G_{\alpha/2}) \le \int_{G_{\alpha/2}} |f(x) - g(x)|\, dx < \epsilon$, so by the maximal theorem,

  $$m(F_\alpha) \le \frac{2\epsilon}{\alpha} + \frac{2A\epsilon}{\alpha}.$$

  Since $\epsilon > 0$ is arbitrary, $m(F_\alpha) = 0$ for all $\alpha > 0$. But $\lim_{r\to 0} A_r f(x) = f(x)$ for all $x \in \bigcup_{n=1}^{\infty} E_{1/n}$, so we have done. ∎

- **Definition** [Stein and Shakarchi, 2009]
  If $f \in L^1_{loc}(\mathbb{R}^d)$, the ***Lebesgue set*** of f consists of all points $\overline{x} \in \mathbb{R}^d$ for which $f(\overline{x})$ is ***finite*** and

  $$\lim_{\substack{m(B)\to 0 \\ \overline{x}\in B}} \frac{1}{m\left(B\right)} \int_B |f(z) - f(\overline{x})|\, dz = 0.$$

or *equivalently*, [Folland, 2013],

$$Lf \equiv \left\{ x \in \mathbb{R}^d : \lim_{r \to 0} \frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} |f(z) - f(x)|\, dz = 0 \right\}.$$

- **Corollary 1.12** *Suppose E is a measureable set in $\mathbb{R}^d$. Then*

  1. *Almost every $x \in E$ is* ***a point of Lebesgue density*** *of E;*
  2. *Almost every $x \notin E$ is* ***not a point of Lebesgue density*** *of E.*

- **Corollary 1.13** *If f is* ***locally integrable*** *on $\mathbb{R}^d$, then* ***almost every point*** *belongs to* ***the Lebesgue set*** *of f.*

  Proof: Apply the theorem above to $|f(z) - q|$ shows that for each rational q, there exists a set E_q of measure 0 such that

  $$\lim_{r \to 0} \frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} |f(z) - q|\, dz = |f(x) - q|$$

  for all $x \notin E_q$.

  If $E = \bigcup_{q \in \mathbb{Q}} E_q$, then $m(E) = 0$. Now suppose that $x \notin E$. Given $\epsilon > 0$, there exists a rational q such that $|f(x) - q| < \epsilon$. Since

  $$\frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} |f(z) - f(x)|\, dz \le \frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} |f(z) - q|\, dz + |f(x) - q|,$$

  we must have

  $$\limsup_{r \to 0} \frac{1}{m\left(B(r,x)\right)} \int_{B(r,x)} |f(z) - f(x)|\, dz \le 2\epsilon,$$

  and thus x is in the Lebesgue set of f. ∎

- **Definition** A collection of sets $\{U_\alpha\}$ is said to ***shrink regularly*** to $\overline{x}$ or has ***bounded eccentricity*** at $\overline{x}$ if there is a constant $c > 0$ such that for each U_α there is a ball B with

  $$\overline{x} \in B, \qquad U_\alpha \subset B, \qquad m(U_\alpha) \ge c\, m(B).$$

- **Theorem 1.14** ***(Lebesgue Differentiation Theorem (Local Integrable version))*** *[Stein and Shakarchi, 2009, Folland, 2013]*
  Suppose f is ***locally integrable*** *on $\mathbb{R}^d$. For every x in the Lebesgue set of f, i.e. for almost every x, we have*

  $$\lim_{\substack{m(U_\alpha) \to 0 \\ x \in U_\alpha}} \frac{1}{m(U_\alpha)} \int_{U_\alpha} |f(z) - f(x)|\, dz = 0$$

  $$\text{and} \quad \lim_{\substack{m(U_\alpha) \to 0 \\ x \in U_\alpha}} \frac{1}{m(U_\alpha)} \int_{U_\alpha} f(z) dz = f(x),$$

  for every family $\{U_\alpha\}$ that ***shrinks regularly*** *to x.*

  Proof: See that if $x \in B$ with $U_\alpha \subset B$ and $m(U_\alpha) \ge c m(B)$, then

  $$\frac{1}{m(U_\alpha)} \int_{U_\alpha} |f(z) - f(x)|\, dz \le \frac{1}{c\, m(B)} \int_B |f(z) - f(x)|\, dz < \epsilon$$

  which follows from the fact that x is in the Lebesgue set of f. ∎

1.3 Lebesgue Density and Radon-Nikodym Derivative

- Now we turn to consequences of *the Lebesgue differentiation theorem.*

  Definition [Stein and Shakarchi, 2009]
  If E is a measureable set in $\mathbb{R}^d$, $x \in \mathbb{R}^d$ is a ***point of Lebesgue density*** of E if

  $$\lim_{\substack{m(B)\to 0\\ x\in B}} \frac{m(B\cap E)}{m(B)} = 1.$$

  Loosely speaking, it says that a small ball contains x are *almost entirely covered by* E. Then for any $\alpha < 1$ close to 1, and *every ball of* ***sufficiently small radius*** containing x, we have

  $$m(E \cap B) \geq \alpha\, m(B).$$

- **Definition** A Borel *measure* ν on $\mathbb{R}^d$ will be called ***regular*** if

  1. $\nu(K) < \infty$ for every ***compact*** K;
  2. $\nu(E) = \inf\{\nu(U) : U \text{ open}, E \subseteq U\}$ for every $E \in \mathcal{B}[\mathbb{R}^d]$.

  (Condition (2) is actually implied by condition (1). *A* ***signed*** *or* ***complex*** *Borel measure* ν will be called ***regular*** if $|\nu|$ is *regular.*

- **Theorem 1.15** ***(Lebesgue Density from Radon-Nikodym derivative)*** *[Folland, 2013]* *Let ν be a* ***regular signed measure*** *on $\mathbb{R}^d$, and let $d\nu = d\lambda + f dm$ be its Lebesgue-Radon-Nikodym decomposition, where $\lambda \perp m$. Then for m-almost every $x \in \mathbb{R}^d$,*

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

  where E_r ***shrinks regularly*** *to x.*

  Proof: It is easy to verify that $d\,|v| = d\,|\lambda| + f dm$, so the regularity of ν implies the regularity of λ and $f dm$. In particular, $f \in L^1_{loc}$. From differentiation theorem, for m-almost every $x \in \mathbb{R}^d$,

  $$\lim_{r\to 0} \frac{\int \mathbb{1}\{E_r\} f dm}{m(E_r)} = f(x).$$

  Then it suffice to show that for $\lambda \perp m$, λ regular,

  $$\lim_{r\to 0} \frac{\lambda(E_r)}{m(E_r)} = 0.$$

  Note that $E_r \subset B(r,x)$ with $m(E_r) \geq cB(r,x)$, so

  $$\frac{\lambda(E_r)}{m(E_r)} \leq \frac{|\lambda|(E_r)}{m(E_r)} \leq \frac{|\lambda|(B(r,x))}{m(E_r)} \leq \frac{|\lambda|(B(r,x))}{cm(B(r,x))}.$$

  Therefore, it suffice to assume that λ is positive measure, with E_r replaced by $B(r,x)$. Let A be the Borel set that $\lambda(A) = m(A^c) = 0$, and

  $$F_k \equiv \left\{x \in A : \limsup_{r\to 0} \frac{\lambda(B(r,x))}{m(B(r,x))} \geq \frac{1}{k}\right\}.$$

It suffice to show that $m(F_k) = 0$ for all k.

By regularity of λ, there is an open $U_\epsilon \supset A$ such that $\lambda(U_\epsilon) < \epsilon$. Each $x \in F_k$ is the center of a ball $B_x \subset U_\epsilon$ such that $\lambda(B_x) > \frac{1}{k} m(B_x)$. Let $V_\epsilon = \bigcup_{x \in F_k} B_x$ and $c < m(V_\epsilon)$, there exists $x_1, \ldots, x_J$ such that $B_{x_1}, \ldots, B_{x_J}$ are disjoint and

$$c < m(V_\epsilon) \le 3^d \sum_{k=1}^{J} m(B_{x_k}) \le 3^d k \sum_{k=1}^{J} \lambda(B_{x_k}) \le 3^d k \lambda(V_\epsilon) \le 3^d k \lambda(U_\epsilon) \le 3^d k \epsilon.$$

So $m(V_\epsilon) \le 3^d k \epsilon$, and since $F_k \subset V_\epsilon$, ϵ is arbitrary, then $m(F_k) = 0$. ∎

2 The Fundamental Theorem of Calculus for Lebesgue Integral

2.1 Functions of Bounded Variations

- **Theorem 2.1** *(**Monotone Differentiation Theorem**). [Tao, 2011]*
  *Any function $F : \mathbb{R} \to \mathbb{R}$ which is **monotone** (either monotone non-decreasing or monotone non-increasing) is **differentiable almost everywhere**.*

- **Definition** (***Jump function***). [Tao, 2011]
  *A **basic jump function** J* is a function of the form

  $$J(x) := \begin{cases} 0 & \text{when } x < x_0 \\ \theta & \text{when } x = x_0 \\ 1 & \text{when } x > x_0 \end{cases}$$

  for some real numbers $x_0 \in \mathbb{R}$ and $0 \le \theta \le 1$; we call x_0 ***the point of discontinuity*** for J and θ ***the fraction***. Observe that such functions are ***monotone non-decreasing***, but have a ***discontinuity*** at one point.

  *A **jump function*** is any ***absolutely convergent*** *combination of basic jump functions*, i.e. a function of the form $F = \sum_n c_n J_n$, where n ranges over an *at most countable set*, each J_n is *a basic jump function*, and the c_n are ***positive*** *reals* with $\sum_n c_n < \infty$. If there are *only **finitely many*** n involved, we say that F is a ***piecewise constant jump function***.

  Example If $q_1, q_2, q_3, \ldots$ is any enumeration of the *rationals*, then $\sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{[q_n, +\infty)}$ is *a jump function*.

- **Remark** ***All jump functions are monotone non-decreasing***.

  From the absolute convergence of the c_n we see that ***every jump function is the uniform limit of piecewise constant jump functions***, for instance $\sum_{n=1}^{\infty} c_n J_n$ is the uniform limit of $\sum_{n=1}^{N} c_n J_n$. One consequence of this is that the *points of discontinuity of a jump function* $\sum_{n=1}^{\infty} c_n J_n$ are *precisely those of the individual summands* $c_n J_n$, i.e. *of the points x_n where each J_n jumps*.

- The key fact is that *these **Jump functions***, together with ***the continuous monotone functions***, ***essentially generate all monotone functions***, at least *in the bounded case*:

  Lemma 2.2 *(**Continuous-singular decomposition for monotone functions**).*
  *Let $F : \mathbb{R} \to \mathbb{R}$ be a **monotone non-decreasing** function.*

1. *The only **discontinuities** of F are **jump discontinuities**. More precisely, if x is a point where F is discontinuous, then the limits $\lim_{y\to x^-} F(y)$ and $\lim_{y\to x^+} F(y)$ both exist, but are **unequal**, with $\lim_{y\to x^-} F(y) < \lim_{y\to x^+} F(y)$.*

2. *There are at most **countably** many **discontinuities** of F.*

3. *If F is **bounded**, then F can be expressed as the **sum** of a **continuous monotone non-decreasing function** F_c and a **jump function** F_{pp}.*

- **Exercise 2.3** *Show that the decomposition of a bounded monotone non-decreasing function F into continuous F_c and jump components F_{pp} given by the above lemma is unique.*

- **Remark** Just as the *integration theory* of *unsigned functions* can be used to develop *the integration theory* of the *absolutely convergent functions*, the ***differentiation theory*** of ***monotone functions*** can be used to develop a parallel *differentiation theory* for the class of ***functions of bounded variation***:

- **Definition** (***Bounded variation***).
Let $F : \mathbb{R} \to \mathbb{R}$ be a function. ***The total variation*** $\|F\|_{TV(\mathbb{R})}$ (or $\|F\|_{TV}$ for short) of F is defined to be the ***supremum***

$$\|F\|_{TV(\mathbb{R})} := \sup_{x_0 < \ldots < x_n} \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})|$$

where the supremum ranges over all ***finite increasing sequences*** $x_0, \ldots, x_n$ of real numbers with $n \geq 0$; this is a quantity in $[0, +\infty]$. We say that ***F has bounded variation (on $\mathbb{R}$)*** if $\|F\|_{TV(\mathbb{R})}$ is ***finite***. (In this case, $\|F\|_{TV(\mathbb{R})}$ is often written as $\|F\|_{BV(\mathbb{R})}$ or just $\|F\|_{BV}$.)

- **Remark** Given any ***interval*** $[a, b]$, we define ***the total variation*** $\|F\|_{TV([a,b])}$ of F on $[a, b]$ as

$$\|F\|_{TV(\mathbb{R})} := \sup_{a \leq x_0 < \ldots < x_n \leq b} \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})|$$

We say that a function F has ***bounded variation on*** $[a, b]$ if $\|F\|_{BV([a,b])}$ is *finite*. Note that $\|F\|_{TV(\mathbb{R})} = \lim_{N\to\infty} \|F\|_{TV([-N,N])}$.

- **Proposition 2.4** *If $F : \mathbb{R} \to \mathbb{R}$ is a **monotone function**, $\|F\|_{TV([a,b])} = |F(b) - F(a)|$ for any interval $[a, b]$. Thus F has **bounded variation** on $\mathbb{R}$ if and only if it is **bounded**.*

- **Proposition 2.5** *For any functions $F, G : \mathbb{R} \to \mathbb{R}$, the total variation $\|\cdot\|_{TV(\mathbb{R})}$ satisfies the following property:*

  1. *(**Non-Negativity**): $\|F\|_{TV(\mathbb{R})} \geq 0$;*
  2. *(**Positive Definiteness**): $\|F\|_{TV(\mathbb{R})} = 0$ if and only if F is constant.*
  3. *(**Homogeneity**): $\|c\,F\|_{TV(\mathbb{R})} = |c|\ \|F\|_{TV(\mathbb{R})}$ for any $c \in \mathbb{R}$.*
  4. *(**Triangle Inequality**): $\|F + G\|_{TV(\mathbb{R})} \leq \|F\|_{TV(\mathbb{R})} + \|G\|_{TV(\mathbb{R})}$*

  *Thus $\|\cdot\|_{TV(\mathbb{R})}$ is a **norm**.*

- **Exercise 2.6** *(Bounded Variation is **Stronger** than Bounded)*

  1. *Show that every function $f : \mathbb{R} \to \mathbb{R}$ of **bounded variation** is **bounded**, and that the limits $\lim_{x\to+\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$, are well-defined.*

2. *Give a counterexample of a* ***bounded, continuous, compactly supported*** *function f that is* ***not*** *of* ***bounded variation.***

- **Proposition 2.7** *A function $F : \mathbb{R} \to \mathbb{R}$ is of* ***bounded variation*** *if and only if it is the* ***difference*** *of two* ***bounded monotone functions.***

- **Remark** Much as *an absolutely integrable function* can be expressed as the difference of its *positive* and *negative parts*, ***a bounded variation function*** can be expressed as ***the difference of two bounded monotone functions***. Let

$$
\begin{aligned}
F^{+}(x) &= \sup_{x_0 < \ldots < x_n \le x} \sum_{i=1}^{n} \max\{F(x_i) - F(x_{i-1}), 0\} \\
F^{-}(x) &= \sup_{x_0 < \ldots < x_n \le x} \sum_{i=1}^{n} \max\{-F(x_i) + F(x_{i-1}), 0\}
\end{aligned}
$$

We have

$$
\begin{aligned}
F(x) &= F(-\infty) + F^{+}(x) - F^{-}(x) \\
\|F\|_{TV([a,b])} &= F^{+}(b) - F^{+}(a) + F^{-}(b) - F^{-}(a) \\
\|F\|_{TV(\mathbb{R})} &= F^{+}(+\infty) + F^{-}(+\infty)
\end{aligned}
$$

for every interval $[a, b]$, where $F(-\infty) := \lim_{x\to-\infty} F(x)$, $F^{+}(+\infty) := \lim_{x\to+\infty} F^{+}(x)$, and $F^{-}(+\infty) := \lim_{x\to+\infty} F^{-}(x)$. (Hint: The main difficulty comes from the fact that a partition $x_0 < \ldots < x_n \le x$ that is good for F^{+} need not be good for F^{-}, and vice versa. However, this can be fixed by taking a good partition for F^{+} and a good partition for F^{-} and *combining* them together into *a common refinement*.)

- **Corollary 2.8** *(**Bounded Variation Differentiation Theorem**).*
Every ***bounded variation*** *function is* ***differentiable almost everywhere.***

- **Definition** (***Locally Bounded Variation***)
A function is ***locally of bounded variation*** if it is of *bounded variation* on *every* **compact** *interval* $[a, b]$.

  Corollary 2.9 *(**Locally Bounded Variation Differentiation Theorem**).*
Every ***locally bounded variation*** *function is* ***differentiable almost everywhere.***

- **Definition** (***Lipschitz Continuous Function***)
A function $f : \mathbb{R} \to \mathbb{R}$ is said to be ***Lipschitz continuous*** if there exists a constant $C > 0$ such that

$$
|f(x) - f(y)| \le C\, |x - y|
$$

for all $x, y \in \mathbb{R}$; the *smallest* C with this property is known as ***the Lipschitz constant*** of f.

  Corollary 2.10 *(**Lipschitz Differentiation Theorem**, one-dimensional case).*
Every Lipschitz continuous function F is ***locally*** *of* ***bounded variation****, and hence* ***differentiable almost everywhere****. Furthermore, the* ***derivative*** *F', when it exists, is* ***bounded*** *in magnitude by the Lipschitz constant of F.*

  Remark The same result is true in *higher dimensions*, and is known as ***the Rademacher differentiation theorem***.

- **Definition** (***Convex Function***)
  A function $f : \mathbb{R} \to \mathbb{R}$ is said to be ***convex*** if one has $f((1-t)x + ty) \le (1-t)f(x) + tf(y)$ for all $x < y$ and $0 < t < 1$.

  Corollary 2.11 *(**Convex Differentiation Theorem**, one-dimensional case)*
  *If f is convex, then it is **continuous** and **almost everywhere differentiable**, and its **derivative** f' is **equal almost everywhere** to a **monotone non-decreasing function**, and so is itself **almost everywhere differentiable**.*

  (Hint: Drawing the graph of f, together with a number of chords and tangent lines, is likely to be very helpful in providing visual intuition.)

  Remark Thus we see that in some sense, ***convex functions*** are "***almost everywhere twice differentiable***". Similar claims also hold for *concave functions*, of course.

- **Remark** From above, we see that *the class* of ***functions of locally bounded variations*** contains the following sub-classes:

  1. ***Bounded Monotone Functions***
  2. ***Lipschitz Continuous Functions***
  3. ***Convex (Concave) Function***
  4. ***Absolute Continuous Function*** thus includes *Uniformly Continuous Function* too

2.2 The Second Fundamental Theorem of Calculus for Lebesgue Integral

- **Proposition 2.12** *(**Upper bound for second fundamental theorem**).*
  *Let $F : [a, b] \to \mathbb{R}$ be **monotone non-decreasing** (so that, as discussed above, F' is defined almost everywhere, is unsigned, and is measurable). Then*

$$\int_{[a,b]} F'(x)dx \le F(b) - F(a).$$

  *In particular, F' is **absolutely integrable**.*

- For function of bounded variation, the derivative is also absolutely integrable

  Proposition 2.13 *Any function of bounded variation has an (almost everywhere defined) derivative that is **absolutely integrable**.*

- For Lipschitz continuous function, we can directly prove the second fundamental theorem of calculus:

  Theorem 2.14 *(**Second fundamental theorem for Lipschitz functions**).*
  *Let $F : [a, b] \to \mathbb{R}$ be **Lipschitz continuous**.*

$$\int_{[a,b]} F'(x)dx = F(b) - F(a).$$

  (Hint: Argue as in the proof of Proposition above, but use ***the dominated convergence theorem*** in place of *Fatous lemma*)

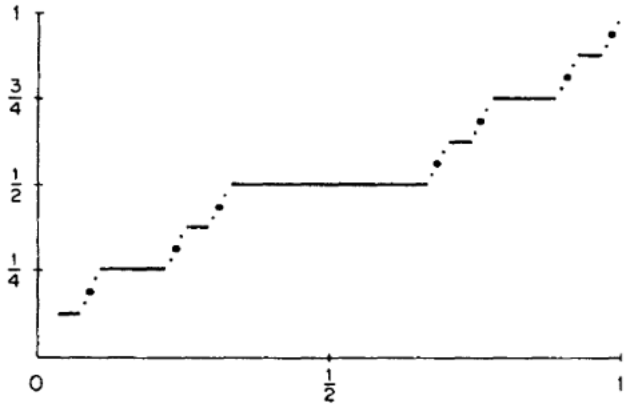

Figure 1: The Cantor Function [Reed and Simon, 1980]

- **Remark** One of the main ***challenge*** to show the second fundament theorem of calculus for *all monotone function* (i.e. to show the equality condition holds above) is that *all the variation* of F may be ***concentrated in a set of measure zero***, and thus *undetectable* by the *Lebesgue integral* of F'. The following is one of example

  Example ***The Heaviside function*** is defined as $F := \mathbb{1}\{[0, +\infty)\}$. It is clear that F' vanishes almost everywhere, but $F(b)F(a)$ is *not equal to* $\int_{[a,b]} F'(x)dx$ if b and a lie on ***opposite*** sides of the discontinuity at 0.

- Moreover, we have

  Proposition 2.15 *If F is a jump function, then F'* ***vanishes*** *almost everywhere.*

  Thus *the second fundamental theorem of calculus does not hold for any jump functions.*

- **Remark** Even only consider ***the continuous monotone function***, it is still possible for *all the fluctuation to now be* ***concentrated****, not in a countable collection of jump discontinuities, but instead* ***in an uncountable set of zero measure***, such as *the middle thirds* ***Cantor set****.* This can be illustrated by the key counterexample of ***the Cantor function***, also known as ***the Devil's staircase function****.*

  This example shows that the classical derivative $F'(x) := \lim_{h\to 0} \frac{F(x+h)F(x)}{h}$ of a function has some defects; *it cannot "see" some of the variation of a continuous monotone function* such as the Cantor function.

- **Remark** In view of this counterexample, we see that we need to add *an additional hypothesis* to ***the continuous monotone*** *non-increasing function* F before we can recover the second fundamental theorem. One such hypothesis is ***absolute continuity****.*

- **Definition** A function $F : \mathbb{R} \to \mathbb{R}$ is ***continuous*** if, for every $\epsilon > 0$ and $x_0 \in \mathbb{R}$, there exists a $\delta > 0$ such that $|F(b) - F(a)| \leq \epsilon$ whenever (a, b) is an interval of length at most δ that contains x_0.

  Definition A function $F : \mathbb{R} \to \mathbb{R}$ is ***uniformly continuous*** if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|F(b) - F(a)| \leq \epsilon$ whenever (a, b) is an interval of length at most δ.

- **Definition** (***Absolute Continuity***)
  A function $F : \mathbb{R} \to \mathbb{R}$ is said to be ***absolutely continuous*** if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that $\sum_{j=1}^{n} |F(b_j) - F(a_j)| \leq \epsilon$ whenever $(a_1, b_1), \ldots, (a_n, b_n)$ is a **finite collection of disjoint intervals** of ***total length*** $\sum_{j=1}^{n} |b_j - a_j|$ ***at most*** δ.

- **Proposition 2.16** *The followings statements are true:*

  1. *Every* ***absolutely continuous*** *function is* ***uniformly continuous*** *and therefore* ***continuous***.
  2. *Every* ***absolutely continuous*** *function is of* ***bounded variation*** *on every* ***compact*** *interval* $[a, b]$. *(Hint: first show this is true for any sufficiently small interval.) Thus, by the Local Bounded Variation Differentiation Theorem, absolutely continuous functions are* ***differentiable almost everywhere***.
  3. *Every* ***Lipschitz continuous*** *function is* ***absolutely continuous***.
  4. *The function* $x \mapsto \sqrt{x}$ *is absolutely continuous, but not Lipschitz continuous, on the interval* $[0, 1]$.
  5. ***The Cantor function*** *is continuous,* ***monotone***, *and* ***uniformly continuous***, *but* ***not absolutely continuous***, *on* $[0, 1]$.
  6. *If* $f : \mathbb{R} \to \mathbb{R}$ *is* ***absolutely integrable***, *then the indefinite integral* $F(x) := \int_{[-\infty, x]} f(y)dy$ *is* ***absolutely continuous***, *and* F *is differentiable almost everywhere with* $F'(x) = f(x)$ *for almost every* x.
  7. *The* ***sum*** *or* ***product*** *of two absolutely continuous functions on an interval* $[a, b]$ *remains absolutely continuous.*

- **Remark** We can draw the relative strength of different concepts on a *compact interval* $[a, b]$.

$$
\begin{tikzcd}
\textit{Lipschitz continuous} \arrow[r] \arrow[dr] & \textit{absolutely continuous} \arrow[r] \arrow[d] & \textit{uniformly continuous} \arrow[r] & \textit{continuous} \\
\textit{convex} \arrow[u] \arrow[ur] \arrow[dr] & \textit{locally bounded variation} \arrow[dr] & & \\
\textit{monotone} \arrow[r, "\textit{bounded}"] \arrow[rr, bend right] & \textit{bounded variation} \arrow[r] \arrow[u] & \textit{differentiable a.e.} &
\end{tikzcd}
$$

  - ***uniformly continuous*** $\not\to$ ***absolutely continuous***: See Cantor function example [Tao, 2011].
  - ***absolutely continuous*** $\not\to$ ***Lipschitz continuous***: $x \mapsto \sqrt{x}$

- **Proposition 2.17** *Absolutely continuous functions map* ***null sets*** *to* ***null sets***, *i.e. if* $F : \mathbb{R} \to \mathbb{R}$ *is* ***absolutely continuous*** *and* E *is a null set then* $F(E) := \{F(x) : x \in E\}$ *is also a null set.*

  Exercise 2.18 *Show that the Cantor function does not have this property above.*

- For absolutely continuous functions, we can recover the second fundamental theorem of calculus:

  Theorem 2.19 ***(Second Fundamental Theorem for Absolutely Continuous Functions).***
  Let $F : [a, b] \to \mathbb{R}$ be ***absolutely continuous****. Then*

  $$\int_{[a,b]} F'(x)dx = F(b) - F(a).$$

  Note that F' is ***absolutely integrable****.*

- **Proposition 2.20** ***(Classification of Absolute Continuous Function)***
  A function $F : [a, b] \to \mathbb{R}$ is ***absolutely continuous*** *if and only if it takes the form*

  $$F(x) = \int_{[a,x]} f(y)dy + C$$

  for some ***absolutely integrable*** *$f : [a, b] \to \mathbb{R}$ and a constant C.*

- **Remark** We see that the ***absolute continuity*** was used primarily in *two ways*:

  1. firstly, to ensure ***the almost everywhere existence*** of F'
  2. to control *an exceptional* ***null set*** E.

  It turns out that one can achieve the latter control by making *a different hypothesis*, namely that *the function F is everywhere differentiable* rather than merely *almost everywhere differentiable*. More precisely, we have

- **Theorem 2.21** ***(Second Fundamental Theorem of Calculus, again).***
  Let $[a, b]$ be a compact interval of positive length, let $F : [a, b] \to \mathbb{R}$ be a ***differentiable*** *function, such that F'* ***is absolutely integrable****. Then the Lebesgue integral*

  $$\int_{[a,b]} F'(x)dx = F(b) - F(a).$$

- **Exercise 2.22** *Let $F : [-1, 1] \to \mathbb{R}$ be the function defined by setting*

  $$F(x) := x^2 \sin\left(\frac{1}{x^3}\right)$$

  when x is non-zero, and $F(0) := 0$. Show that F is everywhere differentiable, but the deriative F' is not absolutely integrable, and so the second fundamental theorem of calculus does not apply in this case (at least if we interpret $\int_{[a,b]} F'(x)dx$ using the absolutely convergent Lebesgue integral).

References

Gerald B Folland. *Real analysis: modern techniques and their applications.* John Wiley & Sons, 2013.

Michael Reed and Barry Simon. *Methods of modern mathematical physics: Functional analysis*, volume 1. Gulf Professional Publishing, 1980.

Halsey Lawrence Royden and Patrick Fitzpatrick. *Real analysis*, volume 198. Prentice Hall, Macmillan New York, 1988.

Elias M Stein and Rami Shakarchi. *Real analysis: measure theory, integration, and Hilbert spaces.* Princeton University Press, 2009.

Terence Tao. *An introduction to measure theory*, volume 126. American Mathematical Soc., 2011.