Mathematical Physics

Chapter I

Sets and structures

1) Naive set theory.

1.1. Sets and subsets.

$S = \{x \mid P(x)\}$, where $P(x)$ is a proposition and x is a variable

- Although all mathematics can be reduced to set theory, set theory itself is not reducable to pure logic.
- Naive set theory leads to a collection of self-referential paradoxes, e.g. Russell's paradox:

$$R = \{A \mid A \notin A\} \qquad \begin{array}{l} R \in R \Rightarrow R \notin R \\ R \notin R \Rightarrow R \in R \end{array} \quad \text{contradiction}$$

These kinds of paradoxes are resolved in Zermelo-Fraenkel axiomatic system, e.g. some of the axioms:

$\forall P(x)\, \forall a\, (a \in \{x \mid P(x)\} \Leftrightarrow P(a))$ main axiom of set theory

$A = B \Leftrightarrow \forall a\, (a \in A \Leftrightarrow a \in B)$ axiom of extensionality

- We will be working with naive set theory, since its paradoxes supposedly are irrelevant to physics.

Examples:

1) $\{a\} = \{x \mid x = a\}$ singleton
2) $\{a_1, \dots, a_n\} = \{x \mid x = a_1 \vee \dots \vee x = a_n\}$ finite set
3) $\mathcal{A} = \{A_i \mid i \in I\}$, I is an indexing set family of indexed sets (collection)

$A \subseteq B \overset{\text{def}}{\Longleftrightarrow} (a \in A \Rightarrow a \in B)$ A is a subset of B (or B is a superset of A)

$A = B \Leftrightarrow (a \in A \Leftrightarrow B \in B) \Leftrightarrow A \subseteq B \wedge B \subseteq A$

Examples:

1) $\emptyset$: $\forall a\, (a \notin \emptyset)$ empty set

   $\forall A\, (\emptyset \subseteq A)$

2) 2^A: the set of all subsets of A power set of A (alternative notation: $\mathbb{P}(A) = P(A) = 2^A$)

   e.g. if $A = \{a_1, \dots, a_n\}$, then $|2^A| = \underbrace{1}_{\emptyset} + \underbrace{\binom{n}{1}}_{\{a_1\},\dots,\{a_n\}} + \underbrace{\binom{n}{2}}_{\{a_1,a_2\},\dots \{a_{n-1},a_n\}} + \dots + \underbrace{\binom{n}{n}}_{\{a_1,\dots,a_n\}} = (1+1)^n = \underbrace{2^n}_{\text{hence the symbol}}$

1.2. Unions and intersections.

$A \cup B \overset{\text{def}}{=} \{x \mid x \in A \vee x \in B\}$ union

$A \cap B \overset{\text{def}}{=} \{x \mid x \in A \wedge x \in B\}$ intersection

$A \setminus B \overset{\text{def}}{=} \{x \mid x \in A \wedge x \notin B\}$ difference

Properties:

- A and B are disjoint: $A \cap B = \emptyset$
- $\bigcup \mathcal{A} \overset{\text{def}}{=} \{x \mid \exists A \in \mathcal{A}\,(x \in A)\}$ $\quad$ $\bigcap \mathcal{A} \overset{\text{def}}{=} \{x \mid \forall A \in \mathcal{A}\,(x \in A)\}$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $\quad$ $A \cap \bigcup \mathcal{B} = \bigcup \{A \cap B_i \mid i \in I\}$

  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $\quad$ $A \cup \bigcap \mathcal{B} = \bigcap \{A \cup B_i \mid i \in I\}$
- $A \cap (B \cup C) = (A \cap B) \cup C \iff C \subseteq A$
- $A \cap B = (A^c \cup B^c)^c$

  $A \cup B = (A^c \cap B^c)^c$
- $A \setminus B = A \cap B^c$
- $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ $\quad$ $A \setminus \bigcup \mathcal{B} = \bigcap \{A \setminus B_i \mid i \in I\}$

  $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$ $\quad$ $A \setminus \bigcap \mathcal{B} = \bigcup \{A \setminus B_i \mid i \in I\}$
- $2^A \cap 2^B = 2^{A \cap B}$ $\quad$ $\bigcap_{A \in \mathcal{A}} 2^A = 2^{\bigcap \mathcal{A}}$

  $2^A \cup 2^B \subseteq 2^{A \cup B}$ $\quad$ $\bigcup_{A \in \mathcal{A}} 2^A \subseteq 2^{\bigcup \mathcal{A}}$

② Relations.

2.1. Cartesian products.

Sets are not endowed with the notion of order, i.e. $\{a,b\} = \{b,a\}$, but we can artificially introduce this notion:

$(a,b) \stackrel{\text{def}}{\equiv} \{a, \{a,b\}\}$ ordered pair

$(a_1,\dots,a_n) \stackrel{\text{def}}{\equiv} (a_1, (a_2,\dots,a_n))$ ordered n-tuple

Now we can define the cartesian product of sets:

$A \times B \stackrel{\text{def}}{\equiv} \{(a,b) \mid a \in A \wedge b \in B\}$

$A_1 \times \dots \times A_n \stackrel{\text{def}}{\equiv} \{(a_1,\dots,a_n) \mid a_1 \in A_1 \wedge \dots \wedge a_n \in A_n\}$ (e.g. $\underbrace{A \times \dots \times A}_{n} \equiv A^n$)

Properties:

- $(A \cup B) \times C = (A \times C) \cup (B \times C)$ $\quad \bigcup \mathcal{A} \times \bigcup \mathcal{B} = \bigcup A_i \times B_j$
  $(A \cap B) \times C = (A \times C) \cap (B \times C)$ $\quad \bigcap \mathcal{A} \times \bigcap \mathcal{B} = \bigcap A_i \times B_j$
- $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$

2.2. Relations.

$R \subset A^n$: n-ary relation on A

Consider a binary relation ($n=2$) on A. Common notation: aRb (instead of $(a,b) \in R$).

R is reflexive : $aRa \ \forall a \in A$
R is transitive : $(aRb \wedge bRc) \Rightarrow aRc \ \forall a,b,c \in A$
R is symmetric : $aRb \Rightarrow bRa \ \forall a,b \in A$
R is anti-symmetric : $(aRb \wedge bRa) \Rightarrow a=b \ \forall a,b \in A$

1. Equivalence.

R is an equivalence relation on A: R is reflexive, transitive, and symmetric.

Equivalence relation partitions A into disjoint equiv. classes $[a]_R \stackrel{\text{def}}{\equiv} \{a' \in A \mid aRa'\}$, i.e. $\top(([a]=[b]) \vee ([a] \cap [b] = \emptyset))$. Now we can define the set of all equiv. classes of A w.r.t. R:

$A/R \stackrel{\text{def}}{\equiv} \{[a]_R \mid a \in A\}$ factor space of A

Examples:

1) $A = \mathbb{Z}$ $\quad m = n \pmod p \Leftrightarrow m - n = kp, \ k,m,n \in \mathbb{Z} \ p \in \mathbb{N}$
   $\mathbb{Z}_p = \{[0]_p, [1]_p, \dots, [p-1]_p\}$ residue classes mod p
2) $A = \mathbb{R}^2$ $\quad (x,y) \sim (x',y') \Leftrightarrow x' = x + m \quad y' = y + n, \ n,m \in \mathbb{Z}$
   $T^2 = \mathbb{R}^2/\sim = \{[(x,y)] \mid 0 \leq x < 1 \ 0 \leq y < 1\}$

2. Order relations.

R is a partial order on A : R is reflexive, transitive, and antisymmetric.
R is a total order on A : R is a partial order and $\top(aRb \vee bRa)$

Then (A,R) is a poset (partially ordered set), e.g. $(\mathbb{Q}, \leq), (2^A, \subseteq)$.

③ Mappings.

$\varphi: X \to Y$ is a mapping from X to Y: $\varphi = \{(x,y) \mid \forall x \in X \ \exists! y \in Y\}$

i.e. $((x,y) \in \varphi \wedge (x,y') \in \varphi) \Rightarrow y = y'$

X: domain of φ

Y: codomain of φ
(range)

$\varphi[X] \stackrel{def}{=} \{y \in Y \mid y = \varphi(x),\ x \in X\}$ image of φ

$\varphi^{-1}[Y] \stackrel{def}{=} \{x \in X \mid \varphi(x) \in Y\}$ preimage of φ
(inverse image)

Examples:

1) $\sin: \mathbb{R} \to \mathbb{R}$ $\quad \sin^{-1}[\{0\}] = \{0, \pm\pi, \pm 2\pi, \dots\}$

$\sin^{-1}[\{2\}] = \emptyset$

2) $\varphi: X_1 \times \dots \times X_n \to Y$ n-ary function

3) $pr_i: X_1 \times \dots \times X_n \to X_i$ projection on X_i

$(x_1, \dots, x_n) \mapsto x_i$

4) $\psi \circ \varphi : X \to Z$ composition of $\varphi: X \to Y$ and $\psi: Y \to Z$

$\psi \circ \varphi(x) = \psi(\varphi(x))$ satisfies $\alpha \circ (\beta \circ \gamma) = (\alpha \circ \beta) \circ \gamma$

$\varphi: X \to Y$ is surjective : $\forall y \in Y\ \exists x \in X\ (\varphi(x) = y)$ onto

$\varphi: X \to Y$ is injective : $\varphi(x) = \varphi(x') \Rightarrow x = x'$ one-to-one

$\varphi: X \to Y$ is bijective : $\forall y \in Y\ \exists! x \in X\ (\varphi(x) = y)$ one-to-one correspondence

- $\varphi: X \to X$ is a transformation of X: φ is bijective. e.g. $id_X: X \to X$, $x \mapsto x$
- If $\varphi: X \to Y$ is bijective, then $\exists \varphi^{-1}: Y \to X$ s.t. $\varphi^{-1} \circ \varphi = id_X$ and $\varphi \circ \varphi^{-1} = id_Y$.
- If $\varphi: X \to Y$ and $\psi: Y \to Z$ are bijective, then $(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}$.
- $\varphi|_U$: restriction of $\varphi: X \to Y$ to $U \subset X$.
- $\forall \varphi: X \to Y$ defines equiv. relation on X: $xRx' \stackrel{def}{\Leftrightarrow} \varphi(x) = \varphi(x')$. Then $X/R = \{\varphi^{-1}(\{y\}) \mid y \in Y\}$.

Examples:

1) $i_U = id_X|_U$ inclusion map for $U \subset X$

Then $\forall \varphi: X \to Y$ $\varphi|_U = \varphi \circ i_U$

2) $\chi_U: X \to \{0,1\}$ characteristic function of $U \subset X$

$\chi_U(x) = \begin{cases} 0, & x \notin U \\ 1, & x \in U \end{cases} \Rightarrow 2^X = \{\varphi \mid \varphi: X \to \{0,1\}\}$

3) $\varphi: X \to X/\sim$ canonical map under $\sim$ (surjective)

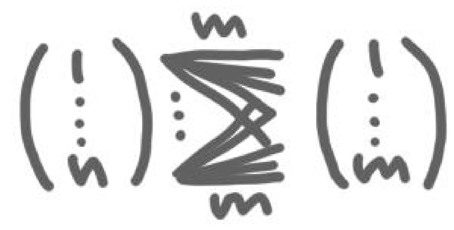
$\varphi(x) = [x]_\sim$

④ Infinite sets.

A is finite : $\exists$ bijection $\varphi: A \to \{1, 2, \dots, n\}$, where $|A| = n$: cardinality of A
A is infinite : A is not finite

Examples:

1) $B^A \stackrel{\text{def}}{=} \{\varphi \mid \varphi: A \to B\}$
   Reasoning for notation: if A and B are finite, then $|B^A| = \underbrace{m \cdot \ldots \cdot m}_{n \text{ times}} = m^n = |B|^{|A|}$



2) $2^A = \{\varphi \mid \varphi: A \to \{0,1\}\} = \{0,1\}^A$
   $|2^A| = 2^{|A|}$

A is countable : $\exists$ bijection $\varphi: A \to \mathbb{N}$
A is uncountable : A is not countable

- If A is countable, then $\forall\, A' \subset A$ is either countable or finite.
- If A and B are countable, then $A \times B$ is countable.
- If A is countable, then 2^A is uncountable.

Examples:

1) $\mathbb{Z}$ is countable, $\mathbb{Q} = \mathbb{Z} \times \mathbb{N}$ is countable
2) $\mathbb{R}$ is uncountable
   $\forall\, x \in [0,1]$ $x = 0.\epsilon_1\epsilon_2\ldots$, where $\epsilon_i \in \{0,1\}$, then $[0,1] = \{0,1\}^{\mathbb{N}} = 2^{\mathbb{N}}$, i.e. $|\mathbb{R}| = 2^{|\mathbb{N}|}$.
3) Cantor set is uncountable
   $x = 0.\epsilon_1\epsilon_2\ldots$, where $\epsilon_i \in \{0,1,2\}$ but let $\epsilon_i \neq 1$ $\Rightarrow$ uncountable
   This set is nowhere dense in $\mathbb{R}$ (i.e. $\forall\, (a,b) \subset \mathbb{R}$ it's not dense)

4) $S = \{f \mid f: \mathbb{R} \to \mathbb{R}\}$ is uncountable and $|S| > |\mathbb{R}|$

⑤ Physics.

Physical theories have two aspects: static and dynamic.

1. Static.
   Background in which the theory is set, i.e. the mathematical structure.

   Math. structures:
   - algebraic (set + relations imposed on a set)
   - geometric (set + relations imposed on the power set of a set)

2. Dynamic.
   Laws of physics are usually written as differential equations.

⑥ Category theory.

C is a **category**: 1. $Ob(C)$ class of **objects**

2. $Hom(C)$ class of **morphisms** between the objects

3. $\circ : Hom(X,Y) \times Hom(Y,Z) \to Hom(X,Z)$ **composition** of morphisms

(assoc.) $h \circ (g \circ f) = (h \circ g) \circ f \quad \forall f,g,h$

(ident.) $\forall X \in Ob(C)\ \exists i_X \in Hom(X,X)$ s.t. $\forall f \in Hom(X,Y) \quad f \circ i_X = f$
$\forall g \in Hom(Y,X) \quad i_X \circ g = g$

Examples:

sets	mappings between sets
groups	homomorphisms
vector spaces	linear maps
algebras	algebra homomorphisms
topological spaces	homeomorphisms
smooth manifolds	diffeomorphisms

$f: X \to Y$ is a **monomorphism** : $\forall A \in Ob(C)\ \forall \alpha, \alpha' \in Hom(A,X)$
$f \circ \alpha = f \circ \alpha' \Rightarrow \alpha = \alpha'$

$f: X \to Y$ is an **epimorphism** : $\forall B \in Ob(C)\ \forall \beta, \beta' \in Hom(Y,B)$
$\beta \circ f = \beta' \circ f \Rightarrow \beta = \beta'$

$f: X \to Y$ is an **isomorphism** : $\exists f' \in Hom(Y,X)$ s.t. $f' \circ f = i_X$, $f \circ f' = i_Y$

$f: X \to X$ is an **endomorphism** : $f \in Hom(X,X)$

$f: X \to X$ is an **automorphism** : $f \in Hom(X,X)$ an isomorphic

$$\begin{tikzcd} & A \arrow[dl, "\alpha"'] \arrow[dr, "f\circ\alpha"] & \\ X \arrow[rr, "f"'] & & Y \end{tikzcd}$$

If f is monic, then α is uniquely defined

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "\beta\circ f"'] & & Y \arrow[dl, "\beta"] \\ & B & \end{tikzcd}$$

If f is epic, then β is uniquely defined

- f is an isomorphism $\Rightarrow$ f is a monomorphism and an epimorphism
- f is a monomorphism and an epimorphism $\not\Rightarrow$ f is an isomorphism
- For the category of sets:

  f is a monomorphism $\iff$ f is injective

  f is an epimorphism $\iff$ f is surjective

  f is an isomorphism $\iff$ f is bijective

Chapter V
Measure theory and integration

① Measurable spaces.

1.1. Overview.

- Topology ignores the notion of "size", measure theory is the area of mathematics concerned with these sorts of properties. A measure space is the structure that defines which sets are measurable (analogous to a topology, telling which sets are open).
- Firstly a measure space requires a σ-algebra imposed on the power set of the underlying space. Then a measure can be defined, it's a positive $\mathbb{R}$-valued function on the σ-algebra that is countably additive (i.e. the measure of a union of disjoint sets is the sum of their measures).
- There are other restrictions imposed on a measure. By general reckoning the broadest useful measure on $\mathbb{R}^n$ is the Lebesgue measure.

1.2. Measurable spaces.

$\mathcal{M} \subseteq 2^X$ is a σ-algebra on X: (1) $\emptyset \in \mathcal{M}$ — $\emptyset$ is measurable

collection of measurable subsets of X

(2) $A \in \mathcal{M} \Rightarrow A^c = X \setminus A \in \mathcal{M}$ — $\mathcal{M}$ is closed under complementation in X

(3) $A_i \in \mathcal{M},\ i \in I \Rightarrow \bigcup_i A_i \in \mathcal{M}$ — $\mathcal{M}$ is closed under countable unions

$(X, \mathcal{M})$ is a measurable space: $\mathcal{M}$ is a σ-algebra on X.

Properties:

- (1) and (2) $\Rightarrow X = \emptyset^c \in \mathcal{M}$
- $A \cap B = (A^c \cup B^c)^c \Rightarrow A \cap B \in \mathcal{M}$
- $A \setminus B = A \cap B^c = (A^c \cup B)^c \Rightarrow A \setminus B \in \mathcal{M}$
- difference between (meas. 2) and (top. 2): compl. of an open set is closed (most of the times)
- If $\mathcal{M}_i$, $i \in I$, are σ-algebras on X, then $\bigcap_{i \in I} \mathcal{M}_i$ is also a σ-algebra. Therefore, given $\mathcal{A} \subset 2^X$, there is a unique "smallest" σ-algebra $S \supseteq \mathcal{A}$, i.e. the intersection of all σ-algebras that contain $\mathcal{A}$. It's called the σ-algebra generated by $\mathcal{A}$.

  Borel sets on (X, τ): σ-algebra on X generated by open sets. (includes open, closed, and clopen sets)
- If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable, then $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ is measurable, where $\mathcal{M} \otimes \mathcal{N}$ is a σ-algebra generated by all sets of the form $A \times B$, $A \in \mathcal{M}$ $B \in \mathcal{N}$.

Examples:

1) $\mathcal{M} = \{\emptyset, X\}$ $\mathcal{N} = 2^X$ are the smallest and the largest σ-algebras of X

2) $X = \mathbb{R}$ with the standard topology, i.e. open sets are countable unions of open intervals. Hence Borel sets are generated by $\{(a,b) \mid a < b\}$. Using the axioms (1)-(3) one can show that $(-\infty, a), (a, +\infty), (-\infty, a], [a, +\infty), [a, b], \{a\}$ are also Borel sets.

② Measurable functions.

2.1. Measurable functions.

Consider measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$, then:

$f: X \to Y$ is **measurable** : $A \in \mathcal{N} \Rightarrow f^{-1}(A) \in \mathcal{M}$ i.e. preimage of every measurable set is measurable

- This definition mirrors the definition of a continuous function between topological spaces.
- If (X, τ_X) and (Y, τ_Y) are topological spaces and $\mathcal{M}$ and $\mathcal{N}$ are σ-algebras of Borel sets on them, then $\forall f \in C^0(X,Y)$ is measurable.
- Consider $f: X \to \mathbb{R}$. Since the family of Borel sets on $\mathbb{R}$ is generated by the intervals $(a, +\infty)$, we can formulate a criterion for measurability:

  $f: X \to \mathbb{R}$ is measurable $\Leftrightarrow$ $\forall a \in \mathbb{R}$ $f^{-1}((a, +\infty)) = \{x \mid f(x) > a\}$ is measurable.

  Example:

  $\chi_A: X \to \mathbb{R}$, $A \subset X$, is measurable $\Leftrightarrow$ A is measurable since $\{x \mid \chi_A(x) > a\} = \begin{cases} X, & a < 0 \\ A, & 0 \leq a < 1 \\ \emptyset, & a \geq 1 \end{cases}$

Properties:

- If $f: X \to Y$ and $g: Y \to Z$ are measurable, then $g \circ f: X \to Z$ is measurable.
- If $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ are measurable, then $f + g$ and $f \cdot g$ are measurable.
- If $f: \mathbb{R} \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}$ are measurable, then $h: \mathbb{R}^2 \to \mathbb{R}$ s.t. $(x,y) \mapsto f(x) \cdot g(y)$ is measurable.
- If $f: X \to \mathbb{R}$ is measurable, then $|f|$, f^a $(a > 0)$, $1/f$ $(f(x) \neq 0)$ are measurable.
  (since each of these functions is a composition of a cont. function and a measurable one, e.g. $|f| = |\cdot| \circ f$)
- If X and Y are measurable spaces, then $pr_X: X \times Y \to X$ and $pr_Y: X \times Y \to Y$ are measurable.
  (since $pr_X^{-1}(A) = A \times Y$ $\forall A \subseteq X$ and $pr_Y^{-1}(B) = X \times B$ $\forall B \subseteq Y$)
- If $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ are measurable and $E \subset X$ is measurable, then $h(x) = \begin{cases} f(x), & x \in E \\ g(x), & x \notin E \end{cases}$ is measurable. (since $\forall A \subset X$ $h^{-1}(A) = (f^{-1}(A) \cap E) \cup (g^{-1}(A) \cap E^c)$)

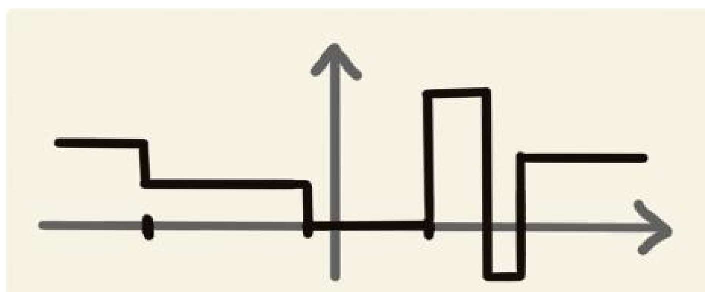
2.2. Simple functions.

$h: X \to \overline{\mathbb{R}}$ is a **simple function**: h is measurable and $h(X) = \{a_1, \dots, a_n\}$ $\quad \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$

$h = a_1 \chi_{A_1} + \dots + a_n \chi_{A_n}$, where $A_i = h^{-1}(\{a_i\})$ and $a_i \neq 0$ $\forall i \in I$

Example:

$h: \mathbb{R} \to \mathbb{R}$ step function ($\forall A_i$ is a union of disjoint intervals)



- If $f: X \to \overline{\mathbb{R}}$ and $g: X \to \overline{\mathbb{R}}$ are measurable, then $\overline{h} = \sup(f,g)$ and $\underline{h} = \inf(f,g)$ are measurable.

  $\overline{h}(x) = \begin{cases} f(x), & f(x) \geq g(x) \\ g(x), & g(x) > f(x) \end{cases}$ $\quad \underline{h}(x) = \begin{cases} f(x), & f(x) < g(x) \\ g(x), & g(x) \leq f(x) \end{cases}$ $\quad \left(\begin{aligned} \{x \mid \overline{h}(x) > a\} &= \{x \mid f(x) > a \vee g(x) > a\} = \\ &= \{x \mid f(x) > a\} \cup \{x \mid g(x) > a\} \end{aligned} \right)$

- Consider $f: X \to \overline{\mathbb{R}}$. $f = f^+ - f^-$ and $|f| = f^+ + f^-$, where $f^+ = \sup(f, 0)$ and $f^- = -\inf(f, 0)$.
- If $\{f_n \mid f_n: X \to \overline{\mathbb{R}}, n \in \mathbb{N}\}$ are measurable, then $\sup f_n$ and $\inf f_n$ are measurable.
- If $\{f_n \mid f_n: X \to \overline{\mathbb{R}}, n \in \mathbb{N}\}$ are measurable and $f_n \to f$, then f is measurable.

  $f_n \to f \iff \lim\limits_{n\to\infty} f_n = \overline{\lim\limits_{n\to\infty}} f_n = \underline{\lim\limits_{n\to\infty}} f_n$, where $\begin{aligned} \overline{\lim\limits_{n\to\infty}} f_n &\overset{\text{def}}{=} \lim\limits_{n\to\infty} \sup f_n = \inf\limits_{n \geq 1} \left(\sup\limits_{k \geq n} f_k \right) \\ \underline{\lim\limits_{n\to\infty}} f_n &\overset{\text{def}}{=} \lim\limits_{n\to\infty} \inf f_n = \sup\limits_{n \geq 1} \left(\inf\limits_{k \geq n} f_k \right) \end{aligned}$

- If $f: X \to \mathbb{R}$ is measurable, then $\exists \{h_n\}$, h_n are simple, s.t. $h_n \to f$.

③ Measure spaces.

3.1. Measure.

$\mu : \mathcal{M} \to \overline{\mathbb{R}}$ is a **measure** on $(X, \mathcal{M})$: (4) $\mu(\emptyset) = 0$

(5) $\mu(A) \geq 0 \quad \forall A \subset \mathcal{M}$

(6) $\mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i), \quad A_i \cap A_j = \emptyset \quad \forall i \neq j$ disjoint sets

i.e. μ is countably additive

$(X, \mathcal{M}, \mu)$ is a **measure space**: $\mathcal{M}$ is a σ-algebra on X and μ is a measure on $(X, \mathcal{M})$.

- If $B \subset A$, then $\mu(B) \leq \mu(A)$. (Since $\mu(A) = \mu(B) + \mu(A \setminus B) \geq \mu(B)$)
- If $\sum_{i \in I} \mu(A_i)$ diverges, then $\sum_{i \in I} \mu(A_i) = \infty$.
- Modularity: $\mu(A) + \mu(B) = \mu(A \cap B) + \mu(A \cup B)$ (since $A \cup B = (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B)$)
- If $\{E_n \mid E_n \subset E_{n+1} \ \forall n \in \mathbb{N}\}$ are measurable, then $\mu(E) = \lim_{n \to \infty} \mu(E_n)$, where $E = \bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$.

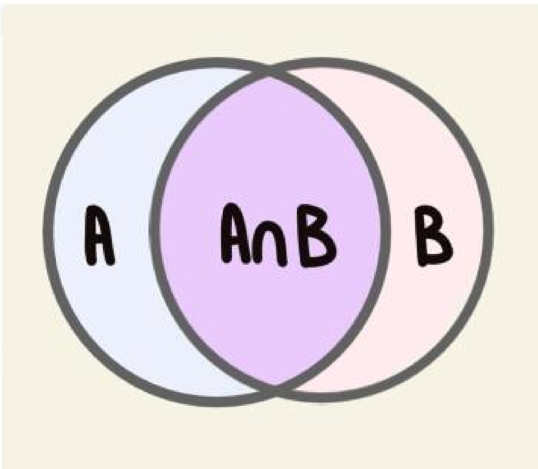
Example:

$\delta_a(A) = \begin{cases} 1, & a \in A \\ 0, & a \notin A \end{cases}$, $a \in X$ Dirac measure

3.2. Probability theory.

$(\Omega, \mathcal{M}, P)$ is a **probability space**: $P(\Omega) = 1$.

Ω : a sample space (the set of all possible outcomes)
$\mathcal{M}$: an event space (σ-algebra on Ω)
P : a probability function (measure on $(\Omega, \mathcal{M})$)
$P(A)$: the probability of an event A ($0 \leq P(A) \leq 1$)



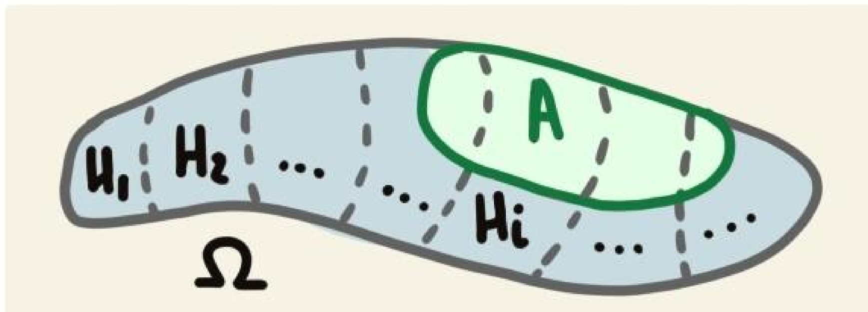
$P(A) = 1$: certainty event
$P(A) = 0$: impossible event
$P(A) + P(B) = P(A \cap B) + P(A \cup B)$

conditional probability of B given A: $P(B \mid A) \overset{\text{def}}{=} \dfrac{P(A \cap B)}{P(A)}$

events A and B are **independent**: $P(A \cap B) = P(A) \cdot P(B)$ or $P(B \mid A) = P(B)$

$\{H_i \in \mathcal{M} \mid i \in I\}$ is a **hypothesis** : $H_i \cap H_j = \emptyset \ \forall i \neq j$ and $\Omega = \bigcup_{i \in I} H_i$ i.e. partitioning of Ω

- $\forall A \in \mathcal{M} \quad A = \bigcup_{i \in I} (H_i \cap A) \quad P(A) = \sum_{i \in I} P(H_i \cap A) = \sum_{i \in I} P(A \mid H_i) P(H_i)$



- Bayes' formula: $P(H_i \mid A) = \dfrac{P(A \cap H_i)}{P(A)} = \dfrac{P(A \mid H_i) P(H_i)}{\sum_{k \in I} P(A \mid H_k) P(H_k)}$

4. Lebesgue measure.

4.1. Open intervals on $\mathbb{R}^n$.

First consider $\mathbb{R}$. Let's define a notion of lenght for open intervals:

$\ell(I) \overset{\text{def}}{=} b - a$, where $I = (a,b)$

- Notice that it works for closed and half-open intervals too:

$(a-\varepsilon, b) = (a-\varepsilon, a) \cup [a,b) \Rightarrow b - a + \varepsilon = a - a + \varepsilon + \ell([a,b)) \Rightarrow \ell([a,b)) = b - a$

$[a,b) = \{a\} \cup (a,b) \Rightarrow \ell(\{a\}) = 0$

$(a,b] = (a,b) \cup \{b\} \Rightarrow \ell((a,b]) = b - a$

$[a,b] = (a,b) \cup \{a\} \cup \{b\} \Rightarrow \ell([a,b]) = b - a$

- Any open set on $\mathbb{R}$ can be represented as a countable union of disjoint open intervals (since $\forall r \in \mathbb{Q}$ s.t. $r \in U$ lies in a maximal open interval $(a,b) \subseteq U$). And open sets generate Borel sets on $\mathbb{R}$.
- If $A \subset \mathbb{R}$ is s.t. $|A| = |\mathbb{N}|$, then $\ell(A) = 0$. (since $A = \bigcup_{k=1}^{\infty} \{a_k\}$)

In the case of $\mathbb{R}^n$ an open interval is a cartesian product of open intervals on $\mathbb{R}$. And we now define a notion of volume instead:

$\mathrm{vol}_n(I) \overset{\text{def}}{=} \prod_{k=1}^{n} \ell(I_k) = \prod_{k=1}^{n} (b_k - a_k)$, where $I = I_1 \times \ldots \times I_n$, $I_k = (a_k, b_k)$

- Arguing as before one can show that the same definition works for closed and partly-open intervals too.
- Any open set on $\mathbb{R}^n$ can be also seen as a coutable union of disjoint open intervals, and open sets generate Borel sets on $\mathbb{R}^n$.

4.2. Outer measure.

To define a measure on $\mathbb{R}^n$ we will need an additional notion first:

$$\boxed{\mu^*(A) \overset{\text{def}}{=} \inf \{ \mathrm{vol}_n(U) \mid A \subseteq U,\ U \text{ is open} \}}$$

outer measure of $A \subset \mathbb{R}^n$

- μ^* is not a measure, it doesn't satisfy countable additivity (meas. 6), but it satisfies countable subadditivity:

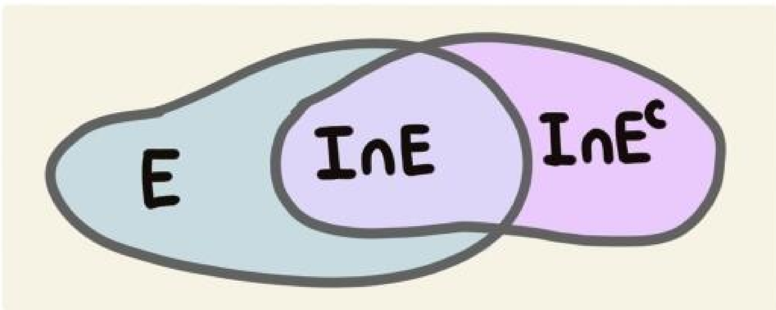
$\mu^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leqslant \sum_{k=1}^{\infty} \mu^*(A_k)$, where $A_i \cap A_j = \emptyset \ \forall i \neq j$ (disjoint sets)

- $\mu^*(\emptyset) = 0$
- $A \subseteq B \Rightarrow \mu^*(A) \leqslant \mu^*(B)$
- $\mu^*(I) = \mathrm{vol}_n(I)$, where I is an interval on $\mathbb{R}^n$

4.3. Lebesgue measure.

Now we can define a measure on $\mathbb{R}^n$:

E is **Lebesgue-measurable**: $\mu^*(I) = \mu^*(I \cap E) + \mu^*(I \cap E^c) \quad \forall$ open interval $I \subset \mathbb{R}^n$



i.e. when we cover mutually disjoint sets $I \cap E$ and $I \cap E^c$ with open intervals, the overlap between them can be made arbitrarily small

The set of all Lebesgue-measurable sets $\mathcal{L}$ form a σ-algebra on $\mathbb{R}^n$, and an outer measure μ^* defines a measure on it, i.e. $(\mathbb{R}^n, \mathcal{L}, \mu = \mu^*|_{\mathcal{L}})$ is a measure space:

1. $\emptyset \in \mathcal{L}$	4. $\mu(\emptyset) = 0$
2. $E \in \mathcal{L} \Rightarrow E^c \in \mathcal{L}$	5. $\mu(E) \geqslant 0$
3. $E_i \in \mathcal{L},\ i \in I \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{L}$	6. $\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$, $E_i \cap E_j = \emptyset \ \forall i \neq j$

- All open sets are Lebesgue-measurable (since any open set is a countable union of disjoint open intervals), therefore, Borel sets are included in $\mathcal{L}$ (but not vice versa!).
- $E \in \mathcal{L} \Leftrightarrow \forall \varepsilon > 0 \ \exists$ open G and closed F s.t. $F \subset E \subset G$ and $\mu(G \setminus F) < \varepsilon$.
- If $\dim(E) < n$, then $\mu(E) = 0$ (e.g. any hyperplane in $\mathbb{R}^n$ has a measure zero).
- If $\mu(E_i) = 0 \ \ \forall i \in I$, then $\mu(\bigcup_{i=1}^{\infty} E_i) = 0$.
- If $\mu(N) = 0$, then $\forall E \ \ \mu(E \cup N) = \mu(E \setminus N) = \mu(E)$
- Properties of the Lebesgue measure:
  - Lebesgue measure is complete (i.e. $\forall N' \subset N$, where $\mu(N) = 0$, is s.t. $\mu(N') = 0$).
  - If $E \in \mathcal{L}$, then $T_a(E) = \{x + a \mid x \in E,\ a \in \mathbb{R}^n\} \in \mathcal{L}$ and $\mu(T_a(E)) = \mu(E)$.
  - If $E \in \mathcal{L}$, then $\alpha E = \{\alpha x \mid x \in E,\ \alpha > 0\} \in \mathcal{L}$ and $\mu(\alpha E) = \alpha^n \mu(E)$.
  - If $E \in \mathcal{L}$ and $A \in GL_n(\mathbb{R})$, then $A(E) \in \mathcal{L}$ and $\mu(A(E)) = |\det A| \cdot \mu(E)$.
  - Lebesgue measure is a unique complete translation-invariant measure on $\mathbb{R}^n$ s.t. $\mu((0,1) \times \ldots \times (0,1)) = 1$.
- There are sets that are not Lebesgue-measurable.
- There is no infinite-dimensional analogue of Lebesgue measure.
- Inner measure:

$$\boxed{\mu_*(E) \stackrel{\text{def}}{=} \sup\{\mathrm{vol}_n(U) \mid U \subset E,\ U \text{ is open}\}}$$

inner measure of $E \subset \mathbb{R}^n$

- $\mu_*(E) \leqslant \mu^*(E)$
- $\mu(U) = \mu_*(U \cap E) + \mu^*(U \cap E^c) \quad \forall$ open $U \supset E$
- $\mu(E) < \infty \iff \mu_*(E) = \mu^*(E) < \infty$

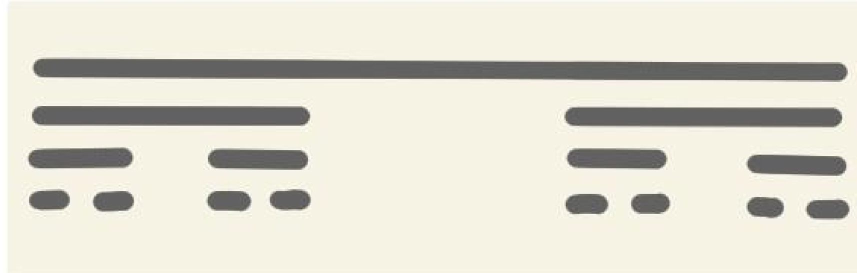
A property holds **almost everywhere**: holds everywhere except on N s.t. $\mu(N) = 0$.

__Example:__

$f = g$ a.e.: $N = \{x \mid f(x) \neq g(x)\}$ is s.t. $\mu(N) = 0$.

Examples:

1) The Cantor set (see Ch. I §4)



C is closed (since C^c is open) $\Rightarrow$ C is a Borel set $\Rightarrow$ $C \in \mathcal{L}$

$$\mu(C_n) = 1 - \frac{1}{3} - 2\left(\frac{1}{3}\right)^2 - 2^2\left(\frac{1}{3}\right)^3 - \ldots - 2^{n-1}\left(\frac{1}{3}\right)^n$$

$$\mu(C) = \lim_{n\to\infty} \mu(C_n) = 1 - \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^k = 1 - \frac{1}{2}\left(\frac{1}{1-2/3} - 1\right) = 1 - \frac{1}{2}\cdot 2 = 0$$

i.e. $\mu(C) = 0$, but it's uncountable

2) The existence of sets that are not Lebesgue-measurable relies on the axiom of choice.

Let's define a relation on $I = (0,1)$:

$xQy \stackrel{\text{def}}{=} \{(x,y) \mid x - y \in \mathbb{Q},\ x,y \in I\}$ equiv. relation

$Q_x = \{y \mid y - x \in \mathbb{Q}\}$ equiv. classes

Axiom of choice $\Rightarrow$ $\exists T$ s.t. it has exactly one representative from each Q_x.

$T_r \stackrel{\text{def}}{=} \{x + r \mid x \in T,\ r \in \mathbb{Q}\ r \in (-1,1)\}$

$\forall y \in I\ \exists T_r$ s.t. $y \in T_r \stackrel{|r|<1}{\Longrightarrow} (0,1) \subset \bigcup_r T_r \subset (-1,2)$

$$1 \leq \mu\left(\bigcup_r T_r\right) \leq 3$$

$$1 \leq \sum_{k=1}^{\infty} \mu(T) \leq 3$$

(all the T_r are disjoint and have the same measure)

If $\mu(T) = 0$, then $\sum_{k=1}^{\infty} \mu(T) = 0$.

If $\mu(T) > 0$, then $\sum_{k=1}^{\infty} \mu(T) = \infty$.

$\Rightarrow$ Contradiction $\Rightarrow$ T is non-measurable

5. Lebesgue integration.

5.1. Simple functions.

Consider a simple function $h: X \to \mathbb{R}$ on a measure space $(X, \mathcal{M}, \mu)$.

$\boxed{\int h \, d\mu \overset{\text{def}}{=} \sum_{i=1}^{n} a_i \mu(A_i)}$, where $h = \sum_{i=1}^{n} a_i \chi_{A_i}$, $a_i \in \mathbb{R}$
$A_i = h^{-1}(a_i) \subset X$ are measurable

- If $h_1: X \to \mathbb{R}$ and $h_2: X \to \mathbb{R}$ are simple, then $\alpha h_1 + \beta h_2$ $(\alpha, \beta \in \mathbb{R})$ is also simple and:
  1. $\alpha h_1 + \beta h_2 = \sum_{i,j} (\alpha a_i + \beta b_j) \chi_{A_i \cap B_j}$, $\alpha a_i + \beta b_j \neq 0$
  2. $\int (\alpha h_1 + \beta h_2) d\mu = \alpha \int h_1 d\mu + \beta \int h_2 d\mu$
- If $h_1: X \to \mathbb{R}$ and $h_2: X \to \mathbb{R}$ are simple and $h_1 \leq h_2$, then $\int h_1 d\mu \leq \int h_2 d\mu$.
  (since $h_2 - h_1 > 0$ implies that $\int (h_2 - h_1) d\mu > 0$)

Example:

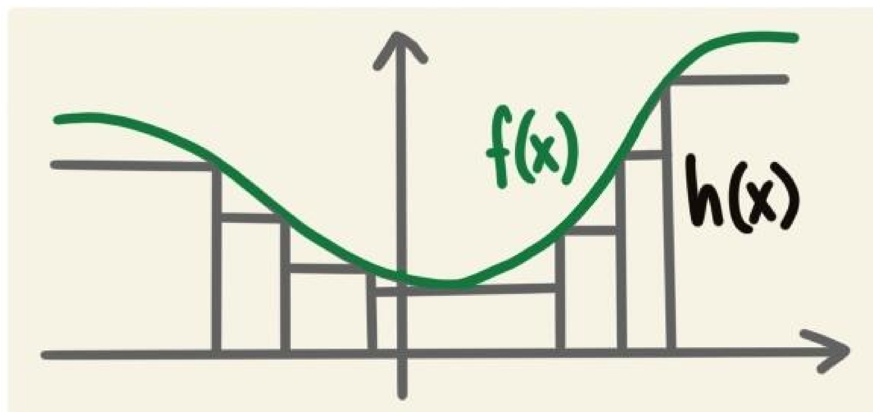
$h(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$ $\qquad \int h \, d\mu = \infty + (-\infty)$ is not well-defined

5.2. Lebesgue integral.

Consider $\mathbb{R}$ with Lebesgue measure on it. We define the integral of a non-negative measurable function $f^+: \mathbb{R} \to \mathbb{R}$ as:

$\boxed{\int f^+ d\mu \overset{\text{def}}{=} \sup \left\{ \int h \, d\mu \;\middle|\; h = \sum_{i=1}^{n} a_i \chi_{A_i} \text{ and } 0 \leq h \leq f^+ \right\}}$

Lebesgue integral of f^+



Th. 5.1. (Levi; monotone convergence th.):
If $\{f_n^+ \mid f_n^+ \geq 0 \; \forall n\}$ is an increasing sequence of measurable functions s.t. $f_n^+ \to f^+$, then $\lim_{n \to \infty} \int f_n^+ d\mu = \int f^+ d\mu$.

Th. 5.2.:
Suppose $f^+ \geq 0$ is measurable. Then $\int f^+ d\mu = 0 \iff f^+ = 0$ a.e.

Now we can extend integration to measurable functions that take negative values too:

$f: X \to \mathbb{R}$ is integrable w.r.t. μ: $f^+ = \sup(f, 0)$ and $f^- = -\inf(f, 0)$ are integrable w.r.t. μ.

$\boxed{\int f \, d\mu \overset{\text{def}}{=} \int f^+ d\mu - \int f^- d\mu}$

Lebesgue integral of f

$\boxed{\int_E f \, d\mu \overset{\text{def}}{=} \int f \chi_E \, d\mu}$, $f\chi_E = \begin{cases} f, & x \in E \\ 0, & x \notin E \end{cases}$, where $E \in \mathcal{M}$

integral of f over $E \subset X$

Properties:

- If $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ are Lebesgue-integrable, then $\int(\alpha f+\beta g)d\mu = \alpha\int f d\mu + \beta\int g d\mu$.
- If $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ are Lebesgue-integrable and $f\le g$, then $\int f d\mu \le \int g d\mu$.
- If $f:\mathbb{R}\to\mathbb{R}$ is Lebesgue-integrable, then $|f| = f^+ + f^-$ is Lebesgue-integrable and $\left|\int f d\mu\right| \le \int |f| d\mu$. (since $\left|\int f d\mu\right| = \left|\int (f^+ - f^-) d\mu\right| \le \left|\int (f^+ + f^-) d\mu\right| = \int |f| d\mu$)
- If $f:\mathbb{R}\to\mathbb{R}$ is Lebesgue-integrable, then $\psi(a) \overset{\text{def}}{=} \mu(\{x\in E \mid |f(x)|>a\}) \overset{a\to\infty}{=} O(a^{-1})$.
- Relation to Riemann integration:

  Lebesgue: the simple functions used to approximate f are arbitrary.
  Riemann : the simple functions used to approximate f are step functions.

  ($\forall \varepsilon>0\ \exists$ step functions $h_1\le f\le h_2$ s.t. $\int_a^b (h_2(x)-h_1(x))dx < \varepsilon$.)

  - If f is Riemann-integrable, then f is Lebesgue-integrable and $\int_E f d\mu = \int_a^b f(x)dx$.
  - Lebesgue integration is defined for bounded and unbounded sets in the same way, and in the same way for bounded and unbounded integrands. This contrasts with Riemann integration.

  Example:

  $f:[0,1]\to\{0,1\}$ $\quad$ 1. $\int f d\mu = 1\cdot\mu([0,1]) = 1$ (since $f=1$ a.e.)

  $f(x) = \begin{cases} 1, & x\notin\mathbb{Q} \\ 0, & x\in\mathbb{Q}\end{cases}$ $\quad$ 2. $\int_0^1 f(x)dx$ is not defined (f cannot be approximated by step functions)

Lemma 5.3. (Fatou):
If $\{f_n \mid f_n\ge 0\}$ are measurable, then $\int \varliminf_{n\to\infty} f_n\, d\mu \le \varliminf_{n\to\infty}\int f_n\, d\mu$.

Th. 5.4. (Lebesgue; dominated convergence th.):
Suppose $\{f_n\}$ are measurable and $f_n\to f$ a.e. If $\exists$ integrable $g: X\to\mathbb{R}$ s.t. $g>|f_n|\ \forall n$, then $\lim_{n\to\infty}\int f_n\, d\mu = \int f d\mu$.

Th. 5.5. (Fubini):
If $f:\mathbb{R}^2\to\mathbb{R}^2$ is Lebesgue-measurable, then $\forall x\in\mathbb{R}$ $f_x(y)=f(x,y)$ is Lebesgue-measurable, and $\forall y\in\mathbb{R}$ $f_y(x)=f(x,y)$ is Lebesgue-measurable, i.e.:
$\int f d\mu^2 = \int f(x,y)\,dx\,dy = \int\left(\int f_y(x)dx\right)dy = \int\left(\int f_x(y)dy\right)dx$

Example:

$\{\chi_{[n,n+1]}\}$ sequence of "unit humps" drifting to the right

$\lim_{n\to\infty}\int \chi_{[n,n+1]}\, d\mu = \lim_{n\to\infty}(1\cdot\mu([n,n+1])) = 1$

$\int \lim_{n\to\infty}\chi_{[n,n+1]}\, d\mu = \int 0\, d\mu = 0 \ne 1$ (since $\{\chi_{[n,n+1]}\}$ has no dominating function)