Lecture 0: Summary (part 2)

Tianpei Xie

Nov. 16th., 2022

Contents

1 σ-Algebra, Measure and Integration on Product Space

1.1 Product σ-Algebra

- **Definition** (***Product Topology***)
  Let X and Y be topological spaces. **_The product topology_** on $X \times Y$ is the topology having as basis the collection $\mathscr{B}$ of all sets of the form $U \times V$, where U is an open subset of X and V is an open subset of Y.

  Definition Let $\pi_X : X \times Y \to X$ be defined by the equation
  $$\pi_X(x, y) = x;$$
  $\pi_Y : X \times Y \to Y$ he defined by the equation
  $$\pi_Y(x, y) = y.$$

  The maps π_X and π_Y are called *the **projections** of $X \times Y$ **onto** its **first** and **second** factors*, respectively.

- **Definition** (***Product σ-Algebra***)
  Suppose that $(X, \mathscr{B}_X)$ and $(Y, \mathscr{B}_Y)$ are measurable spaces. We can form the **_pullback σ-algebras_**
  $$\pi_X^*(\mathscr{B}_X) := \left\{ \pi_X^{-1}(E) : E \in \mathscr{B}_X \right\} = \{E \times Y : E \in \mathscr{B}_X\}$$
  $$\pi_Y^*(\mathscr{B}_Y) := \left\{ \pi_Y^{-1}(F) : F \in \mathscr{B}_Y \right\} = \{X \times F : F \in \mathscr{B}_Y\}$$

  We then define **_the product σ-algebra $\mathscr{B}_X \times \mathscr{B}_Y$_** to be the **$\sigma$-algebra generated by the union of these two pull-back σ-algebras**:
  $$\mathscr{B}_X \times \mathscr{B}_Y := \langle \pi_X^*(\mathscr{B}_X) \cup \pi_Y^*(\mathscr{B}_Y) \rangle.$$

  This definition has several equivalent formulations:

- **Proposition 1.1** *(**Equivalent Definition of Product σ-Algebra**) [Tao, 2011]*
  Let $(X, \mathscr{B}_X)$ and $(Y, \mathscr{B}_Y)$ be measurable spaces.

  1. $\mathscr{B}_X \times \mathscr{B}_Y$ *is the **σ-algebra generated by the sets** $E \times F$ **with** $E \in \mathscr{B}_X$, $Y \in \mathscr{B}_Y$. In other words, $\mathscr{B}_X \times \mathscr{B}_Y$ is the **coarsest σ-algebra** on $X \times Y$ with the property that **the product** of a $\mathscr{B}_X$-measurable set and a $\mathscr{B}_Y$-measurable set is always $\mathscr{B}_X \times \mathscr{B}_Y$ measurable.*

  2. $\mathscr{B}_X \times \mathscr{B}_Y$ *is the **coarsest σ-algebra** on $X \times Y$ that makes the **projection maps** π_X, π_Y both measurable.*

- **Proposition 1.2** *(**Property of Product σ-Algebra**) [Tao, 2011]*
  Let $(X, \mathscr{B}_X)$ and $(Y, \mathscr{B}_Y)$ be measurable spaces.

  1. *(**Slices of Meaurable Set**) If $E \in \mathscr{B}_X \times \mathscr{B}_Y$, then sets*
  $$E_x := \{y \in Y : (x, y) \in E\} \in \mathscr{B}_Y$$
  for every $x \in X$, and similarly that the sets
  $$E^y := \{x \in X : (x, y) \in E\} \in \mathscr{B}_X$$
  for every $y \in Y$.

2. (**Slices of Meaurable Function**) *If* $f : X \times Y \to [0, +\infty]$ *is measurable (with respect to* $\mathscr{B}_X \times \mathscr{B}_Y$*), then the function*

$$f_x : y \to f(x, y)$$

is $\mathscr{B}_Y$*-measurable for every* $x \in X$*, and similarly that the function*

$$f^y : x \to f(x, y)$$

is $\mathscr{B}_X$*-measurable for every* $y \in Y$*.*

3. *The product of two* **trivial** σ**-algebras** *(on two different spaces* X, Y*) is again* **trivial**.

4. *The product of two* **atomic** σ**-algebras** *is again* **atomic**.

5. *The product of two* **finite** σ**-algebras** *is again* **finite**.

6. *The product of two* **Borel** σ**-algebras** *(on two Euclidean spaces* $\mathbb{R}^d$*,* $R^{d'}$ *with* $d, d' \geq 1$*) is again the* **Borel** σ**-algebra** *(on* $\mathbb{R}^d \times \mathbb{R}^{d'} \equiv \mathbb{R}^{d+d'}$*).*

7. *The product of two* **Lebesgue** σ**-algebras** *(on two Euclidean spaces* $\mathbb{R}^d$*,* $R^{d'}$ *with* $d, d' \geq 1$*) is* **not** *the* **Lebesgue** σ**-algebra**. *(Hint: argue by* **contradiction** *and use slices of measurable set as above proposition.)*

8. *However, the Lebesgue* σ*-algebra on* $\mathbb{R}^{d+d'}$ *is the* **completion** *of the product of the Lebesgue* σ*-algebras of* $\mathbb{R}^d$ *and* $\mathbb{R}^{d'}$ *with respect to* $(d+d')$*-dimensional Lebesgue measure.*

- **Exercise 1.3** *[Tao, 2011]*
  If $E \in \mathscr{B}_X \times \mathscr{B}_Y$*, show that the slices* $E_x := \{y \in Y : (x, y) \in E\}$ *lie in a countably generated* σ*-algebra. In other words, show that there exists an at most countable collection* $\mathscr{A} = \mathscr{A}_E$ *of sets (which can depend on* E*) such that* $s\{E_x : x \in X\} \subseteq \langle \mathscr{A} \rangle$*. Conclude in particular that the number of* **distinct** *slices* E_x *is at most* c*, the* **cardinality** *of the continuum.*

- **Exercise 1.4** *[Tao, 2011]*
  Give an example to show that the product of two **discrete** σ**-algebras** *is* **not** *necessarily* **discrete**.

  On the other hand, show that the product of two **discrete** σ**-algebras** 2^X*,* 2^Y *is again a* **discrete** σ**-algebra** *if at least one of the domains* X, Y *is* **at most countably infinite**.

1.2 Product Measure

- **Definition** (σ**-Finite**).
  A measure space $(X, \mathscr{B}, \mu)$ is σ**-finite** if X can be expressed as the **countable union** of sets of **finite** measure, i.e. $X = \bigcup_n X_n$, $\mu(X_n) < \infty$ for all n.

- **Example** ($\mathbb{R}^d$)
  $\mathbb{R}^d$ with **Lebesgue measure** is σ-finite, as $\mathbb{R}^d$ can be expressed as the union of (for instance) the balls $B(0, n)$ for $n = 1, 2, 3, \ldots$, each of which has finite measure.

  On the other hand, $\mathbb{R}^d$ with **counting measure is not** σ**-finite**.

- **Proposition 1.5** (**Existence and Uniqueness of Product Measure**) *[Tao, 2011]*
  Let $(X, \mathscr{B}_X, \mu_X)$ *and* $(Y, \mathscr{B}_Y, \mu_Y)$ *be* σ**-finite measure spaces**. *Then there exists a* **unique**

measure $\mu_X \times \mu_Y : \mathscr{B}_X \times \mathscr{B}_Y \to [0, \infty]$ on product σ-algebra $\mathscr{B}_X \times \mathscr{B}_Y$ that obeys

$$\mu_X \times \mu_Y (E \times F) = \mu_X(E) \, \mu_Y(F)$$

whenever $E \in \mathscr{B}_X$ and $F \in \mathscr{B}_Y$.

- **Remark** When X, Y are **not both** σ-finite, then one can still construct **at least one product measure**, but it will, in general, **not be unique**.

- **Remark** (***Product Measure of Lebesgue Measures***)
  The ***product*** $m^d \times m^{d'}$ of the ***Lebesgue measures*** $m^d, m^{d'}$ on $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d])$ and $(\mathbb{R}^{d'}, \mathcal{L}[\mathbb{R}^{d'}])$ respectively will ***agree*** with Lebesgue measure $m^{d+d'}$ on the product space $\mathcal{L}[\mathbb{R}^d] \times \mathcal{L}[\mathbb{R}^{d'}]$, which is a ***subalgebra*** of $\mathcal{L}[\mathbb{R}^{d+d'}]$. After taking the ***completion*** $\overline{m^d \times m^{d'}}$ of this product measure, one obtains the *full Lebesgue measure $m^{d+d'}$*.

- **Proposition 1.6** *[Tao, 2011]*
  Let $(X, \mathscr{B}_X)$ and $(Y, \mathscr{B}_Y)$ be measurable spaces.

  1. The product of two ***Dirac measures*** on $(X, \mathscr{B}_X)$, $(Y, \mathscr{B}_Y)$ is a ***Dirac measure*** on $(X \times Y, \mathscr{B}_X \times \mathscr{B}_Y)$.

  2. If X, Y are ***at most countable***, the product of the two ***counting measures*** on $(X, \mathscr{B}_X)$, $(Y, \mathscr{B}_Y)$ is the ***counting measure*** on $(X \times Y, \mathscr{B}_X \times \mathscr{B}_Y)$.

- **Proposition 1.7** (***Associativity of Product***). *[Tao, 2011]*
  Let $(X, \mathscr{B}_X, \mu_X)$, $(Y, \mathscr{B}_Y, \mu_Y)$, $(Z, \mathscr{B}_Z, \mu_Z)$ be σ-***finite sets***. We may identify the Cartesian products $(X \times Y) \times Z$ and $X \times (Y \times Z)$ with each other in the obvious manner. If we do so, then

$$(\mathscr{B}_X \times \mathscr{B}_Y) \times \mathscr{B}_Z = \mathscr{B}_X \times (\mathscr{B}_Y \times \mathscr{B}_Z)$$

and

$$(\mu_X \times \mu_Y) \times \mu_Z = \mu_X \times (\mu_Y \times \mu_Z).$$

1.3 Tonelli's Theorem

- **Definition** (***Monotone Class***)
  Define a ***monotone class*** *in* X is a collection $\mathscr{B}$ of subsets of X with the following two **closure properties**:

  1. If $E_1 \subset E_2 \subset \ldots$ are a ***countable increasing*** sequence of sets in $\mathscr{B}$, then $\bigcup_{n=1}^{\infty} E_n \in \mathscr{B}$.

  2. If $E_1 \supset E_2 \supset \ldots$ are a ***countable decreasing*** sequence of sets in $\mathscr{B}$, then $\bigcap_{n=1}^{\infty} E_n \in \mathscr{B}$.

- **Lemma 1.8** (***Monotone Class Lemma***). *[Tao, 2011]*
  Let $\mathscr{A}$ be a ***Boolean algebra*** on X. Then $\langle \mathscr{A} \rangle$ is the ***smallest monotone class*** that contains $\mathscr{A}$.

- **Theorem 1.9** (***Tonelli's Theorem, Incomplete Version***). *[Tao, 2011]*
  Let $(X, \mathscr{B}_X, \mu_X)$ and $(Y, \mathscr{B}_Y, \mu_Y)$ be σ-***finite measure spaces***, and let $f : X \times Y \to [0, +\infty]$ be measurable with respect to $\mathscr{B}_X \times \mathscr{B}_Y$. Then:

1. The functions

$$x \to \int_Y f(x,y)d\mu_Y(y)$$

$$y \to \int_X f(x,y)d\mu_X(x)$$

(which are well-defined) are **measurable** with respect to $\mathscr{B}_X$ and $\mathscr{B}_Y$ respectively.

2. We have

$$\int_{X\times Y} f(x,y)\, d\left(\mu_X \times \mu_Y\right)(x,y) = \int_X \left(\int_Y f(x,y)d\mu_Y(y)\right)d\mu_X(x)$$
$$= \int_Y \left(\int_X f(x,y)d\mu_X(x)\right)d\mu_Y(y) \qquad (1)$$

- **Corollary 1.10** *(**Slice of Null Set**)[Tao, 2011]*
  *Let $(X,\mathscr{B}_X,\mu_X)$ and $(Y,\mathscr{B}_Y,\mu_Y)$ be σ-**finite measure spaces**, and let $E \in \mathscr{B}_X \times \mathscr{B}_Y$ be a **null** set with respect to $\mu_X \times \mu_Y$. Then for μ_X-almost every $x \in X$, the set $E_x := \{y \in Y : (x,y) \in E\}$ is a μ_Y-**null set**; and similarly, for μ_Y-almost every $y \in Y$, the set $E^y := \{x \in X : (x,y) \in E\}$ is a μ_X-**null set**.*

  With this corollary, we can extend *Tonelli's theorem* to the completion $(X\times Y,\overline{\mathscr{B}_X \times \mathscr{B}_Y},\overline{\mu_X \times \mu_Y})$ of the product space $(X \times Y,\mathscr{B}_X \times \mathscr{B}_Y,\mu_X \times \mu_Y)$.

- **Theorem 1.11** *(**Tonelli's Theorem, Complete Version**). [Tao, 2011]*
  *Let $(X,\mathscr{B}_X,\mu_X)$ and $(Y,\mathscr{B}_Y,\mu_Y)$ be **complete** σ-**finite measure spaces**, and let $f : X \times Y \to [0,+\infty]$ be measurable with respect to $\mathscr{B}_X \times \mathscr{B}_Y$. Then:*

  1. *For μ_X-almost every $x \in X$, the function*

     $$y \to f(x,y)$$

     *is $\mathscr{B}_Y$-**measurable** and in particular, $\int_Y f(x,y)d\mu_Y(y)$ exists. Furthermore, the (μ_X-almost everywhere defined) map*

     $$x \to \int_Y f(x,y)d\mu_Y(y)$$

     *is $\mathscr{B}_X$-**measurable**.*

  2. *For μ_Y-almost every $y \in Y$, the function*

     $$x \to f(x,y)$$

     *is $\mathscr{B}_X$-**measurable** and in particular, $\int_X f(x,y)d\mu_X(x)$ exists. Furthermore, the (μ_Y-almost everywhere defined) map*

     $$y \to \int_X f(x,y)d\mu_X(x)$$

     *is $\mathscr{B}_Y$-**measurable**.*

3. We have

$$\int_{X\times Y} f(x,y)\, d\left(\overline{\mu_X \times \mu_Y}\right)(x,y) = \int_X \left(\int_Y f(x,y)d\mu_Y(y)\right)d\mu_X(x)$$
$$= \int_Y \left(\int_X f(x,y)d\mu_X(x)\right)d\mu_Y(y) \qquad (2)$$

- Specialising to the case when f is an indicator function $f = \mathbb{1}_E$, we conclude

  Corollary 1.12 *(**Tonellis Theorem for Sets**)*. *[Tao, 2011]*
  *Let $(X,\mathscr{B}_X,\mu_X)$ and $(Y,\mathscr{B}_Y,\mu_Y)$ be **complete σ-finite measure spaces**, and let $E \in \overline{\mathscr{B}_X \times \mathscr{B}_Y}$. Then:*

  1. *For μ_X-almost every $x \in X$, the set*

  $$E_x := \{y \in Y : (x,y) \in E\} \in \mathscr{B}_Y$$

  and the (μ_X-almost everywhere defined) map

  $$x \to \mu_Y(E_x)$$

  *is $\mathscr{B}_X$-**measurable**.*

  2. *For μ_Y-almost every $y \in Y$, the set*

  $$E^y := \{x \in X : (x,y) \in E\} \in \mathscr{B}_X$$

  and the (μ_Y-almost everywhere defined) map

  $$y \to \mu_X(E^y)$$

  *is $\mathscr{B}_Y$-**measurable**.*

  3. *We have*

  $$\overline{\mu_X \times \mu_Y}(E) = \int_X \mu_Y(E_x)\, d\mu_X(x)$$
  $$= \int_Y \mu_X(E^y)\, d\mu_Y(y) \qquad (3)$$

- **Remark** Tonellis theorem can **fail** if **the σ-finite hypothesis is removed**, and also that *product measure need not be unique.*

1.4 Fubini's Theorem

- **Remark** *Tonelli's theorem* is for *the **unsigned integral**,* but it leads to an important analogue for *the absolutely integral,* known as *Fubini's theorem:*

- **Theorem 1.13** *(**Fubinis Theorem**)*. *[Tao, 2011]*
  *Let $(X,\mathscr{B}_X,\mu_X)$ and $(Y,\mathscr{B}_Y,\mu_Y)$ be **complete σ-finite measure spaces**, and let $f : X \times Y \to \mathbb{C}$ be **absolutely integrable with respect to** $\overline{\mathscr{B}_X \times \mathscr{B}_Y}$. Then:*

1. For μ_X-almost every $x \in X$, the function

$$y \to f(x,y)$$

is **absolutely integrable with respect to** μ_Y and in particular, $\int_Y f(x,y)d\mu_Y(y)$ exists. Furthermore, the (μ_X-almost everywhere defined) map

$$x \to \int_Y f(x,y)d\mu_Y(y)$$

is **absolutely integrable with respect to** μ_X.

2. For μ_Y-almost every $y \in Y$, the function

$$x \to f(x,y)$$

is **absolutely integrable with respect to** μ_X and in particular, $\int_X f(x,y)d\mu_X(x)$ exists. Furthermore, the (μ_Y-almost everywhere defined) map

$$y \to \int_X f(x,y)d\mu_X(x)$$

is **absolutely integrable with respect to** μ_Y.

3. We have

$$\int_{X \times Y} f(x,y)\, d\left(\overline{\mu_X \times \mu_Y}\right)(x,y) = \int_X \left(\int_Y f(x,y)d\mu_Y(y) \right) d\mu_X(x)$$
$$= \int_Y \left(\int_X f(x,y)d\mu_X(x) \right) d\mu_Y(y) \qquad (4)$$

- **Remark** Fubini's theorem fails when one drops the hypothesis that f is *absolutely integrable* **with respect to the product space**.

- **Remark** Despite the failure of Tonelli's theorem in the σ-finite setting, it is possible to *(carefully) extend Fubini's theorem to* **the non-σ-finite setting**, as *the absolute integrability hypotheses*, when combined with *Markov's inequality*, can provide a substitute for *the σ-finite property*.

- **Remark** Informally, *Fubini's theorem* allows one to *always* **interchange the order of two integrals**, as long as the *integrand is absolutely integrable* **in the product space** *(or its completion)*. In particular, specialising to Lebesgue measure, we have

$$\int_{\mathbb{R}^{d+d'}} f(x,y)d(x,y) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^{d'}} f(x,y)dy \right) dx = \int_{\mathbb{R}^{d'}} \left(\int_{\mathbb{R}^d} f(x,y)dx \right) dy$$

whenever $f : \mathbb{R}^{d+d'} \to \mathbb{C}$ is *absolutely integrable*. In view of this, we often write $dxdy$ (or $dydx$) for $d(x,y)$.

- By combining *Fubini's theorem* with *Tonelli's theorem*, we can recast the *absolute integrability hypothesis*:

Corollary 1.14 (*Fubini-Tonelli Theorem*). *[Tao, 2011]*
*Let $(X, \mathscr{B}_X, \mu_X)$ and $(Y, \mathscr{B}_Y, \mu_Y)$ be **complete σ-finite measure spaces**, and let $f : X \times Y \to \mathbb{C}$ be **measurable with respect to** $\overline{\mathscr{B}_X \times \mathscr{B}_Y}$. If*

$$\int_X \left(\int_Y |f(x, y)| \, d\mu_Y(y) \right) d\mu_X(x) < \infty$$

*then f is **absolutely integrable** with respect to $\overline{\mathscr{B}_X \times \mathscr{B}_Y}$, and in particular the conclusions of **Fubini's theorem** hold.*

Similarly if we use $\int_Y \left(\int_X |f(x, y)| \, d\mu_X(x) \right) d\mu_Y(y)$ instead of $\int_X \left(\int_Y |f(x, y)| \, d\mu_Y(y) \right) d\mu_X(x)$.

- **Proposition 1.15** (***Area Interpretation of Integral***). *[Tao, 2011]*
  *Let $(X, \mathscr{B}, \mu)$ be a σ-finite measure space, and let $\mathbb{R}$ be equipped with Lebesgue measure m and the Borel σ-algebra $\mathcal{B}[\mathbb{R}]$. Then $f : X \to [0, +\infty]$ is **measurable if and only if** its **subgraph***

$$\{(x, t) \in X \times \mathbb{R} : 0 \leq t \leq f(x)\}$$

  *is **measurable** in $\mathscr{B} \times \mathcal{B}[\mathbb{R}]$, in which case we have*

$$(\mu \times m) \{(x, t) \in X \times \mathbb{R} : 0 \leq t \leq f(x)\} = \int_X f(x) d\mu(x).$$

  Similarly if we replace $\{(x, t) \in X \times \mathbb{R} : 0 \leq t \leq f(x)\}$ by $\{(x, t) \in X \times \mathbb{R} : 0 \leq t < f(x)\}$.

- **Proposition 1.16** (***Distribution Formula***). *[Tao, 2011]*
  *Let $(X, \mathscr{B}, \mu)$ be a σ-finite measure space, and let $f : X \to [0, +\infty]$ be **measurable**. Then*

$$\int_X f(x) d\mu(x) = \int_{[0,\infty]} \mu \{x \in X : f(x) \geq \lambda\} \, d\lambda \qquad (5)$$

  (Note that the integrand on the right-hand side is monotone and thus Lebesgue measurable.)
  Similarly if we replace $\{x \in X : f(x) \geq \lambda\}$ by $\{x \in X : f(x) > \lambda\}$.

2 Signed Measures and Radon-Nikodym Derivative

2.1 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***) ν 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 μ is ***finite***, if $\mu(X) < \infty$; μ is ***σ-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 μ_+, μ_- is a finite measure;

    2. μ is measure on $\mathscr{B}$, and $f : X \to [-\infty, +\infty]$ is *extended μ-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 2.1 *Let ν be a **signed measure** on $(X, \mathscr{B})$.*

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

    $$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \nu(E_n) = \sup_n \nu(E_n). \tag{6}$$

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

    $$\nu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \nu(E_n) = \inf_n \nu(E_n). \tag{7}$$

- **Definition** (*Positive Measure*)
  If ν is a signed measure on $(X, \mathscr{B})$, a **set $E \in \mathscr{B}$** is called **positive** (resp. **negative**, **null**) for ν 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 ν-**positive**, ν-**negative**, ν-**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 μ-**almost everywhere** $x \in E$.

- **Lemma 2.2** *[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 μ, ν 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 $(\mu - nu)$.

2.2 Decomposition of Signed Measure

- **Remark** Given a signed measure ν, 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 2.3** *(**The Hahn Decomposition Theorem**)[Folland, 2013]*
  *If ν is a **signed measure** on $(X, \mathscr{B})$, there exists a **positive set** P and a **negative set** N for ν 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. ν.*

9

- **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 ν***.

- **Remark** Note that the Hahn decomposition is usually ***not unique*** as the ν-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* μ, ν on $(X, \mathscr{B})$ are ***mutually singular***, or that ν is ***singular*** w.r.t. to μ, 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*** μ and F ***is null for*** ν. Informal speaking, ***mutual singular*** means that μ and ν "***live on disjoint sets***". We describe it using perpendicular sign

$$\mu \perp \nu$$

- **Theorem 2.4** *(**The Jordan Decomposition Theorem**)[Folland, 2013]*
  *If ν is a signed measure on $(X, \mathscr{B})$, there exists **unique positive measure** ν_+ and ν_- such that*

$$\nu = \nu_+ - \nu_- \qquad and \qquad \nu_+ \perp \nu_-.$$

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

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

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

- **Proposition 2.5** *Let ν, μ be signed measures on $(X, \mathscr{B})$ and $|\nu|$ is the total variations of ν. Then*

  1. *$E \in \mathscr{B}$ is ν-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 2.6** *If ν_1, ν_2 are signed measures that both omit $\pm\infty$, then $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$*

- **Exercise 2.7** *Let ν be a signed measure on $(X, \mathscr{B})$.*

  1. *$L^1(\nu) = L^1(|\nu|)$;*

  2. *If $f \in L^1(\nu)$, then*

$$\left| \int_X f \, d\nu \right| \leq \int_X |f| \, d|\nu|$$

  3. *If $E \in \mathscr{B}$, then*

$$|\nu|(E) = \sup \left\{ \left| \int_E f \, d\nu \right| : \ |f| \leq 1 \right\}$$

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

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

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

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

- **Remark** We observe that ν 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 ν.

- **Remark** (*Integration with respect to Signed Measure*)
  Let ν be signed measures on $(X, \mathscr{B})$ and $\nu = \nu_+ - \nu_-$ is *the Jordan decomposition* of ν 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* ν is called σ-**finite** if $|\nu|$ is σ-finite.

2.3   Lebesgue-Radon-Nikodym Theorem

- **Definition** [Folland, 2013]
  Suppose ν is a **signed measure** on $(X, \mathscr{B})$ and μ is a **positive measure** on $(X, \mathscr{B})$. Then ν is said to be **absolutely continuous w.r.t.** μ and write

$$\nu \ll \mu$$

  if $\nu(E) = 0$ for *every $E \in \mathscr{B}$ for which $\mu(E) = 0$*.

- **Proposition 2.8** *Suppose ν is a signed measure on $(X, \mathscr{B})$, ν_+, ν_- are positive and negative variation of ν 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 μ *is a signed measure* (namely, $\nu \ll \mu$ iff $\nu \ll |\mu|$), but we shall have no need of this more general definition.

- **Theorem 2.9** *(ϵ-δ Language of Absolute Continuity of Measures)*
  *Let ν is a **finite signed measure** and μ 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 μ is a *measure* and f is **extended μ-integrable**, then **the signed measure** ν *defined via* $\nu(E) = \int_E f d\mu$ *is **absolutely continuous** w.r.t.* μ; 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 2.10** *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** ν 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 2.11** *[Folland, 2013]*
  *Suppose that ν and μ 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 2.12** *(**Lebesgue-Radon-Nikodym Theorem**)[Folland, 2013]*
  *Let ν be a σ-**finite signed** measure and μ be a σ-**finite positive** measure on $(X, \mathscr{B})$. There
  exists **unique σ-finite signed measure** λ, ρ on $(X, \mathscr{B})$ such that*

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

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

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

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

- **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 ν w.r.t. μ 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 ν* 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 ν*.

  2. Both of these two decompositions seperate ν into two ***mutually signular*** sub-measures
     of ν.

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

  On the other hand,

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

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

  3. Note that λ, ρ from *the Lebesgue decomposition* is **not** *necessarily **positive**.* But both
     ν and μ need to be σ-***finite*** which is *not required* for *the Jordan decomposition*.

- **Proposition 2.13** *[Folland, 2013]*
  Suppose ν is σ-**finite signed measure** and λ, μ are σ-**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 2.14** If $\mu \ll \lambda$ and $\lambda \ll \mu$, then $(d\lambda/d\mu)(d\mu/d\lambda) = 1$ a.e. (with respect to either λ or μ).

- **Proposition 2.15** If $\mu_1, \ldots, \mu_n$ are measures on $(X, \mathscr{B})$, then there exists a measure μ such that $\mu_i \ll \mu$ for all $i = 1, \ldots, n$, namely, $\mu = \sum_{i=1}^n \mu_i$.

- **Exercise 2.16** *(Conditional Expectation)*
  Let $(X, \mathscr{B}, \mu)$ be a **finite measure space**, $\mathscr{F}$ is a sub-σ-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 $g = g'$ ν-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.

3   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*.

3.1   The Lebesgue Differentiation Theorem in One Dimension

- **Theorem 3.1** *(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 3.2** *(**Lebesgue differentiation theorem, second formulation**).*
  Let $f : \mathbb{R} \to \mathbb{C}$ *be an **absolutely integrable** function. Then*

$$\lim_{h \to 0+} \frac{1}{h} \int_{[x, x+h]} f(t)dt = f(x) \tag{8}$$

for almost every $x \in \mathbb{R}$, and

$$\lim_{h \to 0+} \frac{1}{h} \int_{[x-h, x]} f(t)dt = f(x) \tag{9}$$

for almost every $x \in \mathbb{R}$.

- **Remark** (***Density Argument***) [Tao, 2011]
  The conclusion (8) 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.

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

3.2.1   Absolute Integrable Version

- **Theorem 3.3** *(**Lebesgue Differentiation Theorem (Absolute Integrable version)**)*
  [Tao, 2011]

14

*Suppose $f : \mathbb{R}^d \to \mathbb{C}$ is **absolutely integrable**. Then for almost every x, we have*

$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(z) - f(x)| \, dz = 0 \tag{10}$$

$$and \quad \lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(z) dz = f(x),$$

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 (10) 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 $Hf \leq H^* f \leq 2^d Hf$

- We need to prove the following theorem for *Lebesgue differentiation theorem*:

  Theorem 3.4 (The Hardy-Littlewood Maximal Theorem) *[Stein and Shakarchi, 2009, Folland, 2013]*
  Suppose f is integrable, then

  1.

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

  is measurable.

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

3. $H^* f$ satisfies **<u>the Hardy-Littlewood maximal inequality</u>**:

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

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

Note that $H^* f \geq |f|$, $a.e.x$, 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 **<u>Vitali covering</u>** *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 3.5** (*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 3.6** (*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 3.7** (*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 3.8** [Stein and Shakarchi, 2009, Royden and Fitzpatrick, 1988]
    Follwing the setting above, we can arrange the choice of balls so that

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

3.2.2 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 3.9 *[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

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

  or *equivalently*, [Folland, 2013],

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

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

  Corollary 3.10 *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 3.11** *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_α there is a ball B with

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

- **Theorem 3.12** *(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

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

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

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

3.3 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 3.13 (*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.*

4 The Fundamental Theorem of Calculus for Lebesgue Integral

4.1 Functions of Bounded Variations

- **Theorem 4.1 (*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 4.2 *(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 4.3** *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. **<u>The total variation</u>** $\|F\|_{TV(\mathbb{R})}$ (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 4.4** *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 4.5** *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**): $\|cF\|_{TV(\mathbb{R})} = |c| \, \|F\|_{TV(\mathbb{R})}$ 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 4.6** *(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 4.7** *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

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

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

We have

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

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 4.8** *(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 4.9 *(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 4.10 *(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 4.11 *(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

4.2   The Second Fundamental Theorem of Calculus for Lebesgue Integral

- **Proposition 4.12** *(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*

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

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

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

  Proposition 4.13 *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 4.14 *(**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 4.15 *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 $|F(b) - F(a)| \le \epsilon$ 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 $|F(b) - F(a)| \le \epsilon$ 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 $\sum_{j=1}^{n} |F(b_j) - F(a_j)| \le \epsilon$ 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 4.16** *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* $[a, b]$.
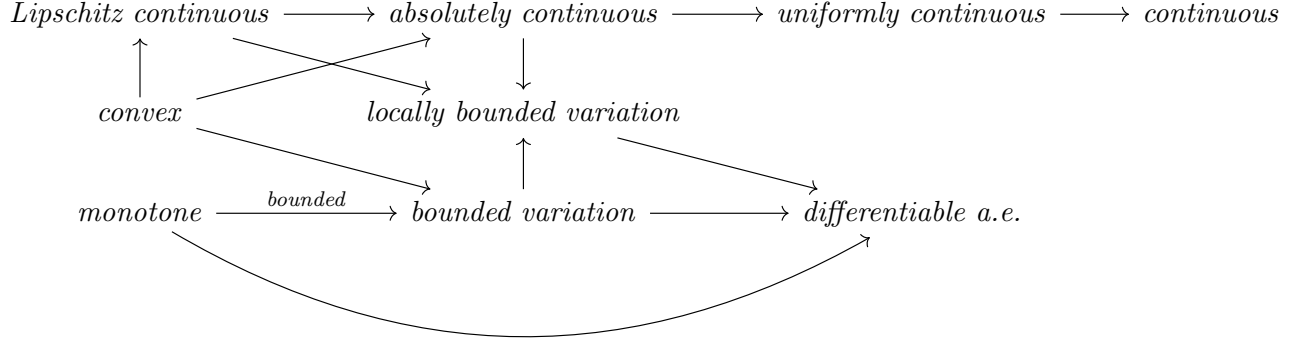
    – **uniformly continuous $\not\to$ absolutely continuous**: See Cantor function example [Tao, 2011].

    – **absolutely continuous $\not\to$ Lipschitz continuous**: $x \mapsto \sqrt{x}$

- **Proposition 4.17** *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 4.18 *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 4.19 (Second Fundamental Theorem for Absolutely Continuous Functions).
  *Let $F : [a, b] \to \mathbb{R}$ be **absolutely continuous**. Then*

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

- **Proposition 4.20** *(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 4.21** *(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 4.22** *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 derivative 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.