University of Cambridge Mathematical Tripos Part II

Probability and Measure

Lecturer:
Pierre Raphaël
`pr463@dpmms.cam.ac.uk`

Notes by:
Robert Frederik Uy
`rfu20@cam.ac.uk`

Course schedule

Measure spaces, σ-algebras, π-systems and uniqueness of extension, statement *and proof* of Carathéodory's extension theorem. Construction of Lebesgue measure on $\mathbb{R}$. The Borel σ-algebra of $\mathbb{R}$. Existence of non-measurable subsets of $\mathbb{R}$. Lebesgue-Stieltjes measures and probability distribution functions. Independence of events, independence of σ-algebras. The Borel–Cantelli lemmas. Kolmogorov's zero-one law. [6]

Measurable functions, random variables, independence of random variables. Construction of the integral, expectation. Convergence in measure and convergence almost everywhere. Fatou's lemma, monotone and dominated convergence, differentiation under the integral sign. Discussion of product measure and Fubini's theorem. [6]

Chebyshev's inequality, tail estimates. Jensen's inequality. Completeness of L^p for $1 \leq p \leq \infty$. The Hölder and Minkowski inequalities, uniform integrability. [4]

L^2 as a Hilbert space. Orthogonal projection, relation with elementary conditional probability. Variance and covariance. Gaussian random variables, the multivariate normal distribution. [2]

The strong law of large numbers, proof for independent random variables with bounded fourth moments. Measure preserving transformations, Bernoulli shifts. Statements *and proofs* of maximal ergodic theorem and Birkhoff's almost everywhere ergodic theorem, proof of the strong law. [4]

The Fourier transform of a finite measure, characteristic functions, uniqueness and inversion. Weak convergence, statement of Lévy's convergence theorem for characteristic functions. The central limit theorem. [2]

Recommended books

Jean-François Le Gall *Measure Theory, Probability, and Stochastic Processes*. Graduate Texts in Mathematics. Springer 2022.

Contents

1 Measure spaces

1.1 Measurable spaces

Definition σ-algebra
Let E be a set. A σ-algebra on E is a family $\mathcal{A} \subset \mathcal{P}(E)$ such that
(i) $E \in \mathcal{A}$
(ii) $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
(iii) *for any sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$*

The elements of $\mathcal{A}$ are called *measurable sets* or sometimes *$\mathcal{A}$-measurable sets.* We then say that $(E, \mathcal{A})$ is a *measurable space.*

Remarks
From the given definition, it is immediate that
- $\varnothing \in \mathcal{A}$
- $\mathcal{A}$ is closed under *finite* unions
- $\mathcal{A}$ is closed under countable *intersections*

Example *Some trivial examples*
- $\mathcal{A} = \mathcal{P}(E)$
- $\mathcal{A} = \{\varnothing, E\}$

The examples above are rather uninteresting. How might we go about generating more interesting σ-algebras?

Lemma 1.1
If $\mathcal{A}_1$ and $\mathcal{A}_2$ are σ-algebras on E, then $\mathcal{A}_1 \cap \mathcal{A}_2$ is also a σ-algebra on E.

Proof. Easy to check. ■

Now, let $\mathcal{C} \subset \mathcal{P}(E)$. From the above definition, we have that $\bigcap_{\substack{\mathcal{A}\ \sigma\text{-algebra} \\ \mathcal{C}\subset\mathcal{A}}} \mathcal{A}$ is a σ-algebra and contains $\mathcal{C}$. In fact, it is the smallest one that contains $\mathcal{C}$.

Definition σ-algebra generated by $\mathcal{C}$
The family

$$\sigma(\mathcal{C}) := \bigcap_{\substack{\mathcal{A}\ \sigma\text{-algebra} \\ \mathcal{C}\subset\mathcal{A}}} \mathcal{A}$$

is called the σ-algebra generated by $\mathcal{C}$.

Example
Let (E, τ) be a topological space. Then $\sigma(\tau)$ is called the *Borel σ-algebra* on E and is denoted by $\mathcal{B}(E)$.

Exercise. Show that $\mathcal{B}(\mathbb{R})$ is generated by the family of intervals $\{(-\infty, a) \colon a \in \mathbb{Q}\}$.

Definition Product σ-algebra
Let $(E_1, \mathcal{A}_1)$ and $(E_2, \mathcal{A}_2)$ be measurable spaces. The product σ-algebra on $E_1 \times E_2$ is

$$\mathcal{A}_1 \times \mathcal{A}_2 := \{A_1 \times A_2 \colon A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$$

Exercise. In the case that E_1 and E_2 are separable topological spaces, show that $\mathcal{B}(E_1) \times \mathcal{B}(E_2) = \mathcal{B}(E_1 \times E_2)$.

1.2 Measures

Definition Measures and measure spaces
Let $(E, \mathcal{A})$ be a measurable space. A measure on E is a map $\mu\colon \mathcal{A} \to [0, \infty]$ such that
(i) $\mu(\varnothing) = 0$
(ii) *for any sequence $(A_n)_{n\in\mathbb{N}}$ of pairwise disjoint sets in $\mathcal{A}$,*

$$\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) \qquad (\sigma\text{-additivity})$$

We then say that $(E, \mathcal{A}, \mu)$ is a measure space.

Remark. Above, we've introduced the *extended non-negative real axis* $[0, \infty]$. By convention, the addition, multiplication and order structures on $[0, \infty)$ are partially extended to $[0, \infty]$ by declaring that

$$\begin{aligned} \infty + x = x + \infty &= \infty & x &\in [0, \infty] \\ \infty \cdot x = x \cdot \infty &= \infty & x &\in (0, \infty] \\ \infty \cdot 0 = 0 \cdot \infty &= 0 & & \\ x &< \infty & x &\in [0, \infty) \end{aligned}$$

Proposition 1.2 Basic properties of μ
(i) *If $A \subset B$, then $\mu(A) \le \mu(B)$. If in addition $\mu(A) < \infty$, then $\mu(B\backslash A) = \mu(B) - \mu(A)$.*
(ii) *For any $A, B \in \mathcal{A}$, $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$.*
(iii) *If $(A_n)_{n\in\mathbb{N}}$ is an increasing sequence in $\mathcal{A}$, then*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) \qquad \text{(continuity from below)}$$

(iv) *If $(B_n)_{n\in\mathbb{N}}$ is a decreasing sequence in $\mathcal{B}$ with $\mu(B_1) < \infty$, then*

$$\mu\left(\bigcap_{n\in\mathbb{N}} B_n\right) = \lim_{n\to\infty} \mu(B_n) \qquad \text{(continuity from above)}$$

(v) *For any sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$,*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n\in\mathbb{N}} \mu(A_n) \qquad (\sigma\text{-subadditivity})$$

Proof. (i): $B = A \sqcup (B\backslash A)$ so $\mu(B) = \mu(A) + \mu(B\backslash A)$.
(ii): $A = (A\backslash B) \sqcup (A \cap B)$, $B = (B\backslash A) \sqcup (A \cap B)$, $A \cup B = (A\backslash B) \sqcup (B\backslash A) \sqcup (A \cap B)$.
(iii): Consider the family $\{C_n\}_{n\in\mathbb{N}}$ in $\mathcal{A}$ given by $C_1 = A_1, C_n = A_n\backslash A_{n-1}$ for $n \ge 2$. Note that $A_n = \sqcup_{k=1}^n C_k$, so $\mu(A_n) = \sum_{k=1}^n \mu(C_k)$. Thus,

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mu\left(\bigsqcup_{k\in\mathbb{N}} C_k\right) = \sum_{k\in\mathbb{N}} \mu(C_k) = \lim_{n\to\infty} \mu(A_n)$$

(iv): Set $A_n = B_1\backslash B_n$. Note that $(A_n)_{n\in\mathbb{N}}$ is increasing. By (iii) and the fact that $\mu(B_1) < \infty$, we have

$$\mu(B_1) - \mu\left(\bigcap_{n\in\mathbb{N}} B_n\right) = \mu\left(B_1\backslash \bigcap_{n\in\mathbb{N}} B_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

But note that $\lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} [\mu(B_1) - \mu(B_n)] = \mu(B_1) - \lim_{n\to\infty} \mu(B_n)$.

(v): Set $C_1 = A_1$ and $C_n = A_1 \setminus \bigcup_{k=1}^{n-1} A_k$ for $n \geq 2$. Note that $(C_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint measurable sets and $\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n} C_k$. In the limit $n \to \infty$, we have

$$\bigcup_{n\in\mathbb{N}} A_n = \bigsqcup_{k\in\mathbb{N}} C_k \Longrightarrow \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{k\in\mathbb{N}} \mu(C_k) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$$

where we have used the fact that, for each $k \in \mathbb{N}$, $\mu(C_k) \leq \mu(A_k)$. ■

Example. As we will show later, there exists a unique measure λ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that, for every closed interval $[a, b]$, we have $\lambda([a, b]) = b - a$. This is called the *Lebesgue measure.*

Now, let's introduce some terminology. Let μ be a measure on $(E, \mathcal{A})$.

- We say that μ is *finite* if $\mu(E) < \infty$.
- We say that μ is a *probability measure* if $\mu(E) = 1$.
- We say that μ is *σ-finite* if there exists a sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ such that $E = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(E_n) < \infty$ for all $n \in \mathbb{N}$.
- We say that $x \in E$ is an *atom* of μ if $\{x\} \in \mathcal{A}$ and $\mu(\{x\}) > 0$.
- We say that μ is *diffuse* if it has no atom
- We say that $B \subset E$ is *negligible* if there exists $A \in \mathcal{A}$ for which $B \subset A$ and $\mu(A) = 0$. We say that $\mathcal{A}$ is *complete* if all negligible sets belong to $\mathcal{A}$.

1.3 Uniqueness of measures

Definition π-system
A family $\mathcal{A} \subset \mathcal{P}(E)$ is a π-system if
 (i) $\varnothing \in \mathcal{A}$
 (ii) *for any $A, B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$*

Definition λ-system
A family $\mathcal{A} \subset \mathcal{P}(E)$ is a λ-system[1] if
 (i) $E \in \mathcal{A}$
 (ii) *for any $A, B \in \mathcal{A}$ with $A \subset B$, $B \setminus A \in \mathcal{A}$*
 (iii) *for any increasing sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$*

Lemma 1.3
$\mathcal{A} \subset \mathcal{P}(E)$ is a σ-algebra if and only if it is both a π-system and a λ-system.

Proof. ($\Longrightarrow$) A σ-algebra is both a π-system and a λ-system.
($\Longleftarrow$) Note that $\mathcal{A}$ contains E and is closed under taking complements. Since it is also closed under finite intersections, we deduce that $\mathcal{A}$ is closed under finite unions. To show that $\mathcal{A}$ is closed under countable unions, fix a sequence $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$, and let $A_1 = B_1$, $A_n = A_{n-1} \cup B_n$ for $n \geq 2$. By closure under finite unions, $A_n \in \mathcal{A}$ for all $n \in \mathbb{N}$. Since $(A_n)_{n\in\mathbb{N}}$ is an increasing sequence in $\mathcal{A}$, we have $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$. Hence, $\mathcal{A}$ is a σ-algebra. ■

Manifestly, if $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{P}(E)$ are both λ-systems, then $\mathcal{A}_1 \cap \mathcal{A}_2$ is also a λ-system. We can thus define the *λ-system generated by* $\mathcal{C} \subset \mathcal{P}(E)$ as

$$\lambda(\mathcal{C}) := \bigcap_{\substack{\mathcal{A}\ \lambda\text{-system} \\ \mathcal{C}\subset\mathcal{A}}} \mathcal{A}$$

[1] Other names include *d-system* (cf. Prof Norris' notes) and *monotone class* (cf. Prof Raphaël's lectures).

Theorem 1.4 Dynkin's π-λ theorem[2]
Let $\mathcal{C}$ be a π-system. Then $\lambda(\mathcal{C}) = \sigma(\mathcal{C})$.

Proof. A σ-algebra is a λ-system, so $\lambda(\mathcal{C}) \subset \sigma(\mathcal{C})$. It thus suffices to show that $\lambda(\mathcal{C})$ is a σ-algebra. To prove this, we need to show that $\lambda(\mathcal{C})$ is closed under finite intersections. By Lemma 1.3, it suffices to show that $\lambda(\mathcal{C})$ is a π-system.

For any $A \in \lambda(\mathcal{C})$, define

$$\lambda_A := \{B \in \lambda(\mathcal{C}) \colon A \cap B \in \lambda(\mathcal{C})\}.$$

Fix $\mathcal{A} \in \mathcal{C}$. Since $\mathcal{C}$ is a π-system, we have $\mathcal{C} \subset \lambda_A$. Now, let us show that λ_A is a λ-system:

- $A \cap E = A$ so $E \in \lambda_A$
- for any $B, B' \in \lambda_A$ with $B \subset B'$, we have $A \cap B, A \cap B' \in \lambda(\mathcal{C})$ with $A \cap B \subset A \cap B'$, so

  $$A \cap (B' \backslash B) = (A \cap B') \backslash (A \cap B) \in \lambda(\mathcal{C})$$

  and thus $B' \backslash B \in \lambda_A$.
- Fix an increasing sequence $(B_n)_{n \in \mathbb{N}}$ in λ_A. Then $(A \cap B_n)_{n \in \mathbb{N}}$ is an increasing sequence in $\lambda(\mathcal{C})$. It then follows that

  $$A \cap \left(\bigcup_{n \in \mathbb{N}} B_n \right) = \bigcup_{n \in \mathbb{N}} (A \cap B_n) \in \lambda(\mathcal{C})$$

  so $\bigcup_{n \in \mathbb{N}} B_n \in \lambda_A$.

Since λ_A is a λ-system with $\mathcal{C} \subset \lambda_A$, we deduce that $\lambda(\mathcal{C}) \subset \lambda_A$ and thus

$$\forall\, A \in \mathcal{C}\ \forall\, B \in \lambda(\mathcal{C})\ A \cap B \in \lambda(\mathcal{C}) \qquad (*)$$

Now, fix $A \in \lambda(\mathcal{C})$. Note that $(*)$ above tells us that $\mathcal{C} \subset \lambda_A$. By the same arguments above, λ_A is still a λ-system, from which we deduce that $\lambda(\mathcal{C}) \subset \lambda_A$. Since $\lambda_A \subset \lambda(\mathcal{C})$ manifestly, we have $\lambda(\mathcal{C}) = \lambda_A$. ■

Corollary 1.5 Uniqueness of measure
Let μ, ν be two measures on $(E, \mathcal{A})$. Suppose that

(i) *there exists a π-system $\mathcal{C} \subset \mathcal{A}$ such that $\sigma(\mathcal{C}) = \mathcal{A}$*
(ii) *$\mu(A) = \nu(A)$ for all $A \in \mathcal{C}$*

Then $\mu = \nu$ if one of the following conditions is satisfied:

(1) *$\mu(E), \nu(E) < \infty$*
(2) *there exists an increasing sequence $(E_n)_{n \in \mathbb{N}}$ in $\mathcal{C}$ such that $E = \bigcup_{n \in \mathbb{N}} E_n$ and $\mu(E_n) = \nu(E_n)$ for all $n \in \mathbb{N}$*

Proof. We consider each case separately.
Case (1) Consider the 'good' set $\mathcal{G} = \{A \in \mathcal{A} \colon \mu(A) = \nu(A)\}$. By hypothesis, $\mathcal{C} \subset \mathcal{G}$. Since $\mathcal{C}$ is a π-system, Dynkin's π-λ theorem implies that it suffices to show that $\mathcal{G}$ is a λ-system. Indeed,

- $E \in \mathcal{G}$
- for any $A, B \in \mathcal{G}$ with $A \subset B$,

  $$\begin{cases} \mu(B \backslash A) = \mu(B) - \mu(A) \\ \nu(B \backslash A) = \nu(B) - \nu(A) \end{cases} \implies \mu(B \backslash A) = \nu(B \backslash A) \implies B \backslash A \in \mathcal{G}$$

[2] Prof Raphaël referred to this as the *monotone class lemma*.

- for any increasing sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{G}$, we have

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \nu(A_n) = \nu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \Longrightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{G}$$

Case (2) Consider the measures

$$\begin{cases} \mu_n(A) = \mu(A \cap E_n) \\ \nu_n(A) = \nu(A \cap E_n) \end{cases}$$

For any $A \in \mathcal{C}$, we have $A \cap E_n \in \mathcal{C}$ and so $\mu_n(A) = \nu_n(A) < \infty$. Now, μ_n and ν_n satisfy condition (1), implying that $\mu_n(A) = \nu_n(A)$ for all $A \in \mathcal{A}$. Finally, since $(E_n)_{n\in\mathbb{N}}$ is increasing and $E = \bigcup_{n\in\mathbb{N}} E_n$, we have

$$\mu(A) = \lim_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \nu_n(A) = \nu(A)$$

Hence, $\mu = \nu$. ■

Remark. A direct consequence of Corollary 1.5 is the uniqueness of the Lebesgue measure, i.e., assuming it exists, it is the only measure for which $\mu([a, b]) = b - a$. By taking $\mathcal{C} = \{[a, b] \colon a < b\}$ (which generates $\mathcal{B}(\mathbb{R})$) and $E_n = [-n, n]$.

1.4 Measurable functions

Definition Measurable function
Let $(E, \mathcal{A})$ and $(F, \mathcal{B})$ be two measurable spaces. A map $f \colon E \to F$ is measurable if, for every $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.

In the case that E and F are both topological spaces, with $\mathcal{A}$ and $\mathcal{B}$ the corresponding Borel σ-algebras, we also call f a *Borel measurable function.*

Proposition 1.6
The composition of two measurable functions is measurable.

Proof. This is immediate from $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. ■

Proposition 1.7
Let $f \colon (E, \mathcal{A}) \to (F, \mathcal{B})$ be a function between measurable spaces. If $\mathcal{B} = \sigma(\mathcal{C})$ for some $\mathcal{C} \subset \mathcal{P}(F)$, then f is measurable if $f^{-1}(C) \in \mathcal{A}$ for all $C \in \mathcal{C}$.

Proof. Consider the 'good' family $\mathcal{G} = \{B \in \mathcal{B} \colon f^{-1}(B) \in \mathcal{A}\}$. Note that $\mathcal{G}$ is a σ-algebra:
- $f^{-1}(F) = E \in \mathcal{A}$ so $F \in \mathcal{G}$
- for any $B \in \mathcal{G}$, $f^{-1}(B^c) = [f^{-1}(B)]^c \in \mathcal{A}$, so $B^c \in \mathcal{G}$
- for any sequence $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{G}$, we have

$$f^{-1}\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \bigcup_{n\in\mathbb{N}} f^{-1}(B_n) \in \mathcal{A} \Longrightarrow \bigcup_{n\in\mathbb{N}} f^{-1}(B_n) \in \mathcal{G}$$

Since $\mathcal{C} \subset \mathcal{G}$, it follows that $\mathcal{B} = \sigma(\mathcal{C}) \subset \mathcal{G}$. ■

Example. To show that a function $f \colon (E, \mathcal{A}) \to \mathbb{R}$ is Borel measurable, it suffices to check that $f^{-1}((a, \infty)) \in \mathcal{A}$ for all $a \in \mathbb{Q}$.

Corollary 1.8
Let E, F be topological spaces equipped with the corresponding Borel σ-algebras. Then every continuous map $f\colon E \to F$ is measurable.

Proof. Note that, by definition of continuity, $f^{-1}(V)$ is open in E for every V open in F. Since the open sets of F generate $\mathcal{B}(F)$, we are done by Proposition 1.7. ■

Proposition 1.9
Let $(E, \mathcal{A})$, $(F_1, \mathcal{B}_1)$ and $(F_2, \mathcal{B}_2)$ be measurable spaces. The maps $f_1\colon (E, \mathcal{A}) \to (F_1, \mathcal{B}_1)$ and $f_2\colon (E, \mathcal{A}) \to (F_2, \mathcal{B}_2)$ are measurable if and only if the map

$$\begin{aligned} f\colon (E, \mathcal{A}) &\to (F_1 \times F_2, \mathcal{B}_1 \otimes \mathcal{B}_2) \\ x &\mapsto (f_1(x), f_2(x)) \end{aligned}$$

is measurable.

Proof. ($\Longrightarrow$) For any $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$, we have

$$\begin{aligned} f^{-1}(B_1 \times B_2) &= \{x \in E\colon f(x) \in B_1 \times B_2\} \\ &= \{x \in E\colon f_1(x) \in B_1 \wedge f_2(x) \in B_2\} \\ &= f_1^{-1}(B_1) \cap f_2^{-1}(B_2) \in \mathcal{A} \end{aligned}$$

so f is measurable.

($\Longleftarrow$) For any $B_1 \in \mathcal{B}_1$, we have

$$f^{-1}(B_1) = f^{-1}(B_1) \cap f^{-2}(F_2) = f^{-1}(B_1 \times F_2) \in \mathcal{A}$$

so f_1 is measurable. By a similar argument, f_2 is measurable. ■

Corollary 1.10
Let $f, g\colon (E, \mathcal{A}) \to \mathbb{R}$ be measurable. For any $\alpha, \beta \in \mathbb{R}$, the maps $\alpha f + \beta g$, fg, $f^+ := \max\{f, 0\}$ and $f^- := \max\{-f, 0\}$ are measurable.

Proof. The first two are immediate from noting that $(a, b) \mapsto a\lambda + b\beta$ and $(a, b) \mapsto ab$ are continuous maps $\mathbb{R} \to \mathbb{R}$. For the other two, note that $|f|$ is measurable and that $f^+ = (|f| + f)/2$ and $f^- = (|f| - f)/2$. ■

Next, we consider sequences of measurable functions. It would be useful to allow functions to take values in

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\} \qquad \text{(extended real line)}$$

Remarks
- $\overline{\mathbb{R}}$ is homeomorphic to $[-1, 1]$
- The topology on $\overline{\mathbb{R}}$ contains the topology on $\mathbb{R}$ but also includes sets of the form $[-\infty, a)$ and $(a, \infty]$.

Exercise. Show that $\mathcal{B}(\overline{\mathbb{R}}) = \sigma\left(\{[-\infty, a)\colon a \in \mathbb{Q}\}\right)$.

Definition Limit infimum and limit supremum
Let $(a_n)_{n \in \mathbb{N}}$ be a sequence in $\overline{\mathbb{R}}$. Then

$$\limsup_{n\to\infty} a_n := \inf_{n \in \mathbb{N}} \sup_{k \geq n} a_k, \qquad \liminf_{n\to\infty} a_n := \sup_{n \in \mathbb{N}} \inf_{k \geq n} a_k$$

Proposition 1.11
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions $(E,\mathcal{A}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$. Then the functions

$$\sup_{n\in\mathbb{N}} f_n, \qquad \inf_{n\in\mathbb{N}} f_n, \qquad \limsup_{n\to\infty} f_n, \qquad \liminf_{n\to\infty} f_n$$

are measurable. In particular, if $f_n \to f$ for some $f\colon (E,\mathcal{A}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, then f is measurable. Moreover, the set

$$\{x \in E\colon f_n(x) \text{ has a limit as } n \to \infty\}$$

is measurable.

Proof. See Example Sheet 1 Question 9. ■

Definition Pushforward measure
Let $f\colon (E,\mathcal{A}) \to (F,\mathcal{B})$ be measurable with μ a measure on $(E,\mathcal{A})$. The pushforward of μ by f is the measure on $(F,\mathcal{B})$ given by

$$\forall\, B \in \mathcal{B} \;\; f_*\mu(B) = \mu(f^{-1}(B))$$

Exercise. Check that $f_*\mu$ is indeed a measure on $(F,\mathcal{B})$.

2 Integration

2.1 Integration of non-negative functions

Definition Simple function
Let $(E, \mathcal{A}, \mu)$ be a measure space. A function $f\colon E \to \mathbb{R}$ is simple if it takes on a finite number of distinct values $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ and, for each value α_k, the set $A_k = f^{-1}(\{\alpha_k\})$ is measurable. In that case, we can write

$$f = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k},$$

which is called the canonical representation of f.

Remark. As mentioned in a previous remark, we adopt the convention that $\infty \cdot 0 = 0 \cdot \infty = 0$.

Definition Integral of a positive simple function
Let f be a simple function $E \to \mathbb{R}_+$ with canonical representation $f = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k}$. Then the integral of f with respect to μ is defined as

$$\int f \, d\mu := \sum_{k=1}^{n} \alpha_k \mu(A_k).$$

Remark. Suppose we have another representation $f = \sum_{k=1}^{N} \beta_k \mathbb{1}_{B_k}$ with $E = \bigcup_{k=1}^{N} B_k$ but β_k not distinct. We can put f into canonical form by appropriately gluing the sets B_k. It is easy to check that

$$\int f \, d\mu = \sum_{k=1}^{N} \beta_k \mu(B_k).$$

Proposition 2.1 Basic properties
Let f, g be positive simple functions on $(E, \mathcal{A}, \mu)$. Then

(i) *for every $a, b \in \mathbb{R}_+$, $af + bg$ is positive and simple, with $\int af + bg \, d\mu = a \int f \, d\mu + b \int g \, d\mu$*

(ii) *if $f \geq g$, then $\int f \, d\mu \geq \int g \, d\mu$*

Proof. (i) Suppose

$$f = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k}, \qquad g = \sum_{k=1}^{n'} \alpha'_k \mathbb{1}_{A'_k}$$

be canonical representations of f and g, respectively. Note that

$$\{B_{ij} = A_i \cap A'_j \colon 1 \leq i \leq n, 1 \leq j \leq n'\}$$

partitions E and that we can represent

$$f = \sum_{i,j} \alpha_i \mathbb{1}_{B_{ij}}, \qquad g = \sum_{i,j} \alpha'_j \mathbb{1}_{B_{ij}}.$$

Then we have

$$af + bg = \sum_{i,j} \tilde{\alpha}_{ij} \mathbb{1}_{B_{ij}}, \qquad \text{where } \tilde{\alpha}_{ij} = a\alpha_i + b\alpha'_j,$$

so $af + bg$ is a positive simple function. Finally, we compute

$$\int (af + bg) \, d\mu = \sum_{i,j} \tilde{\alpha}_{ij} \mu(B_{ij})$$

$$
\begin{aligned}
&= a\sum_{i,j}\alpha_i\mu(B_{ij}) + b\sum_{i,j}\alpha'_j\mu(B_{ij}) \\
&= a\sum_{i=1}^{n}\alpha_i\mu(A_i) + b\sum_{j=1}^{n'}\alpha'_j\mu(A_i) \\
&= a\int f\,d\mu + b\int g\,d\mu
\end{aligned}
$$

(ii) Note that $f - g$ is a positive simple function. Then

$$\int f\,d\mu = \int g\,d\mu + \int f - g\,d\mu \geq \int g\,d\mu$$

as required. ■

Notation. Let $\mathcal{E}_+$ denote the space of positive simple functions.

We will say a function f on $(E, \mathcal{A})$ is *positive measurable* if it is measurable and takes values in $\overline{\mathbb{R}_+} = [0, \infty]$.

Definition Integral of a positive measurable function
Let $f\colon E \to [0, \infty]$ be positive measurable. The integral of f is

$$\int f\,d\mu := \sup_{\substack{h\in\mathcal{E}_+\\ h\leq f}}\int h\,d\mu.$$

Remark. Let $f \in \mathcal{E}_+$. For any $h \in \mathcal{E}_+$ with $h \leq f$, we have $\int h\,d\mu \leq \int f\,d\mu$, so the definitions are consistent.

Exercise. Let f, g be positive measurable functions on E. Show that
(i) if $f \leq g$, then $\int f\,d\mu \leq \int g\,d\mu$
(ii) if $\mu(\{x \in E\colon f(x) > 0\}) = 0$, then $\int f\,d\mu = 0$

Theorem 2.2 Monotone convergence theorem
Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence of positive measurable functions. Suppose that $f_n \uparrow f$. Then

$$\int f\,d\mu = \lim_{n\to\infty}\int f_n\,d\mu$$

Proof. Note that $f = \lim_{n\to\infty}\uparrow f_n$ is a positive measurable function, so $\int f\,d\mu$ is well-defined. Since $(f_n)_{n\in\mathbb{N}}$ is increasing, the sequence $(\int f_n\,d\mu)_{n\in\mathbb{N}}$ is also increasing. Since $f \geq f_n$ for all $n \in \mathbb{N}$, we have

$$\int f\,d\mu \geq \lim_{n\to\infty}\int f_n\,d\mu.$$

Now, pick $h = \sum_{k=1}^{K}\alpha_k\mathbb{1}_{A_k}$ a positive simple function with $h \leq f$. Pick $0 < \varepsilon \ll 1$. Then consider the sets

$$E_n = \{x \in E\colon (1-\varepsilon)h(x) \leq f_n(x)\}$$

Since $(f_n)_{n\in\mathbb{N}}$ is increasing, $(E_n)_{n\in\mathbb{N}}$ is also increasing. Moreover, since $\varepsilon > 0$, we have $E = \bigcup_{n\in\mathbb{N}} E_n$. By the definition of E_n, we have $f_n \geq (1-\varepsilon)\mathbb{1}_{E_n}h$ and thus

$$\int f_n\,d\mu \geq (1-\varepsilon)\int \mathbb{1}_{E_n}h\,d\mu = (1-\varepsilon)\sum_{k=1}^{K}\alpha_k\mu(A_k\cap E_n)$$

Since $(E_n)_{n\in\mathbb{N}}$ is increasing with $E = \bigcup_{n\in\mathbb{N}} E_n$, we have $\mu(A_k) = \lim_{n\to\infty} \mu(A_k \cap E_n)$, so it follows that

$$\lim_{n\to\infty} \int f_n \, d\mu \geq (1-\varepsilon) \lim_{n\to\infty} \sum_{k=1}^{K} \alpha_k \mu(A_k \cap E_n) = (1-\varepsilon) \sum_{k=1}^{K} \alpha_k \mu(A_k) = (1-\varepsilon) \int h \, d\mu$$

This holds for all $0 < \varepsilon \ll 1$, so

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \int h \, d\mu$$

for every $h \in \mathcal{E}_+$ with $h \leq f$. Taking the supremum gives us the result. ■