Course: Theory of Probability I
Term: Fall 2013
Instructor: Gordan Zitkovic

Lecture 1
Measurable spaces

Families of Sets

Definition 1.1 (Order properties)**.** A countable[1] family $\{A_n\}_{n\in\mathbb{N}}$ of subsets of a non-empty set S is said to be

1. **increasing** if $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$,
2. **decreasing** if $A_n \supseteq A_{n+1}$ for all $n \in \mathbb{N}$,
3. **pairwise disjoint** if $A_n \cap A_m = \emptyset$ for $m \neq n$,
4. a **partition** of S if $\{A_n\}_{n\in\mathbb{N}}$ is pairwise disjoint and $\cup_n A_n = S$.

[1] To make sure we are all on the same page, let us fix some notation and terminology:

- $\subseteq$ denotes a subset (not necessarily proper).
- A set A is said to be **countable** if there exists an injection (one-to-one mapping) from A into $\mathbb{N}$. Note that finite sets are also countable. Sets which are not countable are called **uncountable**.
- For two functions $f : B \to C$, $g : A \to B$, the **composition** $f \circ g : A \to C$ of f and g is given by
$$(f \circ g)(x) = f(g(x)),$$
for all $x \in A$.
- $\{A_n\}_{n\in\mathbb{N}}$ denotes a sequence. More generally, $(A_\gamma)_{\gamma\in\Gamma}$ denotes a collection indexed by the set Γ.
- We use the notation $A_n \nearrow A$ to denote that the sequence $\{A_n\}_{n\in\mathbb{N}}$ is increasing and $A = \cup_n A_n$. Similarly, $A_n \searrow A$ means that $\{A_n\}_{n\in\mathbb{N}}$ is decreasing and $A = \cap_n A_n$.

Here is a list of some properties that a family $\mathcal{S}$ of subsets of a nonempty set S can have:

(A1) $\emptyset \in \mathcal{S}$,

(A2) $S \in \mathcal{S}$,

(A3) $A \in \mathcal{S} \Rightarrow A^c \in \mathcal{S}$,

(A4) $A, B \in \mathcal{S} \Rightarrow A \cup B \in \mathcal{S}$,

(A5) $A, B \in \mathcal{S}$, $A \subseteq B \Rightarrow B \setminus A \in \mathcal{S}$,

(A6) $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$,

(A7) $A_n \in \mathcal{S}$ for all $n \in \mathbb{N} \Rightarrow \cup_n A_n \in \mathcal{S}$,

(A8) $A_n \in \mathcal{S}$, for all $n \in \mathbb{N}$ and $A_n \nearrow A$ implies $A \in \mathcal{S}$,

(A9) $A_n \in \mathcal{S}$, for all $n \in \mathbb{N}$ and $\{A_n\}_{n\in\mathbb{N}}$ is pairwise disjoint implies $\cup_n A_n \in \mathcal{S}$,

Definition 1.2 (Families of sets)**.** A family $\mathcal{S}$ of subsets of a non-empty set S is called an

1. **algebra** if it satisfies (A1),(A3) and (A4),

2. σ**-algebra** if it satisfies (A1), (A3) and (A7)
3. π**-system** if it satisfies (A6),
4. λ**-system** if it satisfies (A2), (A5) and (A8).

Problem 1.1. Show that:

1. Every σ-algebra is an algebra.
2. Each algebra is a π-system and each σ-algebra is an algebra and a λ-system.
3. A family $\mathcal{S}$ is a σ-algebra if and only if it satisfies (A1), (A3), (A6) and (A9).
4. A λ-system which is a π-system is also a σ-algebra.
5. There are π-systems which are not algebras.
6. There are algebras which are not σ-algebras
7. There are λ-systems which are not π-systems.

Hint: Pick all finite subsets of an infinite set. That is not an algebra yet, but sets can be added to it so as to become an algebra which is not a σ-algebra.

Definition 1.3 (Generated σ-algebras)**.** For a family $\mathcal{A}$ of subsets of a non-empty set S, the intersection of all σ-algebras on S that contain $\mathcal{A}$ is denoted by $\sigma(\mathcal{A})$ and is called the σ**-algebra generated** [2] **by** $\mathcal{A}$.

Problem 1.2. Show, by means of an example, that the *union* of a family of algebras (on the same S) does not need to be an algebra. Repeat for σ-algebras, π-systems and λ-systems.

Definition 1.4 (Topology)**.** A **topology**[3] on a set S is a family τ of subsets of S which contains $\emptyset$ and S and is closed under finite intersections and arbitrary (countable or uncountable!) unions. The elements of τ are often called the **open sets**. A set S on which a topology is chosen (i.e., a pair (S, τ) of a set and a topology on it) is called a **topological space**.

Definition 1.5 (Borel σ-algebras)**.** If (S, τ) is a topological space, then the σ-algebra $\sigma(\tau)$, generated by all open sets, is called the **Borel** σ-algebra on (S, τ) or, less precisely[4], the Borel σ-algebra on S.

Example 1.6. Some important σ-algebras. Let S be a non-empty set:

1. The set $\mathcal{S} = 2^S$ (also denoted by $\mathcal{P}(S)$) consisting of all subsets of $\mathcal{S}$ is a σ-algebra.
2. At the other extreme, the family $\mathcal{S} = \{\emptyset, S\}$ is the smallest σ-algebra on S. It is called the **trivial** σ-algebra on $\mathcal{S}$.

[2] Since the family 2^S of *all* subsets of S is a σ-algebra, the concept of a generated σ-algebra is well defined: there is always at least one σ-algebra containing $\mathcal{A}$ - namely 2^S. $\sigma(\mathcal{A})$ is itself a σ-algebra (why?) and it is the smallest (in the sense of set inclusion) σ-algebra that contains $\mathcal{A}$. In the same vein, one can define the algebra, the π-system and the λ-system generated by $\mathcal{A}$. The only important property is that intersections of σ-algebras, π-systems and λ-systems are themselves σ-algebras, π-systems and λ-systems.

[3] Almost all topologies in these notes will be generated by a metric, i.e., a set $A \subset S$ will be open if and only if for each $x \in A$ there exists $\varepsilon > 0$ such that $\{y \in S : d(x, y) < \varepsilon\} \subseteq A$. The prime example is $\mathbb{R}$ where a set is declared open if it can be represented as a union of open intervals.

[4] We often abuse terminology and call S itself a topological space, if the topology τ on it is clear from the context. In the same vein, we often speak of the Borel σ-algebra on a set S.

3. The set $\mathcal{S}$ of all subsets of S which are either countable or whose complements are countable is a σ-algebra. It is called the **countable-cocountable σ-algebra** and is the smallest σ-algebra on S which contains all singletons, i.e., for which $\{x\} \in \mathcal{S}$ for all $x \in S$.

4. The Borel σ-algebra on $\mathbb{R}$ (generated by all open sets as defined by the Euclidean metric on $\mathbb{R}$), is denoted by $\mathcal{B}(\mathbb{R})$.

Problem 1.3. Show that the $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{A})$, for any of the following choices of the family $\mathcal{A}$:

1. $\mathcal{A} = \{\text{all open subsets of } \mathbb{R}\}$,
2. $\mathcal{A} = \{\text{all closed subsets of } \mathbb{R}\}$,
3. $\mathcal{A} = \{\text{all open intervals in } \mathbb{R}\}$,
4. $\mathcal{A} = \{\text{all closed intervals in } \mathbb{R}\}$,
5. $\mathcal{A} = \{\text{all left-closed right-open intervals in } \mathbb{R}\}$,
6. $\mathcal{A} = \{\text{all left-open right-closed intervals in } \mathbb{R}\}$,
7. $\mathcal{A} = \{\text{all open intervals in } \mathbb{R} \text{ with rational end-points}\}$, and
8. $\mathcal{A} = \{\text{all intervals of the form } (-\infty, r], \text{ where } r \text{ is rational}\}$.

Hint: An arbitrary open interval $I = (a, b)$ in $\mathbb{R}$ can be written as $I = \cup_{n\in\mathbb{N}}[a + n^{-1}, b - n^{-1}]$.

Measurable mappings

Definition 1.7 (Measurable spaces)**.** A pair $(S, \mathcal{S})$ consisting of a non-empty set S and a σ-algebra $\mathcal{S}$ of its subsets is called a **measurable space**[5].

[5] If $(S, \mathcal{S})$ is a measurable space, and $A \in \mathcal{S}$, we often say that A **is measurable in** $\mathcal{S}$.

Definition 1.8 (Pull-backs and push-forwards)**.** For a function $f : S \to T$ and subsets $A \subseteq S$, $B \subseteq T$, we define the

1. **push-forward** $f(A)$ of $A \subseteq S$ as
$$f(A) = \{f(x) : x \in A\} \subseteq T,$$
2. **pull-back**[6] $f^{-1}(B)$ of $B \subseteq T$ as
$$f^{-1}(B) = \{x \in S : f(x) \in B\} \subseteq S.$$

[6] It is often the case that the notation is abused and the pull-back of B under f is denoted simply by $\{f \in B\}$. This notation presupposes, however, that the domain of f is clear from the context.

Problem 1.4. Show that the pull-back operation preserves the elementary set operations, i.e., for $f : S \to T$, and $B, \{B_n\}_{n\in\mathbb{N}} \subseteq T$,

1. $f^{-1}(T) = S$, $f^{-1}(\emptyset) = \emptyset$,
2. $f^{-1}(\cup_n B_n) = \cup_n f^{-1}(B_n)$,
3. $f^{-1}(\cap_n B_n) = \cap_n f^{-1}(B_n)$, and
4. $f^{-1}(B^c) = [f^{-1}(B)]^c$.

Note: The assumption that the families in 2. and 3. above are countable is not necessary. Uncountable unions or intersections commute with the pull-back, too.

Give examples showing that the push-forward analogues of the statements 1., 3. and 4. above are not true.

Definition 1.9 (Measurability). A mapping $f : S \to T$, where $(S, \mathcal{S})$ and $(T, \mathcal{T})$ are measurable spaces, is said to be $(\mathcal{S}, \mathcal{T})$-**measurable**[7] if $f^{-1}(B) \in \mathcal{S}$ for each $B \in \mathcal{T}$.

[7] When $T = \mathbb{R}$, we tacitly assume that the Borel σ-algebra is defined on T, and we simply call f *measurable*. In particular, a function $f : \mathbb{R} \to \mathbb{R}$, which is measurable with respect to the pair of the Borel σ-algebras is often called a **Borel function**.

Proposition 1.10 (A measurability criterion). *Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ be two measurable spaces, and let $\mathcal{C}$ be a subset of $\mathcal{T}$ such that $\mathcal{T} = \sigma(\mathcal{C})$. If $f : S \to T$ is a mapping with the property that $f^{-1}(C) \in \mathcal{S}$, for any $C \in \mathcal{C}$, then f is $(\mathcal{S}, \mathcal{T})$-measurable.*

Proof. Let $\mathcal{D}$ be the family of subsets of T defined by

$$\mathcal{D} = \{B \subset T : f^{-1}(B) \in \mathcal{S}\}.$$

By the assumptions of the proposition, we have $\mathcal{C} \subseteq \mathcal{D}$. On the other hand, by Problem 1.4, the family $\mathcal{D}$ has the structure of the σ-algebra, i.e., $\mathcal{D}$ is a σ-algebra that contains $\mathcal{C}$. Remembering that $\mathcal{T} = \sigma(\mathcal{C})$ is the *smallest* σ-algebra that contains $\mathcal{C}$, we conclude that $\mathcal{T} \subseteq \mathcal{D}$. Consequently, $f^{-1}(B) \in \mathcal{S}$ for all $B \in \mathcal{T}$. □

Problem 1.5. Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ be measurable spaces.

1. Suppose that S and T are topological spaces, and that $\mathcal{S}$ and $\mathcal{T}$ are the corresponding Borel σ-algebras. Show that each continuous function $f : S \to T$ is $(\mathcal{S}, \mathcal{T})$-measurable.
2. For a function $f : S \to \mathbb{R}$, show that f is measurable if and only if
$$\{x \in S : f(x) \leq q\} \in \mathcal{S}, \text{ for all rational } q.$$
3. Find an example of $(S, \mathcal{S})$, $(T, \mathcal{T})$ and a measurable function $f : S \to T$ such that $f(A) = \{f(x) : x \in A\} \notin \mathcal{T}$ for all nonempty $A \in \mathcal{S}$.

Hint: Remember that the function f is continuous if the pull-backs of open sets are open.

Proposition 1.11 (Compositions of measurable maps). *Let $(S, \mathcal{S})$, $(T, \mathcal{T})$ and $(U, \mathcal{U})$ be measurable spaces, and let $f : S \to T$ and $g : T \to U$ be measurable functions. Then the composition $h = g \circ f : S \to U$, given by $h(x) = g(f(x))$ is $(\mathcal{S}, \mathcal{U})$-measurable.*

Proof. It is enough to observe that $h^{-1}(B) = f^{-1}(g^{-1}(B))$, for any $B \subseteq U$. □

Corollary 1.12 (Compositions with a continuous maps). *Let $(S, \mathcal{S})$ be a measurable space, T be a topological space and $\mathcal{T}$ the Borel σ-algebra on T. Let $g : T \to \mathbb{R}$ be a continuous function. Then the map $g \circ f : S \to \mathbb{R}$ is measurable for each measurable function $f : S \to T$.*

Definition 1.13 (Generation by several functions)**.** Let[8] $(f_\gamma)_{\gamma\in\Gamma}$ be a family of maps from a set S into a measurable space $(T,\mathcal{T})$. The **σ-algebra generated by** $(f_\gamma)_{\gamma\in\Gamma}$, denoted by $\sigma\Big((f_\gamma)_{\gamma\in\Gamma}\Big)$, is the intersection of all σ-algebras on S which make each f_γ, $\gamma\in\Gamma$, measurable.

[8] The letter Γ will typically be used to denote an abstract index set - we only assume that it is nonempty, but make no other assumptions about its cardinality.

Problem 1.6. In the setting of Definition 1.13, show that

$$\sigma\Big((f_\gamma)_{\gamma\in\Gamma}\Big) = \sigma\Big(\bigcup_{\gamma\in\Gamma} f_\gamma^{-1}(\mathcal{T})\Big),$$

where $f_\gamma^{-1}(\mathcal{T}) = \{f_\gamma^{-1}(B) : B\in\mathcal{T}\}$.

Products of measurable spaces

Definition 1.14 (Products, choice functions)**.** Let $(S_\gamma)_{\gamma\in\Gamma}$ be a family of sets, parametrized by some (possibly uncountable) index set Γ. The **product**[9] $\prod_{\gamma\in\Gamma} S_\gamma$ is the set of all functions $\boldsymbol{s}:\Gamma\to\cup_\gamma S_\gamma$ (called **choice functions**[10]) with the property that $\boldsymbol{s}(\gamma)\in S_\gamma$.

[9] When Γ is finite, each function $\boldsymbol{s}:\Gamma\to\cup_{\gamma\in\Gamma}S_\gamma$ can be identified with an ordered "tuple" $(\boldsymbol{s}(\gamma_1),\dots,\boldsymbol{s}(\gamma_n))$, where n is the cardinality (number of elements) of Γ, and $\gamma_1,\dots,\gamma_n$ is some ordering of its elements. With this identification, it is clear that our definition of a product coincides with the well-known definition in the finite case.

[10] The celebrated *Axiom of Choice* in set theory postulates that no matter what the family $(S_\gamma)_{\gamma\in\Gamma}$ is, there exists at least one choice function. In other words, axiom of choice simply asserts that products of sets are non-empty.

Definition 1.15 (Natural projections)**.** For $\gamma_0\in\Gamma$, the function $\pi_{\gamma_0}:\prod_{\gamma\in\Gamma}S_\gamma\to S_{\gamma_0}$ defined by

$$\pi_{\gamma_0}(\boldsymbol{s}) = \boldsymbol{s}(\gamma_0),\ \text{for } \boldsymbol{s}\in\prod_{\gamma\in\Gamma}S_\gamma,$$

is called the **(natural) projection onto the coordinate** γ_0.

Definition 1.16 (Products of measurable spaces)**.** Let $\{(S_\gamma,\mathcal{S}_\gamma)\}_{\gamma\in\Gamma}$ be a family of measurable spaces. The **product** $\otimes_{\gamma\in\Gamma}(S_\gamma,\mathcal{S}_\gamma)$ is a measurable space $(\prod_{\gamma\in\Gamma}S_\gamma,\otimes_{\gamma\in\Gamma}\mathcal{S}_\gamma)$, where $\otimes_{\gamma\in\Gamma}\mathcal{S}_\gamma$ is the smallest σ-algebra that makes all natural projections $(\pi_\gamma)_{\gamma\in\Gamma}$ measurable.

Example 1.17. When Γ is finite, the above definition can be made more intuitive. Suppose, just for simplicity, that $\Gamma=\{1,2\}$, so that $(S_1,\mathcal{S}_1)\otimes(S_2,\mathcal{S}_2)$ is a measurable space of the form $(S_1\times S_2,\mathcal{S}_1\otimes\mathcal{S}_2)$, where $\mathcal{S}_1\otimes\mathcal{S}_2$ is the smallest σ-algebra on $S_1\times S_2$ which makes both π_1 and π_2 measurable. The pull-backs under π_1 of sets in $\mathcal{S}_1$ are given by

$$\pi^{-1}(B_1) = \{(x,y)\in S_1\times S_2 : x\in B_1\} = B_1\times S_2,\ \text{for } B_1\in\mathcal{S}_1.$$

Similarly

$$\pi^{-1}(B_2) = S_1\times B_2,\ \text{for } B_2\in\mathcal{S}_2.$$

Therefore, by Problem 1.6,

$$\mathcal{S}_1\otimes\mathcal{S}_2 = \sigma\Big(\{B_1\times S_2, S_1\times B_2 : B_1\in\mathcal{S}_1, B_2\in\mathcal{S}_2\}\Big).$$

Equivalently (why?)

$$\mathcal{S}_1 \otimes \mathcal{S}_2 = \sigma\Big(\{B_1 \times B_2 \,:\, B_1 \in \mathcal{S}_1, B_2 \in \mathcal{S}_2\}\Big).$$

In a completely analogous fashion, we can show that, for finitely many measurable spaces $(S_1, \mathcal{S}_1), \dots, (S_n, \mathcal{S}_n)$, we have

$$\bigotimes_{i=1}^{n} \mathcal{S}_i = \sigma\Big(\{B_1 \times B_2 \times \dots \times B_n \,:\, B_1 \in \mathcal{S}_1, B_2 \in \mathcal{S}_2, \dots, B_n \in \mathcal{S}_n\}\Big)$$

The same goes for countable products[11].

[11] Uncountable products, however, behave very differently.

Problem 1.7. We know that the Borel σ-algebra (based on the usual Euclidean topology) can be constructed on each $\mathbb{R}^n$. A σ-algebra on $\mathbb{R}^n$ (for $n > 1$), can also be constructed as a product σ-algebra $\otimes_{i=1}^n \mathcal{B}(\mathbb{R})$. A third possibility is to consider the mixed case where $1 < m < n$ is picked and the σ-algebra $\mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^{n-m})$ is constructed on $\mathbb{R}^n$ (which is now interpreted as a product of $\mathbb{R}^m$ and $\mathbb{R}^{n-m}$). Show that we get the same σ-algebra in all three cases.

Problem 1.8. Let $(P, \mathcal{P})$, $\{(S_\gamma, \mathcal{S}_\gamma)\}_{\gamma\in\Gamma}$ be measurable spaces and set $S = \prod_{\gamma\in\Gamma} S_\gamma$, $\mathcal{S} = \otimes_{\gamma\in\Gamma}\mathcal{S}_\gamma$. Prove that a map $f : P \to S$ is $(\mathcal{P}, \mathcal{S})$-measurable if and only if the composition $\pi_\gamma \circ f : P \to S_\gamma$ is $(\mathcal{P}, \mathcal{S}_\gamma)$ measurable for each $\gamma \in \Gamma$.

Note: Loosely speaking, this result states that a "vector"-valued mapping is measurable if and only if all of its components are measurable.

Definition 1.18 (Cylinder sets)**.** Let $\{(S_\gamma, \mathcal{S}_\gamma)\}_{\gamma\in\Gamma}$ be a family of measurable spaces, and let $(\prod_{\gamma\in\Gamma} S_\gamma, \otimes_{\gamma\in\Gamma}\mathcal{S}_\gamma)$ be its product. A subset $C \subseteq \prod_{\gamma\in\Gamma} S_\gamma$ is called a **cylinder set** if there exist a finite subset $\{\gamma_1, \dots, \gamma_n\}$ of Γ, as well as a measurable set $B \in \mathcal{S}_{\gamma_1} \otimes \mathcal{S}_{\gamma_2} \otimes \dots \otimes \mathcal{S}_{\gamma_n}$ such that

$$C = \{\boldsymbol{s} \in \prod_{\gamma\in\Gamma} S_\gamma \,:\, (\boldsymbol{s}(\gamma_1), \dots, \boldsymbol{s}(\gamma_n)) \in B\}.$$

A cylinder set for which the set B can be chosen of the form $B = B_1 \times \dots \times B_n$, for some $B_1 \in \mathcal{S}_1, \dots, B_n \in \mathcal{S}_n$ is called a **product cylinder set**. In that case

$$C = \{\boldsymbol{s} \in \prod_{\gamma\in\Gamma} S_\gamma \,:\, (\boldsymbol{s}(\gamma_1) \in B_1, \boldsymbol{s}(\gamma_2) \in B_2, \dots, \boldsymbol{s}(\gamma_n) \in B_n\}.$$

Problem 1.9.

1. Show that the family of product cylinder sets generates the product σ-algebra.
2. Show that (not-necessarily-product) cylinders are measurable in the product σ-algebra.

3. Which of the 4 families of sets from Definition 1.2 does the collection of all product cylinders belong to in general? How about (not-necessarily-product) cylinders?

Example 1.19. The following example will play a major role in probability theory. Hence the name **coin-toss space**. Here $\Gamma = \mathbb{N}$ and for $i \in \mathbb{N}$, $(S_i, \mathcal{S}_i)$ is the discrete two-element space $S_i = \{-1, 1\}$, $\mathcal{S}_i = 2^{S_1}$. The product $\prod_{i\in\mathbb{N}} S_i = \{-1,1\}^{\mathbb{N}}$ can be identified with the set of all sequences $\boldsymbol{s} = (s_1, s_2, \dots)$, where $s_i \in \{-1, 1\}$, $i \in \mathbb{N}$. For each cylinder set C, there exists (why?) $n \in \mathbb{N}$ and a subset B of $\{-1,1\}^n$ such that

$$C = \{\boldsymbol{s} = (s_1, \dots, s_n, s_{n+1}, \dots) \in \{-1,1\}^{\mathbb{N}} : (s_1, \dots, s_n) \in B\}.$$

The product cylinders are even simpler - they are always of the form $C = \{-1,1\}^{\mathbb{N}}$ or $C = C_{n_1,\dots,n_k;b_1,\dots,b_k}$, where

$$C_{n_1,\dots,n_k;b_1,\dots,b_k} = \left\{\boldsymbol{s} = (s_1, s_2, \dots) \in \{-1,1\}^{\mathbb{N}} : s_{n_1} = b_1, \dots, s_{n_k} = b_k\right\},$$

for some $k \in \mathbb{N}$, $1 \le n_1 < n_2 < \dots < n_k \in \mathbb{N}$ and $b_1, b_2, \dots, b_k \in \{-1, 1\}$.

We know that the σ-algebra $\mathcal{S} = \otimes_{i\in\mathbb{N}} \mathcal{S}_i$ is generated by all projections $\pi_i : \{-1,1\}^{\mathbb{N}} \to \{-1,1\}$, $i \in \mathbb{N}$, where $\pi_i(\boldsymbol{s}) = s_i$. Equivalently, by Problem 1.9, $\mathcal{S}$ is generated by the collection of all cylinder sets.

Problem 1.10. One can obtain the product σ-algebra $\mathcal{S}$ on $\{-1,1\}^{\mathbb{N}}$ as the Borel σ-algebra corresponding to a particular topology which makes $\{-1,1\}^{\mathbb{N}}$ compact. Here is how. Start by defining a mapping $d : \{-1,1\}^{\mathbb{N}} \times \{-1,1\}^{\mathbb{N}} \to [0, \infty)$ by

$$d(\boldsymbol{s}^1, \boldsymbol{s}^2) = 2^{-i(\boldsymbol{s}^1, \boldsymbol{s}^2)}, \text{ where } i(\boldsymbol{s}^1, \boldsymbol{s}^2) = \inf\{i \in \mathbb{N} : s^1_i \neq s^2_i\}, \quad (1.1)$$

for $\boldsymbol{s}^j = (s^j_1, s^j_2, \dots)$, $j = 1, 2$.

1. Show that d is a metric on $\{-1,1\}^{\mathbb{N}}$.
2. Show that $\{-1,1\}^{\mathbb{N}}$ is compact under d.

Hint: Use the diagonal argument.

3. Show that each cylinder of $\{-1,1\}^{\mathbb{N}}$ is both open and closed under d.
4. Show that each open ball is a cylinder.
5. Show that $\{-1,1\}^{\mathbb{N}}$ is separable, i.e., it admits a countable dense subset.
6. Conclude that $\mathcal{S}$ coincides with the Borel σ-algebra on $\{-1,1\}^{\mathbb{N}}$ under the metric d.

Real-valued measurable functions

Let $\mathcal{L}^0(S, \mathcal{S}; \mathbb{R})$ (or, simply, $\mathcal{L}^0(S; \mathbb{R})$ or $\mathcal{L}^0(\mathbb{R})$ or $\mathcal{L}^0$ when the domain $(S, \mathcal{S})$ or the co-domain $\mathbb{R}$ are clear from the context) be the set of all $\mathcal{S}$-measurable functions $f : S \to \mathbb{R}$. The set of non-negative measurable functions is denoted by $\mathcal{L}^0_+$ or $\mathcal{L}^0([0, \infty))$.

Proposition 1.20 (Measurable functions form a vector space)**.** $\mathcal{L}^0$ *is a vector space, i.e.*

$$\alpha f + \beta g \in \mathcal{L}^0, \text{ whenever } \alpha, \beta \in \mathbb{R},\ f, g \in \mathcal{L}^0.$$

Proof. Let us define a mapping $F : S \to \mathbb{R}^2$ by $F(x) = (f(x), g(x))$. By Problem 1.7, the Borel σ-algebra on $\mathbb{R}^2$ is the same as the product σ-algebra when we interpret $\mathbb{R}^2$ as a product of two copies of $\mathbb{R}$. Therefore, since its compositions with the coordinate projections are precisely the functions f and g, Problem 1.8 implies that F is $(\mathcal{S}, \mathcal{B}(\mathbb{R}^2))$-measurable.

Consider the function $\phi : \mathbb{R}^2 \to \mathbb{R}$ given by $\phi(x, y) = \alpha x + \beta y$. It is linear, and, therefore, continuous. By Corollary 1.12, the composition $\phi \circ F : \mathcal{S} \to \mathbb{R}$ is $(\mathcal{S}, \mathcal{B}(\mathbb{R}))$-measurable, and it only remains to note that

$$(\phi \circ F)(x) = \phi(F(x)) = \alpha f(x) + \beta g(x), \text{ i.e., } \phi \circ F = \alpha f + \beta g. \qquad \square$$

Proposition 1.21 (Products and maxima preserve measurability)**.** *Let* f, g *be in* $\mathcal{L}^0$*. Then* fg, $\max(f, g)$ *and* $\min(f, g)$ *belong to* $\mathcal{L}^0$.

Proof. The functions $(x, y) \mapsto \max(x, y)$ and $(x, y) \to xy$ are continuous from $\mathbb{R}^2$ to $\mathbb{R}$. $\square$

Problem 1.11. Suppose that $f \in \mathcal{L}^0$ has the property that $f(x) \neq 0$ for all $x \in S$. Then the function $1/f$ is also in $\mathcal{L}^0$.

Hint: Find a measurable function g, defined on $\mathbb{R}$ such that $1/f(x) = g(f(x))$, for all $x \in S$.

Definition 1.22 (The indicator function of a set)**.** For $A \subseteq S$, the **indicator function** $\mathbf{1}_A$ is defined by

$$\mathbf{1}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

Despite their simplicity, indicators will be extremely useful throughout these notes.

Problem 1.12. Show that for $A \subset S$, we have $A \in \mathcal{S}$ if and only if $\mathbf{1}_A \in \mathcal{L}^0$.

Mathematical structures and measurability

A measurable structure on a set gives relates to various topological and algebraic structures encountered throughout mathematics.

Since it contains the function 1 and the products of pairs of its elements, the set $\mathcal{L}^0$ has the structure of an *unital algebra* (not to be confused with the algebra of sets defined above). It is true, however, that any algebra $\mathcal{A}$ of subsets of a non-empty set S, together with the operations of union, intersection and complement forms a *Boolean algebra*. Alternatively, it can be given the (algebraic) structure of a *commutative ring with a unit*. Indeed, under the operation $\triangle$ of symmetric difference, $\mathcal{A}$ is an Abelian group (prove that!). If, in addition, the operation of intersection is introduced in lieu of multiplication, the resulting structure is, indeed, the one of a commutative ring.

Additionally, a natural partial order given by $f \preceq g$ if $f(x) \leq g(x)$, for all $x \in S$, can be introduced on $\mathcal{L}^0$. This order is compatible with the operations of addition and multiplication and has the additional property that each pair $\{f, g\} \subseteq \mathcal{L}^0$ admits a *least upper bound*, i.e., the element $h \in \mathcal{L}^0$ such that $f \preceq h$, $g \preceq h$ and $h \preceq k$, for any other k with the property that $f, g \preceq k$. Indeed, we simply take $h(x) = \max(f(x), g(x))$. A similar statement can be made for a *greatest lower bound*. A vector space with a partial order which satisfies the above properties is called a *vector lattice*.

As we move through the theory, we will see how many other mathematical structures appear naturally in measure theory.

Extended real-valued measurable functions

Since a limit of a sequence of real numbers does not necessarily belong to $\mathbb{R}$, it is often necessary to consider functions which are allowed to take the values ∞ and $-\infty$. The set $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ is called the **extended** set of real numbers. Most (but not all) of the algebraic and topological structure from $\mathbb{R}$ can be lifted to $\bar{\mathbb{R}}$. In some cases there is no unique way to do that, so we choose one of them as a matter of convention.

1. **Arithmetic operations.** For $x, y \in \bar{\mathbb{R}}$, all the arithmetic operations are defined in the usual way when $x, y \in \mathbb{R}$. When one or both are in $\{\infty, -\infty\}$, we use the following convention, where $\oplus \in \{+, -, *, /\}$:

   We define $x \oplus y = z$ if all pairs of sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}$ such that $x = \lim_n x_n$, $y = \lim_n y_n$ and $x_n \oplus y_n$ is well-defined for all $n \in \mathbb{N}$, we have

   $$z = \lim_n (x_n \oplus y_n).$$

Otherwise, $x \oplus y$ is not defined. This basically means that all intuitively obvious conventions (such as $\infty + \infty = \infty$ and $\frac{a}{\infty} = 0$ for $a \in \mathbb{R}$ hold). In measure theory, however, we do make one important *exception* to the above rule. We set

$$0 \times \infty = \infty \times 0 = 0 \times (-\infty) = (-\infty) \times 0 = 0.$$

2. **Order.** $-\infty < x < \infty$, for all $x \in \mathbb{R}$. Also, *each* non-empty subset of $\bar{\mathbb{R}}$ admits a supremum and an infimum in $\bar{\mathbb{R}}$ in an obvious way.

3. **Convergence.** It is impossible to extend the usual (Euclidean) metric from $\mathbb{R}$ to $\bar{\mathbb{R}}$, but a metric $d' : \mathbb{R} \times \mathbb{R} \to [0, \pi)$ given by

$$d'(x, y) = |\arctan(y) - \arctan(x)|,$$

extends readily to a metric on $\bar{\mathbb{R}}$ if we set $\arctan(\infty) = \pi/2$ and $\arctan(-\infty) = -\pi/2$. We define convergence (and topology) on $\mathbb{R}$ using d'. For example, a sequence $\{x_n\}_{n\in\mathbb{N}}$ in $\bar{\mathbb{R}}$ converges to $+\infty$ if

(a) It contains only a finite number of terms equal to $-\infty$,

(b) Every subsequence of $\{x_n\}_{n\in\mathbb{N}}$ whose elements are in $\mathbb{R}$ converges to $+\infty$ (in the usual sense).

We define the notions of **limit superior** and **limit inferior** on $\bar{\mathbb{R}}$ for a sequence $\{x_n\}_{n\in\mathbb{N}}$ in the following manner:

$$\limsup_n x_n = \inf_n S_n, \text{ where } S_n = \sup_{k\geq n} x_k,$$

and

$$\liminf_n x_n = \sup_n I_n, \text{ where } I_n = \inf_{k\geq n} x_k.$$

If you have forgotten how to manipulate limits inferior and superior, here is an exercise to remind you:

Problem 1.13. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in $\bar{\mathbb{R}}$. Prove the following statements:

1. $a \in \bar{\mathbb{R}}$ satisfies $a \geq \limsup_n x_n$ if and only if for any $\varepsilon \in (0, \infty)$ there exists $n_\varepsilon \in \mathbb{N}$ such that $x_n \leq a + \varepsilon$ for $n \geq n_\varepsilon$.

2. $\liminf_n x_n \leq \limsup_n x_n$.

3. Define

$$A = \{\lim_k x_{n_k} : x_{n_k} \text{ is a convergent (in } \bar{\mathbb{R}}) \text{ subsequence of } \{x_n\}_{n\in\mathbb{N}}\}.$$

Show that

$$\{\liminf_n x_n, \limsup_n x_n\} \subseteq A \subseteq [\liminf_n x_n, \limsup_n x_n].$$

Give an example in which both inclusion above are strict.

Having introduced a topology on $\bar{\mathbb{R}}$ we immediately have the σ-algebra $\mathcal{B}(\bar{\mathbb{R}})$ of Borel sets there and the notion of measurability for functions mapping a measurable space $(S, \mathcal{S})$ into $\bar{\mathbb{R}}$.

Problem 1.14. Show that a subset $A \subseteq \bar{\mathbb{R}}$ is in $\mathcal{B}(\bar{\mathbb{R}})$ if and only if $A \setminus \{\infty, -\infty\}$ is Borel in $\mathbb{R}$. Show that a function $f : S \to \bar{\mathbb{R}}$ is measurable in the pair $(\mathcal{S}, \mathcal{B}(\bar{\mathbb{R}}))$ if and only if the sets $f^{-1}(\{\infty\})$, $f^{-1}(\{-\infty\})$ and $f^{-1}(A)$ are in $\mathcal{S}$ for all $A \in \mathcal{B}(\mathbb{R})$ (equivalently and more succinctly, $f \in \mathcal{L}^0(\bar{\mathbb{R}})$ iff $\{f = \infty\}, \{f = -\infty\} \in \mathcal{S}$ and $f\mathbf{1}_{\{f \in \mathbb{R}\}} \in \mathcal{L}^0$).

The set of all measurable functions $f : S \to \bar{\mathbb{R}}$ is denoted by $\mathcal{L}^0(S, \mathcal{S}; \bar{\mathbb{R}})$, and, as always we leave out S and $\mathcal{S}$ when no confusion can arise. The set of extended non-negative measurable functions often plays a role, so we denote it by $\mathcal{L}^0([0, \infty])$ or $\mathcal{L}^0_+(\bar{\mathbb{R}})$. Unlike $\mathcal{L}^0(\mathbb{R})$, $\mathcal{L}^0(\bar{\mathbb{R}})$ is not a vector space, but it retains all the order structure. Moreover, it is particularly useful because, unlike $\mathcal{L}^0(\mathbb{R})$, it is closed with respect to the limiting operations. More precisely, for a sequence $\{f_n\}_{n \in \mathbb{N}}$ in $\mathcal{L}^0(\bar{\mathbb{R}})$, we define the functions $\limsup_n f_n : S \to [-\infty, \infty]$ and $\liminf_n f_n : S \to [-\infty, \infty]$ by

$$(\limsup_n f_n)(x) = \limsup_n f_n(x) = \inf_n \left(\sup_{k \geq n} f_k(x) \right),$$

and

$$(\liminf_n f_n)(x) = \liminf_n f_n(x) = \sup_n \left(\inf_{k \geq n} f_k(x) \right).$$

Then, we have the following result, where the supremum and infimum of a sequence of functions are defined pointwise (just like the limits superior and inferior).

Proposition 1.23 (Limiting operations preserve measurability). *Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence in $\mathcal{L}^0(\bar{\mathbb{R}})$. Then*

1. $\sup_n f_n, \inf_n f_n \in \mathcal{L}^0(\bar{\mathbb{R}})$,
2. $\limsup_n f_n, \liminf_n f_n \in \mathcal{L}^0(\bar{\mathbb{R}})$,
3. *if $f(x) = \lim_n f_n(x)$ exists in $\bar{\mathbb{R}}$ for each $x \in S$, then $f \in \mathcal{L}^0(\bar{\mathbb{R}})$, and*
4. *the set $A = \{\lim_n f_n$ exists in $\bar{\mathbb{R}}\}$ is in $\mathcal{S}$.*

Proof.

1. We show only the statement for the supremum. It is clear that it is enough to show that the set $\{\sup_n f_n \leq a\}$ is in $\mathcal{S}$ for all $a \in (-\infty, \infty]$ (why?). This follows, however, directly from the simple identity

$$\{\sup_n f_n \leq a\} = \cap_n \{f_n \leq a\},$$

and the fact that σ-algebras are closed with respect to countable intersections.

2. Define $g_n = \sup_{k \geq n} f_k$ and use part 1. above to conclude that $g_n \in \mathcal{L}^0(\bar{\mathbb{R}})$ for each $n \in \mathbb{N}$. Another appeal to part 1. yields that $\limsup_n f_n = \inf_n g_n$ is in $\mathcal{L}^0(\bar{\mathbb{R}})$. The statement about the limit inferior follows in the same manner.

3. If the limit $f(x) = \lim_n f_n(x)$ exists for all $x \in S$, then $f = \liminf_n f_n$ which is measurable by part 2. above.

4. The statement follows from the fact that $A = f^{-1}(\{0\})$, where

$$f(x) = \arctan\left(\limsup_n f_n(s)\right) - \arctan\left(\liminf_n f_n(x)\right).$$

Note: The unexpected use of the function arctan is really noting to be puzzled by. The only property needed is its measurability (it is continuous) and monotonicity+bijectivity from $[-\infty, \infty]$ to $[-\pi/2, \pi/2]$. We compose the limits superior and inferior with it so that we don't run into problems while trying to subtract $+\infty$ from itself.

□

Additional Problems

Problem 1.15. Which of the following are σ-algebras on $\mathbb{R}$?

1. $\mathcal{S} = \{A \subseteq \mathbb{R} : 0 \in A\}$.
2. $\mathcal{S} = \{A \subseteq \mathbb{R} : A \text{ is finite}\}$.
3. $\mathcal{S} = \{A \subseteq \mathbb{R} : A \text{ is finite, or } A^c \text{ is finite}\}$.
4. $\mathcal{S} = \{A \subseteq \mathbb{R} : A \text{ is countable or } A^c \text{ is countable}\}$.
5. $\mathcal{S} = \{A \subseteq \mathbb{R} : A \text{ is open}\}$.
6. $\mathcal{S} = \{A \subset \mathbb{R} : A \text{ is open or } A \text{ is closed}\}$.

Problem 1.16. A **partition** of a set S is a family $\mathcal{P}$ of non-empty subsets of S with the property that each $\omega \in S$ belongs to exactly one $A \in \mathcal{P}$.

1. Show that the number of different algebras on a finite set S is equal to the number of different partitions of S.
2. How many algebras are there on the set $S = \{1, 2, 3\}$?
3. Does there exist an algebra with 754 elements?
4. For $N \in \mathbb{N}$, let a_n be the number of different algebras on the set $\{1, 2, \dots, n\}$. Show that $a_1 = 1$, $a_2 = 2$, $a_3 = 5$, and that the following recursion holds (where $a_0 = 1$ by definition),

$$a_{n+1} = \sum_{k=0}^{n} \binom{n}{k} a_k.$$

Note: This number for $S_n = \{1, 2, \dots, n\}$ is called the n^{th} **Bell number** B_n, and no nice closed-form expression for it is known. See below, though.

5. Show that the exponential generating function for the sequence $\{a_n\}_{n\in\mathbb{N}}$ is $f(x) = e^{e^x-1}$, i.e., that

$$\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = e^{e^x-1} \text{ or, equivalently, } a_n = \left(\tfrac{d^n}{dx^n} e^{e^x-1}\right)\Big|_{x=0}.$$

Problem 1.17. Let $(S, \mathcal{S})$ be a measurable space. For $f, g \in \mathcal{L}^0$ show that the sets $\{f = g\} = \{x \in S : f(x) = g(x)\}$, $\{f < g\} = \{x \in S : f(x) < g(x)\}$ are in $\mathcal{S}$.

Problem 1.18. Show that all

1. monotone,
2. convex

functions $f : \mathbb{R} \to \mathbb{R}$ are measurable.

Problem 1.19. Let $(S, \mathcal{S})$ be a measurable space and let $f : S \to \mathbb{R}$ be a Borel-measurable function. Show that the graph

$$G_f = \{(x, y) \in S \times \mathbb{R} : f(x) = y\},$$

of f is a measurable subset in the product space $(S \times \mathbb{R}, \mathcal{S} \otimes \mathcal{B}(\mathbb{R}))$.