Advanced probability

STAT6010/7610

Marius Hofert

Statistics and Actuarial Science
School of Computing and Data Science

SCHOOL OF COMPUTING & DATA SCIENCE
The University of Hong Kong

Course information

Contact, times, locations

Instructor: HOFERT Marius (`mhofert@hku.hk`)

Lectures: Mon 09:30–12:20, in MWC T7

Office hours: on appointment, via Zoom: `https://hku.zoom.us/j/98597702554`

Questions: Best asked during the lecture breaks, after the lectures and during the office hours. Emails to the instructor will only be answered if they concern personal circumstances or emergency cases.

TA(s): YAO Gan (`ganyao@connect.hku.hk`; responsible for tutorials (starting in second week of classes) and grading assignments)

Representative: Any volunteer?

Course objectives

- Introduction to measure theory and probability
- Basic concepts in theoretical probability
- For students interested in research (in AS, STAT, probability)

- Mathematical proofs and the underlying ideas

References

It is recommended to study from the course material rather than additional sources. The course material is compiled from various sources, including:

- Achim Klenke, “Probability Theory: A comprehensive course”, 2008;
- René Schilling, “Measures, integrals and martingales”, 2006;
- Heinz Bauer, “Measure and Integration Theory”, 2001;
- Gerald B. Folland, “Real Analysis: Modern Techniques and Their Applications”, 1999;
- Rick Durrett, “Probability: Theory and Examples”, edition 5, 2019;
- Sidney I. Resnick, “A Probability Path”, 2014;
- Allan Gut, “Probability: A Graduate Course”, 2005;
- David Williams, “Probability with Martingales”, 1991;
- Patrick Billingsley, “Probability and Measure”, 1995;
- Jean Jacod, Philip Protter, “Probability Essentials”, 2003;

- Richard L. Wheeden, Antoni Zygmund, “Measure and Integral”, 1977.

Teaching and assessment

- 2025-02-24: No lecture
- Course-relevant material is shared on Moodle (`https://moodle.hku.hk/course/view.php?id=122012`).
- **Assessment:**
  - 3 assignments (each 5%): to be handed in as Moodle Assignment (single file) before the tutorial on the due date (late hand-ins are marked as 0)
  - 1 midterm (25%): March 17, 2025, 09:30–10:20, MWC T7 (lecture thereafter)
  - 1 2 h final (60%): t.b.a. in course outline once available ($\approx$ mid term)
- **Absence from assessments:** See course outline. Avoid missing the final.
- See course outline for additional details, e.g. rules for regrading requests.

Absence from assessments

See course outline.

General advice

- Come to every class (even if you cannot follow much). Fill gaps in follow-up course work.
- Perform follow-up course work (after each lecture) to learn the material continuously throughout the term. For a proper learning effect and to prepare you for the exams, you should write by hand (definitions, main results, formulas, doing the examples and exercises again, etc.), so writing your own summary notes is advisable.
- Try every assignment question on your own first before collaborating with others. If you cannot solve them in a reasonable amount of time, discuss ideas with others. As in exams, the final write-up to be handed in must be your own.
- Try your best to avoid getting ill before your finals. It is highly recommended to participate in final exams, as supplementary finals are more difficult due to the longer preparation time, which would otherwise be unfair towards all other students.
- Regularly check your university email (especially before exams, lectures in bad weather).

FAQ

- *Is there a single book for this course we follow?*
  No (as all have their drawbacks and the lecture material is compiled from various sources).
- *Can I have more space on the slides?*
  In LaTeX:

```
\documentclass{article}
\usepackage{pdfpages}
\begin{document}
\includepdf[pages={1-}, scale=0.75, offset=0 180]{mySlides.pdf}% see https://mirror-hk.koddos.net/CTAN/macros/
      latex/contrib/pdfpages/pdfpages.pdf for the arguments; play with 'offset' to put slides in the location
      you want
\end{document}
```

- *What will the exam(s) look like?*
  Closed book. A mix of "Show...", "Calculate...", "Provide a definition of...", "Explain in words why...", etc. Important is to justify your answers (give reasons for your answers; providing an answer without derivation or without stating assumptions will only give a minor fraction of the marks).
- *How can I get more practice?*
  You can. . .

- rework the course slides (definitions, statements, examples);
- change the distributions in the examples and redo them;
- redo the assignment and tutorial problems (also with different numbers);
- team up with a colleague and pass each other (modified) questions; and
- google the topics you struggle with to find more exercises.

- *Do I need to be able to replicate all the proofs?*

  Proofs are an essential part of mathematical learning and are covered to explain why statements are correct, so they help us learn and understand. Important arguments and rough ideas (for longer proofs) may appear or be asked for, but longer proofs do not need to be replicated precisely.

Overview

1 Introduction

1.1 History

1.2 Basics of set theory

1.1 History

- Intuitively, we understand the concept of probability (a number in $[0,1]$ reflecting the chance that a random event happens) and that is important for modelling future events (gambling, weather, life insurance, success of a new drug).
- Early references include:
  - Fair value of an insurance contract, e.g. against crop failure (Code of Hammurabi from Babylon; legal text from 1754 BCE):

    > "A farmer who has a mortgage on his property is required to make annual interest payments in the form of grain. However, in the event of a crop failure, this farmer has the right not to pay anything, and the creditor has no alternative but to forgive the interest due."

    This essentially describes a put option (right but not the obligation of the holder [here: farmer] to sell [here: not to pay interest in the form of crop] the underlying asset [here: crop] at a specific price).
  - Gerolamo Cardano (1501–1576, "Book on Games of Chance" (on gambling; written ≈ 1564; considered throwing dice to understand basic probabilities; considered the ratio of favourable to unfavourable outcomes as probabilities)

- Galileo Galilei (1564–1642) observed that some numbers in $\{2,\dots,12\}$ appear more often as sum when throwing two dice since there are more ways to create them.
- Blaise Pascal (1623–1662) and Pierre de Fermat (1607–1665) exchanged letters to solve gambling "paradoxes", commonly viewed as the birth of probability theory.
  - Antoine Gombaud (as self-styled Chevalier de Méré) considered two games (Chevalier de Méré paradox):
    1) Roll a fair die 4x and note whether at least one 6 occurs.
    2) Roll two fair dice 24x and note whether at least one double 6 occurs.

    He believed the chance of the two games to be the same (they are $1-(1-1/6)^4 \approx 0.5177$ and $1-(1-1/36)^{24} \approx 0.4914$, respectively).
  - Pascal and Fermat spotted that Gombaud believed that the probability of success in n throws is n times that of a single throw; e.g. in Game 1 Gombaud thought that since the probability of success in one throw is $1/6$, the probability of rolling at least one 6 in four throws is $4/6 = 2/3$.
- Christiaan Huygens (1629–1695) wrote a book about the ideas of Pascal

and Fermat.

- Jacob Bernoulli (1655–1705) pointed out the necessity to develop a theory to answer interesting probability problems he proposed in 1685.
- More and more mathematicians worked on probability problems: Abraham de Moivre (1667–1754), Daniel Bernoulli (1700–1782), Leonhard Euler (1707–1783), Carl Friedrich Gauss (1777–1855), Pierre-Simon Laplace (1749–1827).
- Andrey Nikolaevich Kolmogorov (1903–1987) saw the usefulness of measure theory in properly establishing a modern theory of probability in his book "Grundbegriffe der Wahrscheinlichkeitsrechnung" in 1933. Measure theory is indispensable for studying probability theory.
- Paul Lévy (1886–1971) worked on stochastic processes, characteristic functions and limit theorems.

- Today we think of probability as a mathematical theory for modelling random events.
- Random events are described through sets. We thus first need to learn about basic set theory and families of sets with certain properties. We can then define probabilities on such families of sets in a consistent (non-contradicting) way.

1.2 Basics of set theory

- Under a *collection* we understand several objects (*elements*) grouped together.
- Initially, a "set" was believed to be an arbitrary collection of objects, its elements.
- *Russell's paradox* (1901) showed that every set theory that allows *unrestricted comprehension* (i.e. that for any well-defined property we can construct a "set" containing the elements with that property) leads to contradictions since such a general "set" can be an element of itself, e.g. $\{A : |A| \geq 1\}$ is a "set" which contains itself (e.g. $\{e, \pi\} \in \mathcal{A} := \{A \subseteq \mathbb{R} : |A| \geq 1\} \Rightarrow \mathcal{A} \in \mathcal{A}$).
- Russell considered the "set" $R := \{A : A \notin A\}$ of all "sets" that are not elements of themselves. If $R \notin R$, then $R \in R$. And if $R \in R$, then $R \notin R$. So $R \in R \Leftrightarrow R \notin R$ ↯. A more colloquial expression of this problem is: "This statement is false."

  Barber paradox: Suppose a barber shaves precisely those men who do not shave themselves, does the barber shave himself?

  - If yes, then he shaves himself, so the barber (he) does not shave himself ↯.
  - If no, then he doesn't shave himself, so the barber (he) shaves himself ↯.

- Ernst Zermelo (1871–1953), with later additions by Abraham Fraenkel (1891–1965), developed a system of axioms in order to formulate a set theory free of paradoxes, the Zermelo–Fraenkel (ZF) set theory; it does not allow for the existence of a universal set (a set containing all sets; this often leads to problems) nor for unrestricted comprehension (so it also avoids Russell's paradox). With the *axiom of choice* (i.e. for any collection of non-empty sets $\mathcal{A}$ there is a *choice function* f such that $f(A) \in A$ for all $A \in \mathcal{A}$; or: the Cartesian product of a collection of non-empty sets is non-empty), ZF is abbreviated ZFC. We assume to work with ZFC.
- A *set* is a collection of objects (*elements*) that satisfy the ZFC axioms.
- Let $\Omega \neq \emptyset$. A is a *subset* of Ω ($A \subseteq \Omega$) or Ω is a *superset* of A ($\Omega \supseteq A$), if $\omega \in A \Rightarrow \omega \in \Omega \ \forall \omega \in A$. And $A = \Omega$ if $A \subseteq \Omega$ and $\Omega \subseteq A$.
- $\mathcal{P}(\Omega) := \{A : A \subseteq \Omega\}$ is the *power set* of Ω, i.e. the set of all subsets of Ω (including $\emptyset, \Omega$). If $|\Omega| < \infty$, one can show by induction that $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$.
- For $I \subseteq \mathbb{R}$, a *family* $\{A_i\}_{i \in I} \subseteq \mathcal{P}(\Omega)$ of sets is a collection of subsets of a set Ω. In contrast to a set of sets, a family of sets can contain repeated copies of its elements.

- Basic set operations:

Complementation: If $A \subseteq \Omega$, then $A^c := \{\omega \in \Omega : \omega \notin A\}$. We have $(A^c)^c = A$, $\emptyset^c = \Omega$, $\Omega^c = \emptyset$.

Intersection: If $A_i \subseteq \Omega$, $i \in I$, then

$$\bigcap_{i \in I} A_i := \{\omega \in \Omega : \omega \in A_i \ \forall\, i \in I\}.$$

The family $\{A_i\}_{i \in I}$ is *pairwise disjoint* (or *mutually exclusive*) if $A_{i_1} \cap A_{i_2} = \emptyset \ \forall\, i_1, i_2 \in I : i_1 \neq i_2$. Intersections satisfy $A \cap B = B \cap A$ (*commutativity*) and $(A \cap B) \cap C = A \cap (B \cap C)$ (*associativity*).

Union: If $A_i \subseteq \Omega$, $i \in I$, then

$$\bigcup_{i \in I} A_i := \{\omega \in \Omega : \exists\, i \in I : \omega \in A_i\}.$$

If the family $\{A_i\}_{i \in I}$ is pairwise disjoint, one often writes $\biguplus_{i \in I} A_i$. Unions satisfy $A \cup B = B \cup A$ (*commutativity*) and $(A \cup B) \cup C = A \cup (B \cup C)$ (*associativity*).

Set difference: If $A, B \subseteq \Omega$, then $A \backslash B := A \cap B^c$. We thus have $A^c = \Omega \backslash A$.

- Identities involving more than one set operation:

  Distributivity: $A\cap\bigcup_{i\in I} A_i = \bigcup_{i\in I}(A\cap A_i)$ and $A\cup\bigcap_{i\in I} A_i = \bigcap_{i\in I}(A\cup A_i)$

  De Morgan's laws: $(\bigcup_{i\in I} A_i)^c = \bigcap_{i\in I} A_i^c$ and $(\bigcap_{i\in I} A_i)^c = \bigcup_{i\in I} A_i^c$

- Limits of sequences of sets:

  Infimum: $\inf_{k\ge n} A_k := \bigcap_{k\ge n} A_k$

  Supremum: $\sup_{k\ge n} A_k := \bigcup_{k\ge n} A_k$

  Limit inferior: $\liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty}\bigcap_{k\ge n} A_k \underset{\text{def.}}{=} \sup_{n\ge 1}\inf_{k\ge n} A_k$

  Limit superior: $\limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty}\bigcup_{k\ge n} A_k \underset{\text{def.}}{=} \inf_{n\ge 1}\sup_{k\ge n} A_k$

  Limit: If $\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n =: A$, A is the *limit* of A_n for $n\to\infty$ and we write $A = \lim_{n\to\infty} A_n$ or $A_n \underset{n\to\infty}{\to} A$.

- Monotone sequences of sets:

  $\{A_n\}_{n\in\mathbb{N}}$ is *increasing* ($A_n \nearrow$) if $A_1 \subseteq A_2 \subseteq \dots$.

  $\{A_n\}_{n\in\mathbb{N}}$ is *decreasing* ($A_n \searrow$) if $A_1 \supseteq A_2 \supseteq \dots$.

Lemma 1.1 (Interpretation limit inferior, limit superior of sets)

1) $\liminf_{n\to\infty} A_n = \{\omega \in \Omega : \omega \in A_n \text{ for all but finitely many } n\} =: \{\omega \in A_n \text{ abfm}\}$

2) $\limsup_{n\to\infty} A_n = \{\omega \in \Omega : \omega \in A_n \text{ for infinitely many } n\} =: \{\omega \in A_n \text{ io}\}$

Proof.

1) We have

$$\begin{aligned}\liminf_{n\to\infty} A_n &\underset{\text{def.}}{=} \bigcup_{n=1}^{\infty} \bigcap_{k\ge n} A_k \\ &= \{\omega \in \Omega : \exists\, n \in \mathbb{N} : \omega \in A_k \;\forall\, k \ge n\} \\ &= \{\omega \in \Omega : \omega \in A_n \text{ for all but finitely many } n\}.\end{aligned}$$

2) We have

$$\begin{aligned}\limsup_{n\to\infty} A_n &\underset{\text{def.}}{=} \bigcap_{n=1}^{\infty} \bigcup_{k\ge n} A_k \\ &= \{\omega \in \Omega : \forall\, n \in \mathbb{N}\; \exists\, k \ge n : \omega \in A_k\} \\ &= \{\omega \in \Omega : \omega \in A_n \text{ for infinitely many } n\}.\end{aligned}$$

□

Lemma 1.2 (Properties of limit inferior, limit superior of sets)

We have

1) $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$.
2) $(\liminf_{n\to\infty} A_n)^c = \limsup_{n\to\infty} A_n^c$.

Proof.

1) $\liminf_{n\to\infty} A_n = \{\omega \in A_n \text{ abfm}\} \subseteq \{\omega \in A_n \text{ io}\} = \limsup_{n\to\infty} A_n$.
2) We have $(\liminf_{n\to\infty} A_n)^c \underset{\text{def.}}{=} (\bigcup_{n=1}^\infty \bigcap_{k\geq n} A_k)^c \underset{\text{De Morgan}}{=} \bigcap_{n=1}^\infty \big(\bigcap_{k\geq n} A_k \big)^c \underset{\text{De Morgan}}{=} \bigcap_{n=1}^\infty \big(\bigcup_{k\geq n} A_k^c \big) \underset{\text{def.}}{=} \limsup_{n\to\infty} A_n^c$. □

Lemma 1.3 (Monotone sequences of sets)

1) If $A_n \nearrow$, then $\lim_{n\to\infty} A_n$ exists and $\lim_{n\to\infty} A_n = \bigcup_{k=1}^\infty A_k$. Similarly, if $A_n \searrow$, then $\lim_{n\to\infty} A_n$ exists and $\lim_{n\to\infty} A_n = \bigcap_{k=1}^\infty A_k$.
2) For all $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(\Omega)$, $\liminf_{n\to\infty} A_n = \lim_{n\to\infty} (\inf_{k\geq n} A_k)$ and $\limsup_{n\to\infty} A_n = \lim_{n\to\infty} (\sup_{k\geq n} A_k)$.

Proof.

1) $\limsup_{n\to\infty} A_n \underset{\text{def.}}{=} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \underset{n=1}{\subseteq} \bigcup_{k=1}^{\infty} A_k \underset{A_n \nearrow}{=} \bigcup_{k=1}^{\infty} \bigcap_{i=k}^{\infty} A_i \underset{\text{def.}}{=} \liminf_{n\to\infty} A_n \underset{\text{L. 1.2 1)}}{\subseteq} \limsup_{n\to\infty} A_n$, so that $\lim_{n\to\infty} A_n$ exists and $\lim_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} A_k$. Similarly for $A_n \searrow$.

2) For $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(\Omega)$, $\liminf_{n\to\infty} A_n \underset{\text{def.}}{=} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \underset{\text{def.}}{=} \bigcup_{n=1}^{\infty} \inf_{k\geq n} A_k \overset{\inf_{k\geq n} A_k \nearrow}{\underset{1)}{=}} \lim_{n\to\infty} (\inf_{k\geq n} A_k)$. Similarly for $\sup_{k\geq n} A_k \searrow$. □

- Equivalence relations: An *equivalence relation* is a binary relation $\sim$ on a set A (a set of ordered pairs from A) that is reflexive ($a \sim a$), symmetric ($a \sim b \Leftrightarrow b \sim a$) and transitive ($a \sim b$, $b \sim c \Rightarrow a \sim c$).
- The *equivalence class* of $a \in A$ under $\sim$ is $[a] := \{x \in A : x \sim a\}$. Example: "$=$" (is equal to), e.g. $1/2 = 2/4$, and both belong to the same equivalence class of cancelled fractions on $\mathbb{Q}$.
- Indicator functions: The *indicator function* of $A \subseteq \Omega$ is $\mathbb{1}_A(\omega) := \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$
  We have:
  1) $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$

2) $\mathbb{1}_A \le \mathbb{1}_B$ (i.e. $\mathbb{1}_A(\omega) \le \mathbb{1}_B(\omega)$ for all $\omega \in \Omega$) iff $A \subseteq B$

3) Indicator functions belong to the most useful functions there are. They can be used to count quantities, e.g.:

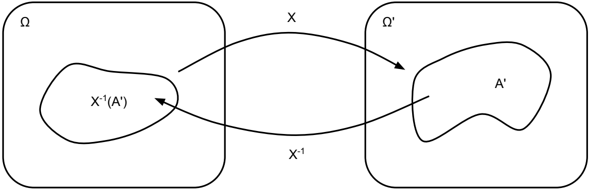
$$\limsup_{n\to\infty} A_n = \{\omega \in A_n \text{ io}\} = \left\{\omega \in \Omega : \sum_{n=1}^{\infty} \mathbb{1}_{A_n}(\omega) = \infty\right\},$$

$$\liminf_{n\to\infty} A_n = \{\omega \in A_n \text{ abfm}\} = \left\{\omega \in \Omega : \sum_{n=1}^{\infty} \mathbb{1}_{A_n^c}(\omega) < \infty\right\}.$$

- Preimages: The *preimage* of a map $X : \Omega \to \Omega'$ between two sets Ω, Ω' is

$$X^{-1}(A') = \{\omega \in \Omega : X(\omega) \in A'\}.$$

Sketch:



Preimages exist even if X is not injective (e.g. for $X(\omega) = \omega^2$, $X^{-1}([1,4]) = [-2,-1] \uplus [1,2]$).

- Preimages are closed with respect to (wrt) the following operations:

Complementation: If $A' \subseteq \Omega'$, then $(X^{-1}(A'))^{c_\Omega} = X^{-1}(A'^{c_{\Omega'}})$.

Proof. $\omega \in (X^{-1}(A'))^{c_\Omega} \Leftrightarrow \omega \notin X^{-1}(A') \Leftrightarrow X(\omega) \notin A' \Leftrightarrow X(\omega) \in A'^{c_{\Omega'}} \Leftrightarrow \omega \in X^{-1}(A'^{c_{\Omega'}})$. □

Union: If $\{A_i'\}_{i\in I} \subseteq \mathcal{P}(\Omega')$, then $X^{-1}(\bigcup_{i\in I} A_i') = \bigcup_{i\in I} X^{-1}(A_i')$.

Proof. $\omega \in X^{-1}(\bigcup_{i\in I} A_i') \Leftrightarrow X(\omega) \in \bigcup_{i\in I} A_i' \Leftrightarrow \exists\, i \in I : X(\omega) \in A_i' \Leftrightarrow \exists\, i \in I : \omega \in X^{-1}(A_i') \Leftrightarrow \omega \in \bigcup_{i\in I} X^{-1}(A_i')$. □

Intersection: If $\{A_i'\}_{i\in I} \subseteq \mathcal{P}(\Omega')$, then $X^{-1}(\bigcap_{i\in I} A_i') = \bigcap_{i\in I} X^{-1}(A_i')$.

Proof. $\omega \in X^{-1}(\bigcap_{i\in I} A_i') \Leftrightarrow X(\omega) \in \bigcap_{i\in I} A_i' \Leftrightarrow X(\omega) \in A_i'\ \forall\, i \in I \Leftrightarrow \omega \in X^{-1}(A_i')\ \forall\, i \in I \Leftrightarrow \omega \in \bigcap_{i\in I} X^{-1}(A_i')$. □

Monotonicity: If $\mathcal{A}', \mathcal{B}' \subseteq \mathcal{P}(\Omega')$, $\mathcal{A}' \subseteq \mathcal{B}'$ (i.e. $A' \in \mathcal{A}' \Rightarrow A' \in \mathcal{B}'$), then $X^{-1}(\mathcal{A}') := \{X^{-1}(A') : A' \in \mathcal{A}'\} \subseteq X^{-1}(\mathcal{B}')$.

Proof. $A \in X^{-1}(\mathcal{A}') \Rightarrow A = X^{-1}(A')$ for some $A' \in \mathcal{A}'$ $\underset{\mathcal{A}' \subseteq \mathcal{B}'}{\Rightarrow} A = X^{-1}(A')$ for some $A' \in \mathrm{B}' \Rightarrow A \in X^{-1}(\mathcal{B}')$. □

2 Measure theory

- **Question:** A fundamental question in measure theory is to measure the volume (or size; or, later, probability) $\lambda(A) \geq 0$ of a set $A \subseteq \mathbb{R}^d (= \Omega)$, $d \geq 1$. Consider $d = 1$ (later also $d > 1$). How can this be done?
- A reasonable such $\lambda : \mathcal{F} \to [0, \infty]$ (some $\mathcal{F}$; λ for now informally called *measure*) should
  1) assign to an interval its length: $\lambda((a, b]) = b - a \; \forall\, a, b \in \mathbb{R} : a \leq b$;
  2) be invariant under translations, rotations and reflections: $\forall\, A, B \subseteq \mathbb{R}$ congruent, $\lambda(A) = \lambda(B)$;
  3) be *σ-additive*: If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{P}(\mathbb{R})$, $A_i \cap A_j = \emptyset \; \forall\, i \neq j$, then $\lambda(\biguplus_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \lambda(A_i)$. At the moment it is not so clear why additivity (without the "$\sigma-$" part) is not sufficient to consider, T. 2.2 will address that.
- Note that λ must be *monotone* since $\lambda(B) = \lambda(A \uplus (B \backslash A)) \underset{3)}{=} \lambda(A) + \lambda(B \backslash A) \geq \lambda(A) \; \forall\, A \subseteq B$, another reasonable property.

2.1 Non-measureable sets

Question: Can we simply take $\mathcal{F} = \mathcal{P}(\mathbb{R}) = \{A : A \subseteq \mathbb{R}\}$ as domain? No!

Theorem 2.1 (Vitali's theorem)

There is no λ defined on $\mathcal{P}(\mathbb{R})$ which satisfies 1)–3).

Proof. We construct a set V, such that, when changed according to 1)–3), we obtain a contradiction.

- Consider $[0,1]$. Then $x \sim y :\Leftrightarrow x - y \in \mathbb{Q}$ defines an equivalence relation (reflexive, symmetric and transitive) on $[0,1]$, with equivalence classes $[x] := \{y \in [0,1] : y \sim x\}$, $x \in [0,1]$.
- The distinct equivalence classes of "$\sim$" partition $[0,1]$.
- The *Vitali set* contains precisely one element of each distinct equivalence class of "$\sim$", i.e.

$$V = \{v \in [0,1] : \ \forall\, x \in [0,1] \ \exists!\, v \sim x\};$$

  the construction requires the axiom of choice (we assumed to have in ZFC).
- Let $\{q_k\}_{k \in \mathbb{N}}$ be a unique enumeration of $\mathbb{Q} \cap [-1,1]$. This can be constructed with *Cantor's first diagonal argument*, here for $\mathbb{Q}$ (skip those numbers already covered to get a unique enumeration):

$$
\begin{array}{cccccccccccc}
0 \to \frac{1}{1} & \to & \pm\frac{1}{2} & & \pm\frac{1}{3} & \to & \pm\frac{1}{4} & & \pm\frac{1}{5} & \to & \pm\frac{1}{6} & \cdots \\
& \swarrow & & \nearrow & & \swarrow & & \nearrow & & \swarrow & & \\
\pm\frac{2}{1} & & \pm\frac{2}{2} & & \pm\frac{2}{3} & & \pm\frac{2}{4} & & \pm\frac{2}{5} & & \pm\frac{2}{6} & \cdots \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & & & \\
\pm\frac{3}{1} & & \pm\frac{3}{2} & & \pm\frac{3}{3} & & \pm\frac{3}{4} & & \pm\frac{3}{5} & & \pm\frac{3}{6} & \cdots \\
& \swarrow & & \nearrow & & \swarrow & & & & & & \\
\pm\frac{4}{1} & & \pm\frac{4}{2} & & \pm\frac{4}{3} & & \pm\frac{4}{4} & & \pm\frac{4}{5} & & \pm\frac{4}{6} & \cdots \\
\downarrow & \nearrow & & \swarrow & & & & & & & & \\
\pm\frac{5}{1} & & \pm\frac{5}{2} & & \pm\frac{5}{3} & & \pm\frac{5}{4} & & \pm\frac{5}{5} & & \pm\frac{5}{6} & \cdots \\
& \swarrow & & & & & & & & & & \\
\pm\frac{6}{1} & & \pm\frac{6}{2} & & \pm\frac{6}{3} & & \pm\frac{6}{4} & & \pm\frac{6}{5} & & \pm\frac{6}{6} & \cdots \\
\downarrow & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots &
\end{array}
$$

- Define the shifted sets $V_k := V + q_k := \{v + q_k : v \in V\}$.
- Then $[0,1] \subseteq \biguplus_{k=1}^{\infty} V_k \subseteq [-1,2]$, since:
  i) $\bigcup_{k=1}^{\infty} V_k \subseteq [-1,2]$, $k \in \mathbb{N}$. *Proof.* $V_k \subseteq [-1,2]$, $k \in \mathbb{N}$.
  ii) $[0,1] \subseteq \bigcup_{k=1}^{\infty} V_k$. *Proof.* Let $x \in [0,1]$. Since "$\sim$" partitions $[0,1]$ into its equivalence classes, we have $x \in [v]$ for some $v \in V \Rightarrow x - v = q_k$ for

some $k \Rightarrow x = v + q_k \in V_k$.

iii) The V_k's are pairwise disjoint.

Proof. If $x \in V_k \cap V_j$ for some $k \neq j$, then $x = v_k + q_k$ and $x = v_j + q_j$ for some $v_k, v_j \in V$ and distinct $q_k, q_j \in \mathbb{Q} \cap [-1, 1]$. Thus $v_k = x - q_k \underset{k \neq j}{\neq} x - q_j = v_j$, so that $\exists\, v_k, v_j \in V : v_k \neq v_j$ but $v_k - x = -q_k \in \mathbb{Q}$ and $v_j - x = -q_j \in \mathbb{Q}$, so $v_k \sim x$ and $v_j \sim x \Rightarrow v_k \sim v_j$ which contradicts the definition of V (V contains precisely one element of each equivalence class). $\square$

- $[0,1] \subseteq \biguplus_{k=1}^{\infty} V_k \subseteq [-1,2] \underset{\text{mon.}}{\Rightarrow} 1 \underset{1)}{=} \lambda([0,1]) \leq \lambda(\biguplus_{k=1}^{\infty} V_k) \leq \lambda([-1,2]) \underset{1)}{=} 3 \underset{3)}{\Rightarrow}$ $1 \leq \sum_{k=1}^{\infty} \lambda(V_k) \leq 3 \underset{2)}{\Rightarrow} 1 \leq \sum_{k=1}^{\infty} \lambda(V) \leq 3$ ↯ $\square$

The Vitali set V is known as a *non-measurable set*, a set we cannot reasonably measure (assign a volume to). In $d > 1$, we can consider $V \times [0,1]^{d-1}$ as a non-measurable set.

Question: How about weakening σ-additivity to finitely-many sets only?

Theorem 2.2 (Banach and Tarski (1924))
Let $d \geq 3$, $A, B \subseteq \mathbb{R}^d$ bounded, non-empty interior. Then $\exists\, k \in \mathbb{N}$ and partitions $A = \biguplus_{i=1}^k A_i$, $B = \biguplus_{i=1}^k B_i$ such that A_i, B_i are congruent $\forall\, i = 1, \dots, k$.

- *Banach–Tarski paradox.* More colloquial, a pea can be chopped up and reassembled into the sun ($\Rightarrow$ buy gold, double it). But the partition elements are not easily constructed, their volumes are impossible to define (since $\lambda(A_i) = \lambda(B_i)$ for all i, at least one of the sets must be non-measurable, otherwise $A = B$).
- For countable Ω, one can always define λ or more general measures μ on $\mathcal{P}(\Omega)$ (see later), but for uncountable Ω, $\mathcal{P}(\Omega)$ can contain non-measurable sets (e.g. Vitali sets). $\mathcal{P}(\Omega)$ is thus too large to be useful for measuring volumes.
- Instead, we need to define λ or more general μ on a family of sets $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$ that is *closed* w.r.t. certain set operations (i.e. performing these operations on sets in $\mathcal{F}$ yields a set in $\mathcal{F}$).

Question: The construction being put aside for now, what are the types of sets $\mathcal{F}$ we can construct measures μ on (and what are their properties)?

2.2 Systems of sets

In the construction of measures, several *systems of sets* (families of sets satisfying certain properties) play a role, including the following (later also π-systems and Dynkin systems).

Definition 2.3 (Semiring)
$\mathcal{A} \subseteq \mathcal{P}(\Omega)$ is a *semiring* on Ω if

i) $\emptyset \in \mathcal{A}$;

ii) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$; and

iii) $A, B \in \mathcal{A} \Rightarrow A \setminus B = \biguplus_{i=1}^{n} A_i$ for some $n \in \mathbb{N}$ and $A_1, \dots, A_n \in \mathcal{A}$ with $A_i \cap A_j = \emptyset \; \forall\, i \neq j$.

Definition 2.4 (Ring)
$\mathcal{A} \subseteq \mathcal{P}(\Omega)$ is a *ring* on Ω if

i) $\emptyset \in \mathcal{A}$;

ii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$; and

iii) $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$.

Definition 2.5 (Algebra)
$\mathcal{A} \subseteq \mathcal{P}(\Omega)$ is an *algebra* (or *field*) on Ω if

i) $\Omega \in \mathcal{A}$;

ii) $A \in \mathcal{A} \Rightarrow A^c = \Omega \backslash A \in \mathcal{A}$; and

iii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

Definition 2.6 (σ-algebra)
$\mathcal{A} \subseteq \mathcal{P}(\Omega)$ is an *σ-algebra* (or *σ-field*) on Ω if

i) $\Omega \in \mathcal{A}$;

ii) $A \in \mathcal{A} \Rightarrow A^c = \Omega \backslash A \in \mathcal{A}$; and

iii) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Proposition 2.7 (σ-algebra $\subsetneq$ algebra $\subsetneq$ ring $\subsetneq$ semiring)
Every σ-algebra is an algebra, every algebra a ring, and every ring a semiring on Ω, with the inclusions being strict. An algebra on a finite set Ω is a σ-algebra. If Ω is an element of a ring, the ring is an algebra. If a semiring is closed wrt the union of two sets, the semiring is a ring.

Proof.

- σ-algebra $\subsetneq$ algebra: Take $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{P}(\Omega)$ with $A_n = \emptyset\ \forall\, n \ge 3 \Rightarrow$ D. 2.5 iii). **Strictness:** Consider $\Omega = (0,1]$, $\mathcal{A} = \{\biguplus_{i=1}^n (a_i, b_i] : 0 \le a_i \le b_i \le 1$ for some $n \in \mathbb{N}\}$ is an algebra but since $(0,1) = \bigcup_{n=1}^\infty (0, 1-1/n] \notin \mathcal{A}$, $\mathcal{A}$ is not a σ-algebra. If $|\Omega| < \infty$, every countable union of sets in $\mathcal{A}$ is a finite union, so an algebra on Ω is also a σ-algebra.
- algebra $\subsetneq$ ring: $\emptyset = \Omega^c \underset{\text{i),ii)}}{\in} \mathcal{A}$. This implies that if $A, B \in \mathcal{A}$, then $A\backslash B = A \cap B^c \underset{\text{De Morgan}}{=} (A^c \cup B)^c \underset{\text{ii),iii)}}{\in} \mathcal{A} \Rightarrow$ D. 2.4 iii). **Strictness:** On $\Omega = \{1,2,3,4\}$, $\mathcal{A} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$ is a ring (check) but not an algebra since $\{1,2,3\}^c = \{4\} \notin \mathcal{A}$. If $\Omega \in \mathcal{A}$, then $A^c = \Omega\backslash A \in \mathcal{A}$, so the ring $\mathcal{A}$ is also an algebra.
- ring $\subsetneq$ semiring: $A \cap B \underset{\text{De Morgan}}{=} A \cap (A \cap B^c)^c = A\backslash(A\backslash B) \underset{A\backslash B\,\in\,\mathcal{A}}{\in} \mathcal{A} \Rightarrow$ D. 2.3 ii). And $A, B \in \mathcal{A} \underset{\mathcal{A}\text{ ring}}{\Rightarrow} A\backslash B \in \mathcal{A}$ and thus $A\backslash B$ satisfies D. 2.3 iii). **Strictness:** On $\Omega = \{1,2,3,4\}$, $\mathcal{A} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2,3\}\}$ is a semiring (check) but not a ring since $\{1\} \cup \{2\} = \{1,2\} \notin \mathcal{A}$. If $\mathcal{A}$ is closed wrt the union of two sets, then by induction also for finitely many, so also for finitely many pairwise disjoint sets $\Rightarrow A\backslash B = \biguplus_{i=1}^n A_i \in \mathcal{A}$. $\square$

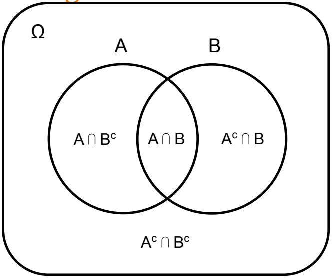
Remark 2.8 (About σ-algebras)

- A σ-algebra $\mathcal{F}$ (typical notation) is a family of subsets of Ω that includes Ω, is closed under complements and countable unions.
- σ-algebras are also closed w.r.t. countable intersections since $\bigcap_{i=1}^{\infty} A_i \underset{\text{De Morgan}}{=} (\bigcup_{i=1}^{\infty} A_i^c)^c \in \mathcal{F}$. So σ-algebras contain complements, countable unions and countable intersections, complements of such sets, etc. Apart from special cases, it seems hopeless to imagine all sets in $\mathcal{F}$.

Example 2.9 (Examples of σ-algebras)

1) The *trivial σ-algebra* $\mathcal{F} = \{\emptyset, \Omega\}$ is the smallest σ-algebra (contained in every σ-algebra) and the power set $\mathcal{F} = \mathcal{P}(\Omega)$ is the largest σ-algebra.
2) Let $A \subseteq \Omega$. Then $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ is a σ-algebra.
3) Let $A, B \subseteq \Omega$, $A \nsubseteq B$, $B \nsubseteq A$, $A \cap B \neq \emptyset$, $A \cup B \neq \Omega$. Then $\mathcal{F} = \{\emptyset, A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c, A, A^c, B, B^c, (A \cap B) \cup (A^c \cap B^c), (A \cap B^c) \cup (A^c \cap B), (A \cap B)^c, (A \cap B^c)^c, (A^c \cap B)^c, (A^c \cap B^c)^c, \Omega\}$ is a σ-algebra.
   Construction: The four disjoint intersections $A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c$ partition Ω. Imagine taking the union of 0 of these 4 elements to form a new

set (so Ω), then of 1 of these 4 (so each at a time), then of 2 (all unions of two of these elements; only 2 not covered yet), then of 3 (so the complement of each), then of all 4 (so Ω); clearly, there are $2^4 = 16$ sets. Such ideas are best imagined with a *Venn diagram*:



With 3 sets, this can be up to $2^8 = 256$ sets already.

4) Let Ω be any set. The *countable-cocountable σ-algebra* $\mathcal{F} = \{A \subseteq \Omega :$ A is countable or A^c is countable$\}$ is a σ-algebra:

i) $\Omega^c = \emptyset$ is countable $\Rightarrow \Omega \in \mathcal{F}$;

ii) $A \in \mathcal{F} \Rightarrow A$ is countable or A^c is countable $\Rightarrow (A^c)^c$ is countable or A^c is countable $\Rightarrow A^c$ is countable or $(A^c)^c$ is countable $\Rightarrow A^c \in \mathcal{F}$.

iii) $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$.

Case 1: A_i countable $\forall\, i \in \mathbb{N} \underset{\text{diag. arg.}}{\overset{\text{Cantor's first}}{\Rightarrow}} \bigcup_{i=1}^{\infty} A_i$ countable $\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Case 2: $\exists\, k \in \mathbb{N} : A_k$ uncountable $\underset{A_k\,\in\,\mathcal{F}}{\Rightarrow} A_k^c$ countable $\Rightarrow (\bigcup_{i=1}^{\infty} A_i)^c \underset{\text{De Morgan}}{=} \bigcap_{i=1}^{\infty} A_i^c \subseteq A_k^c$ countable $\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

5) Let $\mathcal{F}$ be a σ-algebra on Ω and $\Omega' \subseteq \Omega$. The *trace σ-algebra*

$$\mathcal{F}' = \mathcal{F}\big|_{\Omega'} := \{A \cap \Omega' : A \in \mathcal{F}\}$$

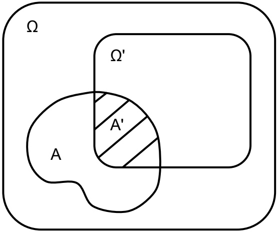
of Ω' in $\mathcal{F}$ is a σ-algebra on Ω':

i) $\Omega' = \Omega \cap \Omega' \underset{\Omega\,\in\,\mathcal{F}}{\in} \mathcal{F}'$;

ii) $A' \in \mathcal{F}' \Rightarrow \exists\, A \in \mathcal{F}$ such that $A' = A \cap \Omega' \Rightarrow (A')^{c_{\Omega'}} = (A \cap \Omega')^{c_{\Omega'}} \underset{\text{De Morgan}}{=} A^{c_{\Omega'}} \cup (\Omega')^{c_{\Omega'}} = A^{c_{\Omega'}} \cup \emptyset = A^{c_{\Omega'}} = A^{c_\Omega} \cap \Omega' \underset{A^{c_\Omega}\,\in\,\mathcal{F}}{\in} \mathcal{F}'$;

iii) $\{A_i'\}_{i\in\mathbb{N}} \subseteq \mathcal{F}' \Rightarrow \forall\, i \in \mathbb{N}\ \exists\, A_i \in \mathcal{F} : A_i' = A_i \cap \Omega' \Rightarrow \bigcup_{i=1}^{\infty} A_i' = \bigcup_{i=1}^{\infty}(A_i \cap \Omega') \underset{\text{distr.}}{=} \big(\bigcup_{i=1}^{\infty} A_i\big) \cap \Omega' \underset{\bigcup_{i=1}^{\infty} A_i\,\in\,\mathcal{F}}{\in} \mathcal{F}'$.

Venn diagram:



6) Let $X : \Omega \to \Omega'$ be a map between two sets Ω, Ω', and let $\mathcal{F}'$ be a σ-algebra on Ω'. Then

$$\mathcal{F} = \sigma(X) := X^{-1}(\mathcal{F}') = \{X^{-1}(A') : A' \in \mathcal{F}'\}$$

is a σ-algebra on Ω, the *preimage σ-algebra* or *σ-algebra generated by X*:

i) $\Omega = X^{-1}(\Omega') \underset{\Omega' \in \mathcal{F}'}{\in} \mathcal{F}$;

ii) $A \in \mathcal{F} \underset{\text{def.}}{\Rightarrow} \exists\, A' \in \mathcal{F}' : A = X^{-1}(A') \Rightarrow A^{c_\Omega} = (X^{-1}(A'))^{c_\Omega} \underset{\text{S. 1.2}}{=} X^{-1}(A'^{c_{\Omega'}}) \underset{A'^{c_{\Omega'}}}{\in} \mathcal{F}$;

iii) $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F} \Rightarrow \forall\, i \in \mathbb{N},\ A_i = X^{-1}(A_i')$ for some $A_i' \in \mathcal{F}' \Rightarrow \bigcup_{i=1}^\infty A_i = \bigcup_{i=1}^\infty X^{-1}(A_i') \underset{\text{S. 1.2}}{=} X^{-1}(\bigcup_{i=1}^\infty A_i') \underset{\bigcup_{i=1}^\infty A_i' \in \mathcal{F}'}{\in} \mathcal{F}$.

7) A *filtration* is an increasing sequence $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$ of σ-algebras and can be used to model *information accrual* over time.

Example: Consider modeling infinite coin tosses with

$$\Omega = \{0,1\}^\infty = \{\boldsymbol{\omega} = (\omega_1, \omega_2, \dots) : \omega_i \in \{0,1\}\ \forall\, i \in \mathbb{N}\}.$$

Let $\mathcal{F}_n = \Big\{\{\boldsymbol{\omega} \in \Omega : (\omega_1, \dots, \omega_n) \in A\}$ for some $A \subseteq \{0,1\}^n\Big\}$ model all events whose occurence can be decided after the first n tosses (e.g. $B = \{\boldsymbol{\omega} \in \Omega : \omega_3 = 1\} \in \mathcal{F}_3$ but $B \notin \mathcal{F}_2$).

Then:

- $\forall\, n \in \mathbb{N}$, $\mathcal{F}_n = \big\{\{\boldsymbol{\omega} \in \Omega : (\omega_1, \dots, \omega_n) \in A\}$ for some $A \subseteq \{0,1\}^n\big\}$ is a σ-algebra on Ω:
  i) $A = \{0,1\}^n \Rightarrow \Omega \in \mathcal{F}_n$;
  ii) If $B \in \mathcal{F}_n$ for some $A \subseteq \{0,1\}^n$, then B^c is obtained for $A^c \subseteq \{0,1\}^n$ $\Rightarrow B^c \in \mathcal{F}_n$;
  iii) If $\{B_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_n$ with corresponding sets $\{A_i\}_{i\in\mathbb{N}} \subseteq \{0,1\}^n \Rightarrow \bigcup_{i=1}^\infty B_i = \{\{\boldsymbol{\omega} \in \Omega : (\omega_1, \dots, \omega_n) \in A\}$ for $A = \bigcup_{i=1}^\infty A_i\} \underset{\bigcup_{i=1}^\infty A_i \subseteq \{0,1\}^n}{\in} \mathcal{F}_n$.
- $\mathcal{F} := \bigcup_{i=1}^\infty \mathcal{F}_i$ is an algebra but not a σ-algebra:
  i) $\mathcal{F}_1$ σ-algebra $\Rightarrow \Omega \in \mathcal{F}_1 \underset{\mathcal{F}_1 \subseteq \mathcal{F}}{\Rightarrow} \Omega \in \mathcal{F}$;
  ii) $A \in \mathcal{F} \Rightarrow \exists\, i \in \mathbb{N} : A \in \mathcal{F}_i \underset{\mathcal{F}_i\ \sigma\text{-algebra}}{\Rightarrow} A^c \in \mathcal{F}_i \Rightarrow A^c \in \mathcal{F}$;
  iii) For $A_1, \dots, A_n \in \mathcal{F}$, $\exists\, j_1, \dots, j_n \in \mathbb{N} : A_i \in \mathcal{F}_{j_i}$, $i = 1, \dots, n \Rightarrow A_1, \dots, A_n \in \mathcal{F}_{\max\{j_1,\dots,j_n\}} \underset{\mathcal{F}_{\max}\ \sigma\text{-algebra}}{\Rightarrow} \bigcup_{i=1}^n A_i \in \mathcal{F}_{\max\{j_1,\dots,j_n\}} \subseteq \mathcal{F}$.

  However, let $A_i = \{\boldsymbol{\omega} \in \Omega : \omega_i = 1\}$, $i \in \mathbb{N}$, and $A_{2\mathbb{N}} = \{\boldsymbol{\omega} \in \Omega : \omega_{2\mathbb{N}} = 1\}$. Then $A_i \in \mathcal{F}_i \subseteq \mathcal{F}\ \forall\, i \in \mathbb{N}$, but $A_{2\mathbb{N}} = \bigcap_{i\in 2\mathbb{N}}^\infty A_i \notin \mathcal{F}$ (since the occurrence of $A_{2\mathbb{N}}$ cannot be determined in n tosses for any finite n).

Question: What is the smallest σ-algebra on Ω containing a given $\mathcal{A} \subseteq \mathcal{P}(\Omega)$?

Proposition 2.10 (σ-algebra generated by $\mathcal{A}$)

Given $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, $\exists!$ minimal σ-algebra on Ω containing $\mathcal{A}$, the *σ-algebra generated by $\mathcal{A}$*, $\sigma(\mathcal{A}) := \bigcap_{\mathcal{F}\ \sigma\text{-alg.}, \mathcal{F}\supseteq\mathcal{A}} \mathcal{F}$. If $\mathcal{G} \subseteq \mathcal{F}'$, $\mathcal{F}'$ σ-algebra, then $\sigma(\mathcal{G}) \subseteq \mathcal{F}'$.

Proof. Let $\mathcal{F}_{\mathcal{A}} := \{\mathcal{F} : \mathcal{F}\ \sigma\text{-algebra},\ \mathcal{F} \supseteq \mathcal{A}\}$, so that $\sigma(\mathcal{A}) \underset{\text{def.}}{=} \bigcap_{\mathcal{F}\in\mathcal{F}_{\mathcal{A}}} \mathcal{F}$.

1) $\sigma(\mathcal{A})$ is a σ-algebra on Ω since

   i) $\mathcal{F} \in \mathcal{F}_{\mathcal{A}}$ is a σ-algebra $\Rightarrow \Omega \in \mathcal{F}\ \forall\, \mathcal{F} \in \mathcal{F}_{\mathcal{A}} \Rightarrow \Omega \in \bigcap_{\mathcal{F}\in\mathcal{F}_{\mathcal{A}}} \mathcal{F} \underset{\text{def.}}{=} \sigma(\mathcal{A})$;

   ii) $A \in \sigma(\mathcal{A}) \Rightarrow A \in \mathcal{F}\ \forall\, \mathcal{F} \in \mathcal{F}_{\mathcal{A}} \underset{\mathcal{F}\ \sigma\text{-alg.}}{\Rightarrow} A^c \in \mathcal{F}\ \forall\, \mathcal{F} \in \mathcal{F}_{\mathcal{A}} \Rightarrow A^c \in \bigcap_{\mathcal{F}\in\mathcal{F}_{\mathcal{A}}} \mathcal{F} \underset{\text{def.}}{=} \sigma(\mathcal{A})$;

   iii) If $\{A_i\}_{i\in\mathbb{N}} \subseteq \sigma(\mathcal{A}) \Rightarrow \{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}\ \forall\, \mathcal{F} \in \mathcal{F}_{\mathcal{A}} \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{F}\ \forall\, \mathcal{F} \in \mathcal{F}_{\mathcal{A}}$ $\Rightarrow \bigcup_{i=1}^\infty A_i \in \sigma(\mathcal{A})$.

2) $\mathcal{A} \subseteq \mathcal{F}\ \forall\, \mathcal{F} \in \mathcal{F}_{\mathcal{A}} \Rightarrow \mathcal{A} \subseteq \bigcap_{\mathcal{F}\in\mathcal{F}_{\mathcal{A}}} \mathcal{F} \underset{\text{def.}}{=} \sigma(\mathcal{A})$, so $\sigma(\mathcal{A}) \supseteq \mathcal{A}$.

3) $\sigma(\mathcal{A})$ is the smallest σ-algebra containing $\mathcal{A}$ since $\forall\, \sigma$-algebras $\mathcal{F}' \supseteq \mathcal{A}$ we have $\mathcal{F}' \in \mathcal{F}_{\mathcal{A}} \Rightarrow \sigma(\mathcal{A}) = \bigcap_{\mathcal{F}\in\mathcal{F}_{\mathcal{A}}} \mathcal{F} \subseteq \mathcal{F}'$.

4) $\mathcal{F}' \underset{\text{ass.}}{\supseteq} \mathcal{G} \underset{\mathcal{F}'\ \sigma\text{-alg.}}{\Rightarrow} \mathcal{F}' \in \mathcal{F}_{\mathcal{G}} \Rightarrow \sigma(\mathcal{G}) \underset{\text{def.}}{=} \bigcap_{\mathcal{F}\in\mathcal{F}_{\mathcal{G}}} \mathcal{F} \subseteq \mathcal{F}'$. $\square$

When considering partitions, we only need to construct unions (not intersections), which simplifies imagining the construction of σ-algebras such as $\sigma(\mathcal{A})$.

Lemma 2.11 (σ-algebra generated by a partition)
If $\mathcal{A} = \{A_i\}_{i\in\mathbb{N}}$ partitions Ω, then $\sigma(\mathcal{A}) = \{\biguplus_{i\in I} A_i : A_i \in \mathcal{A}\ \forall\, i \in I,\ \forall\, I \subseteq \mathbb{N}\}$.

Proof. Let $\mathcal{F} := \{\biguplus_{i\in I} A_i : A_i \in \mathcal{A}\ \forall\, i \in I,\ \forall\, I \subseteq \mathbb{N}\}$. Then:

1) $\mathcal{F}$ is a σ-algebra:
   i) $\Omega \underset{\text{ass.}}{=} \biguplus_{i\in\mathbb{N}} A_i \underset{\text{def. }\mathcal{F}}{\in} \mathcal{F}$.
   ii) Let $A = \biguplus_{i\in I} A_i \in \mathcal{F}$ for some $I \subseteq \mathbb{N}$. Then $A^c = (\biguplus_{i\in I} A_i)^c \underset{\text{def. }()^c}{=} \Omega \setminus \biguplus_{i\in I} A_i$ $\underset{\text{ass.}}{=} \biguplus_{i\in\mathbb{N}\setminus I} A_i \underset{\text{def. }\mathcal{F}}{\in} \mathcal{F}$.
   iii) Let $\{B_k\}_{k\in\mathbb{N}} \subseteq \mathcal{F}$. Then $\forall\, k \in \mathbb{N}$, $\exists\, I_k \subseteq \mathbb{N} : B_k = \biguplus_{i\in I_k} A_i$. Therefore, $\bigcup_{k\in\mathbb{N}} B_k = \bigcup_{k\in\mathbb{N}} \biguplus_{i\in I_k} A_i = \biguplus_{j\in\bigcup_{k\in\mathbb{N}} I_k} A_j \underset{\text{def. }\mathcal{F}}{\in} \mathcal{F}$.
2) $\sigma(\mathcal{A}) \subseteq \mathcal{F}$: Each $A_i \in \mathcal{F}$, $i \in \mathbb{N}$ (take $I = \{i\}$), so $\mathcal{A} \subseteq \mathcal{F}$. By 1), $\mathcal{F}$ is a σ-algebra, so $\sigma(\mathcal{A}) \underset{\text{P. 2.10}}{\subseteq} \mathcal{F}$.
3) $\mathcal{F} \subseteq \sigma(\mathcal{A})$: Let $\biguplus_{i\in I} A_i \in \mathcal{F}$ for some $I \subseteq \mathbb{N}$ and $A_i \in \mathcal{A}\ \forall\, i \in I$. Then $A_i \in \sigma(\mathcal{A})\ \forall\, i \in I \underset{\sigma(\mathcal{A})\ \sigma\text{-alg.}}{\Rightarrow} \biguplus_{i\in I} A_i \in \sigma(\mathcal{A})$. □

Example 2.12 (σ-algebra generated by unions and maps)

- If $\mathcal{F}_1, \mathcal{F}_2$ are σ-algebras on Ω, then, in general, $\mathcal{F}_1 \cup \mathcal{F}_2$ is not a σ-algebra anymore (see exercise). But if $(\mathcal{F}_i)_{i \in I}$ are σ-algebras on Ω, then $\sigma(\mathcal{F}_i, i \in I) := \sigma(\bigcup_{i \in I} \mathcal{F}_i)$ is a σ-algebra on Ω (the smallest that contains the union).
- Let $\Omega \neq \emptyset$ and $X_i : \Omega \to \Omega_i$, $i \in I$, with corresponding σ-algebras $\mathcal{F}_i$ on Ω_i. Then $\sigma(X_i,\ i \in I) := \sigma(\bigcup_{i \in I} \sigma(X_i)) \underset{\text{E. 2.9 6)}}{=} \sigma(\bigcup_{i \in I} X_i^{-1}(\mathcal{F}_i))$ is the *σ-algebra generated by* $(X_i)_{i \in I}$.

Lemma 2.13 (Interpretation of $\sigma(\mathcal{F}_1, \mathcal{F}_2)$)

If $\mathcal{F}_1, \mathcal{F}_2$ are σ-algebras on Ω, then $\sigma(\mathcal{F}_1, \mathcal{F}_2) = \sigma(\{A_1 \cap A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\})$.

Proof. Let $\mathcal{A} := \{A_1 \cap A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$. To show: $\sigma(\mathcal{A}) = \sigma(\mathcal{F}_1, \mathcal{F}_2)$.

"$\subseteq$": $A_1 \cap A_2 \in \mathcal{A} \underset{\sigma(\mathcal{F}_1, \mathcal{F}_2)\ \sigma\text{-alg.}}{\Rightarrow} A_1 \cap A_2 \in \sigma(\mathcal{F}_1, \mathcal{F}_2) \Rightarrow \mathcal{A} \subseteq \sigma(\mathcal{F}_1, \mathcal{F}_2) \underset{\text{P. 2.10}}{\Rightarrow} \sigma(\mathcal{A}) \subseteq \sigma(\mathcal{F}_1, \mathcal{F}_2)$.

"$\supseteq$": $\Omega \in \mathcal{F}_2 \Rightarrow A_1 = A_1 \cap \Omega \in \mathcal{A}\ \forall A_1 \in \mathcal{F}_1 \Rightarrow \mathcal{F}_1 \subseteq \mathcal{A} \underset{\text{def.}}{\subseteq} \sigma(\mathcal{A})$. Similarly, $\mathcal{F}_2 \subseteq \sigma(\mathcal{A})$. Therefore, $\mathcal{F}_1 \cup \mathcal{F}_2 \subseteq \sigma(\mathcal{A}) \Rightarrow \sigma(\mathcal{F}_1, \mathcal{F}_2) \underset{\text{def.}}{=} \sigma(\mathcal{F}_1 \cup \mathcal{F}_2) \underset{\text{P. 2.10}}{\subseteq} \sigma(\mathcal{A})$. □

Question: How can we define σ-algebras on product spaces?

- The *product space* Ω of $(\Omega_i)_{i\in I}$ is $\Omega = \prod_{i\in I}\Omega_i := \{\omega : I \to \bigcup_{i\in I}\Omega_i : \omega(i) \in \Omega_i \;\forall\, i \in I\}$; if $I \subseteq \mathbb{N}$, then $\Omega = \{\omega = (\omega_i)_{i\in\mathbb{N}} : \omega_i \in \Omega_i \;\forall\, i \in I\}$.
- However, if, for $i \in I$, $\mathcal{F}_i$ is a σ-algebra on Ω_i, then $\mathcal{F} = \prod_{i\in I}\mathcal{F}_i$ is in general not a σ-algebra on Ω anymore.

Example 2.14 (Counterexample)

Consider $\Omega = \Omega_1 \times \Omega_2$ for $\Omega_i = \{0,1\}$ with σ-algebra $\mathcal{F}_i = \mathcal{P}(\Omega_i) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$, $i = 1,2$. Then $A_1 = \{(0,0),(0,1),(1,0),(1,1)\} = \{0,1\} \times \{0,1\} = \Omega_1 \times \Omega_2 \in \mathcal{F} := \mathcal{F}_1 \times \mathcal{F}_2$ and $A_2 = \{(1,1)\} = \{1\} \times \{1\} \in \mathcal{F}$, but $A_1 \backslash A_2 = \{(0,0),(0,1),(1,0)\}$ cannot be written as a Cartesian product $A_1' \times A_2'$ with $A_i' \in \mathcal{F}_i$, $i = 1,2$, so $A_1 \backslash A_2 \notin \mathcal{F}$.

Definition 2.15 (Product σ-algebra)

For $i \in I$, let $\mathcal{F}_i$ be a σ-algebra on Ω_i, and let $\Omega = \prod_{i\in I}\Omega_i$ be the product space. For $i \in I$, let $\pi_i : \Omega \to \Omega_i$ denote the *projection onto the ith coordinate* with corresponding preimage $\pi_i^{-1}(A_i) = \{\omega \in \Omega : \pi_i(\omega) \in A_i\}$. Then the *product-$\sigma$-algebra* on Ω is

$$\bigotimes_{i\in I}\mathcal{F}_i := \sigma(\pi_i, i\in I) \overset{\text{def.}}{\underset{\text{E. 2.12}}{=}} \sigma\Big(\bigcup_{i\in I}\sigma(\pi_i)\Big) \overset{\text{def.}}{\underset{\text{E. 2.9 6)}}{=}} \sigma\Big(\bigcup_{i\in I}\pi_i^{-1}(\mathcal{F}_i)\Big),$$

i.e., the σ-algebra generated by all coordinate projections.

Proposition 2.16 (Interpretation for countable I)

If I is countable, then $\bigotimes_{i\in I}\mathcal{F}_i = \sigma(\prod_{i\in I}A_i : A_i\in\mathcal{F}_i\ \forall\, i\in I)$.

Proof.

"$\subseteq$": $\forall\, i\in I$, $\pi_i^{-1}(A_i) \underset{\text{def.}}{=} \{\omega\in\Omega : \pi_i(\omega)\in A_i\} = \prod_{i\in I}A_i$ with $A_k=\Omega_k$ $\forall\, k\neq i \Rightarrow \pi_i^{-1}(A_i)\in\sigma(\prod_{i\in I}A_i : A_i\in\mathcal{F}_i\ \forall\, i\in I) \overset{I\text{ count.}}{\underset{\sigma\text{-alg.}}{\Rightarrow}} \bigcup_{i\in I}\pi_i^{-1}(A_i)\in \sigma(\prod_{i\in I}A_i : A_i\in\mathcal{F}_i) \Rightarrow \bigotimes_{i\in I}\mathcal{F}_i \underset{\text{def.}}{=} \sigma(\bigcup_{i\in I}\pi_i^{-1}(\mathcal{F}_i)) \underset{\text{P. 2.10}}{\subseteq} \sigma(\prod_{i\in I}A_i : A_i\in\mathcal{F}_i\ \forall\, i\in I)$.

"$\supseteq$": Let $A_i\in\mathcal{F}_i$, $i\in I$. Then $\prod_{i\in I}A_i = \bigcap_{i\in I}\pi_i^{-1}(A_i)$ and $\pi_i^{-1}(A_i) \underset{\text{S. 1.2}}{\in} \pi_i^{-1}(\mathcal{F}_i)\in\bigcup_{i\in I}\pi_i^{-1}(\mathcal{F}_i)\in\sigma\Big(\bigcup_{i\in I}\pi_i^{-1}(\mathcal{F}_i)\Big) \underset{\text{def.}}{=} \bigotimes_{i\in I}\mathcal{F}_i \overset{I\text{ count.}}{\underset{\sigma\text{-alg.}}{\Rightarrow}} \prod_{i\in I}A_i\in\bigotimes_{i\in I}\mathcal{F}_i$. $\square$

So for E. 2.14, P. 2.16 implies that $\mathcal{F}_1\otimes\mathcal{F}_2 = \sigma(A_1\times A_2 : A_i\in\mathcal{F}_i) = \sigma(\{\emptyset,\{0\},\{1\},\{0,1\}\}^2) = \sigma(\{\emptyset, \{0\}\times\{0\}, \{0\}\times\{1\}, \dots, \{0,1\}\times\{0,1\}=\Omega\})$.

Question: Are there other systems of sets that can help verifying the properties of σ-algebras or "measures"?

It is often easier to verify closure wrt disjoint unions first, leading to Dynkin systems.

Definition 2.17 (Dynkin system)

$\mathcal{D} \subseteq \mathcal{P}(\Omega)$ is a *Dynkin system* on Ω if

i) $\Omega \in \mathcal{D}$;

ii) $A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$; and

iii) $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$, $A_i \cap A_j = \emptyset \ \forall\, i \neq j \Rightarrow \biguplus_{i=1}^{\infty} A_i \in \mathcal{D}$.

Proposition 2.18 (Properties of Dynkin systems)

1) ii) is equivalent to ii') $A, B \in \mathcal{D}$, $A \subseteq B \Rightarrow B \backslash A \in \mathcal{D}$.

2) ii) and iii) are equivalent to ii') and iii') $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$, $A_i \nearrow \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{D}$.

3) σ-algebra $\subsetneq$ Dynkin system: Every σ-algebra is a Dynkin system. If $\mathcal{D}$ is a Dynkin system and a *π-system* ($A, B \in \mathcal{D} \Rightarrow A \cap B \in \mathcal{D}$), then $\mathcal{D}$ is a σ-algebra.

Proof.

1) "i), ii), iii) $\Rightarrow$ ii')": By i) and ii), $\emptyset \in \mathcal{D}$, so, by iii), finite unions are in $\mathcal{D}$ $(*)$.
Let $A, B \in \mathcal{D}$, $A \subseteq B$. Then $B\backslash A = B \cap A^c \underset{A \subseteq B}{\overset{\text{De Morgan}}{=}} (B^c \uplus A)^c \underset{\text{ii)},(*)}{\in} \mathcal{D}$.
"i), ii'), iii) $\Rightarrow$ ii)": $A^c = \Omega\backslash A \underset{\text{i),ii')}}{\in} \mathcal{D}$.

2) "i), ii), iii) $\Rightarrow$ iii')": By 1), we may also utilize ii'). Let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$, $A_i \nearrow$. With $A_0 := \emptyset \in \mathcal{D}$, and $B_n := A_n\backslash A_{n-1} \underset{\text{ii')}}{\in} \mathcal{D}$, $n \in \mathbb{N}$, we have that $\bigcup_{n=1}^\infty A_n \underset{\text{def.}}{=} \lim_{N\to\infty} \bigcup_{n=1}^N A_n \underset{\text{elementwise}}{=} \lim_{N\to\infty} \biguplus_{n=1}^N B_n \underset{\text{def.}}{=} \biguplus_{n=1}^\infty B_n \underset{\text{iii)}}{\in} \mathcal{D}$.
"i), ii'), iii') $\Rightarrow$ iii)": By 1), ii) and ii') imply ii). Let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$, $A_i \cap A_j = \emptyset \ \forall i \neq j$. For $n \in \mathbb{N}$, let $B_n := \biguplus_{i=1}^n A_i$. Then $B_1 \in \mathcal{D}$ and $B_n = A_n \uplus B_{n-1} \underset{\text{De Morgan}}{=} (A_n^c\backslash B_{n-1})^c \underset{\text{ii), ii'), i.h.}}{\in} \mathcal{D}$, $n \geq 2$. Also, $B_n \nearrow$. Thus $\bigcup_{n=1}^\infty A_n \underset{\text{def.}}{=} \lim_{N\to\infty} \biguplus_{n=1}^N A_n = \lim_{N\to\infty} \bigcup_{n=1}^N B_n \underset{\text{def.}}{=} \bigcup_{n=1}^\infty B_n \underset{\text{iii')}}{\in} \mathcal{D}$.

3)
- Countable unions of any sets from a σ-algebra $\mathcal{F}$ are in $\mathcal{F}$, so also countable unions of pairwise disjoint sets. Hence σ-algebras are Dynkin systems.
- To see the equivalence if $\mathcal{D}$ is a π-system, let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$. Let $B_1 := A_1 \in \mathcal{D}$ and $B_n := A_n\backslash\bigcup_{i=1}^{n-1} A_i = A_n \cap (\bigcup_{i=1}^{n-1} A_i)^c \underset{\text{De Morgan}}{=} A_n \cap \bigcap_{i=1}^{n-1} A_i^c \underset{\text{ii), }\pi\text{-sys.}}{\in} \mathcal{D}$, $n \geq 2$. Then $\bigcup_{n=1}^\infty A_n = \lim_{N\to\infty} \bigcup_{n=1}^N A_n \underset{\text{elementwise}}{=} \lim_{N\to\infty} \biguplus_{n=1}^N B_n = \biguplus_{n=1}^\infty B_n \underset{\text{iii)}}{\in} \mathcal{D}$, so $\mathcal{D}$ is a σ-algebra. $\square$

Similar to σ-algebras, $\delta(\mathcal{A}) := \bigcap_{\mathcal{D} \text{ Dynkin}, \mathcal{D} \supseteq \mathcal{A}} \mathcal{D}$ is the *Dynkin system generated by* $\mathcal{A}$.

Example 2.19 (Examples of Dynkin systems)

1) The *trivial Dynkin system* $\mathcal{D} = \{\emptyset, \Omega\}$ is the smallest Dynkin system (contained in every Dynkin system) and the power set $\mathcal{D} = \mathcal{P}(\Omega)$ is the largest Dynkin system. Both are also σ-algebras.
2) Let $A \subseteq \Omega$. Then $\delta(\{A\}) = \{\emptyset, A, A^c, \Omega\} = \sigma(\{A\})$ is a Dynkin system and σ-algebra.
3) Let $A, B \subseteq \Omega$, $A \not\subseteq B$, $B \not\subseteq A$, $A \cap B \neq \emptyset$, $A \cup B \neq \Omega$. Then

$$\delta(\{A, B\}) = \{\emptyset, A, B, A^c, B^c, \Omega\} \underset{\text{E. 2.9 3)}}{\subsetneq} \sigma(\{A, B\}).$$

   If $A \cap B = \emptyset$ (disjoint) or $A \cap B^c = \emptyset$ $(A \subseteq B)$ or $A^c \cap B = \emptyset$ $(B \subseteq A)$ or $A^c \cap B^c = \emptyset$ $(A \cup B = \Omega)$, then $\delta(\{A, B\}) = \sigma(\{A, B\})$.

Theorem 2.20 (Dynkin's π-λ theorem)
If $\mathcal{D} \subseteq \mathcal{P}(\Omega)$ is a Dynkin system containing a π-system $\mathcal{A}$, then $\sigma(\mathcal{A}) \subseteq \mathcal{D}$. In particular, $\delta(\mathcal{A}) = \sigma(\mathcal{A})$.

Proof.

- For the first statement, we show that $\delta(\mathcal{A})$ is a π-system $\underset{\text{P. 2.18 3)}}{\Rightarrow}$ $\delta(\mathcal{A})$ is a σ-algebra $\Rightarrow \sigma(\mathcal{A}) \underset{\sigma \text{ smallest}}{\subseteq} \delta(\mathcal{A}) \underset{\delta \text{ smallest}}{\subseteq} \mathcal{D}$.
- The second statement follows for $\mathcal{D} = \delta(\mathcal{A})$ since $\sigma(\mathcal{A})$ is a Dynkin system (P. 2.18 3)) containing $\mathcal{A}$, thus $\delta(\mathcal{A}) \underset{\delta \text{ smallest}}{\subseteq} \sigma(\mathcal{A})$ (and the first part implies that $\sigma(\mathcal{A}) \subseteq \delta(\mathcal{A})$).
- To show that $\delta(\mathcal{A})$ is a π-system, consider for any $B \in \delta(\mathcal{A})$ the set

$$\mathcal{D}_B := \{A \in \delta(\mathcal{A}) : A \cap B \in \delta(\mathcal{A})\} \quad \text{(the 'good' sets).}$$

  1) We first show that $\forall\, B \in \delta(\mathcal{A})$, $\mathcal{D}_B$ is a Dynkin system:
     i) $\Omega \in \delta(\mathcal{A})$ and $\Omega \cap B = B \underset{\text{ass.}}{\in} \delta(\mathcal{A}) \underset{\text{def.}}{\Rightarrow} \Omega \in \mathcal{D}_B$;
     ii) $A \in \mathcal{D}_B \underset{\text{def.}}{\Rightarrow} A \in \delta(\mathcal{A})$ and $A \cap B \in \delta(\mathcal{A}) \Rightarrow A^c \in \delta(\mathcal{A})$ and $A^c \cap B = B\backslash A = B\backslash(A \cap B) \underset{B \in \delta(\mathcal{A}),\ A \cap B \subseteq B,\ \text{P. 2.18 1)}}{\overset{A \in \mathcal{D}_B,\ \text{so } A \cap B \in \delta(\mathcal{A})}{\in}} \delta(\mathcal{A}) \Rightarrow A^c \in \mathcal{D}_B$; and

iii) $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}_B$ with $A_i \cap A_j = \emptyset\ \forall\, i \neq j \underset{\text{def. }\mathcal{D}_B}{\Rightarrow} A_i \in \delta(\mathcal{A})$ and $A_i \cap B \in \delta(\mathcal{A})\ \forall\, i \Rightarrow \biguplus_{i=1}^{\infty} A_i \in \delta(\mathcal{A})$ and $(\biguplus_{i=1}^{\infty} A_i) \cap B \underset{\text{distr.}}{=} \biguplus_{i=1}^{\infty}(A_i \cap B) \overset{A_i\cap B\in\delta(\mathcal{A})}{\underset{\text{iii)}}{\in}} \delta(\mathcal{A}) \Rightarrow \biguplus_{i=1}^{\infty} A_i \in \mathcal{D}_B$.

2) $\forall\, B \in \mathcal{A}$ we have $\delta(\mathcal{A}) \subseteq \mathcal{D}_B$, since $\forall\, A, B \in \mathcal{A}$, $A \cap B \underset{\mathcal{A}\ \pi\text{-sys.}}{\in} \mathcal{A} \subseteq \delta(\mathcal{A}) \underset{\text{def. }\mathcal{D}_B}{\Rightarrow} A \in \mathcal{D}_B \underset{\forall A\,\in\,\mathcal{A}}{\Rightarrow} \mathcal{A} \subseteq \mathcal{D}_B \overset{\text{Dynkin by 1)}}{\underset{\delta\text{ smallest}}{\Rightarrow}} \delta(\mathcal{A}) \subseteq \mathcal{D}_B$. So $\forall\, B \in \mathcal{A}$, $\forall\, A \in \delta(\mathcal{A})$, we have $A \in \mathcal{D}_B \underset{\text{def. }\mathcal{D}_B}{\Rightarrow} A \cap B \in \delta(\mathcal{A})$.

3) We now extend 2): $\forall\, B \in \delta(\mathcal{A})$, we have $\delta(\mathcal{A}) \subseteq \mathcal{D}_B$, since $\forall\, B \in \delta(\mathcal{A})$, $A \cap B \overset{\text{2) with}}{\underset{A\,\leftrightarrow\,B}{\in}} \delta(\mathcal{A})\ \forall A \in \mathcal{A} \underset{\forall\, A}{\Rightarrow} \mathcal{A} \subseteq \mathcal{D}_B \overset{\text{1)}}{\underset{\delta\text{ smallest}}{\Rightarrow}} \delta(\mathcal{A}) \subseteq \mathcal{D}_B$. Therefore, $\forall\, B \in \delta(\mathcal{A})$ and $\forall\, A \in \delta(\mathcal{A})$, we have $A \in \mathcal{D}_B \underset{\text{def. }\mathcal{D}_B}{\Rightarrow} A \cap B \in \delta(\mathcal{A})$, so $\delta(\mathcal{A})$ is indeed a π-system. □

The proof used the *principle of good sets*. It is often used to show that a certain property holds for all elements of a σ-algebra $\mathcal{F}$, we can consider the family $\mathcal{G}$ of all 'good' subsets (those which satisfy the property). If $\mathcal{G}$ is a σ-algebra that contains a generator $\mathcal{A}$ of $\mathcal{F}$, then $\mathcal{F} \underset{\mathcal{F}\text{ gen. by }\mathcal{A}}{=} \sigma(\mathcal{A}) \underset{\sigma\text{ smallest}}{\subseteq} \mathcal{G} \Rightarrow$ all sets in $\mathcal{F}$ are 'good', so $\mathcal{F}$ satisfies the property. And if it is easier to verifying that $\mathcal{G}$ is Dynkin, one can apply Dynkin's $\pi - \lambda$ theorem to conclude that $\mathcal{G}$ is a σ-algebra.

Question: What about σ-algebras generated by topologies?

- σ-algebras generated by topologies (sets of open sets) are of particular interest. A *topology* on Ω is a family of subsets $\mathcal{T} \subseteq \mathcal{P}(\Omega)$ such that i) $\emptyset, \Omega \in \mathcal{T}$; ii) $A_i \in \mathcal{T}$, $i \in I \Rightarrow \bigcup_{i\in I} A_i \in \mathcal{T}$; and iii) $A_i \in \mathcal{T}$, $i \in \{1, \dots, n\} \Rightarrow \bigcap_{i=1}^n A_i \in \mathcal{T}$.
- The following σ-algebra can be defined on quite general topologies.

Definition 2.21 (Borel σ-algebra, Borel sets)
If $(\Omega, \mathcal{T})$ is a topological space (e.g. metric space), then $\mathcal{B}(\Omega) := \sigma(\mathcal{T}) = \sigma(\{O : O \subseteq \Omega,\ O \text{ open}\})$ is the *Borel σ-algebra* on Ω and its elements are *Borel sets*.

Borel sets include open sets, closed sets, countable unions and countable intersections of these, etc.

Question: How can we imagine them?

Lemma 2.22 (Characterization of open sets in $\mathbb{R}$)
Every open set in $\mathbb{R}$ is a countable disjoint union of open intervals.

Proof. $O \subseteq \mathbb{R}$ open $\Rightarrow$ For $x \in O$, let $I_x := \bigcup_{I \text{ open interval} \subseteq O:\, x \in I} I$ be the open interval of maximal length containing x. If $x, y \in O$, then either $I_x = I_y$ or $I_x \cap I_y = \emptyset \underset{\text{choice}}{\Rightarrow}$ Let $\mathcal{I} = \{I_x : x \in O\}$ be the set of all distinct intervals of maximal length $\Rightarrow O \underset{\text{elementwise}}{=} \bigcup_{x\in O} I_x \underset{\text{distinct}}{=} \bigcup_{I\in\mathcal{I}} I$, and the union is at most countable (each $I \in \mathcal{I}$ contains a $q_I \in \mathbb{Q}$). $\square$

Proposition 2.23 (Borel σ-algebra is generated by all open intervals)
If $\Omega = \mathbb{R}$, then $\mathcal{B}(\mathbb{R}) = \sigma(\{(a,b) : a < b\})$.

Proof.

"$\subseteq$": By L. 2.22, every open $O \subseteq \mathbb{R}$ is a countable disjoint union of open intervals and thus in $\sigma(\{(a,b) : a < b\}) \Rightarrow \{O : O \subseteq \mathbb{R},\ O \text{ open}\} \subseteq \sigma(\{(a,b) : a < b\}) \Rightarrow \mathcal{B}(\mathbb{R}) \underset{\text{def.}}{=} \sigma(\{O : O \subseteq \mathbb{R},\ O \text{ open}\}) \underset{\sigma \text{ smallest}}{\subseteq} \sigma(\{(a,b) : a < b\})$.

"$\supseteq$": $(a,b) \in \mathcal{B}(\mathbb{R})\ \forall\, a < b \Rightarrow \{(a,b) : a < b\} \subseteq \mathcal{B}(\mathbb{R}) \underset{\sigma \text{ smallest}}{\Rightarrow} \sigma(\{(a,b) : a < b\}) \subseteq \mathcal{B}(\mathbb{R})$ $\square$

Remark 2.24 (Generators of Borel σ-algebras)

1) Often, $\Omega = \mathbb{R}^d$, $d \ge 2$, is of interest, and $\mathcal{B}(\mathbb{R}^d)$ is defined as in D. 2.21. L. 2.22 is then false in general (open ball $\neq$ countable disjoint union of open rectangles), but one can show that any open set is a countable union of rectangles with

rational endpoints. With this one can show that, $\forall\, d \in \mathbb{N}$,

$$\begin{aligned}\mathcal{B}(\mathbb{R}^d) &= \sigma(\{(\boldsymbol{a},\boldsymbol{b}) : \boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d, \boldsymbol{a} < \boldsymbol{b}\}) = \sigma(\{[\boldsymbol{a},\boldsymbol{b}] : \boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d, \boldsymbol{a} < \boldsymbol{b}\})\\ &= \sigma(\{(\boldsymbol{a},\boldsymbol{b}] : \boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d, \boldsymbol{a} < \boldsymbol{b}\}) = \sigma(\{[\boldsymbol{a},\boldsymbol{b}) : \boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d, \boldsymbol{a} < \boldsymbol{b}\})\\ &= \sigma(\{(-\boldsymbol{\infty},\boldsymbol{b}) : \boldsymbol{b} \in \mathbb{R}^d\}) = \sigma(\{(\boldsymbol{a},\boldsymbol{\infty}) : \boldsymbol{a} \in \mathbb{R}^d\})\\ &= \sigma(\{(-\boldsymbol{\infty},\boldsymbol{b}] : \boldsymbol{b} \in \mathbb{R}^d\}) = \sigma(\{[\boldsymbol{a},\boldsymbol{\infty}) : \boldsymbol{a} \in \mathbb{R}^d\});\end{aligned}$$

for a proof of $\sigma(\{(a,b] : a < b\}) = \mathcal{B}(\mathbb{R})$ for $d = 1$, see exercises. One can also show via P. 2.16 that $\mathcal{B}(\mathbb{R}^d) = \bigotimes_{j=1}^d \mathcal{B}(\mathbb{R})$; see also Folland (1999, P. 1.4, P. 1.5).

2) We later (Section 5) also consider $\Omega = \bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$. One can show that $A \subseteq \bar{\mathbb{R}}$ is open iff if A is a countable union of members of $\{(a,b) : a,b \in \mathbb{R}\} \cup \{[-\infty, b) : b \in \mathbb{R}\} \cup \{(a,\infty] : a \in \mathbb{R}\}$. One then obtains $\mathcal{B}(\bar{\mathbb{R}}) = \{B \cup E : B \in \mathcal{B}(\mathbb{R}), E \subseteq \{-\infty, \infty\}\}$ and, e.g., $\mathcal{B}(\bar{\mathbb{R}}) = \sigma(\{(a,\infty] : a \in \mathbb{R}\})$ besides other representations. Furthermore, $\mathcal{B}(\bar{\mathbb{R}}^d) = \bigotimes_{j=1}^d \mathcal{B}(\bar{\mathbb{R}})$.

3) Other Ω, e.g. $\Omega = \bar{\mathbb{R}}_+^d = [\mathbf{0}, \boldsymbol{\infty}]^d$ can be obtained via the respective trace σ-algebra.

2.3 Measures

We first define and investigate 'measures' (of which 'λ', '$\mathbb{P}$' are special cases) and later think about how 'measures' arise (existence, uniqueness), which will answer our initial question about λ.

Definition 2.25 (Measurable space, measurable sets, measure, σ-additivity, measure space, (σ-)finite, Borel measure)
Let $\mathcal{F}$ be a σ-algebra on Ω. Then $(\Omega, \mathcal{F})$ is a *measurable space* and sets in $\mathcal{F}$ are *measurable sets*. A *measure* μ on $\mathcal{F}$ is a function such that

i) $\mu : \mathcal{F} \to [0, \infty]$;

ii) $\mu(\emptyset) = 0$; and

iii) $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$, $A_i \cap A_j = \emptyset\ \forall\, i \neq j \Rightarrow \mu(\biguplus_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ (*σ-additivity*).

The triplet $(\Omega, \mathcal{F}, \mu)$ is a *measure space*. If $\Omega = \bigcup_{i=1}^{\infty} A_i$ for $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$: $\mu(A_i) < \infty\ \forall\, i$, then μ is a *σ-finite* measure. If $\mu(\Omega) < \infty$, then μ is a *finite* measure. A measure μ on $\mathcal{F} = \mathcal{B}(\mathbb{R}^d)$ is a *Borel measure on $\mathbb{R}^d$*, $d \geq 1$.

On σ-additivity and uncountable additivity:

- σ-additivity (in contrast to finite additivity) allows for limits to be included (pointwise limits of 'measurable functions' are 'measureable', 'dominated convergence', construction of the 'Lebesgue integral', etc.; see later).

  Example: For an enumeration $\{q_i\}_{i=1}^{\infty}$ of $\mathbb{Q} \cap [0, 1]$, consider

  $$f_n(x) = \mathbb{1}_{\{q_1,\dots,q_n\}}(x) \underset{n\to\infty}{\overset{\text{pointwise}}{\longrightarrow}} f(x) = \mathbb{1}_{\mathbb{Q}\cap[0,1]}(x).$$

  - Then $\int_0^1 f_n(x)\,\mathrm{d}x = 0\ \forall\, n \in \mathbb{N}$ ($\forall\, n$ we find a small enough partition such that the n times f_n is 1 does not alter the value of the integral by a given small $\varepsilon > 0$).
  - Therefore $\int_0^1 f_n(x)\,\mathrm{d}x \underset{n\to\infty}{\to} 0$, so we expect $\int_0^1 f(x)\,\mathrm{d}x = 0$.
  - However, the Riemann integral $\int_0^1 f(x)\,\mathrm{d}x$ does not exist (there's a rational and an irrational number in each subinterval of $[0, 1]$).

- Requiring uncountable additivity would be too strong, since $\forall\, A \subseteq \mathbb{R}$,

  $$\lambda(A) = \lambda\Big(\bigcup_{x\in A}\{x\}\Big) \underset{\text{ass.}}{\overset{!}{=}} \sum_{x\in A}\lambda(\{x\}) \underset{\text{def.}}{=} \sup_{\substack{B\subseteq A,\\ |B|<\infty}} \sum_{x\in B}\lambda(\{x\}) \underset{\text{iii)}}{\overset{\lambda(\{x\})=0}{=}} 0,$$

  so all sets would have length 0.

Question: What are examples of measures (so that we don't speak about $\emptyset$)?

Example 2.26 (Measures)

1) If $(\Omega, \mathcal{F})$ is a measurable space, then $\mu(A) = 0$, $A \in \mathcal{F}$, and $\mu(A) = \infty \mathbb{1}_{\{A \neq \emptyset\}}$ (with the convention $\infty \cdot 0 = 0$) are *trivial measures*. They are valid measures even if $\mathcal{F} = \mathcal{P}(\Omega)$.
2) For uncountable Ω, consider the countable-cocountable σ-algebra $\mathcal{F} = \{A \subseteq \Omega : A \text{ or } A^c \text{ is countable}\}$. Then $\mu(A) = \mathbb{1}_{\{A \text{ uncountable}\}}$ is a measure on $\mathcal{F}$:
   i) $\mu : \mathcal{F} \to [0, 1] \subseteq [0, \infty]$ ✓;
   ii) $\mu(\emptyset) \underset{\emptyset \text{ count.}}{=} 0$ ✓; and
   iii) Let $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$, $A_i \cap A_j = \emptyset \;\forall\, i \neq j$. Then

      Case 1: A_i countable $\forall\, i \in \mathbb{N} \Rightarrow \biguplus_{i=1}^{\infty} A_i$ countable $\Rightarrow \mu(\biguplus_{i=1}^{\infty} A_i) = 0$. Also, $\mu(A_i) \underset{A_i \text{ count.}}{=} 0$ for all i, so indeed $\sum_{i=1}^{\infty} \mu(A_i) = 0$.

      Case 2: $\exists\, A_k : A_k$ is uncountable. Then $\biguplus_{i=1}^{\infty} A_i$ is uncountable and thus $\mu(\biguplus_{i=1}^{\infty} A_i) = 1$. Also, A_k^c is countable and $A_i \underset{A_i \cap A_k = \emptyset}{\subseteq} A_k^c$ $\forall\, i \neq k$, so A_i must be countable $\forall\, i \neq k \Rightarrow \mu(A_i) = 0 \;\forall\, i \neq k$ and $\mu(A_k) = 1$, so indeed $\sum_{i=1}^{\infty} \mu(A_i) = 1$.

3) For countable Ω consider $\mathcal{F} = \mathcal{P}(\Omega)$. Then $\forall\, f : \Omega \to [0, \infty]$,

$$\mu(A) := \sum_{w \in A} f(w), \quad A \in \mathcal{F},$$

defines a measure on $\mathcal{F}$:

i) $f \geq 0 \Rightarrow \mu : \mathcal{F} \to [0, \infty]$.
ii) $\mu(\emptyset) = \sum_{\omega \in \emptyset} f(\omega) \underset{\text{sum}}{\overset{\text{empty}}{=}} 0$.
iii) If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$, $A_i \cap A_j = \emptyset\ \forall\, i \neq j$, then

$$\mu\Bigl(\biguplus_{i=1}^{\infty} A_i\Bigr) \underset{\text{def.}}{=} \sum_{\omega \in \biguplus_{i=1}^{\infty} A_i} f(\omega) = \sum_{i=1}^{\infty} \sum_{\omega \in A_i} f(\omega) \underset{\text{def.}}{=} \sum_{i=1}^{\infty} \mu(A_i).$$

- If $f \equiv 1$, then $\mu(A) = |A|$ is the *counting measure*.
- If, for $\tilde{\omega} \in \Omega$, $f(w) = \mathbb{1}_{\{\omega = \tilde{\omega}\}}$, then $\mu(A) = \mathbb{1}_{\{\tilde{\omega} \in A\}}$ is the *Dirac measure* or *point mass* or *unit mass* of $\tilde{\omega}$.

Question: What properties do measures have? Some even extend to (semi)rings.

Proposition 2.27 (Basic properties of measures)

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space.

1) $A, B \in \mathcal{F} \Rightarrow \mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$. If μ is finite, then $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.
2) $A, B \in \mathcal{F}$, $A \subseteq B \Rightarrow \mu(A) \le \mu(B)$ (*monotonicity*).
   If $\mu(A) < \infty$, then $\mu(B \backslash A) = \mu(B) - \mu(A)$ (*subtractivity*).
3) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F} \Rightarrow \mu(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \mu(A_i)$ (*σ-subadditivity*).
4) If μ is finite, $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$, $S_{j,n} := \sum_{J \subseteq \{1,\dots,n\}:|J|=j} \mu(\bigcap_{k \in J} A_k)$, then
$$\mu\Big(\bigcup_{j=1}^{n} A_j\Big) = \sum_{j=1}^{n} (-1)^{j-1} S_{j,n} \quad \text{(\textit{inclusion–exclusion principle})}.$$
5) If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$, $A_i \nearrow$, then $\mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{n \to \infty} \mu(A_n)$ (*cont. from below*).
6) If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$, $A_i \searrow$ and $\mu(A_1) < \infty$, then $\mu(\bigcap_{i=1}^{\infty} A_i) = \lim_{n \to \infty} \mu(A_n)$ (*continuity from above*).
7) If $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$ forms a partition of Ω, $A \in \mathcal{F}$, then $\mu(A) = \sum_{i=1}^{\infty} \mu(A \cap B_i)$ (*law of total measure*).

Proof.

1) $A \cup B = A \uplus (B \backslash A)$ and $B = (A \cap B) \uplus (B \backslash A)$. By additivity, $\mu(A \cup B) \underset{(*)}{=} \mu(A) + \mu(B \backslash A)$ and $\mu(B) \underset{(**)}{=} \mu(A \cap B) + \mu(B \backslash A)$. Adding opposite sides gives $\mu(A \cup B) + \mu(A \cap B) + \mu(B \backslash A) = \mu(A) + \mu(B \backslash A) + \mu(B)$.
   - If $\mu(B \backslash A) < \infty$, subtract $\mu(B \backslash A)$ and we are done.
   - And if $\mu(B \backslash A) = \infty$, then $\mu(A \cup B) \underset{(*)}{=} \infty$ and $\mu(B) \underset{(**)}{=} \infty$, so the formula as stated is still valid (it then states "$\infty = \infty$").

   If μ is finite, subtract $\mu(A \cap B)$ from both sides in the just shown first statement.
2) $\mu(B) \underset{(**)}{=} \mu(A \cap B) + \mu(B \backslash A) \overset{A \cap B \underset{\text{ass.}}{=} A}{\underset{(***)}{=}} \mu(A) + \mu(B \backslash A) \geq \mu(A)$. If $\mu(A) < \infty$, subtract it to obtain $\mu(B \backslash A) \underset{(***)}{=} \mu(B) - \mu(A)$ (irresp. of the value of $\mu(B)$).
3) Let $B_1 := A_1$, $B_n := A_n \backslash \bigcup_{i=1}^{n-1} A_i$, $n \geq 2 \Rightarrow B_n$'s are pairwise disjoint $\Rightarrow$ $\bigcup_{i=1}^{\infty} A_i = \lim_{N \to \infty} \bigcup_{i=1}^{N} A_i = \lim_{N \to \infty} \biguplus_{i=1}^{N} B_i = \biguplus_{i=1}^{\infty} B_i \Rightarrow \mu(\bigcup_{i=1}^{\infty} A_i)$ $= \mu(\biguplus_{i=1}^{\infty} B_i) \underset{\sigma\text{-add.}}{=} \sum_{i=1}^{\infty} \mu(B_i) \overset{B_i \subseteq A_i}{\underset{\text{mon.}}{\leq}} \sum_{i=1}^{\infty} \mu(A_i)$.
4) Induction in n based on 1):

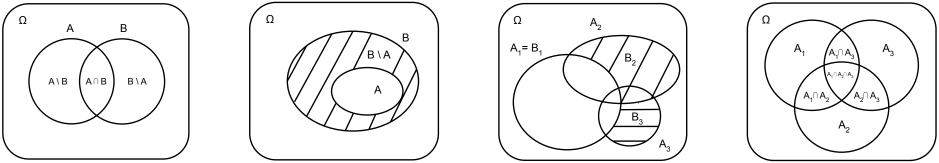
   $n = 2 : \mu(A_1 \cup A_2) \underset{1)}{=} \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2) = \sum_{j=1}^{2} (-1)^{j-1} S_{j,2}$.

   $n \Rightarrow n+1$:

$$
\begin{aligned}
\mu\Big(\bigcup_{j=1}^{n+1} A_j\Big) &= \mu\Big(\Big(\bigcup_{j=1}^{n} A_j\Big) \cup A_{n+1}\Big)\\
&\underset{1)}{=} \mu\Big(\bigcup_{j=1}^{n} A_j\Big) + \mu(A_{n+1}) - \mu\Big(\Big(\bigcup_{j=1}^{n} A_j\Big) \cap A_{n+1}\Big)\\
&\overset{\text{hypo.}}{\underset{\text{distr.}}{=}} \sum_{j=1}^{n}(-1)^{j-1} S_{j,n} + \mu(A_{n+1}) - \mu\Big(\bigcup_{j=1}^{n}(A_j \cap A_{n+1})\Big)\\
&\underset{\text{hypo.}}{=} \sum_{j=2}^{n}(-1)^{j-1} S_{j,n} + \underbrace{S_{1,n}}_{\underset{\text{def.}}{=} \sum_{i=1}^{n} \mu(A_i)} + \mu(A_{n+1}) - \sum_{j=1}^{n}(-1)^{j-1} \sum_{J\subseteq\{1,\dots,n\}:|J|=j} \mu\Big(\Big(\bigcap_{k\in J} A_k\Big) \cap A_{n+1}\Big)\\
&= \sum_{j=2}^{n}(-1)^{j-1} S_{j,n} + S_{1,n+1} - \sum_{j=1}^{n}(-1)^{j-1} \sum_{J\subseteq\{1,\dots,n\}:|J|=j} \mu\Big(\Big(\bigcap_{k\in J} A_k\Big) \cap A_{n+1}\Big)\\
&\underset{(*)}{=} S_{1,n+1} + \sum_{j=2}^{n+1}(-1)^{j-1} S_{j,n+1} = \sum_{j=1}^{n+1}(-1)^{j-1} S_{j,n+1},
\end{aligned}
$$

where $(*)$ holds since the first sum contains all intersections of at least two sets of which none contains A_{n+1}, and the last sum contains all intersections of at least two sets where one is A_{n+1}, so together we obtain the sum containing all intersections of at least two sets among $A_1, \dots, A_{n+1}$.

Reminder: If possible and helpful, draw Venn diagrams. For 1)–4):



5) $A_0 := \emptyset$, $A_i \nearrow \Rightarrow \mu(\bigcup_{i=1}^\infty A_i) \underset{\text{partition}}{\overset{\text{disjoint}}{=}} \mu(\biguplus_{i=1}^\infty (A_i \backslash A_{i-1})) \underset{\sigma\text{-add.}}{=} \sum_{i=1}^\infty \mu(A_i \backslash A_{i-1}) \underset{\text{def.}}{=}$ $\lim_{n\to\infty} \sum_{i=1}^n \mu(A_i \backslash A_{i-1}) \underset{\text{add.}}{=} \lim_{n\to\infty} \mu(\biguplus_{i=1}^n (A_i \backslash A_{i-1})) \underset{A_i \nearrow}{=} \lim_{n\to\infty} \mu(A_n)$.

6) Let $B_i := A_1 \backslash A_i = A_1 \cap A_i^c$, $i \in \mathbb{N}$. Then $B_i \nearrow \bigcup_{i=1}^\infty B_i = \bigcup_{i=1}^\infty (A_1 \cap A_i^c) \underset{\text{distr.}}{=}$ $A_1 \cap \bigcup_{i=1}^\infty A_i^c \underset{\text{De Morgan}}{=} A_1 \cap (\bigcap_{i=1}^\infty A_i)^c = A_1 \backslash \bigcap_{i=1}^\infty A_i$. Since $\mu(A_1) < \infty$,

$$\mu(A_1) - \mu\left(\bigcap_{i=1}^\infty A_i\right) \underset{2)}{=} \mu\left(A_1 \backslash \bigcap_{i=1}^\infty A_i\right) = \mu\left(\bigcup_{i=1}^\infty B_i\right) \underset{5)}{\overset{B_i \nearrow}{=}} \lim_{n\to\infty} \mu(B_n)$$

$$\underset{\text{def.}}{=} \lim_{n\to\infty} \mu(A_1 \backslash A_n) \underset{\mu(A_n) < \infty}{\overset{2)}{=}} \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n).$$

Since $\mu(A_1) < \infty$, subtract $\mu(A_1)$ from both sides to get the result.

7) $\mu(A) = \mu(A \cap \Omega) \underset{\text{part.}}{=} \mu(A \cap \biguplus_{i=1}^\infty B_i) \underset{\text{distr.}}{=} \mu(\biguplus_{i=1}^\infty (A \cap B_i)) \underset{\sigma\text{-add.}}{=} \sum_{i=1}^\infty \mu(A \cap B_i)$. $\square$

Proposition 2.28 (Uniqueness)

Let μ, ν be measures on $(\Omega, \mathcal{F})$ and $\mathcal{A}$ a π-system such that $\exists\,(A_i)_{i\in\mathbb{N}} \subseteq \mathcal{A}$ with $\bigcup_{i=1}^{\infty} A_i = \Omega$, $\mu(A_i) < \infty\ \forall\, i \in \mathbb{N}$. If $\mu|_{\mathcal{A}} = \nu|_{\mathcal{A}}$, then $\mu|_{\sigma(\mathcal{A})} = \nu|_{\sigma(\mathcal{A})}$. In particular, if $\sigma(\mathcal{A}) = \mathcal{F}$, then $\mu = \nu$.

Proof.

1) For $B \in \mathcal{A} : \mu(B) < \infty$, let $\mathcal{D}_B := \{A \in \sigma(\mathcal{A}) : \mu(A \cap B) = \nu(A \cap B)\}$. We first show that $\mathcal{D}_B$ is a Dynkin system:

   i) $\Omega \in \sigma(\mathcal{A}) : \mu(\Omega \cap B) = \mu(B) \underset{B\,\in\,\mathcal{A}}{=} \nu(B) = \nu(\Omega \cap B) \Rightarrow \Omega \in \mathcal{D}_B$;

   ii) $A \in \mathcal{D}_B \Rightarrow A \in \sigma(\mathcal{A}) \Rightarrow A^c \in \sigma(\mathcal{A})$. Furthermore, $\mu(A^c \cap B) + \mu(A \cap B) \underset{\sigma\text{-add.}}{=} \mu((A^c \cap B) \uplus (A \cap B)) \overset{\text{tot.}}{\underset{\text{meas.}}{=}} \mu(B) \underset{B\,\in\,\mathcal{A}}{=} \nu(B) = \dots \overset{\text{same}}{\underset{\text{backwards}}{=}} \nu(A^c \cap B) + \nu(A \cap B)$, with $\mu(A \cap B) \underset{A\,\in\,\mathcal{D}_B}{=} \nu(A \cap B) \overset{\mu(A\cap B)\,\le\,\mu(B)\,<\,\infty}{\underset{\text{subtract}}{\Rightarrow}} \mu(A^c \cap B) = \nu(A^c \cap B) \Rightarrow A^c \in \mathcal{D}_B$; and

   iii) $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}_B : A_i \cap A_j = \emptyset\ \forall\, i \neq j \Rightarrow \mu((\biguplus_{i=1}^{\infty} A_i) \cap B) \underset{\text{distr.}}{=} \mu(\biguplus_{i=1}^{\infty}(A_i \cap B)) \underset{\sigma\text{-add.}}{=} \sum_{i=1}^{\infty} \mu(A_i \cap B) \underset{\text{def. }\mathcal{D}_B}{=} \sum_{i=1}^{\infty} \nu(A_i \cap B) = \dots \overset{\text{same}}{\underset{\text{backwards}}{=}} \nu((\biguplus_{i=1}^{\infty} A_i) \cap B) \Rightarrow \biguplus_{i=1}^{\infty} A_i \in \mathcal{D}_B$.

2) By ass., $\mathcal{A}$ is a π-system, so $\forall\, A_1, A_2 \in \mathcal{A}$, $A_1 \cap A_2 \in \mathcal{A}$ and thus $\mu(A_1 \cap A_2) \overset{\mu\,=\,\nu}{\underset{\text{on }\mathcal{A}}{=}} \nu(A_1 \cap A_2)$, so $\mathcal{A} \subseteq \mathcal{D}_B$. By Dynkin's π-λ T., $\sigma(\mathcal{A}) \subseteq \mathcal{D}_B \underset{\text{def. }\mathcal{D}_B}{\subseteq} \sigma(\mathcal{A})$.

So $\mathcal{D}_B = \sigma(\mathcal{A})\ \forall\, B \in \mathcal{A} : \mu(B) < \infty$, thus $\mu(A \cap B) = \nu(A \cap B)\ \forall\, A \in \sigma(\mathcal{A})$, $\forall\, B \in \mathcal{A} : \mu(B) < \infty$.

3) i) By ass., $\exists\, (A_i)_{i\in\mathbb{N}} \underset{\text{ass.}}{\subseteq} \mathcal{A} : \bigcup_{i=1}^{\infty} A_i = \Omega$, $\mu(A_i) < \infty\ \forall\, i \in \mathbb{N}$. Let $A_0' := \emptyset$ and $A_n' := \bigcup_{i=1}^{n} A_i$, $n \in \mathbb{N}$. Then $A_n' \nearrow \Omega$ and $A_n' = \biguplus_{i=1}^{n}(A_i \backslash A_{i-1}')$ is a disjoint decomposition of A_n' into sets of $\sigma(\mathcal{A})$.

ii) $\forall\, A \in \sigma(\mathcal{A})$,

$$\mu(A \cap A_n') \underset{\text{distr.}}{=} \mu\Bigl(\biguplus_{i=1}^{n} A \cap A_i \cap A_{i-1}'^{c}\Bigr) \underset{\sigma\text{-add.}}{=} \sum_{i=1}^{n} \mu(\underbrace{A \cap A_{i-1}'^{c}}_{\in\,\sigma(\mathcal{A})} \cap \underbrace{A_i}_{\in\,\mathcal{A}})$$

$$\underset{2)}{\overset{\mu(A_i)<\infty}{=}} \sum_{i=1}^{n} \nu(\underbrace{A \cap A_{i-1}'^{c}}_{\in\,\sigma(\mathcal{A})} \cap \underbrace{A_i}_{\in\,\mathcal{A}}) = \dots \underset{\text{backwards}}{\overset{\text{same}}{=}} \nu(A \cap A_n').$$

iii) Therefore, $\forall\, A \in \sigma(\mathcal{A})$, $\mu(A) = \mu(A \cap \Omega) \underset{\text{cont. below}}{\overset{\text{i)}}{=}} \lim_{n\to\infty} \mu(A \cap A_n') \underset{\text{ii)}}{=} \lim_{n\to\infty} \nu(A \cap A_n') = \dots \underset{\text{backwards}}{\overset{\text{same}}{=}} \nu(A)$. $\square$

By i), σ-finiteness on $\mathcal{A}$ can be replaced by the existence of $(A_i)_{i\in\mathbb{N}} \subseteq \mathcal{A}$, $A_i \nearrow \Omega$ with $\mu(A_i) < \infty$, $i \in \mathbb{N}$, (*exhausting sequence*). This trivially holds if μ is finite.

Example 2.29 (Product measure)

For $j = 1, \dots, d$, let Ω_j be equipped with a σ-algebra $\mathcal{F}_j$. The product space $\prod_{j=1}^d \Omega_j$ can then be equipped with the product-σ-algebra $\bigotimes_{j=1}^d \mathcal{F}_j \underset{\text{P. 2.16}}{=} \sigma(\prod_{j=1}^d A_j : A_j \in \mathcal{F}_j \;\forall\, j)$. If μ_j is a σ-finite measure on $(\Omega_j, \mathcal{F}_j) \;\forall\, j$, then

$$\Bigl(\prod_{j=1}^d \mu_j\Bigr)\Bigl(\prod_{j=1}^d A_j\Bigr) := \prod_{j=1}^d \mu_j(A_j), \quad \prod_{j=1}^d A_j \in \bigotimes_{j=1}^d \mathcal{F}_j,$$

is the *product measure* on $(\prod_{j=1}^d \Omega_j, \bigotimes_{j=1}^d \mathcal{F}_j)$; by P. 2.28, it suffices to define $\prod_{j=1}^d \mu_j$ on the π-system $\mathcal{A} = \{\prod_{j=1}^d A_j : A_j \in \mathcal{F}_j \;\forall\, j\}$.

The product measure is indeed a measure since:

i) $\mu_j : \mathcal{F}_j \to [0, \infty]$, $j = 1, \dots, d \Rightarrow \prod_{j=1}^d \mu_j : \bigotimes_{j=1}^d \mathcal{F}_j \to [0, \infty]$.

ii) In $\prod_{j=1}^d \Omega_j$, $\prod_{j=1}^d A_j = \emptyset$ iff $\exists\, k \in \{1, \dots, d\} : A_k = \emptyset$. As such, if $A_k = \emptyset$ for at least one k, then

$$\Bigl(\prod_{j=1}^d \mu_j\Bigr)(\emptyset) = \Bigl(\prod_{j=1}^d \mu_j\Bigr)\Bigl(\prod_{j=1}^d A_j\Bigr) = \prod_{j=1}^d \mu_j(A_j) \underset{\mu_k(A_k)=0}{=} 0.$$

iii) If $\{A_i\}_{i\in\mathbb{N}} \subseteq \bigotimes_{j=1}^d \mathcal{F}_j$, $A_i \cap A_j = \emptyset\ \forall\, i \neq j$, then $A_i = \prod_{j=1}^d A_{i,j}$ for $A_{i,j} \in \mathcal{F}_j$ $\forall\, i, j$. With

$$\biguplus_{i=1}^{\infty} A_i = \biguplus_{i=1}^{\infty} \prod_{j=1}^{d} A_{i,j} \underset{(*)}{=} \prod_{j=1}^{d} \biguplus_{i=1}^{\infty} A_{i,j}$$

we have

$$\Bigl(\prod_{j=1}^{d} \mu_j\Bigr)\Bigl(\biguplus_{i=1}^{\infty} A_i\Bigr) \underset{(*)}{=} \Bigl(\prod_{j=1}^{d} \mu_j\Bigr)\Bigl(\prod_{j=1}^{d} \biguplus_{i=1}^{\infty} A_{i,j}\Bigr) \underset{\text{def.}}{=} \prod_{j=1}^{d} \mu_j\Bigl(\biguplus_{i=1}^{\infty} A_{i,j}\Bigr)$$

$$\underset{\sigma\text{-add.}}{\overset{\mu_j}{=}} \prod_{j=1}^{d} \sum_{i=1}^{\infty} \mu_j(A_{i,j}) \underset{\text{terms} \geq 0}{\overset{\text{multiply out}}{=}} \sum_{i=1}^{\infty} \prod_{j=1}^{d} \mu_j(A_{i,j})$$

$$\underset{\text{def.}}{=} \sum_{i=1}^{\infty} \Bigl(\prod_{j=1}^{d} \mu_j\Bigr)\Bigl(\prod_{j=1}^{d} A_{i,j}\Bigr) \underset{\text{def. } A_i}{=} \sum_{i=1}^{\infty} \Bigl(\prod_{j=1}^{d} \mu_j\Bigr)(A_i).$$

2.4 Null sets

Definition 2.30 (Null set, a.e., a.s., completeness)
If $(\Omega, \mathcal{F}, \mu)$ is a measure space, every $N \in \mathcal{F} : \mu(N) = 0$ is a *(μ-)null set*. If a statement holds $\forall\, \omega \in \Omega \backslash N$ for a null set N, it holds *(μ-)almost everywhere (a.e.)*, or, if μ is a probability measure, the statement holds *(μ-)almost surely (a.s.)*. If $\mathcal{F}$ contains all subsets of null sets, μ is a *complete* measure.

If a statement holds $\forall\, \omega \in \Omega$, one says it holds "everywhere", "surely" (or "pointwise"). Whether a statement holds everywhere/surely or only a.e./a.s. typically does not matter as every (in)equality involving measures holds irrespectively of changes on null sets.

Question: What collection of null sets is still a null set?

Lemma 2.31 (Countable union of null sets)
A countable union of null sets in $\mathcal{F}$ is a null set in $\mathcal{F}$.

Proof. If $\{N_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$ are null sets, then $\bigcup_{i=1}^{\infty} N_i \in \mathcal{F}$ and $0 \le \mu(\bigcup_{i=1}^{\infty} N_i) \underset{\sigma\text{-subadd.}}{\le} \sum_{i=1}^{\infty} \mu(N_i) = \sum_{i=1}^{\infty} 0 = 0$. $\square$

Question: If μ is not complete, can we extend it to a complete measure $\bar{\mu}$ so that all subsets N' of null sets N are measurable sets, with $\bar{\mu}(N') = 0$?

Theorem 2.32 (Completion of a σ-algebra and measure)
Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $\mathcal{N} = \{N \in \mathcal{F} : \mu(N) = 0\}$. Then

1) $\bar{\mathcal{F}} := \{A \cup N' : A \in \mathcal{F},\ N' \subseteq N \text{ for } N \in \mathcal{N}\}$ is a σ-algebra on Ω, the *completion* of $\mathcal{F}$.
2) $\bar{\mu}(A \cup N') := \mu(A)\ \forall\, A \in \mathcal{F},\ \forall\, N' \subseteq N$ for $N \in \mathcal{N}$ uniquely extends μ to a complete measure on $\bar{\mathcal{F}}$.

Proof.

1) i) $\Omega = \Omega \cup \emptyset \in \bar{\mathcal{F}}$;

ii) $\bar{A} \in \bar{\mathcal{F}} \Rightarrow \bar{A} = A \cup N'$ for some $A \in \mathcal{F}$, $N \in \mathcal{N} : N \supseteq N'$. Wlog, assume $A \cap N = \emptyset$; otherwise consider $N' \leftarrow N' \backslash A$ and $N \leftarrow N \backslash A \in \mathcal{N}$. Then $\bar{A} = A \cup N' = A \cup \emptyset \cup \emptyset \cup N' \underset{N' \subseteq N}{\overset{A \cap N = \emptyset}{=}} (A \cap N^c) \cup (A \cap N') \cup (N \cap N^c) \cup (N \cap N') \underset{\text{distr.}}{=} (A \cup N) \cap (N^c \cup N')$, so that $\bar{A}^c = ((A \cup N) \cap (N^c \cup N'))^c \underset{\text{De Morgan}}{=}$

$$\underbrace{(A \cup N)^c}_{\in \mathcal{F} \text{ as } A, N \in \mathcal{F}} \cup \underbrace{(N \backslash N')}_{\subseteq N \in \mathcal{N}} \in \bar{\mathcal{F}}.$$

iii) $\{\bar{A}_i\}_{i\in\mathbb{N}} \subseteq \bar{F} \Rightarrow \bar{A}_i = A_i \cup N_i'$ for $A_i \in \mathcal{F}$, $N_i \in \mathcal{N} : N_i \supseteq N_i'\ \forall\, i \in \mathbb{N} \Rightarrow$
$\bigcup_{i=1}^{\infty} \bar{A}_i = \bigcup_{i=1}^{\infty}(A_i \cup N_i') = \underbrace{(\bigcup_{i=1}^{\infty} A_i)}_{\in \mathcal{F}} \cup \underbrace{(\bigcup_{i=1}^{\infty} N_i')}_{\subseteq \bigcup_{i=1}^{\infty} N_i \underset{N_i \in \mathcal{F}}{\in} \mathcal{N}} \in \bar{\mathcal{F}}$.

2) ▪ $\bar{\mu}$ is well-defined on $\bar{\mathcal{F}}$ since for $A_1 \cup N_1' = A_2 \cup N_2'$, we have $A_1 \subseteq A_1 \cup N_1' = A_2 \cup N_2'$ and so $\mu(A_1) \le \mu(A_2) + 0$, and likewise $\mu(A_2) \le \mu(A_1)$, so $\mu(A_1) = \mu(A_2)$ and thus $\bar{\mu}(A_1) = \mu(A_1) = \mu(A_2) = \bar{\mu}(A_2)$.

▪ $\bar{\mu}$ is a measure on $\bar{F}$ (by definition, we already know that $\bar{\mu}|_{\mathcal{F}} = \mu$):

i) $\mu : \mathcal{F} \to [0, \infty] \Rightarrow \bar{\mu} : \bar{\mathcal{F}} \to [0, \infty]$;

ii) $\emptyset = \emptyset \cup \emptyset \in \bar{\mathcal{F}} \Rightarrow \bar{\mu}(\emptyset) = \bar{\mu}(\emptyset \cup \emptyset) = \mu(\emptyset) = 0$; and

iii) $\forall\, i \in \mathbb{N}$, let $\bar{A}_i = A_i \cup N_i'$ for $A_i \in \mathcal{F}$, $N_i \in \mathcal{N} : N_i \supseteq N_i'$ and $\bar{A}_i \cap \bar{A}_j = \emptyset\ \forall\, i \ne j$. Then $\bar{\mu}(\biguplus_{i=1}^{\infty} \bar{A}_i) \underset{\text{def. } \bar{A}_i}{=} \bar{\mu}((\biguplus_{i=1}^{\infty} A_i) \cup (\biguplus_{i=1}^{\infty} N_i')) \underset{(*)}{=} \mu(\biguplus_{i=1}^{\infty} A_i) \underset{\sigma\text{-add.}}{=} \sum_{i=1}^{\infty} \mu(A_i) \underset{\text{def. } \bar{\mu}}{=} \sum_{i=1}^{\infty} \bar{\mu}(A_i \cup N_i') \underset{\text{def. } \bar{A}_i}{=} \sum_{i=1}^{\infty} \bar{\mu}(\bar{A}_i)$, where $(*)$ holds since $\biguplus_{i=1}^{\infty} A_i \in \mathcal{F}$, $\biguplus_{i=1}^{\infty} N_i' \subseteq \bigcup_{i=1}^{\infty} N_i \underset{\text{L. 2.31}}{\in} \mathcal{N}$.

▪ Uniqueness: Suppose $\exists$ a complete measure $\bar{\nu}$ on $\bar{\mathcal{F}} : \bar{\nu}(A \cup N') = \mu(A)$ $\forall\, A \in \mathcal{F}, \forall\, N' \subseteq N, \forall\, N \in \mathcal{N}$. Then $\bar{\nu}(A \cup N') \underset{\text{subadd.}}{\le} \bar{\nu}(A) + \bar{\nu}(N') = \bar{\nu}(A) + 0 \underset{\bar{\nu} \text{ extends } \mu}{=} \mu(A) \underset{\bar{\mu} \text{ extends } \mu}{=} \bar{\mu}(A) \underset{\text{mon.}}{\le} \bar{\mu}(A \cup N')$. Likewise $\bar{\mu}(A \cup N') \le \bar{\nu}(A \cup N')$, so $\bar{\mu}(A \cup N') = \bar{\nu}(A \cup N')$. □

2.5 Construction of measures

Question: How can we construct measures on general measurable spaces?

- **Main idea:** Start from a 'premeasure' μ_0 (a notion of measure) on a sufficiently simple system of sets $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ and extend μ_0 via an 'outer measure' μ^* to a measure μ on $\sigma(\mathcal{A})$. Results of this form are referred to as *Carathéodory extension theorem*, attributed to Constantin Carathéodory (1873–1950).
- To be consistent (not leading to contradictions), $\mathcal{A}$ should have some structure:
  - Folland (1999, T. 1.13, T. 1.14), Durrett (2019, T. A.1.1, T. A.1.3): $\mathcal{A}$ is an algebra ($\Rightarrow$ Hahn–Kolmogorov theorem)
  - Bauer (2001, T. 5.1, T. 5.3, T. 5.6), Wikipedia: $\mathcal{A}$ is a ring ($\Rightarrow$ Carathéodory's extension theorem)
  - Klenke (2008, T. 1.53), Schilling (2006, T. 6.1): $\mathcal{A}$ is a semiring ($\Rightarrow$ Carathéodory's extension theorem; most general: Klenke)
- The more structure (from semirings to rings to algebras) the easier it is to prove the extension theorem, but typically the harder it is to apply it for a specific construction as more properties need to be verified (but note that one also has less properties available to work with). We consider semirings.

Premeasures have the properties of measures but are defined on a smaller domain.

Definition 2.33 (Premeasure, σ-finite)

Let $\mathcal{A}$ be a semiring on Ω. A *premeasure* μ_0 on $\mathcal{A}$ satisfies

i) $\mu_0 : \mathcal{A} \to [0, \infty]$;

ii) $\mu_0(\emptyset) = 0$; and

iii) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$, $A_i \cap A_j = \emptyset \ \forall\, i \neq j$, and $\biguplus_{i=1}^{\infty} A_i \in \mathcal{A} \Rightarrow \mu_0(\biguplus_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu_0(A_i)$ (*σ-additivity*).

If $\Omega = \bigcup_{i=1}^{\infty} A_i$ for $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} : \mu_0(A_i) < \infty \ \forall\, i$, then μ_0 is *σ-finite*.

Outer measures are used to approximate volumes from 'above' (the 'outside').

Definition 2.34 (Outer measure)

An *outer measure* $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ satisfies

i) $\mu^*(\emptyset) = 0$;

ii) $A, B \subseteq \Omega : A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$ (monotonicity); and

iii) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{P}(\Omega) \Rightarrow \mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ (σ-subadditivity).

Steps of the main idea: Start with a premeasure μ_0 on a semiring $\mathcal{A}$.

1) Show that the premeasure μ_0 defined on the semiring $\mathcal{A}$ induces an outer measure μ^*.
2) Show that μ^* extends μ_0 to $\mathcal{P}(\Omega)$, i.e. $\mu^*|_{\mathcal{A}} = \mu_0|_{\mathcal{A}}$.
3) Show that $\mathcal{A} \subseteq \mathcal{A}^*$ where $\mathcal{A}^*$ is the family of *Carathéodory-measurable sets*
$$\mathcal{A}^* := \{A \subseteq \Omega : \mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c) \;\forall\, B \subseteq \Omega\}.$$
4) Show that $\mathcal{A}^*$ is a σ-algebra on Ω ($\underset{3)}{\Rightarrow} \sigma(\mathcal{A}) \subseteq \mathcal{A}^*$) and μ^* is a measure on $\mathcal{A}^*$.
5) Then $\mu := \mu^*|_{\sigma(\mathcal{A})}$ is a measure. Show that $\mu|_{\mathcal{A}} = \mu_0|_{\mathcal{A}}$, i.e. μ extends μ_0 to $\sigma(\mathcal{A})$.
6) Show that if μ_0 is σ-finite on $\mathcal{A}$, the extension μ of μ_0 to $\sigma(\mathcal{A})$ is unique.

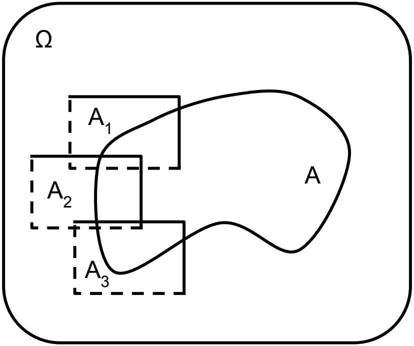
Theorem 2.35 (Carathéodory extension theorem)
Let $\mathcal{A}$ be a semiring on Ω and μ_0 a σ-finite premeasure on $\mathcal{A}$. Then μ_0 has a unique extension to a σ-finite measure μ on $\sigma(\mathcal{A})$.

Proof. Step 1) is instructive. Step 2) is more work (partly because we start with semirings). Steps 3)–5) are the main parts of the theorem. Step 6) is uniqueness.

1) Let $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ be defined by

$$\mu^*(A) := \inf_{\substack{A \subseteq \bigcup_{i=1}^{\infty} A_i, \\ A_i \in \mathcal{A}}} \sum_{i=1}^{\infty} \mu_0(A_i),$$

where $\inf \emptyset = \infty$. In particular, if $\nexists$ covering of A by a countable union of sets from $\mathcal{A}$, then $\mu^*(A) = \infty$, and if $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}$ such that $\mu_0(A_i) = 0$ $\forall\, i$, then $\mu^*(A) = 0$.



We show that μ^* is an outer measure:

i) $A_i = \emptyset \;\forall\, i \in \mathbb{N} \Rightarrow \mu^*(\emptyset) = 0$;

ii) Let $A, B \subseteq \Omega : A \subseteq B$. If $\bigcup_{i=1}^{\infty} A_i \supseteq B$ then $\bigcup_{i=1}^{\infty} A_i \supseteq A \Rightarrow \mu^*(A) \le \mu^*(B)$ (as the inf over a larger number of coverings can only be smaller);

iii) Let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{P}(\Omega)$. Wlog, assume $\mu^*(A_i) < \infty \;\forall\, i \in \mathbb{N}$; otherwise σ-subadditivity trivially holds; "$\infty \underset{\text{ii)}}{\le} \infty$". By definition of the infimum, $\forall\, \varepsilon > 0$, $\exists$ a covering $\bigcup_{k=1}^{\infty} A_{i,k} \supseteq A_i$ with $\sum_{k=1}^{\infty} \mu_0(A_{i,k}) \underset{(*)}{\le} \mu^*(A_i) + \varepsilon/2^i$, $i \in \mathbb{N}$. Therefore $\bigcup_{i,k=1}^{\infty} A_{i,k} \supseteq \bigcup_{i=1}^{\infty} A_i$ is a covering and thus

$$\mu^*\Big(\bigcup_{i=1}^\infty A_i\Big) \underset{\text{all coverings}}{\overset{\mu^*\text{ smallest over}}{\leq}} \sum_{i,k=1}^\infty \mu_0(A_{i,k}) \underset{(*)}{\leq} \sum_{i=1}^\infty (\mu^*(A_i)+\varepsilon/2^i) = \varepsilon + \sum_{i=1}^\infty \mu^*(A_i).$$

As $\varepsilon > 0$ was arbitrary, we obtain (for $\varepsilon \to 0+$) $\mu^*(\bigcup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty \mu^*(A_i)$.

2) To show that $\mu^*|_{\mathcal{A}} = \mu_0|_{\mathcal{A}}$, we proceed in several steps:

2.1) We show that $\mathcal{A}_\uplus := \{\biguplus_{i=1}^n A_i : A_i \in \mathcal{A},\ n \in \mathbb{N}\}$ is a ring, the *ring generated by $\mathcal{A}$*. To this end, we first cover two auxiliary results:

a) For any $A, B_1, \dots, B_m \in \mathcal{A}$ there are disjoint $C_1, \dots, C_k \in \mathcal{A}$: $A \setminus \bigcup_{j=1}^m B_j = \biguplus_{l=1}^k C_l$. Proof by induction:

$m = 1$: $A, B_1 \in \mathcal{A} \underset{\text{D. 2.3 iii)}}{\Rightarrow} A \setminus B_1$ is a finite disjoint union of sets in $\mathcal{A}$.

$m \Rightarrow m+1$: $A \setminus \bigcup_{j=1}^{m+1} B_j = (A \setminus \bigcup_{j=1}^m B_j) \setminus B_{m+1} \underset{\text{hypo.}}{=} (\biguplus_{l=1}^k C_l) \setminus B_{m+1}$ $\underset{\text{distr.}}{=} \biguplus_{l=1}^k (C_l \setminus B_{m+1})$. By D. 2.3 iii), each $C_l \setminus B_{m+1}$ is a finite disjoint union of sets in $\mathcal{A}$, and so is $\biguplus_{l=1}^k (C_l \setminus B_{m+1})$ and thus $A \setminus \bigcup_{j=1}^{m+1} B_j$.

b) $\mathcal{A}_\uplus$ is a π-system: For $A = \biguplus_{i=1}^n A_i \in \mathcal{A}_\uplus$ and $B = \biguplus_{j=1}^m B_j \in \mathcal{A}_\uplus$, we have $A \cap B = \biguplus_{i,j=1}^{n,m} (A_i \cap B_j)$ for $A_i \cap B_j \underset{\text{D. 2.3 ii)}}{\in} \in \mathcal{A}$, so $A \cap B \in \mathcal{A}_\uplus$.

We can now show that $\mathcal{A}_\uplus$ is a ring:

i) $\emptyset \in \mathcal{A} \Rightarrow \emptyset \in \mathcal{A}_\uplus$

iii) For $A=\biguplus_{i=1}^n A_i\in\mathcal{A}_\uplus$ and $B=\biguplus_{j=1}^m B_j\in\mathcal{A}_\uplus$, we have

$$A\backslash B=\Bigl(\biguplus_{i=1}^n A_i\Bigr)\cap\Bigl(\biguplus_{j=1}^m B_j\Bigr)^c \underset{\text{De Morgan}}{=} \Bigl(\biguplus_{i=1}^n A_i\Bigr)\cap\bigcap_{j=1}^m B_j^c$$

$$\underset{\text{distr.}}{=} \biguplus_{i=1}^n\Bigl(A_i\cap\bigcap_{j=1}^m B_j^c\Bigr) \underset{\text{De Morgan}}{=} \biguplus_{i=1}^n\Bigl(A_i\cap\Bigl(\bigcup_{j=1}^m B_j\Bigr)^c\Bigr)$$

$$=\biguplus_{i=1}^n\Bigl(A_i\backslash\bigcup_{j=1}^m B_j\Bigr) \underset{\text{a)}}{=} \biguplus_{i=1}^n\biguplus_{l=1}^{k_i} C_{i,l} \underset{\text{def.}}{\in} \mathcal{A}_\uplus$$

for some disjoint $C_{i,l}\in\mathcal{A}\ \forall i\in\{1,\dots,n\}, l\in\{1,\dots,k_i\}$.

ii) $A,B\in\mathcal{A}_\uplus\Rightarrow A\cup B=(A\backslash B)\uplus(A\cap B)\uplus(B\backslash A)\overset{\text{iii)}}{\underset{\text{b)}}{\in}}\mathcal{A}_\uplus$.

2.2) Extend μ_0 from $\mathcal{A}$ to its generated ring $\mathcal{A}_\uplus$ by $\mu_0(\biguplus_{i=1}^n A_i):=\sum_{i=1}^n\mu_0(A_i)$. This extension is unique as long as it is well-defined. To show well-definedness, let $\biguplus_{i=1}^n A_i=\biguplus_{j=1}^m B_j$, $A_i,B_j\in\mathcal{A}$. Then $A_i=A_i\cap(\biguplus_{j=1}^m B_j)\underset{\text{distr.}}{=}\biguplus_{j=1}^m(A_i\cap B_j)$. By additivity of μ_0 on $\mathcal{A}$, $\mu_0(A_i)=\sum_{j=1}^m\mu_0(A_i\cap B_j)$. Hence, $\sum_{i=1}^n\mu_0(A_i)=\sum_{i=1}^n\sum_{j=1}^m\mu_0(A_i\cap B_j)\overset{\text{switch role of}}{\underset{A_i,\,B_j}{=}}\dots=\sum_{j=1}^m\mu_0(B_j)$.

2.3) We now show that μ_0 is σ-additive on $\mathcal{A}_\uplus$, which implies that μ_0 is a premeasure on $\mathcal{A}_\uplus$.

- Let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}_\uplus$, $A_i \cap A_j = \emptyset \;\forall\, i \neq j$, and $\biguplus_{i=1}^\infty A_i \in \mathcal{A}_\uplus \underset{\text{def. } \mathcal{A}_\uplus}{\Rightarrow} \forall\, i \in \mathbb{N}\; \exists\, n_i \in \mathbb{N} : (A_{i,j})_{j=1}^{n_i} \subseteq \mathcal{A}$ and $A_i = \biguplus_{j=1}^{n_i} A_{i,j}$, where $A_{i,j} \cap A_{k,l} = \emptyset$ $\forall\, i \neq k, \forall\, j, l$. By enumerating $(A_{i,j})$, $\exists$ pairwise disjoint $(\tilde{A}_k)_{k=1}^\infty \subseteq \mathcal{A}$ and a sequence of integers $0 =: k_0 < k_1 < k_2 \dots$ such that $A_i = \biguplus_{j=k_{i-1}+1}^{k_i} \tilde{A}_j$, $i \in \mathbb{N}$. This implies that $\biguplus_{i=1}^\infty A_i = \biguplus_{i=1}^\infty \biguplus_{j=k_{i-1}+1}^{k_i} \tilde{A}_j$.
- Since $\biguplus_{i=1}^\infty A_i \in \mathcal{A}_\uplus$, there must exist $n \in \mathbb{N}$ and a partition $\{I_i\}_{i=1}^n$ of $\mathbb{N}$ such that $\biguplus_{i=1}^\infty A_i = \biguplus_{i=1}^n \biguplus_{j\in I_i} \tilde{A}_j$ for $\biguplus_{j\in I_i} \tilde{A}_j \in \mathcal{A}$.
- Therefore,

$$\mu_0\Big(\biguplus_{i=1}^\infty A_i\Big) = \mu_0\Big(\biguplus_{i=1}^n \biguplus_{j\in I_i} \tilde{A}_j\Big) \underset{\text{by } 2.2)}{\overset{\text{def. } \mu_0|_{\mathcal{A}_\uplus}}{=}} \sum_{i=1}^n \mu_0\Big(\biguplus_{j\in I_i} \tilde{A}_j\Big)$$

$$\underset{\text{by D. 2.33 iii)}}{\overset{\mu_0|_{\mathcal{A}}\ \sigma\text{-add.}}{=}} \sum_{i=1}^n \sum_{j\in I_i} \mu_0(\tilde{A}_j) \underset{\text{partitions } \mathbb{N}}{\overset{\{I_i\}_{i=1}^n}{=}} \sum_{i=1}^\infty \sum_{j=k_{i-1}+1}^{k_i} \mu_0(\tilde{A}_j) \underset{\text{by } 2.2)}{\overset{\text{def. } \mu_0|_{\mathcal{A}_\uplus}}{=}} \sum_{i=1}^\infty \mu_0(A_i).$$

2.4) One can show that the premeasure μ_0 on the ring $\mathcal{A}_\uplus$ is monotone and σ-subadditive; due to the ring properties, this works similarly as the proofs of P. 2.27 2) and 3). We can now show that $\mu^*|_{\mathcal{A}} = \mu_0|_{\mathcal{A}}$.

$\forall\, \{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A} : \bigcup_{i=1}^\infty A_i \supseteq A \in \mathcal{A}$, we have $\mu_0(A) \underset{\text{distr.}}{=} \mu_0(\bigcup_{i=1}^\infty (A_i \cap A)) \underset{\sigma\text{-subadd.}}{\le} \sum_{i=1}^\infty \mu_0(A_i \cap A) \underset{\text{def.}}{=} \lim_{n\to\infty} \sum_{i=1}^n \mu_0(A_i \cap A) \underset{\text{mon.}}{\le} \lim_{n\to\infty} \sum_{i=1}^n \mu_0(A_i) \underset{\text{def.}}{=} \sum_{i=1}^\infty \mu_0(A_i)$. Hence also for the infimum over all such $\{A_i\}_{i\in\mathbb{N}}$, we have

$$\mu_0(A) \underset{\forall\, \{A_i\}_{i\in\mathbb{N}}}{\le} \inf_{\substack{A\subseteq\bigcup_{i=1}^\infty A_i,\\ A_i\in\mathcal{A}}} \sum_{i=1}^\infty \mu_0(A_i) \underset{\text{def.}}{=} \mu^*(A).$$

And the cover $\{A, \emptyset, \emptyset, \dots\} \subseteq \mathcal{A}$ implies that $\mu^*(A) \underset{\text{inf}}{\le} \mu_0(A)$, so $\mu^*(A) = \mu_0(A)\ \forall\, A \in \mathcal{A}$.

3) We show that $\mathcal{A} \subseteq \mathcal{A}^*$, i.e. $\mu^*(B) = \mu^*(B\cap A) + \mu^*(B\cap A^c)\ \forall\, A \in \mathcal{A}, \forall\, B \subseteq \Omega$. Let $A \in \mathcal{A}$, $B \subseteq \Omega$. By definition of the infimum, $\forall\, \varepsilon > 0$, $\exists\, \{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A} : \bigcup_{i=1}^\infty A_i \supseteq B$ and $\sum_{i=1}^\infty \mu_0(A_i) \le \mu^*(B) + \varepsilon$. If $E_i := A_i \cap A \underset{\text{by D. 2.3 ii)}}{\overset{\mathcal{A}\ \pi\text{-sys.}}{\in}} \mathcal{A}$, $i \in \mathbb{N}$, then $A_i \backslash A = A_i \backslash E_i \underset{(*)}{\overset{\mathcal{A}\text{ semiring}}{=}} \biguplus_{k=1}^{n_i} C_{i,k}$ for some $n_i \in \mathbb{N}$ and disjoint $C_{i,1}, \dots, C_{i,n_i} \in \mathcal{A}$. Hence $A_i = E_i \uplus A_i \backslash E_i = E_i \uplus \biguplus_{k=1}^{n_i} C_{i,k}$ and

$$B \cap A \underset{B\,\subseteq\,\bigcup_{i=1}^\infty A_i}{\subseteq} \left(\bigcup_{i=1}^\infty A_i\right) \cap A \underset{\text{distr.}}{=} \bigcup_{i=1}^\infty (A_i \cap A) \underset{\text{def.}}{=} \bigcup_{i=1}^\infty E_i,$$

$$B \cap A^c \underset{B \subseteq \bigcup_{i=1}^\infty A_i}{\subseteq} \Big(\bigcup_{i=1}^\infty A_i\Big) \cap A^c \underset{\text{distr.}}{=} \bigcup_{i=1}^\infty (\underbrace{A_i \cap A^c}_{= A_i \setminus A}) \underset{(*)}{=} \bigcup_{i=1}^\infty \biguplus_{k=1}^{n_i} C_{i,k}.$$

We thus obtain that

$$\begin{aligned}
&\mu^*(B \cap A) + \mu^*(B \cap A^c) \\
&\underset{\text{by D. 2.34 ii)}}{\overset{\mu^* \text{ mon.}}{\leq}} \mu^*\Big(\bigcup_{i=1}^\infty E_i\Big) + \mu^*\Big(\bigcup_{i=1}^\infty \biguplus_{k=1}^{n_i} C_{i,k}\Big) \underset{\text{by D. 2.34 iii)}}{\overset{\mu^* \ \sigma\text{-subadd.}}{\leq}} \sum_{i=1}^\infty \mu^*(E_i) + \sum_{i=1}^\infty \sum_{k=1}^{n_i} \mu^*(C_{i,k}) \\
&\underset{\mu^*|_{\mathcal{A}} \underset{2)}{=} \mu_0|_{\mathcal{A}}}{=} \sum_{i=1}^\infty \mu_0(E_i) + \sum_{i=1}^\infty \sum_{k=1}^{n_i} \mu_0(C_{i,k}) = \sum_{i=1}^\infty \Big(\mu_0(E_i) + \sum_{k=1}^{n_i} \mu_0(C_{i,k})\Big) \\
&\underset{\text{by D. 2.33 iii)}}{\overset{\mu_0 \text{ add.}}{=}} \sum_{i=1}^\infty \mu_0\Big(E_i \uplus \biguplus_{k=1}^{n_i} C_{i,k}\Big) = \sum_{i=1}^\infty \mu_0(A_i) \leq \mu^*(B) + \varepsilon.
\end{aligned}$$

Letting $\varepsilon \to 0+$, we obtain $\mu^*(B \cap A) + \mu^*(B \cap A^c) \leq \mu^*(B)$. Trivially,

$$\mu^*(B) = \mu^*((B \cap A) \uplus (B \cap A^c)) \underset{\text{by D. 2.34 iii)}}{\overset{\mu^* \text{ subadd.}}{\leq}} \mu^*(B \cap A) + \mu^*(B \cap A^c),$$

so we have equality and thus obtain that $A \underset{\text{def.}}{\in} \mathcal{A}^*$, so $\mathcal{A} \subseteq \mathcal{A}^*$.

4) We show that $\mathcal{A}^*$ is a σ-algebra (by showing it is a Dynkin system and a π-system; see P. 2.18 3)) and that μ^* is a measure on $\mathcal{A}^*$.

4.1) π**-system:** Let $A_1, A_2 \in \mathcal{A}^*$ and $B \subseteq \Omega$. Then $\underbrace{(A_1 \cap A_2)^c}_{\text{contained in at most 1}} \underset{\text{Venn}}{=} \underbrace{(A_1^c \cap A_2)}_{\text{strictly in } A_2} \uplus \underbrace{(A_1 \cap A_2^c)}_{\text{strictly in } A_1} \uplus \underbrace{(A_1^c \cap A_2^c)}_{\text{contained in neither}}$ implies that

$$\begin{aligned}
\mu^*(B) &\underset{\text{by D. 2.34 iii)}}{\overset{\mu^* \text{ subadd.}}{\leq}} \mu^*(B \cap (A_1 \cap A_2)) + \mu^*(B \cap (A_1 \cap A_2)^c) \\
&= \mu^*(B \cap A_1 \cap A_2) \\
&\quad + \mu^*((B \cap A_1^c \cap A_2) \uplus (B \cap A_1 \cap A_2^c) \uplus (B \cap A_1^c \cap A_2^c)) \\
&\underset{\text{by D. 2.34 iii)}}{\overset{\mu^* \text{ subadd.}}{\leq}} \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap A_1^c \cap A_2) + \mu^*(B \cap A_1 \cap A_2^c) \\
&\quad + \mu^*(B \cap A_1^c \cap A_2^c) \\
&\underset{\text{reorder}}{=} \mu^*(B \cap A_2 \cap A_1) + \mu^*(B \cap A_2 \cap A_1^c) \\
&\quad + \mu^*(B \cap A_2^c \cap A_1) + \mu^*(B \cap A_2^c \cap A_1^c) \\
&\underset{B \cap A_2,\, B \cap A_2^c \subseteq \Omega}{\overset{A_1 \in \mathcal{A}^*}{=}} \mu^*(B \cap A_2) + \mu^*(B \cap A_2^c) \underset{B \subseteq \Omega}{\overset{A_2 \in \mathcal{A}^*}{=}} \mu^*(B),
\end{aligned}$$

so $A_1 \cap A_2$ satisfies $\mu^*(B) = \mu^*(B \cap (A_1 \cap A_2)) + \mu^*(B \cap (A_1 \cap A_2)^c)$ $\forall\, B \subseteq \Omega$ and thus $A_1 \cap A_2 \in \mathcal{A}^*$.

4.2) **Dynkin system:**

i) $\forall\, B \subseteq \Omega$, we have $\mu^*(B) \underset{\text{by def.}}{\overset{\mu^*(\emptyset)=0}{=}} \mu^*(B \cap \Omega) + \mu^*(B \cap \emptyset)$, so $\Omega \in \mathcal{A}^*$.

ii) By definition of $\mathcal{A}^*$, $A \in \mathcal{A}^*$ iff $A^c \in \mathcal{A}^*$, so $A \in \mathcal{A}^* \Rightarrow A^c \in \mathcal{A}^*$.

iii) Let $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}^*$, $A_i \cap A_j = \emptyset\ \forall\, i \neq j$. With $\tilde{A}_n := \biguplus_{i=1}^n A_i \underset{\text{De Morgan}}{=} (\bigcap_{i=1}^n A_i^c)^c \overset{\text{ii)}}{\underset{4.1)}{\in}} \mathcal{A}^*$, $n \in \mathbb{N}$, and $\tilde{A}_\infty := \biguplus_{i=1}^\infty A_i$, we have $\forall\, B \subseteq \Omega$,

$$\mu^*(B \cap \tilde{A}_{n+1}) \underset{\tilde{A}_n\, \in\, \mathcal{A}^*}{=} \mu^*((B \cap \tilde{A}_{n+1}) \cap \tilde{A}_n) + \mu^*((B \cap \tilde{A}_{n+1}) \cap \tilde{A}_n^c)$$

$$\overset{\tilde{A}_{n+1}\, \cap\, \tilde{A}_n\, =\, \tilde{A}_n}{\underset{\tilde{A}_{n+1}\, \cap\, \tilde{A}_n^c\, =\, A_{n+1}}{=}} \mu^*(B \cap \tilde{A}_n) + \mu^*(B \cap A_{n+1}), \quad n \in \mathbb{N}.$$

This implies, inductively, that $\mu^*(B \cap \tilde{A}_n) \underset{(*)}{=} \sum_{i=1}^n \mu^*(B \cap A_i)$, so

$$\mu^*(B \cap \tilde{A}_\infty) + \mu^*(B \cap \tilde{A}_\infty^c) \overset{\mu^*\text{ subadd.}}{\underset{\text{by D. 2.34 iii)}}{\geq}} \mu^*((B \cap \tilde{A}_\infty) \uplus (B \cap \tilde{A}_\infty^c))$$

$$= \mu^*(B) \underset{\tilde{A}_n\, \in\, \mathcal{A}^*}{=} \mu^*(B \cap \tilde{A}_n) + \mu^*(B \cap \tilde{A}_n^c)$$

$$\overset{\tilde{A}_n\, \subseteq\, \tilde{A}_\infty}{\underset{\mu^*\text{ mon. by def.}}{\geq}} \mu^*(B \cap \tilde{A}_n) + \mu^*(B \cap \tilde{A}_\infty^c) \underset{(*)}{=} \sum_{i=1}^n \mu^*(B \cap A_i) + \mu^*(B \cap \tilde{A}_\infty^c).$$

As this inequality holds $\forall\, n \in \mathbb{N}$, take the limit $n \to \infty$ to see that

$$\mu^*(B\cap\tilde{A}_\infty) + \mu^*(B\cap\tilde{A}_\infty^c) \geq \mu^*(B) \geq \sum_{i=1}^\infty \mu^*(B\cap A_i) + \mu^*(B\cap\tilde{A}_\infty^c)$$

$$\overset{\mu^*\ \sigma\text{-subadd.}}{\underset{\text{by D. 2.34 iii)}}{\geq}} \mu^*\left(B \cap \biguplus_{i=1}^\infty A_i\right) + \mu^*(B \cap \tilde{A}_\infty^c) \underset{\text{def. } \tilde{A}_\infty}{=} \mu^*(B \cap \tilde{A}_\infty) + \mu^*(B \cap \tilde{A}_\infty^c)$$

and thus equality, so $\tilde{A}_\infty \in \mathcal{A}^*$ and thus $\mathcal{A}^*$ is a Dynkin system. By 4.1), $\mathcal{A}^*$ is a σ-algebra. For $B = \tilde{A}_\infty$, we get σ-additivity of μ^*:

$$\mu^*\Big(\biguplus_{i=1}^\infty A_i\Big) = \mu^*(\tilde{A}_\infty) \underset{(*)}{=} \sum_{i=1}^\infty \mu^*(\tilde{A}_\infty \cap A_i) + \mu^*(\tilde{A}_\infty \cap \tilde{A}_\infty^c)$$
$$= \sum_{i=1}^\infty \mu^*\Big(\Big(\biguplus_{i=1}^\infty A_i\Big) \cap A_i\Big) + 0 = \sum_{i=1}^\infty \mu^*(A_i).$$

4.3) μ^* is a measure on the σ-algebra $\mathcal{A}^*$:

i) $\mu^* : \mathcal{P}(\Omega) \to [0,\infty] \Rightarrow \mu^* : \mathcal{A}^* \to [0,\infty]$;

ii) $\mu^*(\emptyset) \underset{\text{D. 2.34 i)}}{=} 0$; and

iii) μ^* is σ-additive on $\mathcal{A}^*$, see the last statement in 4) iii).

5) $\mathcal{A} \underset{3)}{\subseteq} \mathcal{A}^*$. By 4), $\mathcal{A}^*$ is a σ-algebra $\underset{\sigma\text{ smallest}}{\Rightarrow} \sigma(\mathcal{A}) \subseteq \mathcal{A}^* \underset{4)}{\Rightarrow} \mu := \mu^*|_{\sigma(\mathcal{A})}$ is a measure. And $\mu|_{\mathcal{A}} \underset{\text{def.}}{=} \mu^*|_{\mathcal{A}} \underset{2)}{=} \mu_0|_{\mathcal{A}}$, so μ extends μ_0.

6) By ass., μ_0 is σ-finite on $\mathcal{A} \underset{\text{def.}}{\Rightarrow} \mu^*$ and thus μ are σ-finite. Furthermore, as a semiring, $\mathcal{A}$ is a π-system $\underset{\text{P. 2.28}}{\Rightarrow}$ Another σ-finite measure ν on $\sigma(\mathcal{A})$ with $\nu|_{\mathcal{A}} = \mu|_{\mathcal{A}}$ must coincide with μ on $\sigma(\mathcal{A})$, so μ is unique on $\sigma(\mathcal{A})$. $\square$

By T. 2.32 we can always assume the extension μ to be complete, which we do.

2.6 Borel measures on $\mathbb{R}^d$

Question: How can we use Carathéodory's extension theorem to construct Borel measures on $\mathbb{R}^d$?

- At the core of the construction lie functions $F : \mathbb{R}^d \to \mathbb{R}$ that are right-continuous and increasing in a specific way.
- $F : \mathbb{R}^d \to \mathbb{R}$ is *right-continuous* if $F(\boldsymbol{x}) = \lim_{\boldsymbol{h}\to\boldsymbol{0}+} F(\boldsymbol{x}+\boldsymbol{h}) =: F(\boldsymbol{x}+)$ $\forall\, \boldsymbol{x} \in \mathbb{R}^d$.
- One also frequently utilizes F that are *grounded*, i.e. $\forall\, j \in \{1,\dots,d\}$ one has $\lim_{x_j\to-\infty} F(\boldsymbol{x}) = 0$. This will be crucial for constructing Borel probability measures on $\mathbb{R}^d$.

Definition 2.36 (d-increasing)
$F : \mathbb{R}^d \to \mathbb{R}$ is *d-increasing* if $\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F \geq 0 \;\forall\, \boldsymbol{a} \leq \boldsymbol{b}$, where the *$F$-volume* is

$$\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F := \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} F(\underbrace{a_1^{i_1} b_1^{1-i_1}}_{=\begin{cases} b_1, & i_1 = 0, \\ a_1, & i_1 = 1. \end{cases}}, \dots, a_d^{i_d} b_d^{1-i_d}).$$

- For $d=1$: $\Delta_{(a_1,b_1]}F = F(b_1)-F(a_1)$, i.e. $\Delta_{(a_1,b_1]}F \ge 0$ implies that F is *increasing* on $\mathbb{R}$ ($F \nearrow$) in the usual sense.
- For $d=2$: $\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F = \underbrace{F(b_1,b_2)-F(a_1,b_2)}_{=\Delta_{(a_1,b_1]}F(x_1,b_2)} - \big(\underbrace{F(b_1,a_2)-F(a_1,a_2)}_{=\Delta_{(a_1,b_1]}F(x_1,a_2)}\big)$ $= \Delta_{(a_2,b_2]}\Delta_{(a_1,b_1]}F$.

Question: How can we better understand F-volumes?

For $J \subseteq \{1,\dots,d\}$, let $\boldsymbol{x}_J = (x_j)_{j\in J}$ and $\boldsymbol{x}_{J^c} = (x_j)_{j\notin J}$. For $x \in \mathbb{R}$, let ${}_d x := (x,\dots,x) \in \mathbb{R}^d$. And for $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d$, let

$$\boldsymbol{x}_{J\leftarrow \boldsymbol{y}_J} = \begin{cases} x_j, & j \notin J, \\ y_j, & j \in J; \end{cases}$$

for $J=\{j\}$, we simply write $\boldsymbol{x}_{j\leftarrow y_j} = (x_1,\dots,x_{j-1},y_j,x_{j+1},\dots,x_d)$.

Lemma 2.37 (Understanding F-volumes)
Let $F:\mathbb{R}^d \to \mathbb{R}$ be d-increasing.

1) Then $\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F = \Delta_{(a_d,b_d]}\dots\Delta_{(a_1,b_1]}F$, or any permutation of first-order differences.
2) F-volumes are monotone, i.e. $(\boldsymbol{a},\boldsymbol{b}] \subseteq (\boldsymbol{c},\boldsymbol{d}]$ implies $\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F \le \Delta_{(\boldsymbol{c},\boldsymbol{d}]}F$.

3) If $F(\boldsymbol{x}) = \prod_{j=1}^d F_j(x_j)$, then $\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F = \prod_{j=1}^d (F_j(b_j) - F_j(a_j))$.
4) If F is grounded, then $\Delta_{(-\boldsymbol{\infty},\boldsymbol{x}]}F = F(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^d$, where $\Delta_{(-\boldsymbol{\infty},\boldsymbol{x}]}F := \lim_{\boldsymbol{a}\to-\boldsymbol{\infty}} \Delta_{(\boldsymbol{a},\boldsymbol{x}]}F$.
5) If F is grounded and $\emptyset \neq J \subsetneq \{1,\dots,d\}$, then $\Delta_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}F(\boldsymbol{x}) \ge 0 \ \forall \boldsymbol{x}_{J^c} \in \mathbb{R}^{|J^c|}$. In particular, for $J = \{j\}$, $j = 1,\dots,d$, F is componentwise increasing.

Proof.

1) Proof by induction:
$d = 1$: $\Delta_{(a,b]}F \underset{\text{def.}}{=} F(b) - F(a) = \Delta_{(a,b]}F$ ✓
$d-1 \Rightarrow d$:

$$\begin{aligned}
\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F &\underset{\text{def.}}{=} \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} F(a_1^{i_1}b_1^{1-i_1},\dots,a_d^{i_d}b_d^{1-i_d}) \\
&= \sum_{\boldsymbol{i}\in\{0,1\}^d:\, i_d=0} (-1)^{\sum_{j=1}^{d-1} i_j} F(a_1^{i_1}b_1^{1-i_1},\dots,a_{d-1}^{i_{d-1}}b_{d-1}^{1-i_{d-1}}, b_d) \\
&\quad - \sum_{\boldsymbol{i}\in\{0,1\}^d:\, i_d=1} (-1)^{\sum_{j=1}^{d-1} i_j} F(a_1^{i_1}b_1^{1-i_1},\dots,a_{d-1}^{i_{d-1}}b_{d-1}^{1-i_{d-1}}, a_d) \\
&\underset{\text{i.h.}}{=} \Delta_{(a_{d-1},b_{d-1}]}\dots\Delta_{(a_1,b_1]}F(x_1\dots,x_{d-1},b_d)
\end{aligned}$$

$$
\begin{aligned}
&- \Delta_{(a_{d-1},b_{d-1}]} \dots \Delta_{(a_1,b_1]} F(x_1 \dots, x_{d-1}, a_d) \\
&= \Delta_{(a_d,b_d]} \Delta_{(a_{d-1},b_{d-1}]} \dots \Delta_{(a_1,b_1]} F.
\end{aligned}
$$

This holds for any permutation of first-order differences since addition is commutative.

2) For all $j \in \{1, \dots, d\}$, we have

$$
\Delta_{(a_j,b_j]} \Delta_{(\boldsymbol{a}_{j^c},\boldsymbol{b}_{j^c}]} F \underset{1)}{=} \Delta_{(\boldsymbol{a},\boldsymbol{b}]} F \underset{d\text{-incr.}}{\geq} 0, \quad a_j \leq b_j
$$

Therefore, $x_j \mapsto \Delta_{(\boldsymbol{a}_{j^c},\boldsymbol{b}_{j^c}]} F(\boldsymbol{x})$ is $\nearrow$ $(*)$. Since $(\boldsymbol{a}, \boldsymbol{b}] \subseteq (\boldsymbol{c}, \boldsymbol{d}]$, we thus have that

$$
\begin{aligned}
\Delta_{(\boldsymbol{a},\boldsymbol{b}]} F &\underset{1)}{=} \Delta_{(a_d,b_d]} \Delta_{(a_{d-1},b_{d-1}]} \dots \Delta_{(a_1,b_1]} F \underset{(*)}{\leq} \Delta_{(c_d,d_d]} \Delta_{(a_{d-1},b_{d-1}]} \dots \Delta_{(a_1,b_1]} F \\
&\underset{1)}{=} \Delta_{(a_{d-1},b_{d-1}]} \Delta_{(c_d,d_d]} \dots \Delta_{(a_1,b_1]} F \underset{1)}{\overset{(*)}{\leq}} \dots \underset{1)}{\overset{(*)}{\leq}} \Delta_{(\boldsymbol{c},\boldsymbol{d}]} F.
\end{aligned}
$$

3) We have

$$
\begin{aligned}
\Delta_{(\boldsymbol{a},\boldsymbol{b}]} F &\underset{1)}{=} \Delta_{(a_d,b_d]} \dots \Delta_{(a_2,b_2]} \Delta_{(a_1,b_1]} F_1(x_1) F_2(x_2) \cdot \ldots \cdot F_d(x_d) \\
&= \Delta_{(a_d,b_d]} \dots \Delta_{(a_2,b_2]} (F_1(b_1) - F_1(a_1)) \prod_{j=2}^{d} F_j(x_j)
\end{aligned}
$$

$$= (F_1(b_1) - F_1(a_1))\Delta_{(a_d,b_d]} \dots \Delta_{(a_2,b_2]} \prod_{j=2}^{d} F_j(x_j)$$

$$\underset{d-1\text{ times}}{\overset{\text{apply}}{=}} \prod_{j=1}^{d} (F_j(b_j) - F_j(a_j)). \qquad \square$$

4) For $\boldsymbol{x} \in \mathbb{R}^d$, we have

$$\Delta_{(-\boldsymbol{\infty},\boldsymbol{x}]}F = \lim_{\boldsymbol{a}\to-\boldsymbol{\infty}} \Delta_{(\boldsymbol{a},\boldsymbol{x}]}F \underset{\text{def.}}{=} \lim_{\boldsymbol{a}\to-\boldsymbol{\infty}} \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} F(a_1^{i_1}x_1^{1-i_1}, \dots, a_d^{i_d}x_d^{1-i_d})$$

$$= \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} \underbrace{\lim_{\boldsymbol{a}\to-\boldsymbol{\infty}} F(a_1^{i_1}x_1^{1-i_1}, \dots, a_d^{i_d}x_d^{1-i_d})}_{=\begin{cases} 0, & \exists\, j : i_j = 1 \text{ by groundedness,} \\ F(\boldsymbol{x}), & \boldsymbol{i} = \boldsymbol{0}, \end{cases}} = F(\boldsymbol{x}).$$

5) Expanding $\boldsymbol{a}_J$ to $\boldsymbol{a}$ such that $\boldsymbol{a}_{J^c} = -\boldsymbol{\infty}$, and $\boldsymbol{b}_J$ to $\boldsymbol{b}$ such that $\boldsymbol{b}_{J^c} = \boldsymbol{x}_{J^c}$, we have $\Delta_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}F(\boldsymbol{x}) \underset{4)}{=} \Delta_{(\boldsymbol{a},\boldsymbol{b}]}F \underset{d\text{-incr.}}{\geq} 0 \;\forall\, \boldsymbol{x}_{J^c} \in \mathbb{R}^{|J^c|}$.

Before we can construct Borel measures on $\mathbb{R}^d$, we need one more auxiliary result.

Lemma 2.38 (Representation of differences involving multiple sets)
Let $\mathcal{A}$ be a semiring and $A, A_1, \dots, A_n \in \mathcal{A}$. Then $\exists\, m \in \mathbb{N}$ and pairwise disjoint $B_1, \dots, B_m \in \mathcal{A}$ such that $A \setminus \bigcup_{i=1}^n A_i = \biguplus_{j=1}^m B_j$.

Proof. Proof by induction:

$n = 1: A \setminus A_1 \underset{\text{D. 2.3 iii)}}{=} \biguplus_{j=1}^m B_j$ ✓

$n-1 \Rightarrow n$: We have

$$A \setminus \bigcup_{i=1}^n A_i = \biggl(A \setminus \bigcup_{i=1}^{n-1} A_i\biggr) \setminus A_n \underset{\text{hypo.}}{=} \biggl(\biguplus_{k=1}^p C_k\biggr) \setminus A_n \underset{\text{distr.}}{=} \biguplus_{k=1}^p (C_k \setminus A_n) \underset{\text{D. 2.3 iii)}}{=} \biguplus_{k=1}^p \biguplus_{l=1}^{p_k} C_{k,l}$$

for pairwise disjoint $\{C_k\}_{k=1}^p, \{C_{k,l}\}_{k,l=1}^{p,p_k} \subseteq \mathcal{A}$, which is of the required form. □

Lemma
Let μ_0 be an additive premeasure on a semiring $\mathcal{A}$.

1) If $A_1, \dots, A_n \in \mathcal{A}$ are pairwise disjoint and $A \in \mathcal{A}$ such that $\biguplus_{i=1}^n A_i \subseteq A$, then $\sum_{i=1}^n \mu_0(A_i) \le \mu_0(A)$.
2) If $A_1, \dots, A_n, A \in \mathcal{A}$ such that $A \subseteq \bigcup_{i=1}^n A_i$, then $\mu_0(A) \le \sum_{i=1}^n \mu_0(A_i)$.

Proof.

1) By L. 2.38, $A \backslash \biguplus_{i=1}^{n} A_i = \biguplus_{j=1}^{m} B_j$ for pairwise disjoint $B_1, \dots, B_m \in \mathcal{A}$ and $A_1, \dots, A_n, B_1, \dots, B_m$ are pairwise disjoint. Therefore, $\mu_0(A) \underset{\text{add.}}{=} \sum_{i=1}^{n} \mu_0(A_i) + \sum_{j=1}^{m} \mu_0(B_j) \underset{\mu_0 \geq 0}{\geq} \sum_{i=1}^{n} \mu_0(A_i)$.

2) For $i = 1, \dots, n$, let $A_i' = A \cap A_i \in \mathcal{A}$ and $A_i'' = A_i' \backslash \bigcup_{k=1}^{i-1} A_k'$. Then $A \underset{A \subseteq \bigcup_{i=1}^{n} A_i}{=} A \cap \bigcup_{i=1}^{n} A_i \underset{\text{distr.}}{=} \bigcup_{i=1}^{n} A_i' = \biguplus_{i=1}^{n} A_i''$. For all $i = 1, \dots, n$, $A_i'' \underset{\text{L. 2.38}}{=} \biguplus_{k=1}^{n_i} B_{i,k}$ for pairwise disjoint $B_{i,k} \in \mathcal{A}$, $k = 1, \dots, n_i$, and thus $A = \biguplus_{n=1}^{n} \biguplus_{k=1}^{n_i} B_{i,k}$. Since $A_1'', \dots, A_n''$ are pairwise disjoint, so are $B_{i,k}\ \forall\, i, k$. Also, $\biguplus_{k=1}^{n_i} B_{i,k} = A_i'' \subseteq A_i' \subseteq A_i\ \forall\, i$, so $\sum_{k=1}^{n_i} \mu_0(B_{i,k}) \underset{\text{add.}}{=} \mu_0(\biguplus_{k=1}^{n_i} B_{i,k}) \underset{1)}{\leq} \mu_0(A_i)$. Therefore, $\mu_0(A) \underset{\text{add.}}{=} \sum_{i=1}^{n} \sum_{k=1}^{n_i} \mu_0(B_{i,k}) \leq \sum_{i=1}^{n} \mu_0(A_i)$. □

Theorem 2.39 (Construction of Borel measures on $\mathbb{R}^d$)
If $F : \mathbb{R}^d \to \mathbb{R}$ is d-increasing and right-continuous, $\exists!$ Borel measure λ_F such that $\lambda_F((\boldsymbol{a}, \boldsymbol{b}]) = \Delta_{(\boldsymbol{a},\boldsymbol{b}]} F$, $\boldsymbol{a} \leq \boldsymbol{b}$.

Proof.

- $\mathcal{A} := \{(\boldsymbol{a}, \boldsymbol{b}] : -\boldsymbol{\infty} < \boldsymbol{a} \leq \boldsymbol{b} < \boldsymbol{\infty}\}$ is a semiring on $\mathbb{R}^d$:
  i) For any $\boldsymbol{a} \in \mathbb{R}^d$, $\emptyset = (\boldsymbol{a}, \boldsymbol{a}] \in \mathcal{A}$.

ii) For $i=1,2$, $-\infty<\boldsymbol{a}_i\le\boldsymbol{b}_i<\infty$, we have

$$(\boldsymbol{a}_1,\boldsymbol{b}_1]\cap(\boldsymbol{a}_2,\boldsymbol{b}_2]=\left(\begin{pmatrix}\max_{i=1,2}\{a_{i,1}\}\\ \vdots\\ \max_{i=1,2}\{a_{i,d}\}\end{pmatrix},\ \begin{pmatrix}\min_{i=1,2}\{b_{i,1}\}\\ \vdots\\ \min_{i=1,2}\{b_{i,d}\}\end{pmatrix}\right]\in\mathcal{A},$$

interpreted as $\emptyset\in\mathcal{A}$ if $\exists\, j\in\{1,\dots,d\}:\max_{i=1,2}\{a_{i,j}\}\ge\min_{i=1,2}\{b_{i,j}\}$.

iii) The difference of two hypercubes $(\boldsymbol{a}_1,\boldsymbol{b}_1]$, $(\boldsymbol{a}_2,\boldsymbol{b}_2]$ is the union of (at most $2d$, so) finitely many hypercubes and thus in $\mathcal{A}$.

- We now show that $\mu_0((\boldsymbol{a},\boldsymbol{b}]):=\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F$ is a σ-finite premeasure on $\mathcal{A}$:

i) F d-increasing $\Rightarrow\mu_0:\mathcal{A}\to[0,\infty]$.

ii) $\mu_0(\emptyset)=\mu_0((\boldsymbol{a},\boldsymbol{a}])\underset{\text{def.}}{=}\Delta_{(\boldsymbol{a},\boldsymbol{a}]}F\underset{\text{def.}}{=}\sum_{\boldsymbol{i}\in\{0,1\}^d}(-1)^{\sum_{j=1}^d i_j}F(a_1^{i_1}a_1^{1-i_1},\dots,a_d^{i_d}a_d^{1-i_d})=F(\boldsymbol{a})\sum_{\boldsymbol{i}\in\{0,1\}^d}(-1)^{\sum_{j=1}^d i_j}=0$.

iii) σ-additivity of μ_0: If $\mathcal{A}\ni(\boldsymbol{a},\boldsymbol{b}]=\biguplus_{i=1}^n(\boldsymbol{a}_i,\boldsymbol{b}_i]$, then $\mu_0(\biguplus_{i=1}^n(\boldsymbol{a}_i,\boldsymbol{b}_i])\underset{\text{def.}}{=}\mu_0((\boldsymbol{a},\boldsymbol{b}])\underset{\text{def.}}{=}\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F\overset{\text{mult.}}{\underset{\text{tele.}}{=}}\sum_{i=1}^n\Delta_{(\boldsymbol{a}_i,\boldsymbol{b}_i]}F\underset{\text{def.}}{=}\sum_{i=1}^n\mu_0((\boldsymbol{a}_i,\boldsymbol{b}_i])$, so μ_0 is additive. Let $\biguplus_{i=1}^\infty(\boldsymbol{a}_i,\boldsymbol{b}_i]=(\boldsymbol{a},\boldsymbol{b}]\in\mathcal{A}$.

  - $\sum_{i=1}^\infty\mu_0((\boldsymbol{a}_i,\boldsymbol{b}_i])\le\mu_0((\boldsymbol{a},\boldsymbol{b}])$: $\forall\, n\in\mathbb{N}$, $\biguplus_{i=1}^n(\boldsymbol{a}_i,\boldsymbol{b}_i]\subseteq\biguplus_{i=1}^\infty(\boldsymbol{a}_i,\boldsymbol{b}_i]=(\boldsymbol{a},\boldsymbol{b}]\underset{\text{L. 1)}}{\Rightarrow}\sum_{i=1}^n\mu_0((\boldsymbol{a}_i,\boldsymbol{b}_i])\le\mu_0((\boldsymbol{a},\boldsymbol{b}])\underset{n\to\infty}{\Rightarrow}\checkmark$.

- $\mu_0((\boldsymbol{a},\boldsymbol{b}]) \le \sum_{i=1}^\infty \mu_0((\boldsymbol{a}_i,\boldsymbol{b}_i])$: F right-continuous $\Rightarrow \boldsymbol{x} \mapsto \Delta_{(\boldsymbol{x},\boldsymbol{b}]}F = \mu_0((\boldsymbol{x},\boldsymbol{b}])$ is right-continuous $\Rightarrow$ For $\varepsilon > 0$, $\exists\, \tilde{\boldsymbol{a}} \in (\boldsymbol{a},\boldsymbol{b}]$: $\mu_0((\boldsymbol{a},\boldsymbol{b}]) \underset{(*)}{\le} \mu_0((\tilde{\boldsymbol{a}},\boldsymbol{b}]) + \varepsilon$. Similarly, $\boldsymbol{x} \mapsto \mu_0((\boldsymbol{a},\boldsymbol{x}])$ is right-continuous, so $\forall\, i \in \mathbb{N}\ \exists\, \tilde{\boldsymbol{b}}_i \in (\boldsymbol{b}_i,\boldsymbol{b}]$: $\mu_0((\boldsymbol{a}_i,\tilde{\boldsymbol{b}}_i]) \underset{(**)}{\le} \mu_0((\boldsymbol{a}_i,\boldsymbol{b}_i]) + \frac{\varepsilon}{2^i}$. We also have $[\tilde{\boldsymbol{a}},\boldsymbol{b}] \subseteq (\boldsymbol{a},\boldsymbol{b}] \underset{\text{ass.}}{=} \biguplus_{i=1}^\infty (\boldsymbol{a}_i,\boldsymbol{b}_i] \subseteq \bigcup_{i=1}^\infty (\boldsymbol{a}_i,\tilde{\boldsymbol{b}}_i) \underset{\text{Heine–Borel}}{\Rightarrow} \exists\, n \in \mathbb{N}$: $[\tilde{\boldsymbol{a}},\boldsymbol{b}] \subseteq \bigcup_{k=1}^n (\boldsymbol{a}_{i_k},\tilde{\boldsymbol{b}}_{i_k})$. This implies that $(\tilde{\boldsymbol{a}},\boldsymbol{b}] \subseteq [\tilde{\boldsymbol{a}},\boldsymbol{b}] \subseteq \bigcup_{k=1}^n (\boldsymbol{a}_{i_k},\tilde{\boldsymbol{b}}_{i_k}) \subseteq \bigcup_{k=1}^n (\boldsymbol{a}_{i_k},\tilde{\boldsymbol{b}}_{i_k}]$, so $\mu_0((\tilde{\boldsymbol{a}},\boldsymbol{b}]) \underset{\text{L. 2)}}{\le} \sum_{k=1}^n \mu_0((\boldsymbol{a}_{i_k},\tilde{\boldsymbol{b}}_{i_k}])$. Hence

$$\mu_0((\boldsymbol{a},\boldsymbol{b}]) \underset{(*)}{\le} \mu_0((\tilde{\boldsymbol{a}},\boldsymbol{b}]) + \varepsilon \le \sum_{k=1}^n \mu_0((\boldsymbol{a}_{i_k},\tilde{\boldsymbol{b}}_{i_k}]) + \varepsilon$$
$$\underset{(**)}{\le} \sum_{k=1}^n \left(\mu_0((\boldsymbol{a}_{i_k},\boldsymbol{b}_{i_k}]) + \frac{\varepsilon}{2^{i_k}}\right) + \varepsilon \underset{\mu_0\,\ge\,0}{\le} \sum_{i=1}^\infty \mu_0((\boldsymbol{a}_i,\boldsymbol{b}_i]) + 2\varepsilon.$$

$\underset{\varepsilon\,\to\,0+}{\Rightarrow} \checkmark \Rightarrow \mu_0$ is σ-additive $\Rightarrow \mu_0$ is a premeasure on $\mathcal{A}$.

- With $A_i = ({}_d i - \mathbf{1}, {}_d i] \in \mathcal{A}$, $i \in \mathbb{N}$, we have $\mathbb{R}^d = \bigcup_{i\in\mathbb{N}} A_i$ and $\mu_0(A_i) = \Delta_{({}_d i-\mathbf{1},{}_d i]}F < \infty\ \forall\, i$, so μ_0 is σ-finite.
- By Carathéodory's extension theorem, $\exists!$ measure λ_F on $\sigma(\mathcal{A}) \underset{\text{R. 2.24 1)}}{=} \mathcal{B}(\mathbb{R}^d)$: $\lambda_F\big|_{\mathcal{A}} = \mu_0$, i.e. $\lambda_F((\boldsymbol{a},\boldsymbol{b}]) = \mu_0((\boldsymbol{a},\boldsymbol{b}]) = \Delta_{(\boldsymbol{a},\boldsymbol{b}]}F$, $-\boldsymbol{\infty} < \boldsymbol{a} \le \boldsymbol{b} < \boldsymbol{\infty}$. $\square$

- As mentioned before, we can assume λ_F to be complete. It is known as *Lebesgue–Stieltjes measure* associated to F.
- If $F(\boldsymbol{x}) = \prod_{j=1}^{d} x_j$, $\boldsymbol{x} \in \mathbb{R}^d$, then $\lambda := \lambda_F$ is the *Lebesgue measure* on $\mathbb{R}^d$. Its domain is the completion $\bar{\mathcal{B}}(\mathbb{R}^d)$, known as *Lebesgue σ-algebra*. Sets in $\bar{\mathcal{B}}(\mathbb{R}^d)$ are *Lebesgue measurable* (or *Lebesgue sets*). Clearly, $\mathcal{B}(\mathbb{R}^d) \subseteq \bar{\mathcal{B}}(\mathbb{R}^d)$. One can also show that $\mathcal{B}(\mathbb{R}^d) \neq \bar{\mathcal{B}}(\mathbb{R}^d)$; see E. 3.9 later.
- The Lebesgue measure satisfies $\lambda((\boldsymbol{a}, \boldsymbol{b}]) \underset{\text{L. 2.37 3)}}{=} \prod_{j=1}^{d}(b_j - a_j)$. It thus assigns hypercubes their volume as we set out to construct; one can also show its invariance wrt translations, rotations and reflections. For general F, λ_F can be interpreted as assigning hypercubes their volume distorted by F.
- Since $\mathbb{R}^d = \prod_{j=1}^{d} \mathbb{R}$ is a product space, we can also equip $\mathbb{R}^d$ with the product-σ-algebra $\bigotimes_{j=1}^{d} \mathcal{F}_j \underset{\text{P. 2.16}}{=} \sigma(\prod_{j=1}^{d} B_j : B_j \in \mathcal{B}(\mathbb{R}))$, which, by R. 2.24 1), is $\mathcal{B}(\mathbb{R}^d)$. On $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, we can then consider the product measure $\mu = \prod_{j=1}^{d} \mu_j$, with μ_j being univariate λ as λ is a measure on $\mathcal{B}(\mathbb{R})$; see E. 2.29. Then $\mu(\prod_{j=1}^{d}(a_j, b_j]) \underset{\text{def.}}{=} \prod_{j=1}^{d} \mu_j((a_j, b_j]) = \prod_{j=1}^{d}(b_j - a_j)$, $-\infty < a_j \leq b_j < \infty$, $j = 1, \dots, d$, so μ and the d-dimensional λ we constructed coincide on the semiring $\mathcal{A}$, which is a π-system by definition. Also, the d-dimensional λ is

σ-finite as seen in the proof of T. 2.39. By P. 2.28, we thus must have $\mu = \lambda$, so the Lebesgue measure on $\mathbb{R}^d$ is the product measure of the Lebesgue measure on $\mathbb{R}$. We could have thus also shown T. 2.39 for $d = 1$ and defined the d-dimensional Lebesgue measure via the product measure.

Remark 2.40

1) ■ $\mathcal{A}$ is a semiring but no ring (for $\boldsymbol{a} < \boldsymbol{c} < \boldsymbol{d} < \boldsymbol{b}$, $(\boldsymbol{a}, \boldsymbol{b}]\backslash(\boldsymbol{c}, \boldsymbol{d}]$ is not a hypercube, but a finite disjoint union of hypercubes). To apply Carathéodory's extension theorem for rings, one would consider the ring $\mathcal{A}_{\uplus} := \{\biguplus_{i=1}^n (\boldsymbol{a}_i, \boldsymbol{b}_i] : -\infty < \boldsymbol{a}_i \le \boldsymbol{b}_i < \infty \ \forall\, i, n \in \mathbb{N}\}$ generated by $\mathcal{A}$ and show that $\mu_0(\biguplus_{i=1}^n (\boldsymbol{a}_i, \boldsymbol{b}_i]) := \sum_{i=1}^n \Delta_{(\boldsymbol{a}_i, \boldsymbol{b}_i]} F$ is a premeasure on $\mathcal{A}_{\uplus}$; see Steps 2.1)–2.3) in the proof of T. 2.35.

   ■ $\mathcal{A}_{\uplus}$ is not an algebra (since $\Omega = \mathbb{R}^d \notin \mathcal{A}_{\uplus}$). To apply Carathéodory's extension theorem for algebras, one would consider the algebra $\mathcal{A}' := \{\biguplus_{i=1}^n I_i : I_i = (\boldsymbol{a}_i, \boldsymbol{b}_i]$ or $I_i = (\boldsymbol{a}_i, \boldsymbol{\infty})$ for $-\boldsymbol{\infty} \le \boldsymbol{a}_i \le \boldsymbol{b}_i < \boldsymbol{\infty}\}$ and show that $\mu_0(\biguplus_{i=1}^n (\boldsymbol{a}_i, \boldsymbol{b}_i]) := \sum_{i=1}^n \Delta_{(\boldsymbol{a}_i, \boldsymbol{b}_i]} F$ is a premeasure on $\mathcal{A}'$.

2) Similarly for left-continuous F and intervals of the form $[\boldsymbol{a}, \boldsymbol{b})$, but less common.

Question: Why does a componentwise increasing F not suffice to construct a measure?

Example 2.41 (Componentwise increasing $\not\Rightarrow$ d-increasing)

Let $F(\boldsymbol{x}) = \max\{(\sum_{j=1}^d x_j) - d + 1, 0\}$, $\boldsymbol{x} \in [0,1]^d$. Then F is componentwise increasing, but not d-increasing for $d \geq 3$ since

$$\begin{aligned}
\Delta_{(\mathbf{1/2},\mathbf{1}]} F &= \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} F((1/2)^{i_1} 1^{1-i_1}, \dots, (1/2)^{i_d} 1^{1-i_d}) \\
&= \max\{1 + \dots + 1 + 1 - d + 1, 0\} \quad (i_j = 0 \;\forall\, j) \\
&\quad - d \max\{1 + \dots + 1 + 1/2 - d + 1, 0\} \quad (\exists!\, j : i_j = 1) \\
&\quad \pm \dots + (-1)^d \max\{d/2 - d + 1, 0\} \quad (i_j = 1 \;\forall\, j) \\
&= 1 - d/2 < 0, \quad d \geq 3,
\end{aligned}$$

so λ_F does not induce a Borel measure on $\mathbb{R}^d$, $d \geq 3$ (since $\lambda_F((\mathbf{1/2},\mathbf{1}]) < 0$).

2.7 Lebesgue null sets

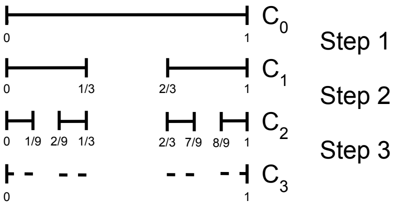
$N \in \bar{\mathcal{B}}(\mathbb{R}^d) : \lambda(N) = 0$ is a *(Lebesgue) null set*. By T. 2.32, $B \in \bar{\mathcal{B}}(\mathbb{R}^d)$ iff $B = A \cup N$ for some (Borel set) $A \in \mathcal{B}(\mathbb{R}^d)$ and N a Lebesgue null set. Other representations are also possible, e.g. $B = A \backslash N$. Lebesgue sets are thus Borel sets modulo Lebesgue null sets.

We now provide some examples of Lebesgue null sets. We start with null sets in $\mathbb{R}$.

Example 2.42 (Lebesgue null sets in $\mathbb{R}$)

1) $\forall\, x \in \mathbb{R}$, $\{x\} \subseteq \mathbb{R}$ is a null set since $\lambda(\{x\}) = \lambda(\bigcap_{i=1}^\infty (x - 1/i, x]) \underset{\text{cont. above}}{\overset{\lambda((x-1,x])<\infty}{=}} \lim_{n\to\infty} \lambda((x - 1/n, x]) = \lim_{n\to\infty}(x - (x - \frac{1}{n})) = \lim_{n\to\infty} \frac{1}{n} = 0$.
2) $\mathbb{Q} \subseteq \mathbb{R}$ is a null set: If $\mathbb{Q} = \{q_i\}_{i\in\mathbb{N}}$ is an enumeration by Cantor's first diagonal argument, then $\lambda(\mathbb{Q}) \underset{\text{L. 2.31}}{=} 0$.

   Question: Are there also uncountable Lebesgue null sets?
3) The *Cantor set* is $\mathcal{C} := \bigcap_{i=1}^\infty C_i$ with $C_0 := [0, 1]$ and $C_i = \frac{C_{i-1}}{3} \cup (\frac{2}{3} + \frac{C_{i-1}}{3})$, $i \in \mathbb{N}$.



One also has $\mathcal{C} = \{x \in [0,1] : x = \sum_{i=1}^{\infty} a_i 3^{-i},\ a_i \in \{0,2\}\ \forall\, i \in \mathbb{N}\}$.

Then:

- $\mathcal{C}$ is uncountable:

  Proof. If $\mathcal{C} = \{c_i\}_{i \in \mathbb{N}}$ is an enumeration with, say,

  $$c_1 = 0.\underline{0}00\dots, \quad c_2 = 0.2\underline{0}0\dots, \quad c_3 = 0.00\underline{2}\dots, \quad \text{etc.},$$

  then $c = 0.220\dots \in \mathcal{C}$ but $c \notin \{c_i\}_{i \in \mathbb{N}}$ ↯ (*Cantor's second diagonal argument*).

- $\mathcal{C}$ is a null set:

  Proof. The ith step removes 2^{i-1} parts of length 3^{-i} from C_{i-1}, so the length of the removed parts is

  $$\lambda([0,1] \backslash \mathcal{C}) = 2^0 \cdot \frac{1}{3^1} + 2^1 \cdot \frac{1}{3^2} + 2^2 \cdot \frac{1}{3^3} + \dots = \sum_{i=1}^{\infty} 2^{i-1} 3^{-i}$$

$$= \frac{1}{2}\sum_{i=1}^{\infty}(2/3)^i = \frac{1}{2}\Big(\frac{1}{1-2/3} - 1\Big) = 1,$$

hence $1 = \lambda([0,1]) = \lambda(([0,1]\setminus\mathcal{C}) \uplus \mathcal{C}) \underset{\text{add.}}{=} \lambda([0,1]\setminus\mathcal{C}) + \lambda(\mathcal{C}) = 1 + \lambda(\mathcal{C})$, so $\lambda(\mathcal{C}) = 0$.

We thus obtain that $\mathcal{C}$ is an uncountable Lebesgue null set.

Now some examples of null sets in $\mathbb{R}^d$.

Example 2.43 (Lebesgue null sets in $\mathbb{R}^d$)

1) A line in $\mathbb{R}^2$ is a Lebesgue null set.
   Proof. Consider wlog $y = 0$. Let $\mathbb{Q} = \{q_i\}_{i\in\mathbb{N}}$ be an enumeration. For $\varepsilon > 0$, let $A_i = (q_i - \frac{\varepsilon}{2}, q_i + \frac{\varepsilon}{2}] \times (-\frac{1}{2^{i+1}}, \frac{1}{2^{i+1}}]$, $i \in \mathbb{N}$. Then $0 \le \lambda(\{(x,y) : y = 0\}) \underset{\text{mon.}}{\le} \lambda(\bigcup_{i=1}^{\infty} A_i) \underset{\text{subadd.}}{\le} \sum_{i=1}^{\infty} \lambda(A_i) = \sum_{i=1}^{\infty} \varepsilon \cdot \frac{1}{2^i} = \varepsilon \underset{\varepsilon \to 0+}{\Rightarrow}$ ✓
2) Similarly, any $\mathbb{R}^k$, $k < d$, or any subset thereof, in $\mathbb{R}^d$, is a null set, e.g. planes in $\mathbb{R}^3$.

2.8 Probability measures

Definition 2.44 (Probability measure, probability space, etc.)
Let $(\Omega, \mathcal{F})$ be a measurable space. A *probability measure* $\mathbb{P}$ on $\mathcal{F}$ is a function such that

i) $\mathbb{P} : \mathcal{F} \to [0, \infty]$;
ii) $\mathbb{P}(\Omega) = 1$; and
iii) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$, $A_i \cap A_j = \emptyset \ \forall i \neq j \Rightarrow \mathbb{P}(\biguplus_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ (*σ-additivity*).

$(\Omega, \mathcal{F}, \mathbb{P})$ is a *probability space*, Ω the *sample space* and $\omega \in \Omega$ a *sample point*. If Ω is countable (finite), $(\Omega, \mathcal{F}, \mathbb{P})$ is *discrete* (*finite*). Any $A \in \mathcal{F}$ is called an *event*. If $A = \{\omega\}$, A is a *simple event*, otherwise a *compound event*.

- Ω is the set of all possible results (*outcomes*) of an experiment we would like to model and $\mathcal{F}$ contains all sets we can assign probabilities to.
- $1 \underset{\text{ii)}}{=} \mathbb{P}(\Omega) = \mathbb{P}(A \uplus A^c \uplus \bigcup_{i=3}^{\infty} \emptyset) \underset{\text{iii)}}{=} \mathbb{P}(A) + \mathbb{P}(A^c) + \sum_{i=3}^{\infty} \mathbb{P}(\emptyset) \underset{\text{i)}}{\Rightarrow} \mathbb{P}(\emptyset) \overset{!}{=} 0$ and $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, $A \in \mathcal{F}$. In particular, $\mathbb{P}$ is a measure.
- $\mathbb{P}$ is finite $\Rightarrow$ all parts of P. 2.27 hold, e.g. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Example 2.45 (Probability measures)

1) If $|\Omega| < \infty$, $\mathcal{F} := \mathcal{P}(\Omega)$ and $\mathbb{P}(A) := |A|/|\Omega|$, then $(\Omega, \mathcal{F}, \mathbb{P})$ is called *Laplace probability space*. The Laplace probability measure assigns each subset $A \subseteq \Omega$ its relative number of elements in comparison to Ω.
2) If Ω is countable and $\mathcal{F} := \mathcal{P}(\Omega)$, then any *probability mass function (pmf)* $f : \Omega \to [0,1] : \sum_{\omega\in\Omega} f(\omega) = 1$ induces a *discrete probability measure* $\mathbb{P}$ on $(\Omega, \mathcal{F})$ via
$$\mathbb{P}(A) := \sum_{\omega\in A} f(\omega), \quad A \in \mathcal{F};$$
see also E. 2.26 3). Conversely, if $\mathbb{P}$ is a probability measure on $\mathcal{F}$, then
$$f(\omega) := \mathbb{P}(\{\omega\}), \quad \omega \in \Omega,$$
defines a pmf such that $\mathbb{P}(A) = \sum_{\omega\in A} f(\omega)$, $A \in \mathcal{F}$; this allows one to define $\mathbb{P}$ by first defining it for simple events.
3) If $\Omega \subseteq \mathbb{R}^d : \lambda(\Omega) < \infty$, $\mathcal{F} := \bar{\mathcal{B}}(\Omega)$ and $\mathbb{P}(A) := \lambda(A)/\lambda(\Omega)\ \forall\, A \in \mathcal{F}$ (relative volume), then $(\Omega, \mathcal{F}, \mathbb{P})$ is called *geometric probability space*. For $\Omega = [0,1]^d$, $([0,1]^d, \bar{\mathcal{B}}([0,1]^d), \lambda)$ is the *standard probability space*, often used in examples.

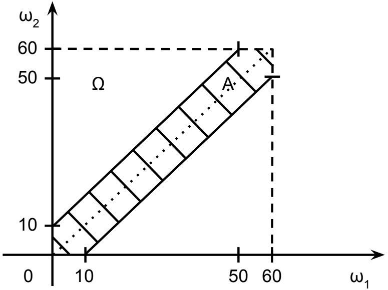
Example 2.46 (Geometric probability space)

Suppose you want to meet with a friend on Zoom during a one-hour lunch break, but you did not fix an exact time in that hour.

1) If you both enter randomly and each waits for max. 10 min, find the probability that you two meet.

Solution. $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = [0,60]^2$ ($(\omega_1, \omega_2) \in [0,60]^2$ is the arrival time in min of you (ω_1) and your colleague (ω_2) since the beginning of the lunch break), $\mathcal{F} = \bar{\mathcal{B}}(\Omega)$, $\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \lambda(A)/3600$, $A \in \mathcal{F}$. Let $A =$ "you and your colleague meet" $= \{(\omega_1, \omega_2) \in \Omega : |\omega_1 - \omega_2| \leq 10\} = \{(\omega_1, \omega_2) \in \Omega : \omega_1 - 10 \leq \omega_2 \leq \omega_1 + 10\}$. We obtain $\mathbb{P}(A) = \frac{2(\frac{1}{2}\cdot 60\cdot 60 - \frac{1}{2}\cdot 50\cdot 50)}{3600} = \frac{11}{36}$.

Sketch:



2) How long would you and your colleague at least have to be willing to wait for each other so that the probability of meeting is at least $p \in [0,1]$?

Solution. Let the smallest waiting time be t (in min since the beginning of the lunchtime break) and $A_t = \{(\omega_1, \omega_2) \in \Omega : |\omega_1 - \omega_2| \le t\}$. As in 1) (just $10 \leftarrow t$), we obtain that $\mathbb{P}(A_t) = \frac{2(\frac{1}{2}\cdot 60\cdot 60 - \frac{1}{2}\cdot(60-t)\cdot(60-t))}{3600} = \frac{120t - t^2}{3600} = p \Leftrightarrow t^2 - 120t + 3600p = 0 \Leftrightarrow t_{1,2} = 60(1 \pm \sqrt{1-p})$. Since $t_1 \notin [0, 60]$, the solution is $t_2 = 60(1 - \sqrt{1-p})$, so you and your colleague would have to be willing to wait at least $60(1 - \sqrt{1-p})$ min to meet with probability $p \in [0, 1]$.

Question: Is there a characterization of probability measures on $\mathbb{R}^d$?

Definition 2.47 (Distribution function on $\mathbb{R}^d$)

$F : \mathbb{R}^d \to [0, 1]$ is a *(multivariate/joint) distribution function (df)* if

i) $\lim_{x_j \to -\infty} F(\boldsymbol{x}) = 0 \ \forall\, j = 1, \dots, d$ (groundedness) and $\lim_{\boldsymbol{x} \to \boldsymbol{\infty}} F(\boldsymbol{x}) = 1$ (normalization);

ii) F is d-increasing; and

iii) F is right-continuous.

Such functions are precisely those that characterize probability measures on $\mathbb{R}^d$.

Theorem 2.48 (Characterization of probability measures on $\mathbb{R}^d$)

1) $\mathbb{P}$ induces F: If $\mathbb{P}: \mathcal{B}(\mathbb{R}^d) \to [0,1]$ is a probability measure, then $F(\boldsymbol{x}) := \mathbb{P}((-\infty, \boldsymbol{x}])$ is a df and satisfies $\lambda_F|_{\mathcal{B}(\mathbb{R}^d)} = \mathbb{P}$.
2) F induces $\mathbb{P}$: If F is a df, then $\mathbb{P} := \lambda_F|_{\mathcal{B}(\mathbb{R}^d)}$ is a probability measure and satisfies $\mathbb{P}((-\infty, \boldsymbol{x}]) = F(\boldsymbol{x})$ on $\mathbb{R}^d$.

Proof.

1) ▪ i) We have

$$\lim_{x_j \to -\infty} F(\boldsymbol{x}) = \lim_{n\to\infty} F(\boldsymbol{x}_{j\leftarrow -n}) \underset{\text{def.}}{=} \lim_{n\to\infty} \mathbb{P}((-\infty, \boldsymbol{x}_{j\leftarrow -n}])$$

$$\overset{\text{cont.}}{\underset{\text{above}}{=}} \mathbb{P}\left(\bigcap_{n=1}^{\infty} (-\infty, \boldsymbol{x}_{j\leftarrow -n}]\right) = \mathbb{P}(\emptyset) = 0.$$

Furthermore,

$$F(\boldsymbol{\infty}) = \lim_{n\to\infty} \mathbb{P}((-\infty, {}_d n]) \overset{\text{cont.}}{\underset{\text{below}}{=}} \mathbb{P}\left(\bigcup_{n=1}^{\infty} (-\infty, {}_d n]\right) = \mathbb{P}(\mathbb{R}^d) \underset{\mathbb{R}^d = \Omega}{=} 1.$$

ii) $\forall -\infty < \boldsymbol{a} \le \boldsymbol{b} < \infty$, we have

$$\begin{aligned}\Delta_{(\boldsymbol{a},\boldsymbol{b}]}F &\underset{\text{L. 2.37 1)}}{=} \Delta_{(a_d,b_d]}\cdots\Delta_{(a_2,b_2]}\underbrace{\Delta_{(a_1,b_1]}F}\\ &= F(\boldsymbol{x}_{1\leftarrow b_1}) - F(\boldsymbol{x}_{1\leftarrow a_1}) \underset{\text{def.}}{=} \mathbb{P}((-\infty,\boldsymbol{x}_{1\leftarrow b_1}]) - \mathbb{P}((-\infty,\boldsymbol{x}_{1\leftarrow a_1}])\\ &\underset{\text{P. 2.27 2)}}{=} \mathbb{P}\Big(\Big(\Big(\begin{smallmatrix} a_1 \\ -\infty \end{smallmatrix}\Big),\Big(\begin{smallmatrix} b_1 \\ \boldsymbol{x}_{\{1\}^c} \end{smallmatrix}\Big)\Big]\Big)\\ &= \Delta_{(a_d,b_d]}\cdots\Delta_{(a_2,b_2]}\mathbb{P}\Big(\Big(\Big(\begin{smallmatrix} a_1 \\ -\infty \end{smallmatrix}\Big),\Big(\begin{smallmatrix} b_1 \\ \boldsymbol{x}_{\{1\}^c} \end{smallmatrix}\Big)\Big]\Big)\\ &= \cdots = \mathbb{P}((\boldsymbol{a},\boldsymbol{b}]) \underset{\text{def.}}{\ge} 0.\end{aligned}$$

iii) Since $F(\boldsymbol{x}) \underset{\text{def.}}{=} \mathbb{P}((-\infty,\boldsymbol{x}]) \underset{\text{mon.}}{\le} \mathbb{P}((-\infty,\boldsymbol{y}]) = F(\boldsymbol{y})$, $-\infty < \boldsymbol{x} \le \boldsymbol{y} < \infty$, we have

$$\begin{aligned}\lim_{\boldsymbol{h}\to\boldsymbol{0}+} F(\boldsymbol{x}+\boldsymbol{h}) &= \lim_{\boldsymbol{n}\to\boldsymbol{\infty}} F\Big(\boldsymbol{x}+\frac{\boldsymbol{1}}{\boldsymbol{n}}\Big) \underset{(*)}{\le} \lim_{\min_j\{n_j\}\to\infty} F\Big(\boldsymbol{x}+\frac{\boldsymbol{1}}{\min_j\{n_j\}}\Big)\\ &= \lim_{m\to\infty} F\Big(\boldsymbol{x}+\frac{\boldsymbol{1}}{_dm}\Big) \underset{\text{def.}}{=} \lim_{m\to\infty} \mathbb{P}\Big(\Big(-\infty,\boldsymbol{x}+\frac{\boldsymbol{1}}{_dm}\Big]\Big)\\ &\overset{\text{cont.}}{\underset{\text{above}}{=}} \mathbb{P}\Big(\bigcap_{m=1}^{\infty}\Big(-\infty,\boldsymbol{x}+\frac{\boldsymbol{1}}{_dm}\Big]\Big) = \mathbb{P}((-\infty,\boldsymbol{x}]) \underset{\text{def.}}{=} F(\boldsymbol{x}),\end{aligned}$$

and flipping the inequality in $(*)$ and changing $\min_j$ to $\max_j$ leads to $\lim_{\boldsymbol{h}\to\mathbf{0}+} F(\boldsymbol{x}+\boldsymbol{h}) \ge F(\boldsymbol{x})$. We thus have $\lim_{\boldsymbol{h}\to\mathbf{0}+} F(\boldsymbol{x}+\boldsymbol{h}) = F(\boldsymbol{x})$, $\boldsymbol{x}\in\mathbb{R}^d$.

- Representation: By P. 2.28, it suffices to show that λ_F coincides with the finite $\mathbb{P}$ on the π-system $\mathcal{A} = \{(-\boldsymbol{\infty}, \boldsymbol{b}] : \boldsymbol{b}\in\mathbb{R}^d\}$, as $\mathcal{A}$ generates $\mathcal{B}(\mathbb{R}^d)$ by R. 2.24 1). And this holds since

$$\lambda_F((-\boldsymbol{\infty},\boldsymbol{b}]) \underset{\text{def.}}{=} \Delta_{(-\boldsymbol{\infty},\boldsymbol{b}]}F \overset{F\text{ grounded}}{\underset{\text{L. 2.37 4)}}{=}} F(\boldsymbol{b}) \underset{\text{def.}}{=} \mathbb{P}((-\boldsymbol{\infty},\boldsymbol{b}]), \quad \boldsymbol{b}\in\mathbb{R}^d.$$

2) By T. 2.39, $\exists!$ Borel measure $\lambda_F : \lambda_F((\boldsymbol{a},\boldsymbol{b}]) \underset{(*)}{=} \Delta_{(\boldsymbol{a},\boldsymbol{b}]}F$, $-\boldsymbol{\infty} < \boldsymbol{a} \le \boldsymbol{b} < \boldsymbol{\infty}$. We need to show that λ_F is a probability measure (left to show: $\lambda_F(\mathbb{R}^d) = 1$) and that it satisfies $\lambda_F((-\boldsymbol{\infty},\boldsymbol{x}]) = F(\boldsymbol{x})\ \forall\,\boldsymbol{x}\in\mathbb{R}^d$.

- Normalization:

$$\begin{aligned}\lambda_F(\mathbb{R}^d) &\overset{\text{cont.}}{\underset{\text{below}}{=}} \lim_{n\to\infty} \lambda_F((-{}_dn, {}_dn]) \underset{(*)}{=} \lim_{n\to\infty} \Delta_{(-{}_dn,{}_dn]}F \\ &\underset{\text{def.}}{=} \lim_{n\to\infty} \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} F((-1)^{i_1}n,\dots,(-1)^{i_d}n)\end{aligned}$$

$$= \sum_{\boldsymbol{i}\in\{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} \underbrace{\lim_{n\to\infty} F((-1)^{i_1}n,\dots,(-1)^{i_d}n)}_{=\begin{cases} 0, & \exists\, j: i_j = 1 \text{ by groundedness,}\\ F(\boldsymbol{\infty}) \underset{F\text{ df}}{=} 1, & \boldsymbol{i} = \boldsymbol{0},\end{cases}} = 1.$$

- Representation: $\lambda_F((-\boldsymbol{\infty},\boldsymbol{x}]) \overset{\text{cont.}}{\underset{\text{below}}{=}} \lim_{n\to\infty} \lambda_F((-_d n,\boldsymbol{x}]) \underset{(*)}{=} \lim_{n\to\infty} \Delta_{(-_d n,\boldsymbol{x}]} F = \Delta_{(-\boldsymbol{\infty},\boldsymbol{x}]} F \overset{F\text{ grounded}}{\underset{\text{L. 2.37 4)}}{=}} F(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^d$. $\square$

Remark 2.49

1) Because of T. 2.48 1), $F(\boldsymbol{x}) := \mathbb{P}((-\boldsymbol{\infty},\boldsymbol{x}])$ is the *df of* $\mathbb{P}$, or, by 2), *of* λ_F.
2) For any probability measure on $\mathcal{B}(\mathbb{R}^d)$ with df F, we have $\mathbb{P}((\boldsymbol{a},\boldsymbol{b}]) \underset{\text{T. 2.48 1)}}{=} \lambda_F((\boldsymbol{a},\boldsymbol{b}]) \underset{\text{def.}}{=} \Delta_{(\boldsymbol{a},\boldsymbol{b}]} F$, so the F-volume $\Delta_{(\boldsymbol{a},\boldsymbol{b}]} F$ is a probability, the probability of $(\boldsymbol{a},\boldsymbol{b}]$ under F. If $d \approx 259$–272, the number of corners of $(\boldsymbol{a},\boldsymbol{b}]$ is $\approx$ the number of atoms in the universe (hence “Monte Carlo simulation” for approximating $\Delta_{(\boldsymbol{a},\boldsymbol{b}]} F$; see later).

3) If $\mathbb{P}$ is a probability measure on $\mathcal{B}(\mathbb{R}^d)$ with df F, then

$$\mathbb{P}(\{\boldsymbol{x}\}) = \mathbb{P}\Bigl(\bigcap_{n=1}^{\infty}\Bigl(\boldsymbol{x}-\frac{\mathbf{1}}{d^n},\boldsymbol{x}\Bigr]\Bigr) \overset{\text{cont.}}{\underset{\text{above}}{=}} \lim_{n\to\infty}\mathbb{P}\Bigl(\Bigl(\boldsymbol{x}-\frac{\mathbf{1}}{d^n},\boldsymbol{x}\Bigr]\Bigr)$$

$$\underset{\text{T. 2.48 1)}}{=} \lim_{n\to\infty}\lambda_F\Bigl(\Bigl(\boldsymbol{x}-\frac{\mathbf{1}}{d^n},\boldsymbol{x}\Bigr]\Bigr) \underset{\text{T. 2.39}}{=} \lim_{n\to\infty}\Delta_{(\boldsymbol{x}-\frac{\mathbf{1}}{d^n},\boldsymbol{x}]}F = \Delta_{(\boldsymbol{x}-,\boldsymbol{x}]}F.$$

In particular, if F is continuous in $\boldsymbol{x}$, then $\mathbb{P}(\{\boldsymbol{x}\}) = 0$, so $\{\boldsymbol{x}\}$ is a null set.

4) The *support* of F is $\operatorname{supp}(F) := \{\boldsymbol{x}\in\mathbb{R}^d : \Delta_{(\boldsymbol{x}-\boldsymbol{h},\boldsymbol{x}]}F > 0\ \forall\,\boldsymbol{h} > \mathbf{0}\}$, and its *range* is $\operatorname{ran}(F) := \{F(\boldsymbol{x}) : \boldsymbol{x}\in\mathbb{R}^d\}$.

5) The domain of a d-dimensional df F is $\mathbb{R}^d$ and F is 1 in the upper-right and 0 in all other regions beyond $\operatorname{supp}(F)$. If not, extend a $\tilde{F}$ with $\boldsymbol{a} := \inf\operatorname{supp}(\tilde{F}) > -\boldsymbol{\infty}$ or $\boldsymbol{b} := \sup\operatorname{supp}(\tilde{F}) < \boldsymbol{\infty}$ to F on $\mathbb{R}^d$ via $F(\boldsymbol{x}) = \tilde{F}(\min\{\max\{x_1,a_1\},b_1\},\dots,\min\{\max\{x_d,a_d\},\,b_d\})$, $\boldsymbol{x}\in\mathbb{R}^d$; see E. 2.41.

6) One can show that every df F has a mixture representation

$$F(\boldsymbol{x}) = p_{\text{ac}}F_{\text{ac}}(\boldsymbol{x}) + p_{\text{d}}F_{\text{d}}(\boldsymbol{x}) + p_{\text{cs}}F_{\text{cs}}(\boldsymbol{x}),$$

for $p_{\text{ac}}, p_{\text{d}}, p_{\text{cs}} \ge 0$, $p_{\text{ac}} + p_{\text{d}} + p_{\text{cs}} = 1$, where

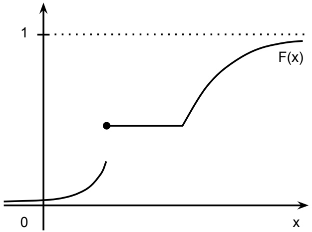
- F_{ac} is an *absolutely continuous* df, i.e. $F_{\text{ac}}(\boldsymbol{x}) = \int_{-\boldsymbol{\infty}}^{\boldsymbol{x}} f_{\text{ac}}(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}$, $\boldsymbol{x}\in\mathbb{R}^d$, for an integrable $f_{\text{ac}} : \mathbb{R}^d \to [0,\infty)$ with $\int_{-\boldsymbol{\infty}}^{\boldsymbol{\infty}} f_{\text{ac}}(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z} = 1$, the *density* of F_{ac}

(if it exists a.e. and is integrable, $f_{\mathsf{ac}}(\boldsymbol{x}) = \frac{\partial}{\partial \boldsymbol{x}} F_{\mathsf{ac}}(\boldsymbol{x})$ is a density candidate, the integration typically being done iteratively via Tonelli's theorem; see later);

- F_{d} is a *discrete* df, i.e. $\operatorname{supp}(F_{\mathsf{d}}) = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots\}$, leading to a multivariate step function, with *probability mass function (pmf)* $\boldsymbol{x} \mapsto \mathbb{P}(\{\boldsymbol{x}\}) \underset{3)}{=} \Delta_{(\boldsymbol{x}-,\boldsymbol{x}]} F_{\mathsf{d}}$ (note that any $F \nearrow \Rightarrow F$ has at most countably-many jumps as each jump gap contains an open interval which contains a different rational number); and
- F_{cs} is a *continuous singular* df, i.e. a continuous df with $\frac{\partial}{\partial \boldsymbol{x}} F_{\mathsf{cs}}(\boldsymbol{x}) = 0$ a.e.

If at least two of $p_{\mathsf{ac}}, p_{\mathsf{d}}, p_{\mathsf{s}}$ are in $(0, 1)$, F is *mixed-type*.

Example: For $d = 1$, $p_{\mathsf{ac}} > 0$, $p_{\mathsf{d}} > 0$, $p_{\mathsf{cs}} = 0$:



7) If F is absolutely continuous with density f, then

$$\begin{aligned}\mathbb{P}((\boldsymbol{a},\boldsymbol{b}]) &\underset{\text{T. 2.39}}{=} \Delta_{(\boldsymbol{a},\boldsymbol{b}]}F \\ &\underset{F\text{ abs. cont.}}{\overset{\text{L. 2.37 1)}}{=}} \Delta_{(a_1,b_1]}\dots\Delta_{(a_d,b_d]}\int_{-\infty}^{x_d}\cdots\int_{-\infty}^{x_1} f(z_1,\dots,z_d)\,\mathrm{d}z_1\dots\mathrm{d}z_d \\ &= \int_{a_d}^{b_d}\cdots\int_{a_1}^{b_1} f(z_1,\dots,z_d)\,\mathrm{d}z_1\dots\mathrm{d}z_d = \int_{(\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}.\end{aligned}$$

8) For $J \subseteq \{1,\dots,d\}$, let $F_J(\boldsymbol{x}_J) = F(\boldsymbol{\infty}_{J\leftarrow \boldsymbol{x}_J}) := \lim_{\boldsymbol{x}_{J^c}\to\boldsymbol{\infty}} F(\boldsymbol{x})$ denote the *J-margin* of F. For $J=\{j\}$, $j=1,\dots,d$, the *jth margin of F* is $F_j(x_j) = F(\boldsymbol{\infty}_{j\leftarrow x_j}) = \lim_{\boldsymbol{x}_{\{j\}^c}\to\boldsymbol{\infty}} F(\boldsymbol{x})$. Similarly, for $A=\prod_{j=1}^d A_j \in \mathcal{B}(\mathbb{R}^d)$, let $\mathbb{P}(A_J) = \mathbb{P}(\prod_{j\in J} A_j) := \mathbb{P}(\prod_{j=1}^d B_j)$ with $B_j = A_j$ for $j\in J$ and $B_j=\Omega$ for $j\notin J$. We thus have

$$\begin{aligned}F_J(\boldsymbol{x}_J) &\underset{\text{def.}}{=} \lim_{\boldsymbol{x}_{J^c}\to\boldsymbol{\infty}} F(\boldsymbol{x}) \underset{\text{T. 2.48}}{=} \lim_{\boldsymbol{x}_{J^c}\to\boldsymbol{\infty}} \mathbb{P}((-\boldsymbol{\infty},\boldsymbol{x}]) \underset{\text{below}}{\overset{\text{cont.}}{=}} \mathbb{P}((-\boldsymbol{\infty},\boldsymbol{x}_{J^c\leftarrow\boldsymbol{\infty}}]) \\ &\underset{\text{def.}}{=} \mathbb{P}((-\boldsymbol{\infty},\boldsymbol{x}_J]);\end{aligned}$$

For $J=\{j\}$, $j=1,\dots,d$, we have $F_j(x_j) = \mathbb{P}((-\infty,x_j])$; in particular, $\mathbb{P}((a_j,b_j]) = \mathbb{P}((-\infty,b_j]\backslash(-\infty,a_j]) \underset{\text{P. 2.27 2)}}{=} \mathbb{P}((-\infty,b_j]) - \mathbb{P}((-\infty,a_j]) = F_j(b_j)-F_j(a_j)$.

Question: Is continuity of F related to that of $F_1, \dots, F_d$?

Lemma 2.50 (Lipschitz inequality)

If F is d-increasing and grounded, then

$$|F(\boldsymbol{b}) - F(\boldsymbol{a})| \le \sum_{j=1}^{d} |F_j(b_j) - F_j(a_j)|, \quad \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d.$$

Proof. $|F(\boldsymbol{b}) - F(\boldsymbol{a})|$ equals

$$\underset{\Delta\text{-ineq.}}{\overset{\text{tele.}}{\le}} \sum_{j=1}^{d} |F(b_1, \dots, b_{j-1}, b_j, a_{j+1}, \dots, a_d) - F(b_1, \dots, b_{j-1}, a_j, a_{j+1}, \dots, a_d)|.$$

Suppose $a_j \le b_j$, then the jth summand satisfies

$$\begin{aligned} &|F(b_1, \dots, b_{j-1}, b_j, a_{j+1}, \dots, a_d) - F(b_1, \dots, b_{j-1}, a_j, a_{j+1}, \dots, a_d)| \\ \underset{J=\{j\}}{\overset{\text{L. 2.37 5)}}{=}} \; &F(b_1, \dots, b_{j-1}, b_j, a_{j+1}, \dots, a_d) - F(b_1, \dots, b_{j-1}, a_j, a_{j+1}, \dots, a_d) \\ \underset{J=\{j,k\}\,\forall k \ne j}{\overset{\text{L. 2.37 5)}}{\le}} &F_j(b_j) - F_j(a_j) \underset{J=\{j\}}{\overset{\text{L. 2.37 5)}}{=}} |F_j(b_j) - F_j(a_j)|; \end{aligned}$$

to see the inequality, apply $b_1 \to \infty, \dots, b_{j-1} \to \infty, a_{j+1} \to \infty, \dots, a_d \to \infty$. Similarly for $a_j > b_j$. $\square$

- In particular, if the margins $F_1, \dots, F_d$ of F are all continuous, then so is F.
- If $F_1, \dots, F_d$ are absolutely continuous, then, since Riemann integrable functions are bounded on compact intervals, we have, $\forall\, \boldsymbol{x} \in \mathbb{R}^d$, $\boldsymbol{h} \geq \boldsymbol{0}$, that

$$
\begin{aligned}
|F(\boldsymbol{x}) - F(\boldsymbol{x} - \boldsymbol{h})| &\underset{\text{L. 2.50}}{\leq} \sum_{j=1}^d |F_j(x_j) - F_j(x_j - h_j)| \\
&\underset{\text{L. 2.37 5)}}{=} \sum_{j=1}^d (F_j(x_j) - F_j(x_j - h_j)) \underset{F_j \text{ abs. cont.}}{=} \sum_{j=1}^d \int_{x_j - h_j}^{x_j} f_j(z_j)\,\mathrm{d}z_j \\
&\leq \sum_{j=1}^d M_{x_j} h_j \underset{\boldsymbol{h} \to \boldsymbol{0}+}{\to} 0,
\end{aligned}
$$

  so F is left-continuous and thus continuous. Therefore, every F with continuous margins is continuous. The mixed-type df F of R. 2.49 6) is not continuous and thus not absolutely continuous (even though F' exists a.e., namely everywhere except in two points).

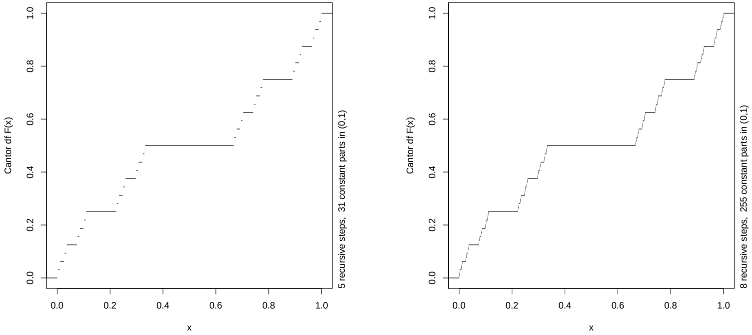
Question: Is every continuous df F also absolutely continuous?

Example 2.51 (Cantor df, the devil's staircase)

Recall that $\mathcal{C} = \{x \in [0,1] : x = \sum_{i=1}^{\infty} a_i 3^{-i},\ a_i \in \{0,2\}\ \forall i \in \mathbb{N}\}$. Let

$$F(x) = \begin{cases} \sum_{i=1}^{\infty} \frac{a_i}{2} 2^{-i}, & x = \sum_{i=1}^{\infty} a_i 3^{-i} \in \mathcal{C},\ a_i \in \{0,2\}, \\ \sup_{z \le x,\, z \in \mathcal{C}} F(z), & x \in \mathcal{C}^c = [0,1]\setminus\mathcal{C}. \end{cases}$$

As the first case contains all base-2 expansions, we have $F(\mathcal{C}) = [0,1]$. **Sketches:**



$F \nearrow \underset{\mathrm{ran}(F)=[0,1]}{\Rightarrow}$ F is continuous. However, F cannot be absolutely continuous as a density candidate f would need to satisfy $f|_{\mathcal{C}^c} = 0$ (by construction) and thus $\int_{[0,1]} f(z)\,\mathrm{d}z = \int_{\mathcal{C}} f(z)\,\mathrm{d}z + \int_{\mathcal{C}^c} f(z)\,\mathrm{d}z \underset{\text{later}}{\overset{\text{null set}}{=}} 0 + 0 = 0 \neq 1$ ↯.

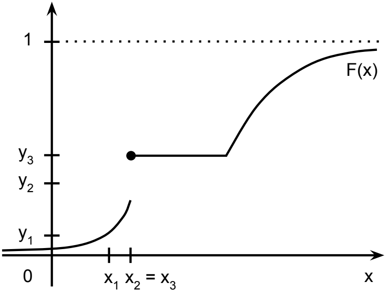
A df F on $\mathbb{R}$ does not necessarily have an ordinary inverse as F neither needs to be strictly increasing, nor continuous, but F always has a generalized inverse, which uniquely characterizes F.

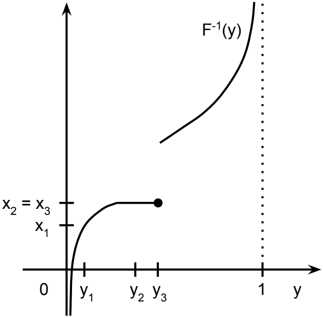
Definition 2.52 (Generalized inverse, quantile function)

If $F : \mathbb{R} \to \mathbb{R}$ is increasing, the *generalized inverse* F^{-1} of F is defined by

$$F^{-1}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}, \quad y \in \mathbb{R}.$$

If F is a df, F^{-1} is the *quantile function (qf)* of F.





- The graph of F^{-1} is obtained by mirroring the graph of F at $y = x$. Jumps (flat parts) of F correspond to flat parts (jumps) of F^{-1}.
- F^{-1} thus uniquely characterizes F.
- One can show (see e.g. Embrechts and Hofert (2013)):
  - $F^{-1} \nearrow$ and left-continuous.
  - If $F \uparrow$, continuous $\Rightarrow F^{-1}$ ordinary inverse of F ($F^{-1}(y) = x$ iff $y = F(x)$).
  - One can often work with F^{-1} as with the ordinary inverse (which we will do), but be careful. E.g. in the above sketch of F and its F^{-1}, we have

$$F(F^{-1}(y_1)) = y_1 \text{ but } \forall\, y \in [F(x_2-), y_3),\ F(F^{-1}(y)) = y_3 > y,$$

so, unlike the ordinary inverse, $F(F^{-1}(y)) = y$ isn't always true.

The following lemma lists useful properties of generalized inverses.

Lemma 2.53 (Properties of generalized inverses)

Let $T : \mathbb{R} \to \mathbb{R}$ be increasing with $T(-\infty) = \lim_{x\to-\infty} T(x)$ and $T(\infty) = \lim_{x\to\infty} T(x)$, and let $x, y \in \mathbb{R}$. Then:

1) $T^{-1}(T(x)) \le x$. If T is strictly increasing, $T^{-1}(T(x)) = x$.

2) Let T be right-continuous. Then $T^{-1}(y) < \infty$ implies $T(T^{-1}(y)) \geq y$. Furthermore, $y \in \operatorname{ran}(T) \cup \{\inf T, \sup T\}$ implies $T(T^{-1}(y)) = y$. Moreover, if $y < \inf T$ then $T(T^{-1}(y)) > y$ and if $y > \sup T$ then $T(T^{-1}(y)) < y$.
3) $T(x) \geq y$ implies $x \geq T^{-1}(y)$. The other implication holds if T is right-continuous. Furthermore, $T(x) < y$ implies $x \leq T^{-1}(y)$.
4) T is continuous if and only if T^{-1} is strictly increasing on $[\inf T, \sup T]$. T is strictly increasing if and only if T^{-1} is continuous on $\operatorname{ran}(T)$.

Proof. See Embrechts and Hofert (2013). □

References

Bauer, H. (2001), Measure and Integration Theory, de Gruyter.

Durrett, R. (2019), Probability: Theory and Examples, 5th ed., Duxbury Press.

Embrechts, P. and Hofert, M. (2013), A note on generalized inverses, *Mathematical Methods of Operations Research*, 77(3), 423–432, `doi:10.1007/s00186-013-0436-7`.

Folland, G. B. (1999), Real Analysis: Modern Techniques and Their Applications, Wiley-Interscience.

Klenke, A. (2008), Probability Theory: A Comprehensive Course, Springer.

Schilling, R. (2006), Measures, integrals and martingales, Cambridge University Press, ISBN 978-0521850155.