Lecture 0: Summary (part 1)

Tianpei Xie

Nov. 14th., 2022

Contents

1 Topology **3**
1.1 Set Theory Basis . . . 3
1.2 Topological Space . . . 5
1.3 Subspaces and Quotients . . . 6
1.4 Connectedness and Compactness . . . 7

2 Analytical Structure of Subsets **8**
2.1 The Limits of Sets . . . 8
2.2 Boolean Algebra . . . 9
2.3 σ-Algebra . . . 11
2.4 Borel σ-Algebra . . . 14
2.5 Comparison between Analytical and Topological Structure of Subsets . . . 16

3 Abstract Measure Theory **17**
3.1 Measure as Non-Negative Function on Algebra of Subsets . . . 17
3.2 Finitely Additive Measure . . . 17
3.3 Countably Additive Measure . . . 19
3.4 Outer Measures and the Carathéodory Extension Theorem . . . 22
3.5 Development of Lebesgue Measure Theory in $\mathbb{R}^d$ . . . 24

4 Measurable Functions and Integration on a Measure Space **25**
4.1 Measurable Functions . . . 25
4.2 Simple Integral of Simple Functions . . . 26
4.3 Unsigned Integral of Measurable Functions . . . 28
4.4 Absolutely Convergent Integral . . . 30
4.5 The Convergence Theorems . . . 31
4.6 Development of Integration of Measurable Functions . . . 34

5 Littlewood's Three Principles **35**
5.1 Every Measurable Set is Nearly a Finite Sum of Intervals . . . 35
5.2 Every Absolutely Integrable Function is Nearly Continuous . . . 35
5.3 Every Pointwise Convergent Sequence of Functions is Nearly Uniformly Convergent . 36

6 Modes of Convergence **37**
6.1 Convergence of Functions in Measure Space . . . 37

6.2 Modes of Convergence via Tail Support and Width . . . 39
6.2.1 Comparison . . . 41
6.3 Relationships between Different Modes of Convergence . . . 42
6.4 Counter Examples . . . 43
6.5 Uniqueness . . . 44
6.6 Modes of Convergence for Step Functions . . . 45
6.7 Modes of Convergence With Additional Conditions . . . 46
6.7.1 Finite Measure Space . . . 46
6.7.2 Fast L^1 Convergence . . . 46
6.7.3 Domination and Uniform Integrability . . . 47
6.8 Convergence in Distribution . . . 48

1 Topology

1.1 Set Theory Basis

- **Definition** Given a set X, the collection of all subsets of X, denoted as 2^X, is defined as

$$2^X := \{E : E \subseteq X\}$$

- **Remark** The followings are basic operation on 2^X: For $A, B \in 2^X$,

  1. ***Inclusion***: $A \subseteq B$ if and only if $\forall x \in A$, $x \in B$.
  2. ***Union***: $A \cup B = \{x : x \in A \vee x \in B\}$.
  3. ***Intersection***: $A \cap B = \{x : x \in A \wedge x \in B\}$.
  4. ***Difference***: $A \setminus B = \{x : x \in A \wedge x \notin B\}$.
  5. ***Complement***: $A^c = X \setminus A = \{x : x \in X \wedge x \notin A\}$.
  6. ***Symmetric Difference***: $A \Delta B = (A \setminus B) \cup (B \setminus A) = \{x \in X : x \notin A \vee x \notin B\}$.

  We have ***deMorgan's laws***:

$$\left(\bigcup_{a\in A} U_a\right)^c = \bigcap_{a\in A} U_a^c, \quad \left(\bigcap_{a\in A} U_a\right)^c = \bigcup_{a\in A} U_a^c$$

- **Remark** Note that the following equality is useful:

$$A \Delta B = (A \cup B) \setminus (A \cap B)$$

- **Definition** ***An equivalence relation*** on X is a relation R on X such that

  1. (***Reflexivity***): xRx for all $x \in X$;
  2. (***Symmetry***): xRy if and only if yRx for all $x, y \in X$;
  3. (***Transitivity***): xRy and yRz then xRz for all $x, y, z \in X$.

  The equivalence class of an element x is denoted as $[x] := \{y \in X : xRy\}$. We usually denote the equivalence relation R as $\sim$. The set of equivalence classes provides ***a partition of the set*** X in that every $z \in X$ can must belong to *only one equivalence class* $[x]$. That is $[x] \cap [y] = \emptyset$ if $x \not\sim y$ and $X = \bigcup_{x\in X}[x]$.

  The set of all equivalence classes of X by $\sim$, denoted $X/\!\sim := \{[x] : x \in X\}$, is ***the quotient set*** of X by $\sim$. $X = \bigcup_{C\in X/\sim} C$.

- **Definition** $f : X \to Y$ is a ***function*** if for each $x \in X$, there exists a unique $y = f(x) \in Y$. X is called the ***domain*** of f and Y is called the ***codomain*** of f. $f(X) = \{y \in Y : y = f(x)\}$ is called the ***range*** of f

  The ***pre-image*** of f is defined as

$$f^{-1}(E) = \{x \in X : f(x) \in E\}.$$

- **Remark** The pre-image operation ***commutes*** with ***all basic set operations***:

$$
\begin{aligned}
A \subseteq B &\Rightarrow f^{-1}(A) \subseteq f^{-1}(B) \\
f^{-1}\left(\bigcup_{\alpha\in A} E_\alpha\right) &= \bigcup_{\alpha\in A} f^{-1}(E_\alpha) \\
f^{-1}\left(\bigcap_{\alpha\in A} E_\alpha\right) &= \bigcap_{\alpha\in A} f^{-1}(E_\alpha) \\
f^{-1}(A \setminus B) &= f^{-1}(A) \setminus f^{-1}(B) \\
f^{-1}(E^c) &= \left(f^{-1}(E)\right)^c
\end{aligned}
$$

- **Remark** The image operation ***commutes*** with only ***inclusion*** *and* ***union*** *operations*:

$$
\begin{aligned}
A \subseteq B &\Rightarrow f(A) \subseteq f(B) \\
f\left(\bigcup_{\alpha\in A} E_\alpha\right) &= \bigcup_{\alpha\in A} f(E_\alpha)
\end{aligned}
$$

  For the other operations:

$$
\begin{aligned}
f\left(\bigcap_{\alpha\in A} E_\alpha\right) &\subseteq \bigcap_{\alpha\in A} f(E_\alpha) \\
f(A \setminus B) &\supseteq f(A) \setminus f(B)
\end{aligned}
$$

- **Definition** A map $f : X \to Y$ is ***surjective, or, onto***, if for every $y \in Y$, there exists a $x \in X$ such that $y = f(x)$. In set theory notation:

$$f : X \to Y \text{ is surjective } \Leftrightarrow f^{-1}(Y) \subseteq X.$$

  A map $f : X \to Y$ is ***injective***, if for every $x_1 \neq x_2 \in X$, their map $f(x_1) \neq f(x_2)$, or equivalently, $f(x_1) = f(x_2)$ only if $x_1 = x_2$.

  If a map $f : X \to Y$ is both *surjective* and *injective*, we say f is a ***bijective***, or there exists an ***one-to-one correspondence*** between X and Y. Thus $Y = f(X)$.

- **Remark**

$$
\begin{aligned}
f^{-1}(f(B)) &\supseteq B, \quad \forall B \subseteq X \\
f(f^{-1}(E)) &\subseteq E, \quad \forall E \subseteq Y \\
f : X \to Y \text{ is surjective } &\Leftrightarrow f^{-1}(Y) \subseteq X. \\
&\Rightarrow f(f^{-1}(E)) = E. \\
f : X \to Y \text{ is injective } &\Rightarrow f^{-1}(f(B)) = B \\
&\Rightarrow f\left(\bigcap_{\alpha\in A} E_\alpha\right) = \bigcap_{\alpha\in A} f(E_\alpha) \\
&\Rightarrow f(A \setminus B) = f(A) \setminus f(B)
\end{aligned}
$$

- **Proposition 1.1** *The following statements for composite functions are true:*

  1. *If f, g are both injective, then $g \circ f$ is injective.*
  2. *If f, g are both surjective, then $g \circ f$ is surjective.*
  3. *Every* ***injective*** *map $f : X \to Y$ can be writen as $f = \iota \circ f_R$ where $f_R : X \to f(X)$ is a* ***bijective*** *map and ι is the* ***inclusion map****.*
  4. *Every* ***surjective*** *map $f : X \to Y$ can be writen as $f = f_p \circ \pi$ where $\pi : X \to (X/\sim)$ is* ***a quotient map*** *(projection $x \mapsto [x]$) for the equivalent relation $x \sim y \Leftrightarrow f(x) = f(y)$ and $f_p : (X/\sim) \to Y$ is defined as $f_p([x]) = f(x)$* ***constant*** *in each coset $[x]$.*
  5. *If $g \circ f$ is* ***injective****, then f is* ***injective****.*
  6. *If $g \circ f$ is* ***surjective****, then g is* ***surjective****.*

- **Principle 1.2** *(**The Axiom of Choice**).*
  If $\{X_\alpha\}_{\alpha \in A}$ is a nonempty collection of nonempty sets, then $\prod_{\alpha \in A} X_\alpha$ is non-empty.

- **Corollary 1.3** *If $\{X_\alpha\}_{\alpha \in A}$ is a* ***disjoint*** *collection of nonempty sets, there is a set $Y \subset \bigcup_{\alpha \in A} X_\alpha$ such that $Y \cap X_\alpha$ contains* ***precisely one element*** *for each $\alpha \in A$.*

1.2 Topological Space

- **Definition** Let X be a set. *A* ***topology*** on X is *a collection $\mathscr{T}$* of *subsets* of X, called ***open subsets***, satisfying

  1. X and $\emptyset$ are *open*.
  2. The ***union*** of ***any family*** of open subsets is open.
  3. The ***intersection*** of *any* ***finite*** *family* of open subsets is open.

  A pair $(X, \mathscr{T})$ consisting of a set X together with a topology $\mathscr{T}$ on X is called ***a topological space***.

- **Definition** A map $F : X \to Y$ is said to be ***continuous*** if for every open subset $U \subseteq Y$, the ***preimage*** $F^{-1}(U)$ is ***open*** in X.

- **Definition** A ***continuous bijective*** map $F : X \to Y$ with ***continuous inverse*** is called a ***homeomorphism***. If there exists a *homeomorphism* from X to Y, we say that X and Y are ***homeomorphic***.

- **Definition** A map $F : X \to Y$ (continuous or not) is said to be ***an open map*** if for every *open* subset $U \subseteq X$, the image set $F(U)$ is *open* in Y, and ***a closed map*** if for every *closed* subset $K \subseteq X$, the image $F(K)$ is *closed* in Y .

- **Definition** A topological space X is said to be a ***Hausdorff space*** if for every pair of ***distinct*** points $p, q \in X$, there exist ***disjoint open subsets*** $U, V \subseteq X$ such that $p \in U$ and $q \in V$.

- **Definition** Suppose X is a topological space. A collection $\mathscr{B}$ of open subsets of X is said to be ***a basis*** for *the topology of X* (plural: ***bases***) if every open subset of X is the *union of some collection of elements* of $\mathscr{B}$.

More generally, suppose X is merely a set, and $\mathscr{B}$ is a collection of *subsets* of X satisfying the following conditions:

1. $X = \bigcup_{B \in \mathscr{B}} B$.
2. If $B_1, B_2 \in \mathscr{B}$ and $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathscr{B}$ such that $x \in B_3 \subseteq B_1 \cap B2$.

Then *the collection of* ***all unions*** *of elements of* $\mathscr{B}$ is a *topology* on X, called ***the topology generated by*** $\mathscr{B}$, and $\mathscr{B}$ is a ***basis*** for this *topology*.

- **Definition** See the following definitions

  1. A set is said to be ***countably infinite*** if it admits a *bijection* with the set of *positive integers*, and
  2. ***countable*** if it is *finite* or *countably infinite*.
  3. A topological space X is said to be ***first-countable*** if there is a ***countable neighborhood basis*** at each point, and
  4. ***second-countable*** if there is ***a countable basis*** for its topology.

1.3 Subspaces and Quotients

- **Definition** If X is a topological space and $S \subseteq X$ is an arbitrary subset, we define ***the subspace topology*** on S (sometimes called *the* ***relative topology***) by declaring a subset $U \subseteq S$ to be *open* in S *if and only* if there exists an open subset $V \subseteq X$ such that $U = V \cap S$.

  Any subset of X endowed with the subspace topology is said to be ***a subspace of*** X.

- **Definition** If X and Y are topological spaces, a continuous injective map $F : X \to Y$ is called a ***topological embedding*** if it is a ***homeomorphism*** onto its image $F(X) \subseteq Y$ in the subspace topology.

- **Definition** If X is a *topological space*, Y is a set, and $\pi : X \to Y$ is a **surjective** map, ***the quotient topology*** on Y determined by π is defined by declaring a subset $U \subseteq Y$ to be *open if and only* if $\pi^{-1}(U)$ is *open* in X.

  If X and Y are topological spaces, a map $\pi : X \to Y$ is called ***a quotient map*** if it is ***surjective*** and ***continuous*** and Y has the quotient topology determined by π.

- **Definition** The following construction is the most common way of producing quotient maps. *A* ***relation*** on a set X is called ***an equivalence relation*** if it is

  1. ***reflexive***: $x \sim x$ for all $x \in X$,
  2. ***symmetric***: $x \sim y$ implies $y \sim x$,
  3. ***transitive***: $x \sim y$ and $y \sim z$ imply $x \sim z$.

  If $R \subseteq X \times X$ is any ***relation*** on X, then ***the intersection of all equivalence relations*** on X ***containing*** R is *an equivalence relation*, called ***the equivalence relation generated by*** R.

  Remark If is an equivalence relation on X, then for each $x \in X$, ***the equivalence class*** *of*

x, denoted by $[x]$, is the *set of all* $y \in X$ *such that* $y \sim x$. The set of *all equivalence classes* is a ***partition*** of X: a collection of disjoint nonempty subsets whose union is X.

- **Definition** Suppose X is a topological space and $\sim$ is an equivalence relation on X. Let $X/\sim$ denote ***the set of equivalence classes*** in X, and let $\pi : X \to X/\sim$ be the ***natural projection*** sending each *point* to its *equivalence class*. Endowed with ***the quotient topology*** determined by π, the space $X/\sim$ is called ***<u>the quotient space</u>*** (or *identification space*) of X determined by π.

- **Definition** If $\pi : X \to Y$ is a map, a subset $U \subseteq X$ is said to be ***saturated*** with respect to π if U is the ***entire preimage*** of its ***image***: $U = \pi^{-1}(\pi(U))$.

  Given $y \in Y$, the ***<u>fiber</u>*** of π over y is the set $\pi^{-1}(y)$.

1.4 Connectedness and Compactness

- **Definition** A topological space X is said to be ***disconnected*** if it has two ***disjoint nonempty open subsets*** whose union is X, and it is ***connected*** otherwise. Equivalently, X is connected if and only if the only subsets of X that are ***both open and closed*** are $\emptyset$ and X itself.

- **Definition** Recall that a topological space X is

  - ***<u>connected</u>*** if there do not exist two *disjoint*, *nonempty*, *open* subsets of X whose union is X;
  - ***<u>path-connected</u>*** if every pair of points in X can be ***joined by a path*** in X, and
  - ***locally path-connected*** if X has a ***basis*** of *path-connected open subsets.*

- **Definition** A ***maximal connected subset*** of X (i.e., a connected subset that is not properly contained in any larger connected subset) is called a ***component*** (or ***connected component***) of X.

- **Definition** A topological space X is said to be ***<u>compact</u>*** if every open cover of X has a ***finite*** *subcover*. A ***compact subset*** of a topological space is one that is a compact space in the subspace topology.

- **Definition** If X and Y are topological spaces, a map $F : X \to Y$ (continuous or not) is said to be ***proper*** if for every ***compact*** set $K \subseteq Y$, the ***preimage*** $F^{-1}(K)$ is ***compact***.

- **Definition** A topological space X is said to be ***<u>locally compact</u>*** if every point has a *neighborhood* contained in a ***compact subset*** of X.

  A subset of X is said to be ***precompact*** in X if its ***closure*** in X is *compact.*

- For a ***Hausdorff space*** X, the following are equivalent:

  1. X is ***locally compact***.
  2. Each point of X has a ***precompact*** neighborhood.
  3. X has a basis of ***precompact*** open subsets.

2 Analytical Structure of Subsets

2.1 The Limits of Sets

- **Definition** A ***nested*** sequence of sets $E_1, E_2, \ldots$ is ***nondecreasing*** if $E_i \subseteq E_{i+1}$, and it is ***nonincreasing*** if $E_i \supseteq E_{i+1}$.

- **Definition** The ***infimum*** and the ***supremum*** of a collection of sets $\{E_n\}_{n \ge k}$ is given by

$$\inf_{n \ge k} E_n = \bigcap_{n=k}^{\infty} E_n, \quad \sup_{n \ge k} E_n = \bigcup_{n=k}^{\infty} E_n,$$

  respectively.

- **Remark** Note that

  1. $\inf_{n \ge 1} E_n, \ldots, \inf_{n \ge k} E_n, \ldots$ is ***monotone increasing*** as k increases since

  $$\inf_{n \ge k} E_n \subseteq \inf_{n \ge k+1} E_n.$$

  The ***more*** sets that are involved in the ***intersection***, *the* ***less*** *cardinality of* the intersection will be. As k increases, *less* sets are involved in the intersection.

  2. $\sup_{n \ge 1} E_n, \ldots, \sup_{n \ge k} E_n, \ldots$ is ***monotone decreasing***. as k increases since

  $$\sup_{n \ge k} E_n \supseteq \sup_{n \ge k+1} E_n.$$

  The ***more*** sets that are involved in the ***union***, *the* ***more*** *cardinality of* the union will be. As k increases, *less* sets are involved in the union.

- **Definition** [Resnick, 2013]
  The ***limit infimum*** and ***limit supremum*** is defined as

$$\liminf_{n \to \infty} E_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n, \quad \limsup_{n \to \infty} E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n, \tag{1}$$

  respectively.

- **Remark** It is clear that for ***nested sequence*** $\{E_n\}_{n \ge 1}$ that is ***nondecreasing***,

$$\liminf_{n \to \infty} E_n = \bigcup_{n=1}^{\infty} E_n = \limsup_{n \to \infty} E_n$$

  so define *the* ***limit*** *of monotone increasing nested sets* as $\lim_{n \to \infty} E_n = \bigcup_{n=1}^{\infty} E_n$.

  Similarly, for ***nonincreasing nested sets*** $\{E_n\}_{n \ge 1}$,

$$\liminf_{n \to \infty} E_n = \bigcap_{n=1}^{\infty} E_n = \limsup_{n \to \infty} E_n$$

  so define *the* ***limit*** *of monotone decreasing nested sets* as $\lim_{n \to \infty} E_n = \bigcap_{n=1}^{\infty} E_n$.

- **Proposition 2.1** *The **interpretation** of limit infimum and limit supremum*

$$\liminf_{n\to\infty} E_n = \{x:\ x \in E_n,\ \textit{for } \textbf{\textit{all but finite}}\ n\} = \{x:\ \exists k, \forall n \ge k,\ x \in E_n\}$$
$$\limsup_{n\to\infty} E_n = \{x:\ x \in E_n,\ \textit{for } \textbf{\textit{infinitely many}}\ n\} = \{x:\ \exists k, \forall n \ge k,\ x \in E_n\}$$

- **Remark** A sequence of sets $E_1, E_2, \ldots$ converges to E if and only if

$$\liminf_{n\to\infty} E_n = \limsup_{n\to\infty} E_n = E.$$

It is equivalent to ***pointwise*** convergence of indicator function $\mathbb{1}\{x \in E_n\}$ to $\mathbb{1}\{x \in E\}$ for each x.

2.2 Boolean Algebra

- **Definition** [Tao, 2011]
Let X be a set. *A (concrete)* ***Boolean algebra (Boolean field)*** on X is a *collection of subsets* $\mathscr{B}$ of X which obeys the following properties:

  1. (***Empty set***) $\emptyset \in \mathscr{B}$;
  2. (***Complements***) For any $E \in \mathscr{B}$, then $E^c \equiv (X \setminus E) \in \mathscr{B}$;
  3. (***Finite unions***) For any $E, F \subset \mathscr{B}$, $E \cup F \in \mathscr{B}$.

  We sometimes say that E is ***$\mathscr{B}$-measurable***, or ***measurable with respect to*** $\mathscr{B}$, if $E \in \mathscr{B}$.

- **Remark** Note that *the finite difference* $A - B$, $A\Delta B$ and *intersections* $A \cap B$ are also ***closed*** under the Boolean algebra.

- **Definition** A ***field (algebra)*** is a *non-empty collection of subsets* in X that is ***closed*** under ***finite union*** and ***complements***.

  It is just a *subset (sub-algebra)* of *Boolean field* $(X, \subset, \cup, \cdot^c)$.

- **Definition** Given two Boolean algebras $\mathscr{B}, (\mathscr{B})'$ on X, we say that $(\mathscr{B})'$ is ***finer*** than, a ***sub-algebra*** of, or a ***refinement*** of $\mathscr{B}$, or that $\mathscr{B}$ is ***coarser*** than or a ***coarsening*** of $(\mathscr{B})'$, if $\mathscr{B} \subset (\mathscr{B})'$.

- **Remark** In ***abstract Boolean algebra***, $\cup$ is replaced by *join* operation $\vee$ and $\cap$ is replaced by *meet* operation $\wedge$.

- **Remark** The definition of Boolean algebra ***does not requires*** X to have a ***topology***. It focus on a collection of subsets that is ***closed*** under *the set union operation* $\cup$ and the set complement $\cdot^c$. In other words, the concerns is the ***set-algebraic property*** not the topological property. Note that the set intersection operation $\cap$ can be obtained from composite of set union and set complement operations.

- **Definition** [Tao, 2011]
Let X be *partitioned* into a union $X = \bigcup_{\alpha \in I} A_\alpha$ of ***disjoint sets*** A_α, which we refer to as ***atoms***. Then this partition *generates* ***a Boolean algebra*** $\mathscr{A}((A_\alpha)_{\alpha\in I})$, defined as *the collection of all the sets E of the form* $E = \bigcup_{\alpha\in J} A_\alpha$ for some $J \subseteq I$, i.e. $\mathscr{A}((A_\alpha)_{\alpha\in I})$ is *the*

collection of all sets that can be represented as ***the union of one or more atoms***. Then $\mathscr{A}((A_\alpha)_{\alpha\in I})$ is ***a Boolean algebra***, and we refer to it as the ***atomic algebra*** with *atoms* $(A_\alpha)_{\alpha\in I}$.

- **Definition** A Boolean algebra is ***finite*** if it only consists of *finite many of subsets* (i.e., its *cardinality* is finite) [Tao, 2011].

- **Remark** The definition of ***atomic algebra*** as *generated* by ***atoms*** resembles the definition of ***topology*** *generated* by ***basis***.
  - In both cases, a subset in the collection of *atomic algebra / topology* is seen as the ***union*** of some subsets in the *atoms / basis*.
  - On the other hand, ***atoms*** *are all* ***disjoint***, while *sets in* ***basis*** are ***not necessarily disjoint***. In fact, by definition, for any two sets in basis that have nonempty intersection, there must exists a third set in basis that is a subset of the intersection.

- **Example** The followings are examples of *Boolean algebra*:
  1. ***The trivial algebra*** $\{X, \emptyset\}$ is *atomic algebra* with atoms $\{X\}$.
  2. ***The discrete algebra*** 2^X is *atomic algebra* generated by collection of ***singletons*** $\{x\}$.

- **Remark** The *non-empty atoms* of an *atomic algebra* are determined up to ***relabeling***. More precisely, if $X = \bigcup_{\alpha\in I} A_\alpha = \bigcup_{\alpha'\in I'} A'_{\alpha'}$ are two partitions of X into non-empty atoms A_α, $A'_{\alpha'}$, then $\bigcup_{\alpha\in I} A_\alpha = \bigcup_{\alpha'\in I'} A'_{\alpha'}$ if and only if exists a ***bijection*** $\phi : \alpha \to \alpha'$ such that $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$. [Tao, 2011]

- **Remark** There is a ***one-to-one correspondence*** between ***finite Boolean algebras*** on X and ***finite partitions*** of X into non-empty sets. (its cardinality is 2^m, for some m). [Tao, 2011]

- **Definition** [Tao, 2011]
  Let n be an integer. The ***dyadic algebra*** $\mathscr{D}_n$ at scale 2^{-n} in $\mathbb{R}^d$ is defined to be the atomic algebra generated by the *half-open dyadic cubes*
  $$\left[\frac{i_1}{2^n}, \frac{i_1+1}{2^n}\right) \times \cdots \left[\frac{i_d}{2^n}, \frac{i_d+1}{2^n}\right)$$
  of length 2^{-n}. Note that $\mathscr{D}_n \subset \mathscr{D}_{n+1}$.

- **Example** Here are some more examples for Boolean algebra [Tao, 2011]
  1. The collection $\overline{\mathcal{E}[\mathbb{R}^d]}$ of ***elementary sets*** (boxes and its finite union and intersections) and co-elementary sets (its complements is elementary) in $\mathbb{R}^d$ forms a Boolean algebra.
  2. The collection $\overline{\mathcal{J}[\mathbb{R}^d]}$ of ***Jordan measureable set*** (contained in finite union of elementary sets) and co-Jordan measureable sets in $\mathbb{R}^d$ forms a Boolean algebra.
  3. The collection $\mathcal{L}[\mathbb{R}^d]$ of ***Lebesgue measureable set*** (contained in countable union of elementary sets) in $\mathbb{R}^d$ forms a Boolean algebra.
  4. The collection $\mathcal{N}[\mathbb{R}^d]$ of ***Lebesgue null sets*** and ***co-null sets*** (its complement is null set) in $\mathbb{R}^d$ forms a Boolean algebra. we refer to it as ***the null algebra*** on $\mathbb{R}^d$.
  5. Given $Y \subset X$, and $\mathscr{B}$ is a Boolean algebra on X, then the ***restriction*** of algebra on Y

is $\mathscr{B}|_Y = \mathscr{B} \cap 2^Y = \{E \cap Y : E \in \mathscr{B}\}$, which is a *sub-algebra*.

6. The ***dyadic algebra*** $\mathscr{D}_n$ at ***scale*** 2^{-n} in $\mathbb{R}^d$ is defined to be ***the atomic algebra*** generated by the *half-open dyadic cubes* of length 2^{-n}.
7. Note that $\left\{\emptyset, \mathbb{R}^d\right\} \subset \mathscr{D}_n \subset \overline{\mathcal{E}[\mathbb{R}^d]} = \bigcup_{n\geq 1} \mathscr{D}_n \subset \overline{J[\mathbb{R}^d]} \subset L[\mathbb{R}^d] \subset 2^{\mathbb{R}^d}$. $N[\mathbb{R}^d] \subset L[\mathbb{R}^d]$. Although $\mathscr{D}_n$ for given n is atomic algebra, $\overline{\mathcal{E}[\mathbb{R}^d]}$ and all its predecessors are ***non-atomic***, since they do not have finite cardinality.
8. $\bigwedge_{\alpha\in I} \mathscr{B}_\alpha \equiv \bigcap_{\alpha \in I} \mathscr{B}_\alpha$ for all $\alpha \in I$ is a Boolean algebra (I is arbitrary), which is ***the finest algebra*** that is ***coarser*** than any $\mathscr{B}_\alpha$.

- **Example** (***Boolean Algebra Generated by*** $\mathcal{F}$)
  - **Definition** Given a collection of sets $\mathcal{F}$, then $\langle \mathcal{F}\rangle_{bool}$ is ***the Boolean algebra generated*** by $\mathcal{F}$, i.e. the ***intersection*** of all the Boolean algebras that *contain* $\mathcal{F}$.

$$\langle \mathcal{F}\rangle_{bool} = \bigwedge_{\mathscr{B}_\alpha \supseteq \mathcal{F}} \mathscr{B}_\alpha.$$

  - **Proposition 2.2** *We have the following results regarding* $\langle \mathcal{F}\rangle_{bool}$
    1. $\langle \mathcal{F}\rangle_{bool}$ *is the* ***coarest*** *Boolean algebra that contains* $\mathcal{F}$.
    2. *Note that* $\mathcal{F}$ *is a Boolean algebra if and only if* $\mathcal{F} = \langle \mathcal{F}\rangle_{bool}$.
    3. *If* $\mathcal{F}$ *is collection of* n ***sets****, then* $\langle \mathcal{F}\rangle_{bool}$ *is* ***a finite Boolean algebra*** *with cardinality* 2^{2^n}.
  - **Proposition 2.3** *(****Recursive description of a generated Boolean algebra****). [Tao, 2011]*
    Let $\mathcal{F}$ *be a collection of sets in a set* X*. Define the sets* $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \ldots$ ***recursively*** *as follows:*
    1. $\mathcal{F}_0 := \mathcal{F}$.
    2. *For each* $n \geq 1$*, we define* $\mathcal{F}_n$ *to be the collection of all sets that* ***either*** *the* ***union*** *of a* ***finite number*** *of sets in* $\mathcal{F}_{n-1}$ *(including the empty union* $\emptyset$*), or the* ***complement*** *of such a union.*

    Then $\langle \mathcal{F}\rangle_{bool} = \bigcup_{n=0}^{\infty} \mathcal{F}_n$.

2.3 σ-Algebra

- **Definition** Given space X, a ***σ-field*** (or, ***σ-algebra***) $\mathscr{F}$ is a non-empty collection of *subsets* in X such that
  1. $\emptyset \in \mathscr{F}$; $X \in \mathscr{F}$;
  2. ***Complements***: For any $B \in \mathscr{F}$, then $B^c \equiv (X - B) \in \mathscr{F}$;

3. ***Countable union***: for any sub-collection $\{B_k\}_{k=1}^{\infty} \subset \mathscr{F}$,
$$\bigcup_{k=1}^{\infty} B_k \in \mathscr{F};$$
Also, ***Countable intersection***: $\bigcap_{k=1}^{\infty} B_k \in \mathscr{F}$, ***de Morgan's law***.

We refer to the pair $(X, \mathscr{F})$ of a set X together with a σ-algebra on that set as ***a measurable space***.

- **Remark** The prefix σ usually denotes "***countable union***". Other instances of this prefix include a ***σ-compact topological space*** (*a countable union of compact sets*), a ***σ-finite measure space*** (*a countable union of sets of finite measure*), or F_σ ***set*** (*a countable union of closed sets*) for other instances of this prefix.

- **Remark** A σ-algebra can be ***equivalently*** defined as an algebra that is closed under ***countable disjoint union***. Using the following transformation, for given $\{E_j\}$,
$$F_j = E_j - \bigcup_{i=1}^{j-1} E_i, \forall\, j \in \mathbb{N}.$$
Then $F_i \cap F_j = \emptyset, i \neq j$ and $\bigcup_{j=1}^{\infty} E_j = \bigcup_{j=1}^{\infty} F_j$.

- **Remark** A field (algebra) may not be a σ-field since it *may not be* ***closure*** *under* ***countable union***.

- **Remark** ***(σ-Algebra vs. Boolean Algebra)***

  1. **Proposition 2.4** *Any σ-algebra is Boolean-algebra.*
  2. **Proposition 2.5** *Any* ***atomic algebra*** *is σ-algebra.*
  3. **Proposition 2.6** *An algebra of* ***finite*** *set X is a σ-algebra of X and it is* ***the power set*** *2^X itself.*

- **Example** Here are some more examples for σ-algebra [Tao, 2011]

  1. ***The trivial algebra*** $\{X, \emptyset\}$ is *σ-algebra* since it is an atomic algebra.
  2. ***The discrete algebra*** 2^X is *σ-algebra* since it is an atomic algebra.
  3. All the ***finite Boolean algebra*** is σ-algebra.
  4. The ***dyadic algebra*** $\mathscr{D}_n$ at ***scale*** 2^{-n} in $\mathbb{R}^d$ is a ***σ-algebra*** since it is an atomic algebra.
  5. The collection $\mathcal{L}[\mathbb{R}^d]$ of ***Lebesgue measureable set*** (contained in countable union of *elementary sets*) in $\mathbb{R}^d$ forms a Boolean algebra.
  6. The collection $\mathcal{N}[\mathbb{R}^d]$ of ***Lebesgue null sets*** and ***co-null sets*** (its complement is null set) in $\mathbb{R}^d$ forms a Boolean algebra. we refer to it as ***the null algebra*** on $\mathbb{R}^d$.
  7. Given $Y \subset X$ as a subspace of X, and $\mathscr{B}$ is a σ-algebra on X, then the ***restriction*** of algebra on Y is $\mathscr{B}|_Y = \mathscr{B} \cap 2^Y = \{E \cap Y : E \in \mathscr{B}\}$, which is a ***$\sigma$-algebra on subspace*** Y.
  8. Note that both the collections of *elementary sets* $\mathcal{E}[\mathbb{R}^d]$ and ***the Jordan measurable sets*** $\mathcal{J}[\mathbb{R}^d]$ ***do not form a σ-algebra***.

9. If $\{\mathscr{B}_\alpha\}$ are σ-algebras, then $\bigwedge_{\alpha\in I}\mathscr{B}_\alpha \equiv \bigcap_{\alpha\in I}\mathscr{B}_\alpha$ for all $\alpha \in I$ is a σ-algebra (I is arbitrary), which is ***the finest σ-algebra*** that is ***coarser*** than any $\mathscr{B}_\alpha$.

- **Example** (***σ-Algebra Generated by $\mathcal{F}$***)
  - **Definition** Denote $\sigma(\mathcal{F}) := \langle\mathcal{F}\rangle$ as ***the σ-algebra generated by $\mathcal{F}$***, given by

$$\sigma(\mathcal{F}) = \langle\mathcal{F}\rangle = \bigwedge_{\mathscr{B}_\alpha \supseteq \mathcal{F}} \mathscr{B}_\alpha.$$

    It is the ***coarsest*** σ-algebra containing $\mathcal{F}$, for any σ-algebra that contains $\mathcal{F}$.
  - It is easy to see that

$$\langle\mathcal{F}\rangle_{bool} \subseteq \langle\mathcal{F}\rangle$$

    The equality holds if and only if $\langle\mathcal{F}\rangle_{bool}$ is a σ-algebra.
  - **Proposition 2.7** *(**Recursive description of a generated σ-algebra**). [Tao, 2011] $\sigma(\mathcal{F})$ is generated according to the following procedure:*
    1. *For every set $A \in \mathcal{F}$, $A \in \sigma(\mathcal{F})$; $\mathcal{F} \subset \sigma(\mathcal{F})$;*
    2. *Take the **finite union** and **finite intersection** of any **subcollections** $\{A_k\} \subset \mathcal{F}$, put $\bigcup_{k=1}^n A_k \in \sigma(\mathcal{F}), n \geq 1$ and $\bigcap_{k=1}^n A_k \in \sigma(\mathcal{F}), n \geq 1$;*
    3. *Put the **countably infinite union** and **intersections** of any **subcollections** $\{A_k\} \subset \mathcal{F}$, put $\bigcup_{k=1}^\infty A_k \in \sigma(\mathcal{F})$ and $\bigcap_{k=1}^\infty A_k \in \sigma(\mathcal{F})$;*
    4. *Put the **complements** $A^c \in \sigma(\mathcal{F}), \forall A \in \sigma(\mathcal{F})$;*
  - Finally we have the ***monotonicity***:
    1. **Proposition 2.8** *If $\mathcal{F}_1 \subset \mathcal{F}_2$, then $\sigma(\mathcal{F}_1) \subset \sigma(\mathcal{F}_2)$.*
    2. **Proposition 2.9** *If $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \sigma(\mathcal{F}_1)$, then $\sigma(\mathcal{F}_2) = \sigma(\mathcal{F}_1)$.*
    3. **Proposition 2.10** *Let $\mathscr{F}$ be a σ-algebra on a set X. Let $S \subset X$ be a subset of X.*

       Then

$$\sigma(\mathscr{F} \cup \{S\}) = \{(E_1 \cap S) \cup (E_2 \cap S^c) : E_1, E_2 \in \mathscr{F}\}$$

       *where σ denotes **generated σ-algebra**.*

- **Remark** Note that $\mathscr{F}_1 \cup \mathscr{F}_2$ is usually not a σ-algebra.
- **Remark** We compare the (open-set) topoloy with σ-algebra:
  - ***The open-set topology*** on X is ***closed*** under ***any union***, or ***finite intersection*** *operation*. It does ***not** consider* the ***complements*** as the complements defines *a **closed set** not in open-set topology*. It contains the open sets as *the basic environment* in investigating the ***infinitesimal behavior*** of functions in ***analysis***.
  - ***A σ-algebra*** concerns more about the ***closure*** under a set of ***operations*** on X: *countable union*, *countable intersection*, ***complementation***. It has nothing to do with *the open set*, *closed set*, or the *continuity*.

- The ***analysis*** replies on ***topology*** on space X; while the ***modern algebra*** replies on ***the closure of operation*** on a space X. A σ-algebra is a collection of subsets in X that endows a ***algebraic structure***.

2.4 Borel σ-Algebra

- **Definition** (***Borel σ-algebra***). [Tao, 2011]
  Let X be a ***metric space***, or more generally ***a topological space***. The ***Borel σ-algebra*** $\mathcal{B}[X]$ of X is defined to be the σ-algebra *generated by the* **open subsets** *of* X.

  Elements of $\mathcal{B}[X]$ will be called ***Borel measurable***.

- **Example** The followings are examples of *Borel measurable subsets* in X:
  1. Any ***the open set*** and ***the closed set*** (which are *complements* of open sets), including ***The arbitrary union*** of *open sets*, and ***arbitrary intersection*** of *closed set*.
  2. The ***countable unions*** of ***closed sets*** (known as F_σ sets),
  3. The ***countable intersections*** of ***open sets*** (known as G_δ sets),
  4. The ***countable intersections*** of F_σ sets, and so forth.

- **Exercise 2.11** *Show that the Borel σ-algebra $\mathcal{B}[\mathbb{R}^d]$ of a Euclidean set is generated by any of the following collections of sets:*
  1. *The open subsets of $\mathbb{R}^d$.*
  2. *The closed subsets of $\mathbb{R}^d$.*
  3. *The compact subsets of $\mathbb{R}^d$.*
  4. *The open balls of $\mathbb{R}^d$.*
  5. *The boxes in $\mathbb{R}^d$.*
  6. *The elementary sets in $\mathbb{R}^d$.*

  (Hint: To show that two families $\mathcal{F}, \mathcal{F}'$ of sets generate the same σ-algebra, it suffices to show that every σ-algebra that contains $\mathcal{F}$, contains $\mathcal{F}'$ also, and conversely.)

- **Remark** $\mathcal{B}[X] \subset \mathcal{L}[X]$, i.e. the Borel σ-algebra is ***coarser*** than the Lebesgue σ-algebra.

- **Remark** There exist ***Jordan measurable*** (and hence Lebesgue measurable) subsets of $\mathbb{R}^d$ which are ***not Borel measurable***. [Tao, 2011]

- **Remark** Despite this demonstration that ***not all Lebesgue measurable subsets are Borel measurable***, it is *remarkably* ***difficult*** *(though not impossible)* to exhibit a specific set that is not Borel measurable. Indeed, a large majority of the explicitly constructible sets that one actually encounters in practice tend to be Borel measurable, and one can view the property of Borel measurability intuitively as a kind of "*constructibility*" property. A Borel σ-algebra is large enough to contain all subsets in X that is of "practical use" in computing measures and integrations within $(0, 1]$.

- **Proposition 2.12** ***(Lebesgue σ-algebra vs. Borel σ-algebra)***

The Lebesgue σ-algebra on $\mathbb{R}^d$ is generated by the union of the Borel σ-algebra and the null σ-algebra.

- **Remark** The ***Borel σ-algebra*** lies in between, which concerns both ***algebraic*** and ***analytical structure***.
  - A ***open set*** U is a *Borel set* in $\mathscr{B}$; also a ***closed set*** $C \equiv U^c$ is a *Borel set* in $\mathscr{B}$.
  - Any ***countable union*** *of* ***closed set***, denoted as "F_σ set", $F_{\sigma,\Lambda} = \bigcup_{\lambda \in \Lambda} C_\lambda \in \mathscr{B}$
  - Any ***countable intersection*** *of* ***open sets***, denoted as "G_δ set", $G_{\delta,\Lambda} = \bigcap_{\lambda \in \Lambda} U_\lambda \in \mathscr{B}$.
  - Note that a F_σ set is ***not closed*** (but could be open) and a G_δ set is ***not open*** *(but could be closed).*

The Borel σ-algebra contains *open sets*, *closed sets*, *G_δ sets*, *F_σ sets*, and their further *countable union and intersections*, according to the topology.

2.5 Comparison between Analytical and Topological Structure of Subsets

Table 1: Comparison between σ-algebra and topology

	Boolean Algebra	***σ-Algebra***	***Borel σ-Algebra***	***Topology***
compatibility		$\Leftarrow \checkmark \Rightarrow$	*σ-algebra generated* by ***open subsets***	*no relation*
collection of subsets	✓	✓	✓	✓
include emptyset	✓	✓	✓	✓
include fullset	✓	✓	✓	✓
finite union	✓	✓	✓	✓
countable union		✓	✓	✓
arbitrary union				✓
finite intersection	✓	✓	✓	✓
countable intersection		✓	✓	
complements	✓	✓	✓	
structure	***analytical***	***analytical***	***analytical*** & ***topological***	***topological***
related ***measure***	✓	✓	✓	
set ***in*** *collection*	***elementary sets***; ***Jordan measurable sets***; ***atomic algebra***; *dyadic algebra*; ***finite union*** of measurable sets; etc.	*Boolean measurable set*; ***Lebesgue measurable sets***, Lebesgue null sets; ***the countable union*** and complements etc.	open sets, ***closed sets***, ***compact sets***, *elementary sets*, G_δ and F_σ sets etc.	***open sets***
set ***not in*** *collection*	some ***Lebesgue measurable sets***	some ***non-measurable sets***	some Jordan measurable set but not Borel measurable	***closed set***, G_δ and F_σ sets
function	***Boolean measurable function***; *Rieman integrable function*,	***Lebesgue measurable function***, *σ-finite function*, *continuous function*	***Borel measurable function***, *continuous function*	***continuous function***

3 Abstract Measure Theory

3.1 Measure as Non-Negative Function on Algebra of Subsets

- **Remark** The concept of measure is a *generalization* of volumes from Euclidean space to arbitrary subsets in 2^X. [Tao, 2011] A set of intuitive axioms for a measure function m defined on power set $2^{\mathbb{R}}$:

  1. The ***unit length*** of interval: $E = (0, 1]$, then $m((0, 1]) = 1$;
  2. If E is ***congruent*** to F: (There exists a proper translation, rotation or reflection from E to F), then $m(E) = m(F)$;
  3. The ***countably additive***: for a countable union of disjoint sets, $\bigcup_{k=1}^{\infty} E_k$, the measure
  $$m\left(\bigcup_{k=1}^{\infty} E_k\right) = \bigcup_{k=1}^{\infty} m\left(E_k\right)$$

  Unfortunately, <u>***these three axioms are inconsistent***</u>: ***no proper definition of measure function*** m ***could satisfies all these three axioms for any subset in*** $\mathbb{R}$. The measure theory should be built on a collection of "ordinary" subsets, which motivates the introduction of σ***-algebra***.

- **Remark** The ***domain*** of measures are *confined* as *the Boolean algebra* $\mathscr{B}$ or σ*-algebra* $\mathscr{F}$ instead of all possible subsets of X. We call the tuple $(X, \mathscr{B})$ ***measurable space***, as it can be used as a domain for *some* ***measure function*** $\mu : \mathscr{B} \to [0, +\infty]$.

- **Remark** The ***measure*** μ defined on a given algebra $\mathscr{B}$ need to be ***compatible*** with the ***analytical structure*** of the ***algebra*** $\mathscr{B}$.
  - If $\mathscr{B}$ is *Boolean-algebra* (closed under *finite union*), then the measure is *finitely additive*.
  - If $\mathscr{B}$ is σ*-algebra* (closed under *countable union*), then the measure is *countably additive*.
  - If $\mathscr{B} = \mathscr{F}|_Y$ is the *restriction* of $\mathscr{F}$ on subspace Y, the corresponding measure on $\mathscr{B}$ should agree with measure on $\mathscr{F}$ for subsets of the subspace Y.
  - If $\mathscr{B} = \mathscr{A}((A_\alpha)_{\alpha \in I})$ is *atomic algebra*, the corresponding measure is also *the finite sum* of *Dirac measures* on each atom.

- **Remark** *The space of finitely additive and countably additive measures* on $\mathscr{B}$ forms ***a vector space*** as it is *closed* under measure addition and scale mutiplication operations.

3.2 Finitely Additive Measure

- **Definition** Let $\mathscr{B}$ be a *Boolean algebra* on a space X. An (*unsigned*) <u>***finitely additive measure***</u> μ on $\mathscr{B}$ is a map $\mu : \mathscr{B} \to [0, +\infty]$ that obeys the following axioms

  1. $\mu(\emptyset) = 0$;
  2. ***Finite union***: for any ***disjoint sets*** $A, B \in \mathscr{B}$,
  $$\mu\left(A \cup B\right) = \mu(A) + \mu(B).$$

- **Proposition 3.1** ***(Properties of Finitely Additive Measure)*** *[Tao, 2011]*
  Let $\mu : \mathscr{B} \to [0, +\infty]$be a finitely additive measure on a Boolean σ-algebra $\mathscr{B}$.

  1. ***(Monotonicity)*** *If E, F are $\mathscr{B}$-measurable and $E \subseteq F$, then*
  $$\mu(E) \leq \mu(F).$$
  2. ***(Finite additivity)*** *If k is a natural number, and $E_1, \ldots, E_k$ are $\mathscr{B}$-measurable and* ***disjoint****, then*
  $$\mu(E_1 \cup \ldots \cup E_k) = \mu(E_1) + \ldots + \mu(E_k).$$
  3. ***(Finite subadditivity)*** *If k is a natural number, and $E_1, \ldots, E_k$ are $\mathscr{B}$-measurable, then*
  $$\mu(E_1 \cup \ldots \cup E_k) \leq \mu(E_1) + \ldots + \mu(E_k).$$
  4. ***(Inclusion-exclusion for two sets)*** *If E, F are $\mathscr{B}$-measurable, then*
  $$\mu(E \cup F) + \mu(E \cap F) = \mu(E) + \mu(F).$$

  (Caution: remember that the cancellation law $a + c = b + c \Rightarrow a = b$ does not hold in [0; +1] if c is infinite, and so the use of cancellation (or subtraction) should be avoided if possible.)

- **Example** See the following examples on finitely additive measures:
  1. ***Lebesgue measure*** m is a *finitely additive measure* on ***the Lebesgue σ-algebra***, and hence on *all sub-algebras* (such as *the null algebra*, *the Jordan algebra*, or *the elementary algebra*).
  2. ***Jordan measure*** and ***elementary measure*** are *finitely additive* (adopting the convention that co-Jordan measurable sets have infinite Jordan measure, and co-elementary sets have infinite elementary measure).
  3. ***Lebesgue outer measure*** is ***not*** *finitely additive* on ***the discrete algebra***.
  4. ***Jordan outer measure*** is ***not*** *finitely additive* on ***the Lebesgue algebra***.

- **Example** (***Dirac measure***).
  Let $x \in X$ and $\mathscr{B}$ be an arbitrary *Boolean algebra* on X. Then ***the Dirac measure*** δ_x at x, defined by setting $\delta_x(E) := \mathbb{1}\{x \in E\}$, is ***finitely additive***.

- **Example** (***Zero measure***).
  The ***zero measure*** $0 : E \mapsto 0$ is a *finitely additive measure* on any Boolean algebra.

- **Example** (***Linear combinations of measures***).
  If $\mathscr{B}$ is a Boolean algebra on X, and $\mu, \nu : \mathscr{B} \to [0, +\infty]$ are *finitely additive measures* on $\mathscr{B}$, then $\mu + \nu : E \mapsto \mu(E) + \nu(E)$ is also a ***finitely additive measure***, as is $c\mu : E \mapsto c \times \mu(E)$ for any $c \in [0, +\infty]$. Thus, for instance, the sum of Lebesgue measure and a Dirac measure is also a finitely additive measure on the Lebesgue algebra (or on any of its sub-algebras).

  In other word, ***the space of all finitely additive measures*** *on $\mathscr{B}$ is a* ***vector space***.

- **Example** (***Restriction of a measure***).
  If $\mathscr{B}$ is a Boolean algebra on X, $\mu : \mathscr{B} \to [0, +\infty]$ is a *finitely additive measure*, and Y is a $\mathscr{B}$-measurable subset of X, then ***the restriction*** $\mu|_Y : \mathscr{B}|_Y \to [0, +\infty]$ of $\mathscr{B}$ to Y, defined by setting $\mu|_Y(E) := \mu(E)$ whenever $E \in \mathscr{B}|_Y$ (i.e. if $E \in \mathscr{B}$ and $E \subseteq Y$), is also a ***finitely additive measure***.

- **Example** (***Counting measure***).
  If $\mathscr{B}$ is a Boolean algebra on X, then the function $\# : \mathscr{B} \to [0, +\infty]$ defined by setting $\#(E)$ to be the ***cardinality*** of E if E is *finite*, and $\#(E) := +\infty$ if E is infinite, is a ***finitely additive measure***, known as ***counting measure***.

- **Proposition 3.2** ***(Finitely Additive Measures on Atomic Algebra)***
  *Let $\mathscr{B}$ be a **finite** Boolean algebra, generated by a finite family $A_1, \ldots, A_k$ of non-empty **atoms**. For every **finitely additive measure** μ on $\mathscr{B}$ there exists $c_1, \ldots, c_k \in [0, +\infty]$ such that*

$$\mu(E) = \sum_{1 \le j \le k : A_j \subseteq E} c_j.$$

  Equivalently, if x_j is a point in A_j for each $1 \le j \le k$, then

$$\mu = \sum_{j=1}^{k} c_j\, \delta_{x_j}.$$

  *where $c_1, \ldots, c_k$ are **uniquely** determined by μ.*

3.3 Countably Additive Measure

- **Definition** Let $(X, \mathscr{B})$ be a measurable space. *An (unsigned)* ***countably additive measure*** μ on $\mathscr{B}$, or ***measure*** for short, is a map $\mu : \mathscr{B} \to [0, +\infty]$ that obeys the following axioms:

  1. (***Empty set***) $\mu(\emptyset) = 0$.
  2. (***Countable additivity***) Whenever $E_1, E_2, \ldots \in \mathscr{B}$ are a ***countable sequence*** of ***disjoint*** *measurable sets*, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

  A triplet $(X, \mathscr{B}, \mu)$, where $(X, \mathscr{B})$ is a ***measurable space*** and $\mu : \mathscr{B} \to [0, +\infty]$ is a ***countably additive measure***, is known as ***a measure space***.

- **Remark** Note the distinction between a ***measure space*** and a ***measurable space***. The latter has the ***capability*** to be equipped with a *measure*, but the former is ***actually*** equipped with a *measure*.

- **Definition** [Folland, 2013]
  Let $(X, \mathscr{B}, \mu)$ be a measure space.

  - If $\mu(X) < \infty$ (which implies that $\mu(E) < \infty$ for all $E \in \mathscr{B}$), then μ is called ***finite***.

- If $X = \cup_{j=1}^{\infty} E_j$ where $E_j \in \mathscr{B}$ and $\mu(E_j) < \infty$, then μ is called ***σ-finite***. More generally, if $E = \cup_{j=1}^{\infty} E_j$ where $E_j \in \mathscr{B}$ and $\mu(E_j) < \infty$, then E is said to be ***σ-finite*** for μ.
- If for each $E \in \mathscr{B}$ with $\mu(E) = \infty$ there exists $F \in \mathscr{B}$ with $F \subseteq E$ and $0 < \mu(F) < \infty$, then μ is called ***semi-finite***.

- **Example** The followings are examples for *countably additive measures*:
  1. ***Lebesgue measure*** is a *countably additive measure* on the ***Lebesgue σ-algebra***, and hence on every sub-σ-algebra (such as the Borel σ-algebra)
  2. The ***Dirac measures*** δ_x are ***countably additive***
  3. The ***counting measure*** $\#$ is *countably additive measure.*
  4. The ***zero measure*** is *countably additive measure.*
  5. Any ***restriction*** of *a countably additive measure* to ***a measurable subspace*** is again *countably additive.*

- **Example** (***Countable combinations of measures***).
  Let $(X, \mathscr{B})$ be a *measurable space.*
  1. If μ is a *countably additive measure* on $\mathscr{B}$, and $c \in [0, +\infty]$, then $c\mu$ is also *countably additive.*
  2. If $\mu_1, \mu_2, \ldots$ are a *sequence of countably additive measures* on $\mathscr{B}$, then *the sum* $\sum_{n=1}^{\infty} \mu_n : E \mapsto \sum_{n=1}^{\infty} \mu_n(E)$ is also a *countably additive measure.*

  That is, ***the space of all countable additive measures*** *on* $\mathscr{B}$ *is a* ***vector space***.

- **Remark** Note that ***countable additivity measures are necessarily finitely additive*** (by padding out a finite union into a countable union using the empty set), and so countably additive measures inherit all the properties of finitely additive properties, such as monotonicity and finite subadditivity. But one also has additional properties:

  Proposition 3.3 *Let $(X, \mathscr{B}, \mu)$ be a* ***measure space****.*
  1. *(**Countable subadditivity**) If $E_1, E_2, \ldots$ are $\mathscr{B}$-measurable, then*
  $$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \le \sum_{n=1}^{\infty} \mu(E_n).$$
  2. *(**Upwards monotone convergence**) If $E_1 \subseteq E_2 \subseteq \ldots$ are $\mathscr{B}$-measurable, then*
  $$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n) = \sup_n \mu(E_n). \tag{2}$$
  3. *(**Downwards monotone convergence**) If $E_1 \supseteq E_2 \supseteq \ldots$ are $\mathscr{B}$-measurable, and $\mu(E_n) < \infty$ for **at least one** n, then*
  $$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n) = \inf_n \mu(E_n). \tag{3}$$

- **Exercise 3.4** *Show that the **downward monotone convergence** claim can **fail** if the hypothesis that $\mu(E_n) < \infty$ for at least one n is **dropped**.*

- **Proposition 3.5** *(**Dominated convergence for sets**). [Tao, 2011]*
*Let $(X, \mathscr{B}, \mu)$ be a measure space. Let $E_1, E_2, \ldots$ be a sequence of $\mathscr{B}$-measurable sets that **converge** to another set E, in the sense that $\mathbb{1}_{E_n}$ converges **pointwise** to $\mathbb{1}_E$. Then*

  1. *E is also $\mathscr{B}$-measurable.*

  2. *If there exists a $\mathscr{B}$-measurable set F of **finite measure** (i.e. $\mu(F) < \infty$) that **contains all of the** E_n, then*
  $$\lim_{n\to\infty} \mu(E_n) = \mu(E).$$
  (Hint: Apply downward monotonicity to the sets $\bigcup_{n>N}(E_n \Delta E)$.)

  3. *The previous part of this proposition can **fail** if the hypothesis that all the E_n are contained in a set of finite measure is **omitted**.*

- **Proposition 3.6** *(**Countably Additive Measures on Countable Set with Discrete σ-Algebra**)*
*Let X be an at most **countable** set with **the discrete σ-algebra**. Then every measure μ on this measurable space can be uniquely represented in the form*
$$\mu = \sum_{x\in X} c_x \, \delta_x$$
for some $c_x \in [0, +\infty]$, thus
$$\mu(E) = \sum_{x\in E} c_x$$
*for all $E \subseteq X$. (This claim fails in the **uncountable** case, although showing this is slightly tricky.)*

- **Definition** (***Completeness***). [Tao, 2011]
A ***null set*** of a measure space $(X, \mathscr{B}, \mu)$ is defined to be a $\mathscr{B}$-measurable set of ***measure zero***. A ***sub-null*** set is any subset of a null set.

  A measure space is said to be ***complete*** if *every sub-null set is a null set.*

- **Theorem 3.7** *The **Lebesgue measure space** $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d], m)$ is **complete**, but **the Borel measure space** $(\mathbb{R}^d, \mathcal{B}[\mathbb{R}^d], m)$ is **not**.*

- **Proposition 3.8** *(**Completion**).*
*Let $(X, \mathscr{B}, \mu)$ be a measure space. There exists a **unique refinement** $(X, \overline{\mathscr{B}}, \overline{\mu})$, known as **the completion** of $(X, \mathscr{B}, \mu)$, which is the **coarsest** refinement of $(X, \mathscr{B}, \mu)$ that is **complete**. Furthermore, $\overline{\mathscr{B}}$ consists precisely of those sets that differ from a $\mathscr{B}$-measurable set by **a $\mathscr{B}$-subnull set**.*

- **Remark** *The Lebesgue measure space* $(\mathbb{R}^d, \mathcal{L}[\mathbb{R}^d], m)$ is the ***completion*** of *the Borel measure space* $(\mathbb{R}^d, \mathcal{B}[\mathbb{R}^d], m)$.

- **Proposition 3.9** *(**Approximation by an algebra**).*
*Let $\mathscr{A}$ be a **Boolean algebra** on X, and let μ be a measure on the σ-algebra generated by $\mathscr{A}$, i.e. $\langle \mathscr{A} \rangle$.*

1. *If $\mu(X) < \infty$, then for every $E \in \langle \mathscr{A} \rangle$ and $\epsilon > 0$ there exists $F \in \mathscr{A}$ such that $\mu(E \Delta F) < \epsilon$.*

2. *More generally, if $X = \cup_{n=1}^{\infty} A_n$ for some $A_1, A_2, \ldots \in \mathscr{A}$ with $\mu(A_n) < \infty$ for all n, $E \in \langle \mathscr{A} \rangle$ has finite measure, and $\epsilon > 0$, then there exists $F \in \mathscr{A}$ such that $\mu(E \Delta F) < \epsilon$.*

3.4 Outer Measures and the Carathéodory Extension Theorem

- **Definition** (***Abstract outer measure***). [Tao, 2011]
  Let X be a set. *An **abstract outer measure*** (or ***outer measure*** for short) is a map $\mu^* : 2^X \to [0, +\infty]$ that assigns an *unsigned extended real number* $\mu^*(E) \in [0, +\infty]$ to every set $E \subseteq X$ which obeys the following axioms:

  1. (***Empty set***) $\mu^*(\emptyset) = 0$.
  2. (***Monotonicity***) If $E \subseteq F$, then $\mu^*(E) \leq \mu^*(F)$.
  3. (***Countable subadditivity***) If $E_1, E_2, \ldots \subseteq X$ is a countable sequence of subsets of X, then

  $$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu^*(E_n).$$

  Outer measures are also known as ***exterior measures***.

- **Remark** ***Lebesgue outer measure*** m^* is *an outer measure*. On the other hand, ***Jordan outer measure*** $m^{*,J}$ is only *finitely subadditive* rather than *countably subadditive* and thus is ***not***, strictly speaking, *an outer measure*.

- **Remark** Note that *outer measures* are ***weaker*** than measures in that they are merely *countably **subadditive***, rather than *countably additive*. On the other hand, they are able to *measure **all subsets** of X*, whereas measures can only measure a σ-algebra of *measurable sets*.

- **Definition** (***Carathéodory measurability***).
  Let μ^* be an *outer measure* on a set X. A *set $E \subseteq X$* is said to be ***Carathéodory measurable*** *with respect to μ^** (or, *μ^***-measurable***) if one has

  $$\mu^*(A) = \mu^*(A \setminus E) + \mu^*(A \cap E)$$

  for every set $A \subseteq X$.

- **Example** (***Null sets are Carathéodory measurable***).
  Suppose that E is a ***null set*** for *an **outer measure*** μ^* (i.e. $\mu^*(E) = 0$). Then that E *is Carathéodory measurable with respect to μ^**.

- **Example** (***Compatibility with Lebesgue measurability***). A set $E \subseteq \mathbb{R}^d$ is *Carathéodory measurable with respect to Lebesgue outer measurable* if and only if it is *Lebesgue measurable*.

- **Theorem 3.10** *(**Carathéodory extension theorem**). [Tao, 2011]*
  *Let $\mu^* : 2^X \to [0, +\infty]$ be an outer measure on a set X, let $\mathscr{B}$ be the collection of all subsets of X that are **Carathéodory measurable with respect to** μ^*, and let $\mu : \mathscr{B} \to [0, +\infty]$ be the **restriction** of μ^* to $\mathscr{B}$ (thus $\mu(E) := \mu^*(E)$ whenever $E \in \mathscr{B}$). Then $\mathscr{B}$ **is a σ-algebra**, and μ **is a measure**.*

- **Remark** The measure μ constructed by *the Carathéodory extension theorem* is automatically ***complete***.

- **Proposition 3.11** *Let $\mathscr{B}$ be a **Boolean algebra** on a set X. Then $\mathscr{B}$ is a σ-algebra **if and only if** it is **closed** under countable* disjoint *unions, which means that $\bigcup_{n=1}^{\infty} E_n \in \mathscr{B}$ whenever $E_1, E_2, E_3, \ldots \in \mathscr{B}$ are a countable sequence of **disjoint** sets in $\mathscr{B}$.*

- **Definition** (***Pre-measure***). [Folland, 2013]
*A **pre-measure*** on a ***Boolean algebra*** $\mathscr{B}_0$ is a function $\mu_0 : \mathscr{B}_0 \to [0, +\infty]$ that satisfies the conditions:

  1. (***Empty Set***): $\mu_0(\emptyset) = 0$
  2. (***Countably Additivity***): If $E_1, E_2, \ldots \in \mathscr{B}_0$ are *disjoint sets* such that $\bigcup_{n=1}^{\infty} E_n$ is in $\mathscr{B}_0$,
  $$\mu_0\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu_0(E_n).$$

- **Remark** A *pre-measure* μ_0 is a ***finitely additive measure*** that ***already*** is *countably additive* ***within*** a Boolean algebra $\mathscr{B}_0$.

- **Remark** *The countably additivity condition* for pre-measure can be releaxed to be *the countably subadditivity* $\mu_0(\cup_{n=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} \mu_0(E_n)$ without affecting the definition of a pre-measure.

- **Proposition 3.12** *Let $\mathscr{B} \subset 2^X$ and $\mu_0 : \mathscr{B} \to [0, +\infty]$ be such that $\emptyset, X \in \mathscr{B}$, and $\mu_0(\emptyset) = 0$. For any $A \subseteq X$, define*
$$\mu^*(A) := \inf\left\{\sum_{j=1}^{\infty} \mu_0(E_j) : E_j \in \mathscr{B},\ \text{and } A \subseteq \bigcup_{j=1}^{\infty} E_j\right\}.$$
Then μ^* is an outer measure.

- **Theorem 3.13** ***(Hahn-Kolmogorov Theorem).***
*Every **pre-measure** $\mu_0 : \mathscr{B}_0 \to [0, +\infty]$ on a Boolean algebra $\mathscr{B}_0$ in X can be **extended** to a **countably additive measure** $\mu : \mathscr{B} \to [0, +\infty]$.*

- **Remark** We can construct an *outer measure* μ^* according to Proposition 3.12. Let $\mathscr{B}$ be the *collection* of all sets $E \subseteq X$ that are *Carathéodory measurable with respect to* * *(μ^*-measurable)*, and let μ be the *restriction* of μ^* to $\mathscr{B}$. The tuple $(X, \mathscr{B}, \mu)$ is what we want in *Hahn-Kolmogorov theorem.*

  The measure μ constructed in this way is called ***the Hahn-Kolmogorov extension of the pre-measure*** μ_0.

- **Proposition 3.14** ***(Uniqueness of the Hahn-Kolmogorov Extension)***
*Let $\mu_0 : \mathscr{B}_0 \to [0, +\infty]$ be a **pre-measure**, let $\mu : \mathscr{B} \to [0, +\infty]$ be the **Hahn-Kolmogorov extension** of μ_0, and let $\mu' : \mathscr{B}' \to [0, +\infty]$ be **another** countably additive extension of μ_0. Suppose also that μ_0 is σ-**finite**, which means that one can express the whole space X as the countable union of sets $E_1, E_2, \ldots \in \mathscr{B}_0$ for which $\mu_0(E_n) < \infty$ for all n. Then μ and μ' agree on their common domain of definition. In other words, show that $\mu(E) = \mu'(E)$ for all $E \in \mathscr{B} \cap \mathscr{B}'$.*

3.5 Development of Lebesgue Measure Theory in $\mathbb{R}^d$

Table 2: Comparison between different measures in measure theory

	Elementary measure	***Jordan measure***	***Lebesgue outer measure***	***Lebesgue measure***
compatibility		$\Leftarrow$ ✓	$\Leftarrow$ ✓	$\Leftarrow$ ✓
non-negative	✓	✓	✓	✓
$m(\emptyset) = 0$	✓	✓	✓	✓
$m([0,1]^d) = 1$	✓	✓	✓	✓
translation-invariant	✓	✓	✓	✓
finitely additive	✓	✓	✓	✓
monotonicity	✓	✓	✓	✓
finitely subadditive	✓	✓	✓	✓
outer regularity			✓	✓
inner regularity			✓	✓
countably subadditivity			✓	✓
countably additivity				✓
measurable set	box $I_1 \times \ldots \times I_d$	All *elementary sets*; any ***compact convex*** *polytope*; any ***open sets*** and ***closed sets***; ***finite union*** of measurable sets; ***graph/epigraph*** *of* ***continous function***;	All *Jordan measurable sets*; ***countable union*** of measurable sets, e.g. G_δ and F_σ	forms a ***σ-algebra*** that includes ***all Borel sets***; sets with Lebesgue outer measure zero (***null sets***).
non-measurable set	*any subsets* other than box	*countable union* of Jordan measurable sets; ***bullet-riddled square*** and ***sets of bullets***; subsets with a lot of "*holes*" or "*fractal*"	same as right	$E = \mathbb{R}/\mathbb{Q} \cap [0,1]$
algebra *for collection of measurable sets*	***boolean algebra*** $\mathscr{A}_0$	*boolean algebra* $\mathscr{A}_1 \supsetneq \mathscr{A}_0$		***σ-algebra*** $\mathscr{A}_2 \supsetneq \mathscr{A}_1$
relation to ***integration***		***Riemann integration***		***Lebesgue integration***

4 Measurable Functions and Integration on a Measure Space

4.1 Measurable Functions

- **Definition** Let $(X, \mathscr{B})$ be a *measurable space*, and let $f : X \to [0, +\infty]$ or $f : X \to \mathbb{C}$ be an ***unsigned*** or ***complex-valued function***. We say that f is ***measurable*** if $f^{-1}(U)$ is ***$\mathscr{B}$-measurable*** for every ***open subset*** U of $[0, +\infty]$ or $\mathbb{C}$.

- **Remark** The inverse image of a Lebesgue measurable set by a *measurable function* need not remain Lebesgue. measurable. This is due to the definition of above measureable function. The pre-image of E is Lebesgue measureable, if if E has a slightly stronger measurability property than Lebesgue measurability, namely ***Borel measurability***.

- We can define *a **measurable mapping*** between *two measurable spaces.*

  Definition For $f : X \to Y$, and $X \equiv (X, \mathscr{F})$, $Y \equiv (Y, \mathscr{B})$ are *measurable spaces*, then f is called ***$(\mathscr{F}, \mathscr{B})$ measureable*** (or $(\mathscr{F}/\mathscr{B})$ measureable or, simply, *measureble*), if $f^{-1}(E) \in \mathscr{F}$ for every $E \in \mathscr{B}$.

- **Definition** Note that if $\{(Y_\alpha, \mathscr{B}_\alpha)\}$ is a family of measureable spaces, and $\{f_\alpha\}$ for $f_\alpha : X \to Y_\alpha$, then there is a ***unique smallest*** σ-algebra on X so that $\{f_\alpha\}$ are *all measureable.* It is generated by $f_\alpha^{-1}(E_\alpha), E_\alpha \in \mathscr{B}_\alpha$, i.e.

$$\mathscr{F} = \left\langle \left\{ f_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathscr{B}_\alpha, \alpha \in I \right\} \right\rangle$$

  It is called the ***σ-algebra generated by*** $\{f_\alpha\}$. In particular, $X = \prod_\alpha Y_\alpha$ has ***product σ-algebra*** that is *generated by coordinate functions* $\{\pi_\alpha\}$.

- **Proposition 4.1** ***(Properties of Measurable Function)***
  Let $(X, \mathscr{B})$ be a measurable space.

  1. *$f : X \to [0, +\infty]$ is **measurable** if and only if the **level sets** $\{x \in X : f(x) > \lambda\}$ are $\mathscr{B}$-measurable.*
  2. *The **indicator function** $\mathbb{1}_E$ of a set $E \subseteq X$ is **measurable** if and only if E itself is **$\mathscr{B}$-measurable**.*
  3. *$f : X \to [0, +\infty]$ or $f : X \to \mathbb{C}$ is measurable if and only if $f^{-1}(E)$ is $\mathscr{B}$-measurable for every **Borel-measurable** subset E of $[0, +\infty]$ or $\mathbb{C}$.*
  4. *$f : X \to \mathbb{C}$ is measurable if and only if its **real** and **imaginary** parts are measurable.*
  5. *$f : X \to \mathbb{R}$ is measurable if and only if the **magnitudes** $f_+ := \max\{f, 0\}$, $f_- := \max\{-f, 0\}$ of its **positive** and **negative** parts are **measurable**.*
  6. *If $f_n : X \to [0, +\infty]$ are a sequence of **measurable** functions that converge **pointwise** to a limit $f : X \to [0, +\infty]$, then f is also **measurable**. The same claim holds if $[0, +\infty]$ is replaced by $\mathbb{C}$.*
  7. *If $f : X \to [0, +\infty]$ is measurable and $\varphi : [0, +\infty] \to [0, +\infty]$ is **continuous**, the composite $\varphi \circ f$ is measurable. The same claim holds if $[0, +\infty]$ is replaced by $\mathbb{C}$.*
  8. *The **sum** or **product** of two **measurable** functions in $[0, +\infty]$ or $\mathbb{C}$ is still measurable.*

- **Definition** A function $f : (X, \mathscr{F}) \to (Y, \mathscr{B})$ is ***simple*** if it only takes ***finitely many*** *different values* $s_1, \ldots, s_k \in Y$.

  Then the σ-algebra $f^{-1}(\mathscr{B})$ reduce to $\sigma\left(\left\{f^{-1}(\{s_\alpha\})\right\}_{\alpha=1}^{k}\right)$, the ***finite σ-algebra*** generated by ***atomic algebra*** with atoms $E_\alpha \equiv f^{-1}(\{s_\alpha\})$. The ***canonical representation*** of f is

$$f = \sum_{\alpha=1}^{k} s_\alpha \mathbb{1}\left\{E_\alpha\right\},$$

  which is determined up to a reordering.

- **Proposition 4.2** ***(Measurable Function with respect to Atomic Algebra is Simple)*** *Let $(X, \mathscr{B})$ be a measurable space that is* ***atomic****, thus $\mathscr{B} = \mathscr{A}((A_\alpha)_{\alpha \in I})$ for some partition $X = \bigcup_{\alpha \in I} A_\alpha$ of X into disjoint non-empty atoms. A function $f : X \to [0, +\infty]$ or $f : X \to \mathbb{C}$ is measurable if and only if it is* ***constant*** *on each atom, or equivalently if one has a* ***representation of the form***

$$f(x) = \sum_{\alpha \in I} c_\alpha \mathbb{1}\left\{x \in A_\alpha\right\},$$

  for some constants $c_\alpha \in [0; +\infty]$ or in $\mathbb{C}$ as appropriate. Furthermore, the c_α are uniquely determined by f.

4.2 Simple Integral of Simple Functions

- **Definition** (***Simple integral***).
  Let $(X, \mathscr{B}, \mu)$ be a measure space with $\mathscr{B}$ ***finite*** (i.e., its *cardinality* is *finite* and there are only *finitely many measurable sets*). X can then be partitioned into a finite number of atoms $A_1, \cdots, A_n$. If $f : X \to [0, +\infty]$ is measurable, it has ***a unique representation*** of the form

$$f(x) = \sum_{\alpha \in I} c_\alpha \mathbb{1}\left\{x \in A_\alpha\right\},$$

  for some constants $c_\alpha \in [0; +\infty]$. We then define the ***simple integral*** $\text{simp}\int_X f d\mu$ of f by the formula

$$\text{simp}\int_X f d\mu \equiv \sum_{\alpha \in I} c_\alpha \mu(A_\alpha)$$

- **Remark** Note that the precise decomposition into atoms *does not affect* the definition of the simple integral.

  Proposition 4.3 ***(Simple integral unaffected by refinements)****. [Tao, 2011]*
  Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $(X, \mathscr{B}', \mu')$ be a ***refinement*** *of $(X, \mathscr{B}, \mu)$, which means that $\mathscr{B}'$ contains $\mathscr{B}$ and $\mu' : \mathscr{B}' \to [0, +\infty]$* ***agrees*** *with $\mu : \mathscr{B} \to [0, +\infty]$ on $\mathscr{B}$. Suppose that both $\mathscr{B}, \mathscr{B}'$ are* ***finite****, and let $f : \mathscr{B} \to [0, +\infty]$ be measurable. We have*

$$\textit{simp}\int_X f d\mu = \textit{simp}\int_X f d\mu'.$$

- The above proposition allows one to extend the *simple integral* to *simple functions*:

  Definition (***Integral of simple functions***).
  An ***(unsigned) simple function*** $f : X \to [0, +\infty]$ on a measurable space $(X, \mathscr{B})$ is a *measurable function* that takes on ***finitely many values*** $a_1, \cdots, a_k$. Note that such a function is then automatically *measurable* with respect to *at least one* ***finite sub-σ-algebra*** $\mathscr{B}'$ of $\mathscr{B}$, namely *the σ-algebra $\mathscr{B}'$* ***generated by the preimages*** $f^{-1}\{a_1\}, \cdots, f^{-1}\{a_k\}$ of $a_1, \cdots, a_k$.

  We then define the ***simple integral*** simp $\int_X f d\mu$ by the formula

$$\begin{aligned} \text{simp} \int_X f d\mu &\equiv \text{simp} \int_X f \; d\mu|_{\mathscr{B}'} \\ &= \sum_{i=1}^{k} a_i \mu\left(f^{-1}\{a_k\}\right) \end{aligned}$$

  where $\mu|_{\mathscr{B}'} : \mathscr{B}' \to [0, +\infty]$ is the ***restriction*** of $\mu : \mathscr{B} \to [0, +\infty]$ to $\mathscr{B}'$.

- **Remark** Note that there could be ***multiple finite σ-algebras*** with respect to which f is *measurable*, but all such extensions will give the same simple integral. Indeed, if f were measurable with respect to two separate finite sub-σ-algebras $\mathscr{B}'$ and $\mathscr{B}''$ of $\mathscr{B}$, then it would also be *measurable* with respect to their ***common refinement*** $\mathscr{B}' \vee \mathscr{B}'' := (\mathscr{B}' \cup \mathscr{B}'')$, which is also *finite* and then by Proposition 4.3, $\int_X f d\mu|_{\mathscr{B}'}$ and $\int_X f d\mu|_{\mathscr{B}''}$ are both equal to $\int_X f d\mu|_{\mathscr{B}' \vee \mathscr{B}''}$, and hence equal to each other.

- **Remark** As with the Lebesgue theory, we say that a property $P(x)$ of an element $x \in X$ of a measure space $(X, \mathscr{B}, \mu)$ ***holds μ-almost everywhere*** if it ***holds outside*** of a ***sub-null set***, i.e. $\mu(\{P(x) \text{ does not hold}\}) = 0$.

- **Proposition 4.4** ***(Property of Simple Integral)***
  Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f, g : X \to [0, +\infty]$ be simple unsigned functions.

  1. ***(Monotonicity)*** *If $f \leq g$ then simp $\int_X f d\mu \leq$ simp $\int_X g d\mu$.*
  2. ***(Compatibility with measure)*** *For every $\mathscr{B}$-measurable set E, we have simp $\int_X \mathbb{1}_E d\mu = \mu(E)$.*
  3. ***(Homogeneity)*** *For every $c \in [0, +\infty]$, one has simp $\int_X (cf) d\mu = c \times$ simp $\int_X f d\mu$.*
  4. ***(Finite additivity)*** *We have simp $\int_X (f + g) d\mu =$ simp $\int_X f d\mu +$ simp $\int_X g d\mu$.*
  5. ***(Insensitivity to refinement)*** *Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $(X, \mathscr{B}', \mu')$ be its refinement, which means that $\mathscr{B}'$ contains $\mathscr{B}$ and $\mu' : \mathscr{B}' \to [0, +\infty]$* ***agrees*** *with $\mu : \mathscr{B} \to [0, +\infty]$ on $\mathscr{B}$. Suppose that both $\mathscr{B}, \mathscr{B}'$ are* ***finite***, *and let $f : \mathscr{B} \to [0, +\infty]$ be measurable. We have*

$$\textit{simp} \int_X f d\mu = \textit{simp} \int_X f d\mu'.$$

  6. ***(Almost everywhere equivalence)*** *If μ-almost everywhere $f = g$, then simp $\int_X f d\mu =$ simp $\int_X g d\mu$*
  7. ***(Finiteness)*** *simp $\int_X f d\mu < \infty$ if and only if f is* ***finite μ-almost everywhere*** *and is* ***supported*** *on a set of* ***finite measure***.

8. *(**Vanishing**) simp $\int_X f d\mu = 0$ if and only if $f = 0$ μ-almost everywhere.*

- **Proposition 4.5** ***(Inclusion-exclusion principle).***
*Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $A_1, \ldots, A_n$ be $\mathscr{B}$-measurable sets of **finite measure**. Show that*

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{J \subseteq [1:n], J \neq \emptyset} (-1)^{|J|-1} \mu\left(\bigcap_{i \in J} A_i\right)$$

(Hint: Compute simp $\int_X (1 - \prod_{i=1}^{n}(1 - \mathbb{1}_{A_i}))d\mu$ in two different ways.)

- **Remark** The simple integral could also be defined on *finitely additive measure spaces*, rather than *countably additive ones*, and all the above properties would still apply. However, on a finitely additive measure space one would have difficulty extending the integral beyond simple functions.

4.3 Unsigned Integral of Measurable Functions

- **Definition** Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f : X \to [0, +\infty]$ be *(unsigned) measurable.* Then we define the ***unsigned integral*** $\int_X f d\mu$ of f by the formula

$$\int_X f d\mu \equiv \sup_{\substack{0 \le g \le f, \\ g \text{ simple}}} \text{simp} \int_X g d\mu$$

- **Proposition 4.6** ***(Properties of the unsigned integral).***
Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f, g : X \to [0, +\infty]$ be measurable.

1. *(**Almost everywhere equivalence**) If $f = g$ μ-almost everywhere, then $\int_X f d\mu = \int_X g d\mu$*
2. *(**Monotonicity**) If $f \le g$ μ-almost everywhere, then $\int_X f d\mu \le \int_X g d\mu$.*
3. *(**Homogeneity**) We have $\int_X (cf) d\mu = c \int_X f d\mu$ for every $c \in [0, +\infty]$.*
4. *(**Superadditivity**) We have $\int_X (f + g) d\mu \ge \int_X f d\mu + \int_X g d\mu$.*
5. *(**Compatibility with the simple integral**) If f is **simple**, then $\int_X f d\mu =$ simp $\int_X f d\mu$.*
6. *(**Markov's inequality**) For any $0 < \lambda < 1$, one has*

$$\mu\left(\{x \in X : f(x) \ge \lambda\}\right) \le \frac{1}{\lambda} \int_X f d\mu$$

   In particular, if $\int_X f d\mu < \infty$, then the sets $\{x \in X : f(x) \ge \lambda\}$ have finite measure for each $\lambda > 0$.

7. *(**Finiteness**) If $\int_X f d\mu < \infty$, then $f(x)$ is **finite** for μ-**almost every** x.*
8. *(**Vanishing**) If $\int_X f d\mu = 0$, then $f(x)$ is zero for μ-almost every x.*
9. *(**Vertical truncation**) We have*

$$\lim_{n \to \infty} \int_X \min\{f, n\} d\mu = \int_X f d\mu$$

10. *(**Horizontal truncation**) If $E_1 \subseteq E_2 \subseteq \dots$ is an **increasing sequence** of $\mathscr{B}$-measurable sets, then*

$$\lim_{n\to\infty} \int_X f \mathbb{1}_{E_n} d\mu = \int_X f \mathbb{1}_{\cup_{n=1}^{\infty} E_n} d\mu.$$

11. *(**Restriction**) If Y is a measurable subset of X, then*

$$\int_X f \mathbb{1}_Y d\mu = \int_Y f|_Y d\mu|_Y,$$

*where $f|_Y : Y \to [0, +\infty]$ is the **restriction** of $f : X \to [0, +\infty]$ to Y, and $\mu|_Y$ is the restriction μ on Y. We will often abbreviate $\int_Y f|_Y d\mu|_Y$ (by slight abuse of notation) as $\int_Y f d\mu$.*

- **Theorem 4.7** *Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f, g : X \to [0, +\infty]$ be measurable. Then*

$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.$$

- **Proposition 4.8** *(**Linearity in** μ).*
Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f : X \to [0, +\infty]$ be measurable.

  1. *$\int_X f d(c\mu) = c \times \int_X f d\mu$ for every $c \in [0, +\infty]$.*
  2. *If $\mu_1, \mu_2, \dots$ are a sequence of measures on $\mathscr{B}$,*

$$\int_X f d\left(\sum_{n=1}^{\infty} \mu_n\right) = \sum_{n=1}^{\infty} \int_X f d\mu_n.$$

- **Proposition 4.9** *(**Pushforward Measure**).*
*Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $\varphi : X \to Y$ be $(\mathscr{B}, \mathscr{C})$ measureable from $(X, \mathscr{B})$ to another measurable space $(Y, \mathscr{C})$. Define the **pushforward** $\phi_* \mu : \mathscr{C} \to [0, +\infty]$ of μ **by** φ by the formula*

$$\varphi_* \mu(E) := \mu(\phi^{-1}(E)).$$

  1. *$\varphi_* \mu$ is a **measure** on $\mathscr{C}$, so that $(Y, \mathscr{C}, \phi_* \mu)$ is a measure space.*
  2. *(**Change of variables formula**). If $f : Y \to [0, +\infty]$ is $\mathscr{C}$-measurable, then*

$$\int_Y f d(\phi_* \mu) = \int_X (f \circ \phi) d\mu.$$

- **Corollary 4.10** *Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be an invertible linear transformation, and let m be Lebesgue measure on $\mathbb{R}^d$. Then $T_* m = \frac{1}{|\det T|} m$, where $T_* m$ is **the pushforward of** m.*

- **Example** (***Sums as integrals***). Let X be an arbitrary set (with the ***discrete σ-algebra***), let $\#$ be ***counting measure***, and let $f : X \to [0, +\infty]$ be an arbitrary unsigned function. Then f is ***measurable*** with

$$\int_X f d\# = \sum_{x \in X} f(x).$$

4.4 Absolutely Convergent Integral

- **Definition** (***Absolutely convergent integral***).
Let $(X, \mathscr{F}, \mu)$ be a measure space. A *measurable function* $f : X \to \mathbb{C}$ is said to be ***absolutely integrable*** if the *unsigned integral*

$$\|f\|_{L^1(X,\mathscr{F},\mu)} \equiv \int_X |f|\,d\mu$$

is ***finite***. We refer to this quantity $\|f\|_{L^1(X)}$ as ***the $L^1(X)$ norm of f***, and use $L^1(X)$ or $L^1(X, \mathscr{F}, \mu)$ or $L^1(\mu)$ to denote the space of absolutely integrable functions. If f is *real-valued* and absolutely integrable, we define ***the Lebesgue integral*** $\int_X f d\mu$ by the formula

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu$$

where $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$ are the magnitudes of the positive and negative components of f. (note that the two unsigned integrals on the right-hand side are finite, as f_+, f_- are pointwise dominated by $|f|$). If f is *complex-valued* and absolutely integrable, we define ***the Lebesgue integral*** $\int_X f(x) d\mu$ by the formula

$$\int_X f d\mu = \int_X \Re(f) d\mu + i \int_X \Im(f) d\mu,$$

where the two integrals on the right are interpreted as real-valued absolutely integrable Lebesgue integrals.

- **Remark** Sometimes $\int_X f d\mu$ is also denoted as $\int_X f(x)\mu(dx)$ or $\int_X f(x) d\mu(x)$, where $X \subseteq \mathbb{R}^d$ and $\mu(E) = \int_E \mu dx$.

- **Proposition 4.11** ***(Properties of absolutly convergent integral)***
Let $(X, \mathscr{B}, \mu)$ be a measure space.

  1. *$L^1(X, \mathscr{B}, \mu)$ is a **complex vector space**.*
  2. *The integration map $f \mapsto \int_X f d\mu$ is a **complex linear map** from $L^1(X, \mathscr{B}, \mu)$ to $\mathbb{C}$.*
  3. *The **triangle inequality***

     $$\|f + g\|_{L^1(\mu)} \le \|f\|_{L^1(\mu)} + \|g\|_{L^1(\mu)}$$

     *and the **homogeneity property***

     $$\|c\,f\|_{L^1(\mu)} = |c|\,\|f\|_{L^1(\mu)}$$

     hold for all $f, g \in L^1(X, \mathscr{B}, \mu)$ and $c \in \mathbb{C}$.
  4. *If $f, g \in L^1(X, \mathscr{B}, \mu)$ are such that $f(x) = g(x)$ for μ-almost every $x \in X$, then $\int_X f d\mu = \int_X g d\mu$.*
  5. *If $f \in L^1(X, \mathscr{B}, \mu)$, and $(X, \mathscr{B}', \mu')$ is a **refinement** of $(X, \mathscr{B}, \mu)$, then $f \in L^1(X, \mathscr{B}', \mu')$, and*

     $$\int_X f d\mu' = \int_X f d\mu.$$

(Hint: it is easy to get one inequality. To get the other inequality, first work in the case when f is both bounded and has finite measure support (i.e. is both vertically and horizontally truncated).)

6. *If $f \in L^1(X, \mathscr{B}, \mu)$, then $\|f\|_{L^1(\mu)} = 0$ if and only if f is zero μ-almost everywhere.*
7. *If $Y \subseteq X$ is $\mathscr{B}$-measurable and $f \in L^1(X, \mathscr{B}, \mu)$, then $f|_Y \in L^1(Y, \mathscr{B}|_Y, \mu|_Y)$ and*

$$\int_Y f|_Y \, d\mu|_Y = \int_X f \mathbb{1}_Y \, d\mu.$$

As before, by abuse of notation we write $\int_Y f d\mu$ for $\int_Y f|_Y \, d\mu|_Y$.

4.5 The Convergence Theorems

- **Proposition 4.12** *(**Uniform Convergence on a Finite Measure Space**). [Tao, 2011]*
*Suppose that $(X, \mathscr{B}, \mu)$ is a **finite measure space** (so $\mu(X) < \infty$), and $f_n : X \to [0, +\infty]$ (resp. $f_n : X \to \mathbb{C}$) are a sequence of unsigned measurable functions (resp. absolutely integrable functions) that **converge uniformly** to a limit f. Then $\int_X f_n d\mu$ **converges** to $\int_X f d\mu$.*

- **Theorem 4.13** *(**Monotone Convergence Theorem**). [Tao, 2011]*
*Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $0 \le f_1 \le f_2 \le \ldots$ be a **monotone non-decreasing** sequence of **unsigned** measurable functions on X. Then we have*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \left(\lim_{n\to\infty} f_n \right) d\mu$$

- **Remark** Note that in the special case when each f_n is an indicator function $f_n = \mathbb{1}\{E_n\}$, this theorem collapses to ***the upwards monotone convergence** property*. Conversely, *the upwards monotone convergence property* will play a key role in the proof of this theorem.

- **Remark** Note that the result still holds if the monotonicity $f_n \le f_{n+1}$ only holds almost everywhere rather than everywhere.

- **Corollary 4.14** *(**Tonelli's Theorem for Sums and Integrals**)*
*Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots$ be a sequence of **unsigned** measurable functions on X. Then*

$$\sum_{k=1}^{\infty} \int_X f_k d\mu = \int_X \left(\sum_{k=1}^{\infty} f_k \right) d\mu$$

- **Lemma 4.15** *(**Borel-Cantelli Lemma**). [Tao, 2011, Resnick, 2013]*
Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $E_1, E_2, \ldots$ be a sequence of $\mathscr{B}$-measurable sets such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then

$$\mu \left\{ \limsup_{n\to\infty} E_n \right\} = 0.$$

*That is, almost every $x \in X$ is contained in **at most finitely many** of the E_n (i.e. $\{n \in \mathbb{N} : x \in E_n\}$ is finite for almost every $x \in X$).*

- **Corollary 4.16** ***(Fatou's Lemma).***
  Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots : X \to [0, \infty]$ be a sequence of unsigned measurable functions. Then
$$\int_X \left(\liminf_{n\to\infty} f_n\right) d\mu \le \liminf_{n\to\infty} \int_X f_n d\mu$$
  Remark Informally, *Fatou's lemma* tells us that when taking ***the pointwise limit*** of ***unsigned functions*** f_n, that mass $\int_X f_n d\mu$ can be ***destroyed in the limit*** (as was the case in the three key moving bump examples), *but* it ***cannot be created in the limit***. Of course the unsigned hypothesis is necessary here.

  While this lemma was stated only for pointwise limits, the same general ***principle*** (*that mass can be destroyed, but not created, by the process of taking limits*) tends to hold for other "*weak*" notions of convergence.

- **Theorem 4.17** ***(Dominated Convergence Theorem).***
  *Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots : X \to \mathbb{C}$ be a sequence of measurable functions that converge **pointwise μ-almost everywhere** to a measurable limit $f : X \to \mathbb{C}$. Suppose that there is an **unsigned absolutely integrable** function $G : X \to [0, +\infty]$ such that $|f_n|$ are pointwise μ-almost everywhere **bounded** by G for each n. Then we have*
$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

- **Remark** From the moving bump examples we see that this statement *fails* if there is no *absolutely integrable dominating function G.*

- **Remark** Note also that when each of the fn is an indicator function $f_n = \mathbb{1}_{E_n}$, the dominated convergence theorem collapses to *dominated convergence for sets* in previous chapter.

- **Exercise 4.18** *Under the hypotheses of the dominated convergence theorem, establish also that $\|f_n - f\|_{L^1} \to 0$ as $n \to \infty$.*

- **Proposition 4.19** ***(Almost Dominated Convergence).***
  Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots : X \to \mathbb{C}$ be a sequence of measurable functions that converge pointwise μ-almost everywhere to a measurable limit $f : X \to \mathbb{C}$. Suppose that there is an unsigned absolutely integrable functions $G, g_1, g_2, \ldots X \to [0, +\infty]$ such that the $|f_n|$ are pointwise μ-almost everywhere bounded by $G + g_n$, and that $\int_X g_n d\mu \to 0$ as $n \to \infty$. Then
$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

- **Exercise 4.20** ***(Defect Version of Fatou's Lemma).***
  *Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $f_1, f_2, \ldots : X \to [0, +\infty]$ be a sequence of **unsigned absolutely integrable functions** that converges **pointwise** to an absolutely integrable limit f. Show that*
$$\int_X f_n d\mu - \int_X f d\mu - \|f - f_n\|_{L^1(\mu)} \to 0$$
  as $n \to \infty$. (Hint: Apply the dominated convergence theorem to $\min(f_n, f)$.) Informally, this result tells us that the gap between the left and right hand sides of Fatous lemma can be measured by the quantity $\|f - f_n\|_{L^1(\mu)}$.

- **Proposition 4.21** *Let $(X, \mathscr{B}, \mu)$ be a measure space, and let $g : X \to [0, +\infty]$ be measurable. Then the function $\mu_g : \mathscr{B} \to [0, +\infty]$ defined by the formula*

$$\mu_g(E) := \int_X g\, \mathbb{1}_E d\mu = \int_E g d\mu$$

is a measure.

- *The monotone convergence theorem* is, in some sense, *a* ***defining property*** *of the unsigned integral*:

Proposition 4.22 *(**Characterisation of the Unsigned Integral**).*
Let $(X, \mathscr{B})$ be a measurable space. $I : f \mapsto I(f)$ be a map from the space $U(X, \mathscr{B})$ of ***unsigned measurable functions*** *$f : X \to [0, +\infty]$ to $[0, +\infty]$ that obeys the following axioms:*

1. *(**Homogeneity**) For every $f \in U(X, \mathscr{B})$ and $c \in [0, +\infty]$, one has $I(c\, f) = c\, I(f)$.*
2. *(**Finite additivity**) For every $f, g \in U(X, \mathscr{B})$, one has $I(f + g) = I(f) + I(g)$.*
3. *(**Monotone convergence**) If $0 \leq f_1 \leq f_2 \leq \dots$ are a* ***nondecreasing*** *sequence of unsigned measurable functions, then $I(\lim_{n\to\infty} f_n) = \lim_{n\to\infty} I(f_n)$.*

Then there exists a ***unique measure*** *μ on $(X, \mathscr{B})$ such that*

$$I(f) = \int_X f d\mu, \quad \text{for all } f \in U(X, \mathscr{B}).$$

Furthermore, μ is given by the formula $\mu(E) := I(\mathbb{1}_E)$ for all $\mathscr{B}$-measurable sets E.

4.6 Development of Integration of Measurable Functions

Table 3: Development on Lebesgue Integration

	Unsigned Simple Function	***Unsigned Measurable Function***	***Abusolute Integrable Function***
Definition	$f = \sum_{i=k}^{m} c_k \mathbb{1}_{E_k}$	$\{f_n\} \to f$ pointwise $\{f_n\}$ unsigned simple	$\\|f\\|_{L^1(\mathbb{R}^d)} < \infty$ $\|f\|$ unsigned
Integration	$\text{simp} \int_{\mathbb{R}^d} f(x)dx = \sum_{i=k}^{m} c_k \mu(E_k)$	$\int_{\mathbb{R}^d} f(x)dx = \underline{\int_{\mathbb{R}^d}} f(x)dx = \sup_{0 \leq g \leq f,\ g \text{ simple}} \text{simp} \int_{\mathbb{R}^d} g(x)dx$	$\int_{\mathbb{R}^d} f(x)dx = \int_{\mathbb{R}^d} f_+(x)dx - \int_{\mathbb{R}^d} f_-(x)dx$
Compatibility		✓	✓
Compatibility to Rieman Integral	✓	✓	✓
Linearity	Unsigned ✓	✓	✓
Equivalence	✓	✓	✓
Vanishing	✓	✓	✓
Monotonicity	✓	✓	✓
Superadditivity		lower Lebesgue integral ✓	
Reflection		lower Lebesgue integral ✓	
Divisibility		lower Lebesgue integral ✓	
Finite additivity	✓	✓	✓
Horizontal truncation		✓	✓
Vertical truncation		✓	✓
Translation Invariance		✓	✓

5 Littlewood's Three Principles

- **Theorem 5.1** *(**Littlewood's Three Principles**)[Royden and Fitzpatrick, 1988, Tao, 2011]:*
  1. *Every (**measurable**) set is nearly a **finite sum** of **intervals**;*
  2. *Every (**absolutely integrable**) function is nearly **continuous**; and*
  3. *Every **(pointwise) convergent sequence of functions** is nearly **uniformly convergent***

- **Remark** ***The Littlewood's 1st and 2nd principles are shown only for Euclidean space*** $\mathbb{R}^d$, since it relies on such concepts as "*elementary set*" or "*continuous function*" defined for *an abstract measure space.* In other word, ***the necessary condition*** these two principles to hold is that *the measure space* $(X, \mathscr{F})$ *is a **topological space** with **Borel σ-algebra** $\mathscr{B}$ included in $\mathscr{F}$.*

  The ***Littlewood's 3rd principles***, i.e., *the Egorov's theorem*, holds for a ***finite measure space*** $(X, \mathscr{F}, \mu)$ *in which* $\mu(X) < \infty$. There are cases in which $m(X) = \infty$ and the theorem does not hold. [Tao, 2011]

5.1 Every Measurable Set is Nearly a Finite Sum of Intervals

- **Proposition 5.2** *(**Criteria for measurability** [Tao, 2011])*
  The followings are equivalent:
  1. *E is Lebesgue measureable.*
  2. *(**Outer approximation by open**) For every $\epsilon > 0$, one can contain E in an open set U with $m^*(U \setminus E) \leq \epsilon$.*
  3. *(**Almost open**) For every $\epsilon > 0$, one can find an open set U such that $m^*(U \Delta E) \leq \epsilon$, where $U \Delta E = (U \setminus E) \cup (E \setminus U) = (U \cup E) \setminus (U \cap E)$ is the symmetric difference. (In other words, E differs from an open set by a set of outer measure at most ϵ.)*

- **Remark** For E ***finite*** Lebesgue measureable, E differs from a ***bounded*** *open set* by a set of ***arbitrarily small*** *Lebesgue outer measure.* This bounded open set can be ***decomposed*** as a finite union of open cubes in $\mathbb{R}^d$. [Royden and Fitzpatrick, 1988].

5.2 Every Absolutely Integrable Function is Nearly Continuous

- **Proposition 5.3** *(**Approximation of L^1 functions**).*
  Let $f \in L^1(\mathbb{R}^d)$ and $\epsilon > 0$.
  1. *There exists an absolutely integrable simple function g such that*
  $$\|f - g\|_{L^1(\mathbb{R}^d)} \leq \epsilon;$$
  2. *There exists a **step function** g (, i.e. g is represented as a finite linear combination of indicator functions of boxes) such that $\|f - g\|_{L^1(\mathbb{R}^d)} \leq \epsilon$;*

3. *There exists a* ***continuous, compactly supported*** *g such that $\|f - g\|_{L^1(\mathbb{R}^d)} \leq \epsilon$.*

- **Theorem 5.4** ***(Lusin's theorem).***
  Let $f : \mathbb{R}^d \to \mathbb{C}$ be ***absolutely integrable****, and let $\epsilon > 0$. Then there exists a Lebesgue measurable set $E \subset \mathbb{R}^d$ of measure at most ϵ such that the* ***restriction*** *of f to the* ***complementary*** *set $\mathbb{R}^d \setminus E$ is* ***continuous*** *on that set.*

  Remark This theorem does not imply that the ***unrestricted*** function f is *continuous* on $\mathbb{R}^d \setminus E$. For instance, the absolutely integrable function $\mathbb{1}\{\mathbb{Q}\} : \mathbb{R} \to \mathbb{C}$ is nowhere continuous, so is certainly not continuous on $\mathbb{R} \setminus E$ for any E of finite measure; but on the other hand, if one deletes the measure zero set $E \equiv \mathbb{Q}$ from the reals, then the restriction of f to $\mathbb{R} \setminus E$ is identically zero and thus continuous. [Tao, 2011]

- **Remark** When dealing with unsigned measurable functions such as $f : \mathbb{R}^d \to [0, +\infty]$, then Lusin's theorem ***does not apply directly*** because f could be ***infinite*** on a set of positive measure, which is clearly in contradiction with the conclusion of Lusin's theorem (unless one allows the continuous function to also take values in the extended non-negative reals $[0, +\infty]$ with the extended topology). However, if one knows already that f ***is almost everywhere finite*** (which is for instance the case when f is absolutely integrable), then *Lusin's theorem applies* (since one can simply zero out f on the null set where it is infinite, and add that null set to the exceptional set of Lusin's theorem).

5.3 Every Pointwise Convergent Sequence of Functions is Nearly Uniformly Convergent

- **Remark** Recall the following convergence definitions:

  1. (***Pointwise convergence***)
     For every $x \in \mathbb{R}^d$, any $\epsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \epsilon$ for all $n \geq N$.

  2. (***Pointwise almost everywhere convergence***)
     For *almost every* $x \in \mathbb{R}^d$, any $\epsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \epsilon$ for all $n \geq N$.

  3. (***Uniform convergence***)
     For any $\epsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \epsilon$ for all $n \geq N$ and $x \in \mathbb{R}^d$.

- **Definition** (***Locally uniform convergence***).
  A sequence of functions $f_n : \mathbb{R}^d \to \mathbb{C}$ converges ***locally uniformly*** to a limit $f : \mathbb{R}^d \to \mathbb{C}$ if, for every *bounded* subset E of $\mathbb{R}^d$, f_n converges *uniformly* to f on E. In other words, for every ***bounded*** $E \subset \mathbb{R}^d$ and any $\epsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \epsilon$ for all $n \geq N$ and $x \in E$.

- **Theorem 5.5** ***(Egorov's theorem).***
  Let $f_n : \mathbb{R}^d \to \mathbb{C}$ be a sequence of measurable functions that ***converge pointwise almost everywhere*** *to another function $f : \mathbb{R}^d \to \mathbb{C}$, and let $\epsilon > 0$. Then there exists a Lebesgue measurable set A of measure at most ϵ, such that f_n* ***converges locally uniformly*** *to f outside of A.*

6 Modes of Convergence

6.1 Convergence of Functions in Measure Space

- **Remark** (***Convergence of Functions vs. Convergence of Numbers and Vectors***)
  Convegence of numbers $a_n \to a$ and convergence of vector $\boldsymbol{v}_n \to \boldsymbol{v}$ are both ***unambiguous***:

  1. $a_n \to a$ means that $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that for $n \geq N$, $|a_n - a| \leq \epsilon$;
  2. $\boldsymbol{v}_n \to \boldsymbol{v}$ means that $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that for $n \geq N$, $\|\boldsymbol{v}_n - \boldsymbol{v}\| \leq \epsilon$; Note that the chioce of norm in Euclidean space will not affect the convergence results: convergence in ℓ_p will implies convergence in ℓ_q norm.

  However, for functions $f_n : X \to \mathbb{C}$ and $f : X \to \mathbb{C}$, there can now be *many different ways* in which the sequence f_n may or may not converge to the limit f. Note that a_n can be thought as f_n with singular domain $X = \{1\}$ and $\boldsymbol{v}_n$ can be thought of f_n with finite set $X = \{1, \ldots, d\}$. On the other hand, once X *becomes* ***infinite***, the functions f_n acquire an ***infinite number of degrees of freedom***, and this allows them to approach f in any number of *inequivalent ways*.

- We have the following modes of convergence

  1. **Definition** (***Pointwise Convergence***)
     We say that f_n converges to f ***pointwise*** if, for any $x \in X$ and $\epsilon > 0$, there exists $N > 0$ (*that* ***depends*** *on ϵ and x*) such that for all $n \geq N$, $|f_n(x) - f(x)| \leq \epsilon$. Denoted as $f_n(x) \to f(x)$.
  2. **Definition** (***Uniform Convergence***)
     We say that f_n converges to f ***uniformly*** if, for any $\epsilon > 0$, there exists $N > 0$ (*that* ***depends*** *on ϵ only*) such that for all $n \geq N$, $|f_n(x) - f(x)| \leq \epsilon$ for every $x \in X$. Denoted as $f_n \to f$, *uniformly*.

     Unlike pointwise convergence, the time N at which $f_n(x)$ must be permanently ϵ-close to $f(x)$ is not permitted to depend on x, but must instead be chosen *uniformly* in x.
  3. **Definition** (***Pointwise Almost Everywhere Convergence***)
     We say that f_n converges to f ***pointwise almost everywhere*** if, for ***μ-almost everywhere*** $x \in X$, $f_n(x)$ converges to $f(x)$. It is denoted as $f_n \stackrel{a.e.}{\to} f$.

     In other words, there exists ***a null set*** E, ($\mu(E) = 0$) such that for *any* $x \in X \setminus E$ and any $\epsilon > 0$, there exists $N > 0$ (*that* ***depends*** *on ϵ and x*) such that for all $n \geq N$, $|f_n(x) - f(x)| \leq \epsilon$.
  4. **Definition** (***Uniformly Almost Everywhere Convergence***) [Tao, 2011]
     We say f_n converges to f ***uniformly almost everywhere***, ***essentially uniformly***, or ***in L^∞ norm*** if, for every $\epsilon > 0$, there exists N such that for every $n \geq N$, $|f_n(x) - f(x)| \leq \epsilon$, ***for μ-almost every*** $x \in X$.

     That is, $f_n \to f$ uniformly in $x \in X \setminus E$, for some E with $\mu(E) = 0$.

     We can also formulate in terms of L^∞ ***norm*** as

     $$\|f_n(x) - f(x)\|_{L^\infty(X)} \stackrel{n\to\infty}{\longrightarrow} 0,$$

where $\|f\|_{L^\infty(X)} = \operatorname{ess\,sup}_x |f(x)| \equiv \inf_{\{E:\mu(E)=0\}} \sup_{x \in X\setminus E} |f(x)|$ is the ***essential bound.*** It is denoted as $f_n \stackrel{L^\infty}{\to} f$.

5. **Definition** (***Almost Uniform Convergence***) [Tao, 2011]
We say that f_n converges to f ***almost uniformly*** if, for every $\epsilon > 0$, there exists an ***exceptional set*** $E \in \mathscr{B}$ of *measure* $\mu(E) \leq \epsilon$ such that f_n converges ***uniformly*** to f on the *complement* of E.

   That is, for arbitrary δ there exists some E with $\mu(E) \leq \delta$ such that $f_n \to f$ *uniformly* in $x \in X \setminus E$.

6. **Definition** (***Convergence in*** L^1 ***Norm***)
We say that f_n converges to f ***in*** L^1 ***norm*** if the quantity

$$\|f_n - f\|_{L^1(X)} = \int_X |f_n(x) - f(x)|\, d\mu \stackrel{n\to\infty}{\longrightarrow} 0.$$

   It is also called the convergence ***in mean.*** Denoted as $f_n \stackrel{L^1}{\to} f$.

7. **Definition** (***Convergence in Measure***)
We say that f_n converges to f ***in measure*** if, for every $\epsilon > 0$, the measures

$$\mu\left(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}\right) \stackrel{n\to\infty}{\longrightarrow} 0.$$

   Denoted as $f_n \stackrel{\mu}{\to} f$.

- **Remark** The difference between the ***uniformly almost everywhere convergence*** vs. ***the almost uniformly convergence*** is that:

  1. the ***former*** corresponds to *uniform convergence* ***outside a null set***, and
  2. the ***latter*** corresponds to *uniform convergence* ***outside an arbitrary small measure set (but still not a null set)***.

- **Remark** Observe that each of *these five modes of convergence* is ***unaffected*** if one ***modifies*** f_n or f on ***a set of measure zero.*** In contrast, the *pointwise* and *uniform* modes of convergence can be ***affected*** if one modifies f_n or f *even on a single point.*

- **Remark** In the context of *probability theory*, in which f_n and f are interpreted as *random variables*, [Billingsley, 2008, Folland, 2013]

convergence in L^1 *norm*	$\Leftrightarrow$	convergence in *mean*
pointwise convergence almost everywhere	$\Leftrightarrow$	***almost sure convergence***
convergence in ***measure***	$\Leftrightarrow$	convergence in ***probability***

- **Proposition 6.1** ***(Linearity of Convergence)****. [Tao, 2011]*
Let $(X, \mathscr{B}, \mu)$ be a measure space, let $f_n, g_n : X \to \mathbb{C}$ be sequences of measurable functions, and let $f, g : X \to \mathbb{C}$ be measurable functions.

  1. *Then f_n converges to f along one of the above seven modes of convergence* ***if and only if*** *$|f_n - f|$ converges to 0 along* ***the same mode****.*

2. *If f_n converges to f along one of the above seven modes of convergence, and g_n converges to g along* ***the same mode****, then $f_n + g_n$ converges to $f + g$ along the same mode, and that $c\, f_n$ converges to $c\, f$ along the same mode for any $c \in \mathbb{C}$.*

3. *(**Squeeze test**) If f_n converges to 0 along one of the above seven modes, and $|g_n| \le f_n$* ***pointwise*** *for each n, then g_n converges to 0 along* ***the same mode****.*

6.2 Modes of Convergence via Tail Support and Width

- **Remark** (***Tail Support and Width***)

Definition Let $E_{n,m} := \{x \in X : |f_n(x) - f(x)| \ge 1/m\}$. Define ***the N-th tail support set***

$$T_{N,m} := \{x \in X : |f_n(x) - f(x)| \ge 1/m,\ \ \exists n \ge N\} = \bigcup_{n \ge N} E_{n,m}.$$

Also let $\mu(E_{n,m})$ be the ***width*** of n-th event $\mathbb{1}\{E_{n,m}\}$. Note that $T_{N,m} \supseteq T_{N+1,m}$ is ***monotone nonincreasing*** and $T_{N,m} \subseteq T_{N,m+1}$ is ***monotone nondecreasing***.

1. The ***pointwise convergence*** of f_n to f indicates that for every x, every $m \ge 1$, there exists some $N \equiv N(m,x) \ge 1$ such that $T^c_{N,m} \ni x$ or $T_{N,m} \not\ni x$. Equivalently, ***the tail support shrinks to emptyset***:

$$\bigcap_{N \in \mathbb{N}} T_{N,m} = \lim_{N \to \infty} T_{N,m} = \limsup_{n \to \infty} E_{n,m} = \emptyset, \quad \text{for all } m.$$

Conversely, to prove ***not pointwise convergence***, we need to find a $x \in X$ and for an arbitrary fixed $m \ge 1$ such that

$$x \in \bigcap_{N \in \mathbb{N}} \bigcup_{n \ge N} \{x \in X : |f_n(x) - f(x)| \ge 1/m\} = \limsup_{n \to \infty} \{x \in X : |f_n(x) - f(x)| \ge 1/m\}.$$

2. The ***pointwise almost everywhere convergence*** indicates that there exists ***a null set*** F with $\mu(F) = 0$ such that for every $x \in X \setminus F$ and any $m \ge 1$, there exists some $N \equiv N(m,x) \ge 1$ such that $(T_{N,m} \setminus F) \not\ni x$. Equivalently, ***the tail support shrinks to a nulll set***. Note that it makes no assumption on $(T_{N,m} \cap F)$.

$$\begin{aligned} & \lim_{N \to \infty} T_{N,m} \setminus F = \limsup_{n \to \infty} E_{n,m} \setminus F = \emptyset, \quad \text{for all } m. \\ \Leftrightarrow\ & \bigcap_{N \in \mathbb{N}} T_{N,m} = \lim_{N \to \infty} T_{N,m} = F \\ \Leftrightarrow\ & \mu\left(\lim_{N \to \infty} T_{N,m}\right) = \mu\left(\bigcap_{N \in \mathbb{N}} T_{N,m}\right) = 0 \end{aligned}$$

Conversely, to prove ***not pointwise almost convergence***, we need to find a $x \in X$ and for an arbitrary fixed $m \ge 1$ such that

$$x \in \bigcap_{N \in \mathbb{N}} \bigcup_{n \ge N} \{x \in X : |f_n(x) - f(x)| \ge 1/m\} \setminus F = \limsup_{n \to \infty} \{x \in X \setminus F : |f_n(x) - f(x)| \ge 1/m\}.$$

3. The ***uniform convergence*** indicates that for each $m \geq 1$, there exists some $N(m) \geq 1$ (not depending on x) such that $T_{N,m} = \emptyset$. (i.e. $T_{N,m} \not\ni x$ for all $x \in X$.) So ***the tail support*** ***<u>is an empty set</u>***

4. The ***uniformly almost everywhere convergence*** indicates that there exists some null set F with $\mu(F) = 0$ such that for each $m \geq 1$, there exists some $N(m) \geq 1$ (not depending on x) such that $(T_{N,m} \backslash F) = \emptyset$. (i.e. $T_{N,m} \not\ni x$ for all $x \in X \backslash F$.) Equivalently, ***the tail support*** ***<u>is a null set</u>***:

$$
\begin{aligned}
& T_{N,m} = F \\
\Leftrightarrow \quad & \mu\left(T_{N,m}\right) = 0
\end{aligned}
$$

5. The ***almost uniform convergence*** indicates that for every δ, there exists *some measurable set* F_δ with $\mu(F_\delta) < \delta$ such that for each $m \geq 1$ there exists some $N(m) \geq 1$ (not depending on x) such that $(T_{N,m} \setminus F_\delta) = \emptyset$. (i.e. $T_{N,m} \not\ni x$ for all $x \in X \setminus F_\delta$.) Equivalently, ***<u>the measure of tail support shrinks to zero</u>***:

$$
\begin{aligned}
\mu\left(T_{N,m}\right) \leq \delta \quad & \Leftrightarrow \quad T_{N,m} = F_\delta \\
\lim_{N \to \infty} \mu\left(T_{N,m}\right) = 0 &
\end{aligned}
$$

6. The ***convergence in measure*** indicates that for any $m \geq 1$ and any $\delta > 0$, there exists $N \equiv N(m, \delta) \geq 1$ such that for all $n \geq N$, *the* ***<u>width</u>*** *of n-th event* ***<u>shrinks to zero</u>***:

$$
\begin{aligned}
& \mu(E_{n,m}) \leq \delta \\
\lim_{n \to \infty} & \mu(E_{n,m}) := \lim_{n \to \infty} \mu\left(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}\right) = 0
\end{aligned}
$$

- **Definition** Define the ***maximum variation*** between (f_n) and f as $\sup_{x \in X} |f_n(x) - f(x)|$. Note that

$$
\sup_{x \in X} |f_n(x) - f(x)| \geq \sup_{x \in X \backslash F, \mu(F) = 0} |f_n(x) - f(x)|.
$$

- **Remark** From *Borel-Cantelli Lemma*, we see that in order to show *the* ***pointwise almost everywhere*** *convergence*, i.e. $\mu\left(\bigcap_N T_{N,\epsilon}\right) = \mu(\limsup_{n \to \infty} E_{n,\epsilon}) = 0$ it suffice to show that ***the measure of the tail support is finite***, $\mu\left(T_{N,\epsilon}\right) = \sum_{n=N}^{\infty} \mu(E_{n,\epsilon}) < \infty$. Note that this condition implies that it not only converges ***in measure*** $\mu(E_{n,\epsilon}) \to 0$ but converge in ***an absolutely summable*** fashion.

6.2.1 Comparison

Table 4: Comparison of Modes of Convergence

	tail support	***width***	***maximum variation***	***subgraph***
definition	$T_{N,\epsilon} = \bigcup_{n \geq N} E_{n,\epsilon}$	$\mu(E_{n,\epsilon})$	$\sup_{x \in X} \{\lvert f_n(x) - f(x) \rvert\}$	$\Gamma(f_n) = \{(x,t) :$ $0 \leq t \leq f_n(x)\}$
pointwise	$\bigcap_{N=1}^{\infty} T_{N,\epsilon} = \emptyset$		*or*, $\to 0$ on X	
point-wise a.e.	$\mu\left(\bigcap_{N=1}^{\infty} T_{N,\epsilon}\right) = 0$		*or*, $\to 0$ on $X \setminus E$	
uniform	$T_{N,\epsilon} = \emptyset$		equivalently, $\to 0$ on X	
uniform a.e. / L^∞ norm	$\mu\left(T_{N,\epsilon}\right) = 0$		equivalently, $\to 0$ on $X \setminus E$	
almost uniform	$\lim_{N\to\infty} \mu\left(T_{N,\epsilon}\right) = 0$		or, $\to 0$ on $X \setminus E$	
in measure		$\lim_{n\to\infty} \mu\left(E_{n,\epsilon}\right) = 0$	or, $\to 0$ on $X \setminus E$	
L^1 ***norm***			$\to 0$ and support fixed or non-increasing	area of $\Gamma(f_n) = \mathcal{A}(\Gamma(f_n))$ $\lim_{n\to\infty} \mathcal{A}(\Gamma(\lvert f_n - f \rvert)) = 0$

6.3 Relationships between Different Modes of Convergence

- **Proposition 6.2** *[Tao, 2011]*
  Let $(X, \mathscr{F}, \mu)$ be a measure space, and let $f_n : X \to \mathbb{C}$ and $f : X \to \mathbb{C}$ be measurable functions

  1. *If f_n converges to f **uniformly**, then f_n converges to f **pointwisely**.*
  2. *If f_n converges to f **uniformly**, then f_n converges to f in L^∞ **norm**. **Conversely**, if f_n converges to f **in L^∞ norm**, then f_n converges to f **uniformly outside of a null set** (i.e. there exists a null set E such that the restriction $f_n|_{X/E}$ of f_n to the complement of E converges to the restriction $f|_{X/E}$ of f).*
  3. *If f_n converges to f in L^∞ **norm**, then f_n converges to f **almost uniformly**.*
  4. *If f_n converges to f **almost uniformly**, then f_n converges to f **pointwise almost everywhere**.*
  5. *If f_n converges to f **pointwise**, then f_n converges to f **pointwise almost everywhere**.*
  6. *If f_n converges to f in L^1 **norm**, then f_n converges to f **in measure**.*
  7. *If f_n converges to f **almost uniformly**, then f_n converges to f **in measure**.*

- **Remark** This diagram shows the *relative strength* of different *modes of convergence*. The direction arrows $A \to B$ means "if A holds, then B holds".

$$\begin{tikzcd}
\textit{uniform} \arrow[r] \arrow[dd] \arrow[rr, bend left] & \textit{uniformly a.e.} \arrow[r, leftrightarrow] \arrow[d] & \textit{in } L^\infty \textit{ norm} \arrow[dl] \\
 & \textit{almost uniform} \arrow[d] \arrow[dr] & \textit{in } L^1 \textit{ norm} \arrow[d] \\
\textit{pointwise} \arrow[r] & \textit{pointwise a.e.} & \textit{in measure}
\end{tikzcd}$$

  Moreover, here are some counter statements:

  - $L^\infty \not\to L^1$: see the "*Escape to Width Infinity*" example below.
  - ***uniform*** $\not\to L^1$: see the "*Escape to Width Infinity*" example below.
  - $L^1 \not\to$ ***uniform*** : see the "*Typewriter Sequence*" example below.
  - ***pointwise*** $\not\to L^1$: see the "*Escape to Horizontal Infinity*" example below.
  - ***pointwise*** $\not\to$ ***uniform***: see the "$f_n = x/n$" example above.
  - For finite measure space, ***pointwise a.e.*** $\to$ ***almost uniform***: see the Egorov's theorem.
  - ***almost uniform*** $\not\to L^1$: see the "*Escape to Vertical Infinity*" example below.
  - ***almost uniform*** $\not\to L^\infty$: see the "*Escape to Vertical Infinity*" example below. The *converse* is true, however.

- For bounded $f_n \leq G, a.e. \forall n$, then ***pointwise a.e.*** $\to L^1$: see *Dominated Convergence Theorem.*
- $L^1 \not\to$ ***pointwise a.e.*** : see the "*Typewriter Sequence*" example below.
- ***in measure*** $\not\to$ ***pointwise a.e.*** : see the "*Typewriter Sequence*" example below.
- $L^1 \to$ ***convergence in integral***: by *triangle inequality.* Note that *the other modes of convergence does* ***not directly*** *lead to convergence in integral.*

6.4 Counter Examples

- **Example** (***Escape to Horizontal Infinity***).
  Let X be the real line with Lebesgue measure, and let

$$f_n(x) \equiv \mathbb{1}_{[n,n+1]}.$$

  Note that the ***height*** and ***width do not shrink to zero***, but ***the tail set*** shrinks to ***the empty set.*** We have the following statements on different modes of convergence:

  1. f_n ***converges pointwise*** *to* $f = 0$, (thus ***pointwise a.e.***)
  2. f_n ***does not*** *converges to* $f = 0$ ***uniformly***,
  3. f_n ***does not*** *converges to* $f = 0$ in L^∞ ***norm***,
  4. f_n ***does not*** *converges to* $f = 0$ ***almost uniformly***
  5. f_n ***does not*** *converges to* $f = 0$ ***in measure***.
  6. $\int_{\mathbb{R}} f_n dx = 1$ ***does not*** *converge* to $\int_{\mathbb{R}} f dx = 0$.
  7. f_n ***does not*** *converges to* $f = 0$ ***in*** L^1 ***norm***.

  Somehow, *all the* ***mass*** in the f_n has *escaped* by *moving off to infinity* in a ***horizontal direction***, leaving none behind for the pointwise limit f. In *frequency domain*, it corresponds to ***escaping to spatial infinity***.

- **Example** (***Escape to Width Infinity***).
  Let X be the real line with Lebesgue measure, and let

$$f_n \equiv \frac{1}{n} \mathbb{1}_{[0,n]}.$$

  See that *the* ***height*** *goes to* ***zero***, but the ***width*** *(and* ***tail support****)* *go to* ***infinity***, causing the L^1 norm to stay ***bounded away from zero***. We have the following statements on different modes of convergence:

  1. f_n ***converges*** *to* $f = 0$ ***uniformly***. (Thus, ***pointwise***, ***pointwise a.e.***, ***uniformly a.e.***, ***almost uniformly***, ***in*** L^∞ ***norm*** and ***in measure***)
  2. $\int_{\mathbb{R}} f_n dx = 1$ ***does not converge*** *to* $\int_{\mathbb{R}} f dx = 0$. This is due to the ***increasingly wide*** *nature* of *the* ***support*** *of the* f_n. If all the f_n were supported *in a single set of finite measure*, this will not happen.
  3. f_n ***does not*** *converges to* $f = 0$ ***in*** L^1 ***norm***.

In *frequency domain*, it corresponds to ***escaping to zero frequency***.

- **Example** (***Escape to Vertical Infinity***).
Let X be the unit interval $[0, 1]$ with Lebesgue measure (restricted from $\mathbb{R}$), and let

$$f_n = n\mathbb{1}_{\left[\frac{1}{n}, \frac{2}{n}\right]}.$$

Note that the ***height*** *goes to* ***infinity***, but *the* ***width*** *(and* ***tail support****) go to* ***zero*** *(or* ***the empty set****)*, causing the L^1 *norm to stay* ***bounded away from zero***. We have the following statements on different modes of convergence:

1. f_n ***converges pointwise*** *to* $f = 0$, (thus ***pointwise a.e.***)
2. f_n ***converges*** *to* $f = 0$ ***almost uniformly***, (thus ***in measure***)
3. f_n ***does not*** *converges to* $f = 0$ ***uniformly***,
4. f_n ***does not*** *converges to* $f = 0$ in L^∞ ***norm***,
5. $\int_{\mathbb{R}} f_n dx = 1$ ***does not*** *converge* to $\int_{\mathbb{R}} f dx = 0$.
6. f_n ***does not*** *converges to* $f = 0$ ***in*** L^1 ***norm***.

Note that we have finite measure on $X = [0, 1]$. This time, the mass has *escaped* ***vertically***, through *the* ***increasingly large*** *values of* f_n. In *frequency domain*, it corresponds to ***escaping to infinity frequency***.

- **Example** (***Typewriter Sequence***).
Let f_n be defined by the formula

$$f_n \equiv \mathbb{1}\left\{x \in \left[\frac{n - 2^k}{2^k}, \frac{n + 1 - 2^k}{2^k}\right]\right\}$$

whenever $k \geq 0$ and $2^k \leq n < 2^k + 1$. This is a sequence of indicator functions of ***intervals*** of ***decreasing length***, ***marching*** *across the unit interval* $[0, 1]$ ***over and over again***. See that *the* ***width*** *goes to* ***zero***, but *the* ***height*** *and the* ***tail support stay fixed*** (and thus ***bounded away from zero***). We have the following statements on different modes of convergence:

1. f_n ***converges*** *to* $f = 0$ ***in*** L^1 ***norm***, (thus ***in measure***)
2. f_n ***does not*** *converges to* $f = 0$ ***pointwise a.e.***, (thus ***not pointwise***, ***not almost uniformly***, ***not uniformly a.e.***, ***not uniformly***, ***not in*** L^∞ ***norm***)

6.5 Uniqueness

- **Proposition 6.3** *Let* $f_n : X \to \mathbb{C}$ *be a sequence of measurable functions, and let* $f, g : X \to \mathbb{C}$ *be two additional measurable functions. Suppose that* f_n *converges to* f ***along one of the seven modes of convergence*** *defined above, and* f_n *converges to* g ***along another of the seven modes of convergence*** *(or perhaps the same mode of convergence as for* f*). Then* f *and* g ***agree almost everywhere***.

- **Remark** It suffice to show that when f_n converges to f ***pointwise almost everywhere***, and f_n converges to g ***in measure***. We need to show that $f = g$ *almost everywhere*.

- **Remark** Even though the modes of convergence all *differ* from each other, they are all ***compatible*** in the sense that they ***never disagree*** about *which function* f a sequence of functions f_n ***converges to***, *outside of a set of measure zero.*

6.6 Modes of Convergence for Step Functions

- **Remark** Consider the ***step function*** f_n as a constant multiple $f_n = A_n \mathbb{1}\{E_n\}$ of a measurable set E_n, which has a limit $f = 0$.
- **Definition** The modes of convergence for step function f_n is determined by the following quantities:
  1. the n-th ***width*** of f_n is $\mu(E_n)$;
  2. the n-th ***height*** of f_n is A_n;
  3. the N-th ***tail support*** $T_N \equiv \bigcup_{n \geq N} E_n$ of the sequence $f_1, f_2, f_3, \ldots$.
- **Remark** Assume the ***height*** A_n exhibit *one of two modes of behaviour*:
  1. $A_n \to 0$, ***converge to zero***;
  2. (A_n) are ***bounded away from zero*** (i.e. there exists $c > 0$ such that $A_n \geq c$ for every n.)
- **Proposition 6.4** *The following regarding the seven modes of convergence of $f_n = A_n \mathbb{1}\{E_n\}$ to $f = 0$:*
  1. *f_n converges **uniformly** to zero if and only if $A_n \to 0$ as $n \to \infty$.*
  2. *f_n converges **in L^∞ norm** to zero if and only if $A_n \to 0$ as $n \to \infty$.*
  3. *f_n converges **almost uniformly** to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\mu(T_N) \to 0$ as $N \to \infty$.*
  4. *f_n converges **pointwise** to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^{\infty} T_N = \emptyset$.*
  5. *f_n converges **pointwise almost everywhere** to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^{\infty} T_N$ is a null set.*
  6. *f_n converges **in measure** to zero if and only if $A_n \to 0$ as $n \to \infty$, or or $\mu(E_n) \to 0$ as $n \to \infty$.*
  7. *f_n converges **in L^1 norm** if and only if $A_n \mu(E_n) \to 0$ as $n \to \infty$.*
- **Remark** We summarize the above proposition:
  - When the ***height*** goes to ***zero***, then one has convergence to zero in ***all modes except possibly for L^1 convergence***, which requires that the ***product*** of the ***height*** and the ***width*** goes to zero.
  - If the ***height** is **bounded away from zero*** *(positive)* and the ***width*** is ***positive*** (finite support), then we ***never*** have ***uniform*** or L^1 convergence.
    * If ***the width goes to zero***, we have convergence in ***measure***.

- ∗ If ***the measure of tail support goes to zero***, we have ***almost uniform*** *convergence.*

- ∗ If ***the tail support shrinks** to a **null set***, we have ***pointwise almost everywhere*** *convergence.*

- ∗ If ***the tail support shrinks** to **the empty set***, we have ***pointwise*** *convergence.*

6.7 Modes of Convergence With Additional Conditions

6.7.1 Finite Measure Space

- **Remark** If we assume that $(X, \mathscr{B}, \mu)$ has ***finite measure***, i.e. $\mu(X) < \infty$, we can shut down two of the four examples (namely, ***escape to horizontal infinity*** or ***escape to width infinity***) and creates a few more equivalences.

- **Example** *A probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ is a finite measure space since $\mathbb{P}(\Omega) = 1$.

- **Theorem 6.5** *(**Egorov's theorem**). [Royden and Fitzpatrick, 1988, Tao, 2011]*
*Let $(X, \mathscr{F}, \mu)$ be a **finite measure space**, that is, $\mu(X) < \infty$ and let $f_n : X \to \mathbb{C}$ be a sequence of measurable functions that converge **pointwise almost everywhere** to another function $f : X \to \mathbb{C}$, and let $\epsilon > 0$. Then there exists a μ-measurable set A of measure at most ϵ, such that f_n **converges uniformly** to f **outside** of A. That is, given finite measure space, convergence pointwise almost everywhere implies **converge almost uniformly**.*

- **Remark** *The **finite measure space*** condition allows us to use ***the downward convergence*** of measure without much concern.

- **Proposition 6.6** *Let X have **finite measure**, and let $f_n : X \to \mathbb{C}$ and $f : X \to \mathbb{C}$ be measurable functions. If f_n converges to f in L^∞ norm, then f_n also converges to f in L^1 norm.*

- **Remark** For finite measure space,

$$
\begin{tikzcd}
\textit{uniform} \arrow[r] \arrow[dd] \arrow[rr, bend left] & \textit{uniformly a.e.} \arrow[r, leftrightarrow] \arrow[d] & \textit{in } L^\infty \textit{ norm} \arrow[dl] \arrow[d, red] \\
 & \textit{almost uniform} \arrow[d] \arrow[dr] & \textit{in } L^1 \textit{ norm} \arrow[d] \\
\textit{pointwise} \arrow[r] & \textit{pointwise a.e.} \arrow[u, red, bend left] & \textit{in measure}
\end{tikzcd}
$$

6.7.2 Fast L^1 Convergence

- **Proposition 6.7** *(**Fast L1 convergence**).*
*Suppose that $f_n, f : X \to \mathbb{C}$ are measurable functions such that $\sum_{n=1}^{\infty} \|f_n f\|_{L^1(\mu)} < \infty$; thus, not only do the quantities $\|f_n f\|_{L^1(\mu)}$ go to zero (which would mean L^1 convergence), but they converge in **an absolutely summable** fashion. Then*

1. *f_n converges **pointwise almost everywhere** to f.*

2. *f_n converges **almost uniformly** to f.*

- **Corollary 6.8** *(**Subsequence Convergence**). [Tao, 2011]*
*Suppose that $f_n : X \to \mathbb{C}$ are a sequence of measurable functions that converge in L^1 **norm** to a limit f. Then there exists a **subsequence** f_{n_j} that converges **almost uniformly** (and hence, **pointwise almost everywhere**) to f (while remaining convergent in L^1 norm, of course).*

- **Corollary 6.9** *(**Subsequence Convergence in Measure**). [Tao, 2011]*
*Suppose that $f_n : X \to \mathbb{C}$ are a sequence of measurable functions that **converge in measure** to a limit f. Then there exists a subsequence f_{n_j} that converges **almost uniformly** (and hence, **pointwise almost everywhere**) to f.*

- **Remark** It is instructive to see how this ***subsequence*** is extracted in the case of *the typewriter sequence*. In general, one can view the operation of passing to a subsequence as being able to ***eliminate*** "***typewriter***" situations in which *the tail support is much larger than the width*.

6.7.3 Domination and Uniform Integrability

- **Remark** Now we turn to the *reverse question*, of *whether **almost uniform** convergence, **pointwise almost everywhere** convergence, or convergence **in measure** can imply L^1 **convergence***. The *escape to vertical* and *width infinity* examples shows that without any further hypotheses, the answer to this question is ***no***.

- **Remark** [Tao, 2011] There are ***two major ways*** to shut down loss of mass via ***escape to infinity***.

  1. One is to enforce ***monotonicity***, which ***prevents each f_n from abandoning the location*** where the mass of the preceding $f_1, \ldots, f_{n-1}$ was concentrated and which thus shuts down the above three escape scenarios. More precisely, we have the monotone convergence theorem.

  2. The other major way is to ***dominate*** *all of the functions involved by an **absolutely convergent one***. This result is known as the dominated convergence theorem.

- **Definition** We say that a sequence $f_n : X \to \mathbb{C}$ is ***dominated*** if there exists an ***absolutely integrable function*** $g : X \to \mathbb{C}$ such that $|f_n(x)| \le g(x)$ for all n and *almost every x*.

- **Definition** (***Uniform integrability***).
A sequence $f_n : X \to \mathbb{C}$ of ***absolutely integrable*** functions is said to be ***uniformly integrable*** if the following three statements hold:

  1. (***Uniform bound on L^1 norm***) One has $\sup_n \|f_n\|_{L^1(\mu)} = \sup_n \int_X |f_n|\, d\mu < +\infty$.

  2. (***No escape to vertical infinity***) One has

$$\lim_{M\to+\infty} \sup_n \int_{|f_n|\ge M} |f_n|\, d\mu \to 0.$$

3. (***No escape to width infinity***) One has

$$\lim_{\delta \to 0} \sup_{n} \int_{|f_n| \le \delta} |f_n| \, d\mu \to 0.$$

- **Proposition 6.10** *(Property of Uniform Integrablility)*

  1. *If f is an **absolutely integrable** function, then the constant sequence $f_n = f$ is **uniformly integrable**. (Hint: use the monotone convergence theorem.)*

  2. *Every **dominated** sequence of measurable functions is **uniformly integrable**.*

- **Exercise 6.11** *Give an example of a sequence f_n of uniformly integrable functions that converge **pointwise almost everywhere** to zero, but do **not converge almost uniformly**, **in measure**, or **in** L^1 **norm**.*

- **Theorem 6.12** ***(Uniformly integrable convergence in measure)***.
  *Let $f_n : X \to \mathbb{C}$ be a **uniformly integrable** sequence of functions, and let $f : X \to \mathbb{C}$ be another function. Then f_n converges in L^1 **norm** to f **if and only** if f_n converges to f **in measure**.*

- **Proposition 6.13** *Suppose that $f_n : X \to \mathbb{C}$ are a **dominated** sequence of measurable functions, and let $f : X \to \mathbb{C}$ be another measurable function. Show that f_n converges **pointwise almost everywhere** to f **if and only if** f_n converges in **almost uniformly** to f.*

6.8 Convergence in Distribution

- **Remark** (***Convergence of Measures Induced by Function***)
  Convergence in distribution is also called ***weak convergence*** in probability theory [Folland, 2013]. In general, it is actually ***not*** *a mode of* ***convergence of functions*** f_n ***itself*** but instead is the ***convergence of measures induced*** *by function f_n on* $\mathscr{B}(\mathbb{R})$.

  In functional analysis, however, ***weak convergence*** is actually reserved for a different mode of convergence, while ***the convergence in distribution*** is ***the weak* convergence***.

  $$\text{weak convergence} \qquad \int f_n d\mu \to \int f d\mu, \quad \forall \mu \in \mathcal{M}(X),$$
  $$\text{convergence in distribution} \qquad \int f d\mu_n \to \int f d\mu, \quad \forall f \in \mathcal{C}_0(X)$$

  Definition (***Weak* Topology on Banach Space***)
  Let X be a *normed vector space* and X^* be its dual space. The ***weak* topology*** on X^* is *the weakest topology* on X^* so that $f(x)$ *is* ***continuous for all*** $x \in X$.

  The weak topology* on space of regular Borel measures $\mathcal{M}(X) \simeq (\mathcal{C}_0(X))^*$ on a ***compact Hausdorff*** space X, is often called ***the vague topology***. Note that $\mu_n \stackrel{w^*}{\to} \mu$ if and only if $\int f d\mu_n \to \int f d\mu$ for all $f \in \mathcal{C}_0(X)$.

- **Definition** (***Cumulative Distribution Function***) [Billingsley, 2008]
  Let $(X, \mathscr{F}, \mu)$ be a measure space. Given any real-valued measurable function $f : X \to \mathbb{R}$,

we define the ***cumulative distribution function*** $F : \mathbb{R} \to [0, \infty]$ of f to be the function $F(\lambda) := \mu_f((-\infty, \lambda]) = \mu(\{x \in X : f(x) \le \lambda\})$ where $\mu_f = \mu \circ f^{-1}$ is a ***measure*** on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ induced by function f.

- **Definition** (***Converge in Distribution***) [Van der Vaart, 2000]
  Let $(X, \mathscr{F}, \mu)$ be a measure space, $f_n : X \to \mathbb{R}$ be a sequence of real-valued *measurable functions*, and $f : X \to \mathbb{R}$ be another measurable function.

  We say that f_n ***converges in distribution*** to f if *the cumulative distribution function* $F_n(\lambda)$ of f_n converges ***pointwise*** to *the cumulative distribution function* $F(\lambda)$ of f at all $\lambda \in \mathbb{R}$ for which F is *continuous*. Denoted as $f_n \xrightarrow{d} f$ or $f_n \rightsquigarrow f$.

  Note that for the distribution $\mu_{f_n} \equiv \mu \circ f_n^{-1}$ is a measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Thus $f_n \xrightarrow{d} f$ if and only if

$$\mu_{f_n}(A) \to \mu_f(A), \quad \forall A \in \mathscr{B}(\mathbb{R}).$$

- **Theorem 6.14** ***(The Portmanteau Theorem)****. [Van der Vaart, 2000]*
  The following statements are equivalent.

  1. $X_n \rightsquigarrow X$.
  2. $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$ *for all* ***continuous functions*** $h : \mathbb{R}^d \to \mathbb{R}$ *that are non-zero only on a* ***closed*** *and* ***bounded*** *set.*
  3. $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$ *for all* ***bounded continuous functions*** $h : \mathbb{R}^d \to \mathbb{R}$.
  4. $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$ *for all* ***bounded measurable functions*** $h : \mathbb{R}^d \to \mathbb{R}$ *for which* $\mathbb{P}(X \in \{x : h \text{ is continuous at } x\}) = 1$.

- We can reformulate the definition of *convergence in distribution* as below:

  Definition [Wellner et al., 2013]
  Let (Ω, d) be a *metric space*, and $(\Omega, \mathscr{B})$ be *a measurable space*, where $\mathscr{B}$ is ***the Borel σ-field on*** Ω, the smallest σ-field containing *all the open balls* (as the basis of *metric topology* on Ω). Let $\{P_n\}$ and P be ***Borel probability measures*** on $(\Omega, \mathscr{B})$.

  Then the sequence P_n ***converges in distribution*** to P, which we write as $P_n \rightsquigarrow P$, if and only if

$$\int_{\Omega} f dP_n \to \int_{\Omega} f dP, \qquad \text{for all } f \in \mathcal{C}_b(\Omega).$$

  Here $\mathcal{C}_b(\Omega)$ denotes the set of *all* ***bounded, continuous****, real functions on* Ω.

  We can see that ***the convergence in distribution*** is actually ***a weak* convergence***. That is, it is ***the weak convergence*** of ***bounded linear functionals*** $I_{\mathcal{P}_n} \xrightarrow{w^*} I_{\mathcal{P}}$ on *the space of all probability measures* $\mathcal{P}(\mathcal{X}) \simeq (\mathcal{C}_b(\mathcal{X}))^*$ on $(\mathcal{X}, \mathscr{B})$ where

$$I_{\mathcal{P}} : f \mapsto \int_{\Omega} f d\mathcal{P}.$$

  Note that the $I_{\mathcal{P}_n} \xrightarrow{w^*} I_{\mathcal{P}}$ is equivalent to $I_{\mathcal{P}_n}(f) \to I_{\mathcal{P}}(f)$ *for all* $f \in \mathcal{C}_b(\mathcal{X})$.

References

Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

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.

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

Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Jon Wellner et al. *Weak convergence and empirical processes: with applications to statistics*. Springer Science & Business Media, 2013.