# Lecture 0: Summary (part 2)

Tianpei Xie

Nov. 16th., 2022

## Contents

# 1 Signed Measures and Radon-Nikodym Derivative

## 1.1 Signed Measure

- **Definition** (***Signed Measure***)
  Let $(X, \mathscr{B})$ be a measure space. A ***signed measure*** on $(X, \mathscr{B})$ is a function $\nu : \mathscr{B} \to [-\infty, +\infty]$ such that

  1. (***Emptyset***) $\nu(\emptyset) = 0$;
  2. (***Finiteness in One Direction***) $\nu$ assumes at most one of the values $\pm\infty$;
  3. (***Countable Additivity***) if $\{E_j\}$ is a sequence of disjoint sets in $\mathscr{B}$, then $\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$, where the latter converges absolutely if $\nu\left(\bigcup_{j=1}^{\infty} E_j\right)$ is finite.

- **Definition** A measure $\mu$ is ***finite***, if $\mu(X) < \infty$; $\mu$ is $\sigma$***-finite***, if $X = \bigcup_{k=1}^{\infty} U_k$, $\mu(U_k) < \infty$.

- **Remark** ***Every signed measure*** can be represented as one of these two forms

  1. $\nu = \mu_+ - \mu_-$, where at least one of $\mu_+, \mu_-$ is a finite measure;
  2. $\mu$ is measure on $\mathscr{B}$, and $f : X \to [-\infty, +\infty]$ is *extended $\mu$-integrable* with at least one of $f_+$ and $f_-$ finite integrable. Then $\nu(A) = \int_X f \mathbb{1}\{A\}\, d\mu$ is a signed measure.

- Like unsigned measure, we have monotone downward and upward convergence:

  **Proposition 1.1** *Let $\nu$ be a **signed measure** on $(X, \mathscr{B})$.*

  1. *(**Upwards monotone convergence**) If $E_1 \subseteq E_2 \subseteq \dots$ are $\mathscr{B}$-measurable, then*

  $$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \nu(E_n) = \sup_n \nu(E_n). \tag{1}$$

  2. *(**Downwards monotone convergence**) If $E_1 \supseteq E_2 \supseteq \dots$ are $\mathscr{B}$-measurable, and $\nu(E_n) < \infty$ for **at least one** $n$, then*

  $$\nu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \nu(E_n) = \inf_n \nu(E_n). \tag{2}$$

- **Definition** (***Positive Measure***)
  If $\nu$ is a signed measure on $(X, \mathscr{B})$, a ***set*** $E \in \mathscr{B}$ is called ***positive*** (resp. ***negative***, ***null***) for $\nu$ if $\nu(F) \geq 0$ (resp. $\nu(F) \leq 0$, $\nu(F) = 0$) for ***all $\mathscr{B}$-measurable subset*** of $E$ (i.e. $F \in \mathscr{B}$ such that $F \subseteq E$).

  In other word, $E$ is $\nu$***-positive***, $\nu$***-negative***, $\nu$***-null*** if and only if $\nu(E \cap M) > 0$, $\nu(E \cap M) < 0$, $\nu(E \cap M) = 0$ ***for any*** $M$. Thus if $\nu(E) = \int_X f \mathbb{1}\{E\}\, d\mu$, then it corresponds to $f \geq 0$, $f \leq 0$ and $f = 0$ for $\mu$***-almost everywhere*** $x \in E$.

- **Lemma 1.2** *[Folland, 2013]*
  *Any **measureable subset** of a positive set is positive, and the **union** of any **countable** positive set is positive.*

- **Remark** For two measures $\mu, \nu$ on $(X, \mathscr{B})$ among which at least one of them is finite, the expression $\mu \geq \nu$ on $E$ means that for every $F \subseteq E \in \mathscr{B}$, $(\mu - nu)(F) \geq 0$. That is, $E$ is *a positive set* of $(\mu - nu)$.

## 1.2 Decomposition of Signed Measure

- **Remark** Given a signed measure $\nu$, we can ***partition*** the space $X$ into positive set (i.e. all of its measurable subsets have positive measure) and negative set (i.e. all of its measurable subsets have negative measure).

- **Theorem 1.3** *(**The Hahn Decomposition Theorem**)[Folland, 2013]*
*If $\nu$ is a **signed measure** on $(X, \mathscr{B})$, there exists a **positive set** $P$ and a **negative set** $N$ for $\nu$ such that $P \cup N = X$ and $P \cap N = \emptyset$. If $P', N'$ is another such pair, then $P \Delta P' = N \Delta N'$ is **null** w.r.t. $\nu$.*

- **Definition** [Folland, 2013, Resnick, 2013]
The decomposition of $X = P \cup N$ as $X$ is a ***disjoint union*** *of a* ***positive set*** *and a* ***negative set*** is called a ***Hahn decomposition for*** $\nu$.

- **Remark** Note that the Hahn decomposition is usually ***not unique*** as the $\nu$-null set can be transferred between subparts $P$ and $N$. To find unique decomposition, we need the following concepts:

- **Definition** [Folland, 2013]
Two *signed measures* $\mu, \nu$ on $(X, \mathscr{B})$ are ***mutually singular***, or that $\nu$ is ***singular*** w.r.t. to $\mu$, or vice versa, if and only if there exists a ***partition*** $E, F \in \mathscr{B}$ of $X$ such that $E \cap F = \emptyset$ and $E \cup F = X$, $E$ ***is null for*** $\mu$ and $F$ ***is null for*** $\nu$. Informal speaking, ***mutual singular*** means that $\mu$ and $\nu$ "***live on disjoint sets***". We describe it using perpendicular sign

$$\mu \perp \nu$$

- **Theorem 1.4** *(**The Jordan Decomposition Theorem**)[Folland, 2013]*
*If $\nu$ is a signed measure on $(X, \mathscr{B})$, there exists **unique positive measure** $\nu_+$ and $\nu_-$ such that*

$$\nu = \nu_+ - \nu_- \qquad \textit{and} \qquad \nu_+ \perp \nu_-.$$

- **Definition** The two positive measures $\nu_+, \nu_-$ are called the ***positive*** and ***negative variations*** of $\nu$, and $\nu = \nu_+ - \nu_-$ is called the ***Jordan decomposition*** of $\nu$.

  Furthermore, define the ***total variations*** of $\nu$ as the measure $|\nu|$ such that

$$|\nu| = \nu_+ + \nu_-.$$

- **Proposition 1.5** *Let $\nu, \mu$ be signed measures on $(X, \mathscr{B})$ and $|\nu|$ is the total variations of $\nu$. Then*

  1. *$E \in \mathscr{B}$ is $\nu$-null if and only if $|\nu|(E) = 0$*
  2. *$\nu \perp \mu$ **if and only if** $|\nu| \perp \mu$ if and only if $(\nu_+ \perp \mu) \wedge (\nu_- \perp \mu)$.*

- **Proposition 1.6** *If $\nu_1, \nu_2$ are signed measures that both omit $\pm\infty$, then* $|\nu_1 + \nu_2| \le |\nu_1| + |\nu_2|$

- **Exercise 1.7** *Let $\nu$ be a signed measure on $(X, \mathscr{B})$.*

  1. $L^1(\nu) = L^1(|\nu|)$;
  2. *If $f \in L^1(\nu)$, then*
  $$\left|\int_X f d\nu\right| \le \int_X |f|\, d\,|\nu|$$
  3. *If $E \in \mathscr{B}$, then*
  $$|\nu|\,(E) = \sup\left\{\left|\int_E f d\nu\right| : \; |f| \le 1\right\}$$

- **Remark** We recall that $\nu$ assume at most one of values on $\pm\infty$:

  1. If $\nu$ does not take $+\infty$, then $\nu_+(X) = \nu(P) < \infty$ ***is a finite measure***;
  2. if $\nu$ does not take $-\infty$, then $\nu_-(X) = -\nu(N) < \infty$ ***is a finite measure***.

  In particular, if the range of $\nu$ is contained in $\mathbb{R}$, then $\nu$ is *bounded*.

- **Remark** We observe that $\nu$ is *of form* $\nu(E) = \int_E f d\mu$ where $|\nu| = \mu$ and $f = \mathbb{1}_P - \mathbb{1}_N$ and $X = P \cup N$ being a *Hahn decomposition* for $\nu$.

- **Remark** (***Integration with respect to Signed Measure***)
  Let $\nu$ be signed measures on $(X, \mathscr{B})$ and $\nu = \nu_+ - \nu_-$ is *the Jordan decomposition* of $\nu$ then
  $$\int_X f d\nu = \int_X f d\nu_+ - \int_X f d\nu_-$$
  for all $f \in L^1(X, \nu)$.

- **Definition** *A signed measure* $\nu$ is called ***$\sigma$-finite*** if $|\nu|$ is *$\sigma$-finite.*

### 1.3 Lebesgue-Radon-Nikodym Theorem

- **Definition** [Folland, 2013]
  Suppose $\nu$ is *a* ***signed measure*** on $(X, \mathscr{B})$ and $\mu$ is *a* ***positive measure*** on $(X, \mathscr{B})$. Then $\nu$ is said to be ***absolutely continuous w.r.t.*** $\mu$ and write
  $$\nu \ll \mu$$
  if $\nu(E) = 0$ for *every $E \in \mathscr{B}$ for which* $\mu(E) = 0$.

- **Proposition 1.8** *Suppose $\nu$ is a signed measure on $(X, \mathscr{B})$, $\nu_+, \nu_-$ are positive and negative variation of $\nu$ and $|\nu|$ is the total variation. Then $\nu \ll \mu$* ***if and only if*** $|\nu| \ll \mu$ ***if and only if*** $(\nu_+ \ll \mu) \wedge (\nu_- \ll \mu)$.

- **Remark** ***Absolutly continuity*** is in a sense ***antithesis*** (i.e. *direct opposite*) of ***mutual singularity***. More precisely, if $\nu \perp \mu$ and $\nu \ll \mu$, then $\nu = 0$, since $E, F$ are disjoint sets such that $E \cup F = X$, and $\mu(E) = |\nu|\,(F) = 0$, then $\nu \ll \mu$ implies that $|\nu|\,(E) = 0$. One can *extend* the notion of absolute continuity to the case where $\mu$ *is a signed measure* (namely, $\nu \ll \mu$ iff $\nu \ll |\mu|$), but we shall have no need of this more general definition.

- **Theorem 1.9** *($\epsilon$-$\delta$ **Language of Absolute Continuity of Measures**)*
  *Let $\nu$ is a **finite signed measure** and $\mu$ is a **positive** measure on $(X, \mathscr{B})$. Then $\nu \ll \mu$ if and only if for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|\nu(E)| < \epsilon$, **whenever** $\mu(E) < \delta$.*

- **Remark** If $\mu$ is a *measure* and $f$ is ***extended $\mu$-integrable***, then ***the signed measure*** $\nu$ *defined via* $\nu(E) = \int_E f d\mu$ *is **absolutely continuous** w.r.t.* $\mu$; it is ***finite*** if and only if $f$ is ***absolutely integrable***. For any complex-valued $f \in L^1(\mu)$, the preceding theorem can be applied to $\Re(f)$ and $\Im(f)$.

- **Corollary 1.10** *If $f \in L^1(X, \mu)$, for every $\epsilon > 0$, there exists a $\delta > 0$, such that $\left|\int_E f d\mu\right| < \epsilon$ whenever $\mu(E) < \delta$.*

- **Definition** For *a **signed measure*** $\nu$ defined via $\nu(E) = \int_E f d\mu$ for all $E \in \mathscr{B}$, we use the notation to express the relationship

$$d\nu = f\, d\mu.$$

  Sometimes, by a slight abuse of language, we shall refer to "***the signed measure*** $f\, d\mu$"

- **Lemma 1.11** *[Folland, 2013]*
  *Suppose that $\nu$ and $\mu$ are **finite measures** on $(X, \mathscr{B})$. Either $\nu \perp \mu$, or there exists $\epsilon > 0$ and $E \in \mathscr{B}$ such that $\mu(E) > 0$ and $\nu \geq \epsilon\mu$ on $E$, i.e. $E$ is a **positive set for** $\nu - \epsilon\mu$.*

- **Theorem 1.12** *(**Lebesgue-Radon-Nikodym Theorem**)[Folland, 2013]*
  *Let $\nu$ be a **$\sigma$-finite signed** measure and $\mu$ be a **$\sigma$-finite positive** measure on $(X, \mathscr{B})$. There exists **unique $\sigma$-finite signed measure** $\lambda, \rho$ on $(X, \mathscr{B})$ such that*

$$\lambda \perp \mu\,, \quad \textit{and} \quad \rho \ll \mu\,, \quad \textit{and} \quad \nu = \lambda + \rho.$$

  *In particular, if $\nu \ll \mu$, then*

$$d\nu = f d\mu, \qquad \textit{for some } f.$$

- **Definition** The decomposition $\nu = \rho + \lambda$, where $\lambda \perp \mu$ and $\rho \ll \mu$, is called the ***Lebesgue decomposition*** *of $\nu$ with respect to $\mu$.*

- **Definition** If $\nu \ll \mu$, then according to *the Lebesgue-Radon-Nikodym theorem*, $d\nu = f d\mu$ for some $f$, where $f$ is called the ***Radon-Nikodym derivative*** *of $\nu$ w.r.t.* $\mu$ and is denoted as

$$f := \frac{d\nu}{d\mu} \quad \Rightarrow \quad d\nu = \frac{d\nu}{d\mu} d\mu.$$

- **Remark** By Lebesgue decomposition, *a signed measure* $\nu$ can be represented as

$$d\nu = d\lambda + f d\mu$$

- **Remark** (***Jordan Decomposition vs. Lebesgue Decomposition***)
  We see ***two unique decompositions***: the Jordan decomposition and the Lebesgue decomposition. We can make a comparison:

  1. Both of these two are *decompositions* of *a **signed** measure* $\nu$.

2. Both of these two decompositions seperate $\nu$ into two ***mutually signular*** sub-measures of $\nu$.

3. Both of these two decompositions are ***unique***

On the other hand,

1. ***The Jordan decomposition*** is to split *a signed measure* $\nu$ ***itself*** into ***two positive measures***, i.e. $\nu_+$ and $\nu_-$ that are ***mutually singular*** $(\nu_+ \perp \nu_-)$.

2. ***The Lebesgue decomposition*** is to split *a signed measure* $\nu$ ***with respect to a postive measure*** $\mu$. The result is *two-fold*: 1) *two mutually singular sub-measures* $\lambda \perp \rho$ 2) their relationship with $\mu$ is ***opposite***: $\lambda \perp \mu$, i.e. their support do not overlap; $\rho \ll \mu$, i.e. its support lies within support of $\mu$.

3. Note that $\lambda, \rho$ from *the Lebesgue decomposition* is ***not*** *necessarily* ***positive***. But both $\nu$ and $\mu$ need to be $\sigma$***-finite*** which is *not required* for *the Jordan decomposition.*

- **Proposition 1.13** *[Folland, 2013]*
*Suppose $\nu$ is $\sigma$**-finite signed measure** and $\lambda, \mu$ are $\sigma$**-finite measure** on $(X, \mathscr{B})$ such that $\nu \ll \mu$ and $\mu \ll \lambda$.*

  1. *If $g \in L^1(X, \nu)$, then $g \left(\frac{d\nu}{d\mu}\right) \in L^1(X, \mu)$ and*
$$\int g d\nu = \int g \frac{d\nu}{d\mu} d\mu$$

  2. *We have $\nu \ll \lambda$, and*
$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}, \quad \lambda\text{-a.e.}$$

- **Corollary 1.14** *If $\mu \ll \lambda$ and $\lambda \ll \mu$, then $(d\lambda/d\mu)(d\mu/d\lambda) = 1$ a.e. (with respect to either $\lambda$ or $\mu$).*

- **Proposition 1.15** *If $\mu_1, \ldots, \mu_n$ are measures on $(X, \mathscr{B})$, then there exists a measure $\mu$ such that $\mu_i \ll \mu$ for all $i = 1, \ldots, n$, namely, $\mu = \sum_{i=1}^n \mu_i$.*

- **Exercise 1.16** ***(Conditional Expectation)***
*Let $(X, \mathscr{B}, \mu)$ be a **finite measure space**, $\mathscr{F}$ is a sub-$\sigma$-algebra of $\mathscr{B}$, and $\nu = \mu|_{\mathscr{F}}$. Show that if $f \in L^1(X, \mu)$, there exists $g \in L^1(X, \nu)$ (thus $g$ is $\mathscr{F}$**-measureable**) such that $\int_E f d\mu = \int_E g d\nu$ for all $E \in \mathscr{F}$; if $g'$ is another such function then $g = g'$ $\nu$-a.e.*

  *In **probability theory**, where $(X, \mathscr{B}) \equiv (\Omega, \mathscr{A})$, $f \equiv X$ is a **random variable**, then $g \equiv \mathbb{E}\left[X | \mathscr{F}\right]$ is called **the conditional expectation of** $X$ **on** $\mathscr{F}$, which is $\mathscr{F}$-measure random variable.*

# 2 Differentiation

- **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.*

## 2.1 The Lebesgue Differentiation Theorem in One Dimension

- **Theorem 2.1** *(**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 2.2** *(**Lebesgue differentiation theorem, second formulation**).*
*Let $f : \mathbb{R} \to \mathbb{C}$ be an **absolutely integrable** function. Then*
$$\lim_{h \to 0+} \frac{1}{h} \int_{[x,x+h]} f(t)dt = f(x) \tag{3}$$
*for almost every $x \in \mathbb{R}$, and*
$$\lim_{h \to 0+} \frac{1}{h} \int_{[x-h,x]} f(t)dt = f(x) \tag{4}$$
*for almost every $x \in \mathbb{R}$.*

- **Remark** (***Density Argument***) [Tao, 2011]
The conclusion (3) 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*).

- **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.

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

### 2.2.1 Absolute Integrable Version

- **Theorem 2.3** *(**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*
$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(z) - f(x)|\, dz = 0 \tag{5}$$
$$\text{and} \quad \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(z)dz = f(x),$$
  *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 (5) 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
$$A_r f(x) = \frac{1}{m\,(B(r,x))} \int_{B(r,x)} f(z)dz.$$

- **Remark** A useful variant of $Hf(x)$ (see [Stein and Shakarchi, 2009]) as
$$H^* f(x) \equiv \sup \left\{ \frac{1}{m\,(B)} \int_B |f(z)|\, dz,\ \ B \text{ is a ball },x \in B \right\}.$$

- **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 for *Lebesgue differentiation theorem*:

  **Theorem 2.4** ***(The Hardy-Littlewood Maximal Theorem)*** *[Stein and Shakarchi, 2009, Folland, 2013]*
  *Suppose $f$ is integrable, then*

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

     *is measurable.*

  2. $H^{*}f(x) < \infty$ *for a.e. $x$.*

  3. $H^{*}f$ *satisfies **the Hardy-Littlewood maximal inequality**:*

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

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

  *Note that $H^{*}f \geq |f|$, 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 show *the Hardy-Littlewood maximal inequality*, we need to first find a dense covering called *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 2.5** ***(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 2.6** ***(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*

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

    *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.*

  - **Lemma 2.7** ***(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*

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

– **Corollary 2.8** *[Stein and Shakarchi, 2009, Royden and Fitzpatrick, 1988]*
*Follwing the setting above, we can arrange the choice of balls so that*

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

### 2.2.2 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 2.9** *[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*

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

- **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

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

or *equivalently*, [Folland, 2013],

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

- These results are from *the Hardy-Littlewood maximal inequality*

**Corollary 2.10** *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 2.11** *If $f$ is* ***locally integrable*** *on $\mathbb{R}^d$, then* ***almost every point*** *belongs to* ***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 $U_\alpha$ there is a ball $B$ with

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

- **Theorem 2.12** *(**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 $x$.*

## 2.3 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* $\nu$ 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 $\nu$* will be called ***regular*** if $|\nu|$ is *regular.*

- **Theorem 2.13** *(**Lebesgue Density from Radon-Nikodym derivative**) [Folland, 2013]*
*Let $\nu$ 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 $x$.*

# 3 The Fundamental Theorem of Calculus for Lebesgue Integral

## 3.1 Functions of Bounded Variations

- **Theorem 3.1** *(**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 $\theta$ ***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 3.2** *(**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 3.3** *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** As *positive measures* on $\mathbb{R}$ are related to *increasing functions*, *complex measures* on $\mathbb{R}$ are related to so-called *functions of bounded variation.*

- **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 3.4** *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 3.5** *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 3.6** *(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 3.7** *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

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

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 $F^-(+\infty) := \lim_{x\to+\infty} F^-(x)$.

- **Corollary 3.8** *(**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 3.9** *(**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 3.10** *(**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 3.11** *(**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

## 3.2 The Second Fundamental Theorem of Calculus for Lebesgue Integral

- **Proposition 3.12** ***(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 3.13** *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 3.14** ***(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*)

- **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 3.15** *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 $\delta$ 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 $\delta$.

- **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*** $\delta$.

- **Proposition 3.16** *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* $[a, b]$.

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

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

- **Proposition 3.17** *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 3.18** *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 3.19** ***(Second Fundamental Theorem for Absolutely Continuous Functions).***
  *Let $F : [a, b] \to \mathbb{R}$ be <u>**absolutely continuous**</u>. Then*

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

- **Proposition 3.20** ***(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 3.21** ***(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 3.22** *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.

Sidney I Resnick. *A probability path.* Springer Science & Business Media, 2013.

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.