# Part III

# Appendices

# A

# Measure and Integration

## A.1 Rings and $\sigma$-Algebras

**Definition A.1.** A collection $\mathcal{F}$ of the elements of a set $\Omega$ is called a *ring* on $\Omega$ if it satisfies the following conditions:

1. $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$,
2. $A, B \in \mathcal{F} \Rightarrow A \setminus B \in \mathcal{F}$.

Furthermore, $\mathcal{F}$ is called an *algebra* if $\mathcal{F}$ is both a ring and $\Omega \in \mathcal{F}$.

**Definition A.2.** A ring $\mathcal{F}$ on $\Omega$ is called a *$\sigma$-ring* if it satisfies the following additional condition:

3. For every countable family $(A_n)_{n\in\mathbb{N}}$ of the subsets of $\mathcal{F}$: $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

A $\sigma$-ring $\mathcal{F}$ on $\Omega$ is called a *$\sigma$-algebra* if $\Omega \in \mathcal{F}$.

**Definition A.3.** Every collection $\mathcal{F}$ of the elements of a set $\Omega$, is called a *semiring* on $\Omega$ if it satisfies the following conditions:

1. $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$,
2. $A, B \in \mathcal{F} \Rightarrow A \subset B \Rightarrow \exists (A_j)_{i\leq j\leq m} \in \mathcal{F}^{\{1,\dots,m\}}$ of disjoint sets such that $B \setminus A = \bigcup_{j=1}^{m} A_j$.

If $\mathcal{F}$ is both a semiring and $\Omega \in \mathcal{F}$, then it is called a *semialgebra*.

**Proposition A.4.** *A set $\Omega$ has the following properties:*

1. *If $\mathcal{F}$ is a $\sigma$-algebra of the subsets of $\Omega$, then it is an algebra.*
2. *If $\mathcal{F}$ is a $\sigma$-algebra of the subsets of $\Omega$, then*
   - $E_1, \dots, E_n, \dots \in \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} E_n \in \mathcal{F}$,
   - $E_1, \dots, E_n \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{n} E_i \in \mathcal{F}$,
   - $B \in \mathcal{F} \Rightarrow \Omega \setminus B \in \mathcal{F}$.
3. *If $\mathcal{F}$ is a ring on $\Omega$, then it is also a semiring.*

**Definition A.5.** Every pair $(\Omega, \mathcal{F})$ consisting of a set $\Omega$ and a $\sigma$-ring $\mathcal{F}$ of the subsets of $\Omega$ is a *measurable space*. Furthermore, if $\mathcal{F}$ is a $\sigma$-algebra, then $(\Omega, \mathcal{F})$ is a *measurable space on which a probability measure can be built.* If $(\Omega, \mathcal{F})$ is a measurable space, then the elements of $\mathcal{F}$ are called $\mathcal{F}$*-measurable* or just *measurable sets.* We will henceforth assume that if a space is measurable, then we can build a probability measure on it.

*Example A.6.*

1. If $\mathcal{B}$ is a $\sigma$-algebra on the set $E$ and $X : \Omega \to E$ a generic mapping, then the set
$$X^{-1}(\mathcal{B}) = \{A \subset \Omega | \exists B \in \mathcal{B} \text{ such that } A = X^{-1}(B)\}$$
is a $\sigma$-algebra on $\Omega$.
2. *Generated $\sigma$-algebra*. If $\mathcal{A}$ is a set of the elements of a set $\Omega$, then there exists a smallest $\sigma$-algebra of subsets of $\Omega$ that contains $\mathcal{A}$. This is the $\sigma$-algebra *generated* by $\mathcal{A}$, denoted $\sigma(\mathcal{A})$. If, now, $\mathcal{G}$ is the set of all $\sigma$-algebras of the subsets of $\Omega$ containing $\mathcal{A}$, then it is not empty because it has $\sigma(\Omega)$ among its elements, so that $\sigma(\mathcal{A}) = \bigcap_{\mathcal{C} \in \mathcal{G}} \mathcal{C}$.
3. *Borel $\sigma$-algebra.* Let $\Omega$ be a topological space. Then the *Borel $\sigma$-algebra* on $\Omega$, denoted by $\mathcal{B}_\Omega$, is the $\sigma$-algebra generated by the set of all open subsets of $\Omega$. Its elements are called Borelian or Borel-measurable.
4. The set of all bounded and unbounded intervals of $\mathbb{R}$ is a semialgebra.
5. If $\mathcal{B}_1$ and $\mathcal{B}_2$ are algebras on $\Omega_1$ and $\Omega_2$, respectively, then the set of rectangles $B_1 \times B_2$, with $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$, is a semialgebra.
6. *Product $\sigma$-algebra.* Let $(\Omega_i, \mathcal{F}_i)_{1 \le i \le n}$ be a family of measurable spaces and let $\Omega = \prod_{i=1}^n \Omega_i$. Defining
$$\mathcal{R} = \left\{ E \subset \Omega | \forall, i = 1, \ldots, n \ \exists E_i \in \mathcal{F}_i \text{ such that } E = \prod_{i=1}^n E_i \right\},$$
then $\mathcal{R}$ is a semialgebra of the elements of $\Omega$. The $\sigma$-algebra generated by $\mathcal{R}$ is called the *product $\sigma$-algebra* of the $\sigma$-algebras $(\mathcal{F}_i)_{1 \le i \le n}$.

**Proposition A.7.** *Let $(\Omega_i)_{1 \le i \le n}$ be a family of topological spaces with a countable base and let $\Omega = \prod_{i=1}^n \Omega_i$. Then the Borel $\sigma$-algebra $\mathcal{B}_\Omega$ is identical to the product $\sigma$-algebra of the family of Borel $\sigma$-algebras $(\mathcal{B}_{\Omega_i})_{1 \le i \le n}$.*

## A.2 Measurable Functions and Measure

**Definition A.8.** Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces. A function $f : \Omega_1 \to \Omega_2$ is *measurable* if
$$\forall E \in \mathcal{F}_2 \colon f^{-1}(E) \in \mathcal{F}_1.$$

*Remark A.9.* If $(\Omega, \mathcal{F})$ is not a measurable space, i.e., $\Omega \notin \mathcal{F}$, then there does not exist a measurable mapping from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, because $\mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ and $f^{-1}(\mathbb{R}) = \Omega \notin \mathcal{F}$.

**Definition A.10.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}^n$ a mapping. If $f$ is measurable with respect to the $\sigma$-algebras $\mathcal{F}$ and $\mathcal{B}_{\mathbb{R}^n}$, the latter being the Borel $\sigma$-algebra on $\mathbb{R}^n$, then $f$ is *Borel-measurable.*

**Proposition A.11.** *Let $(E_1, \mathcal{B}_1)$ and $(E_2, \mathcal{B}_2)$ be two measurable spaces, $\mathcal{U}$ a set of the elements of $E_2$, which generates $\mathcal{B}_2$ and $f : E_1 \to E_2$. The necessary and sufficient condition for $f$ to be measurable is $f^{-1}(\mathcal{U}) \subset \mathcal{B}_1$.*

*Remark A.12.* If a function $f : \mathbb{R}^k \to \mathbb{R}^n$ is continuous, then it is Borel-measurable.

**Definition A.13.** Let $(\Omega, \mathcal{F})$ be a measurable space. Every Borel-measurable mapping $h : \Omega \to \bar{\mathbb{R}}$ that can only have a finite number of distinct values is called an *elementary function.* Equivalently, a function $h : \Omega \to \bar{\mathbb{R}}$ is elementary if and only if it can be written as the finite sum

$$\sum_{i=1}^{r} x_i I_{E_i},$$

where, for every $i = 1, \ldots, r$, the $E_i$ are disjoint sets of $\mathcal{F}$ and $I_{E_i}$ is the indicator function on $E_i$.

**Theorem A.14.** (Approximation of measurable functions through elementary functions.) *Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \bar{\mathbb{R}}$ a nonnegative measurable function. There exists a sequence of measurable elementary functions $(s_n)_{n \in \mathbb{N}}$ such that*

1. $0 \leq s_1 \leq \cdots \leq s_n \leq \cdots \leq f$,
2. $\lim_{n \to \infty} s_n = f$.

**Proposition A.15.** *If $f_1, f_2 : \Omega \to \bar{\mathbb{R}}$ are Borel-measurable functions, then so are the functions $f_1 + f_2$, $f_1 - f_2$, $f_1 f_2$, and $f_1/f_2$, as long as the operations are well defined.*

**Lemma A.16.** *If $f : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)$ and $g : (\Omega_2, \mathcal{F}_2) \to (\Omega_3, \mathcal{F}_3)$ are measurable functions, then so is $g \circ f : (\Omega_1, \mathcal{F}_1) \to (\Omega_3, \mathcal{F}_3)$.*

**Proposition A.17.** *Let $(\Omega_i, \mathcal{F}_i)_{1 \leq i \leq n}$ be a family of measurable spaces, $\Omega = \prod_{i=1}^n \Omega_i$ and $\pi_i : \Omega \to \Omega_i$ for $1 \leq i \leq n$ the $i$th projection. Then the product $\sigma$-algebra $\bigotimes_{i=1}^n \mathcal{F}_i$ of the family of $\sigma$-algebras $(\mathcal{F}_i)_{1 \leq i \leq n}$ is the smallest $\sigma$-algebra on $\Omega$ for which every projection $\pi_i$ is measurable.*

**Proposition A.18.** *If $h : (E, \mathcal{B}) \to (\Omega = \prod_{i=1}^n \Omega_i, \mathcal{F} = \bigotimes_{i=1}^n \mathcal{F}_i)$ is a mapping, then the following statements are equivalent:*

1. *$h$ is measurable;*
2. *for all $i = 1, \dots n$, $h_i = \pi_i \circ h$ is measurable.*

*Proof:* $1 \Rightarrow 2$ follows from Proposition A.17 and Lemma A.16. To prove that $2 \Rightarrow 1$, it is sufficient to see that given $\mathcal{R}$, the set of rectangles on $\Omega$, it follows that, for all $B \in \mathcal{R}: \ h^{-1}(B) \in \mathcal{B}$. Let $B \in \mathcal{R}$. Then for all $i = 1, \dots, n$, there exists a $B_i \in \mathcal{F}_i$ such that $B = \prod_{i=1}^n B_i$. Therefore, by recalling that due to 2 every $h_i$ is measurable, we have that

$$h^{-1}(B) = h^{-1}\left(\prod_{i=1}^n B_i\right) = \bigcap_{i=1}^n h_i^{-1}(B_i) \in \mathcal{B}.$$

□

**Corollary A.19.** Let $(\Omega, \mathcal{F})$ be a measurable space and $h : \Omega \to \mathbb{R}^n$ a function. Defining $h_i = \pi_i \circ h : \Omega \to \mathbb{R}$ for $1 \leq i \leq n$, the following two propositions are equivalent:

1. $h$ is Borel-measurable;
2. for all $i = 1, \dots, n$, $h_i$ is Borel-measurable.

**Definition A.20.** Let $(\Omega, \mathcal{F})$ be a measurable space. Every function $\mu : \Omega \to \bar{\mathbb{R}}$ that

1. for all $E \in \mathcal{F}$: $\mu(E) \geq 0$,
2. for all $E_i, \dots, E_n, \dots \in \mathcal{F}$ such that $E_i \cap E_j = \emptyset$, for $i \neq j$, we have that

$$\mu\left(\bigcup_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty \mu(E_i)$$

is a *measure* on $\mathcal{F}$. Moreover, if $(\Omega, \mathcal{F})$ is a measurable space and if

$$\mu(\Omega) = 1, \tag{A.1}$$

then $\mu$ is a *probability measure* or *probability*. Furthermore, a measure $\mu$ is *finite* if

$$\forall A \in \mathcal{F}: \ \mu(A) < +\infty$$

and $\sigma$-*finite*, if

1. there exists an $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$ such that $\Omega = \bigcup_{n \in \mathbb{N}} A_n$;
2. for all $n \in \mathbb{N}$: $\mu(A_n) < +\infty$.

**Definition A.21.** The ordered triple $(\Omega, \mathcal{F}, \mu)$, where $\Omega$ denotes a set, $\mathcal{F}$ a $\sigma$-ring on $\Omega$, and $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ a measure on $\mathcal{F}$, is a *measure space*. If $\mu$ is a probability measure, then $(\Omega, \mathcal{F}, \mu)$ is a *probability space.*[12]

[12] Henceforth we will call every measurable space that has a probability measure assigned to it a probability space.

**Definition A.22.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $\lambda : \mathcal{F} \to \bar{\mathbb{R}}$ a measure on $\Omega$. Then $\lambda$ is said to be *absolutely continuous* with respect to $\mu$, denoted $\lambda \ll \mu$, if

$$\forall A \in \mathcal{F}: \ \mu(A) = 0 \Rightarrow \lambda(A) = 0.$$

**Proposition A.23.** (Characterization of measure). *Let $\mu$ be additive on an algebra $\mathcal{F}$ and valued in $\mathbb{R}$ (and not everywhere equal to $+\infty$). The following two statements are equivalent:*

1. *$\mu$ is a measure on $\mathcal{F}$.*
2. *For increasing $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, where $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$, we have that*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \sup_{n\in\mathbb{N}} \mu(A_n).$$

*If $\mu$ is finite, then 1 and 2 are equivalent to the following.*

3. *For decreasing $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, where $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$, we have*

$$\mu\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \inf_{n\in\mathbb{N}} \mu(A_n).$$

4. *For decreasing $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, where $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, we have*

$$\lim_{n\to\infty} \mu(A_n) = \inf_{n\in\mathbb{N}} \mu(A_n) = 0.$$

**Proposition A.24.** (Generalization of a measure). *Let $\mathcal{G}$ be a semiring on $E$ and $\mu : \mathcal{G} \to \mathbb{R}_+$ a function that satisfies the following properties:*

1. *$\mu$ is (finitely) additive on $\mathcal{G}$,*
2. *$\mu$ is countably additive on $\mathcal{G}$,*
3. *there exists an $(S_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$ such that $E \subset \bigcup_{n\in\mathbb{N}} S_n$.*

*Under these assumptions*

$$\exists | \bar{\mu} : \mathcal{B} \to \bar{\mathbb{R}}_+ \text{ such that } \bar{\mu}|_{\mathcal{G}} = \mu,$$

*where $\mathcal{B}$ is the $\sigma$-ring generated by $\mathcal{G}$.*[13] *Moreover, if $\mu$ is a probability measure, then so is $\bar{\mu}$.*

**Proposition A.25.** *Let $\mathcal{U}$ be a ring on $E$ and $\mu : \mathcal{U} \to \bar{\mathbb{R}}_+$ (not everywhere equal to $+\infty$) a measure on $\mathcal{U}$. Then, if $\mathcal{B}$ is the $\sigma$-ring generated by $\mathcal{U}$,*

$$\exists | \bar{\mu} : \mathcal{B} \to \bar{\mathbb{R}}_+ \text{ such that } \bar{\mu}|_{\mathcal{U}} = \mu.$$

*Moreover, if $\mu$ is a probability measure, then so is $\bar{\mu}$.*

---

[13] $\mathcal{B}$ is identical to the $\sigma$-ring generated by the ring generated by $\mathcal{G}$.

**Lemma A.26.** (Fatou). *Let* $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$ *be a sequence of random variables and* $(\Omega, \mathcal{F}, P)$ *a probability space. Then*

$$P(\liminf_n A_n) \leq \liminf_n P(A_n) \leq \limsup_n P(A_n) \leq P(\limsup_n A_n).$$

*If* $\liminf_n A_n = \limsup_n A_n = A$*, then* $A_n \to A$*.*

**Corollary A.27.** Under the assumptions of Fatou's Lemma A.26, if $A_n \to A$, then $P(A_n) \to P(A)$.

## A.3 Lebesgue Integration

Let $(\Omega, \mathcal{F})$ be a measurable space. We will denote by $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}})$ (or, respectively, by $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$) the set of measurable functions on $(\Omega, \mathcal{F})$ and valued in $\bar{\mathbb{R}}$ (or $\bar{\mathbb{R}}_+$).

**Proposition A.28.** *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* $\mu$ *a positive measure on* $\mathcal{F}$*. Then there exists a unique mapping* $\Phi$ *from* $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ *to* $\bar{\mathbb{R}}_+$*, such that:*

1. *For every* $\alpha \in \mathbb{R}_+$*,* $f, g \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$*,*
   $\Phi(\alpha f) = \alpha\Phi(f)$*,*
   $\Phi(f + g) = \Phi(f) + \Phi(g)$*,*
   $f \leq g \Rightarrow \Phi(f) \leq \Phi(g)$*.*
2. *For every increasing sequence* $(f_n)_{n\in\mathbb{N}}$ *of elements of* $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ *we have that* $\sup_n \Phi(f_n) = \Phi(\sup_n f_n)$ *(Beppo–Levi property).*
3. *For every* $B \in \mathcal{F}$*,* $\Phi(I_B) = \mu(B)$*.*

**Definition A.29.** If $\Phi$ is the unique functional associated with $\mu$, the measure on the measurable space $(\Omega, \mathcal{F})$, then for every $f \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$:

$$\Phi(f) = \int^* f(x)d\mu(x) \text{ or } \int^* f(x)\mu(dx) \text{ or } \int^* f(x)d\mu$$

the *upper integral* of $\mu$.

*Remark A.30.* Let $(\Omega, \mathcal{F})$ be a measurable space and let $\Phi$ be the functional canonically associated with $\mu$ measure on $\mathcal{F}$.

1. If $s : \Omega \to \bar{\mathbb{R}}_+$ is an elementary function, thus $s = \sum_{i=1}^n x_i I_{E_i}$, then

$$\Phi(s) = \int^* s d\mu = \sum_{i=1}^n x_i \mu(E_i).$$

2. If $f \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ and defining $\Omega_f = \{s : \Omega \to \bar{\mathbb{R}}_+ | s \text{ elementary }, s \leq f\}$, then $\Omega_f$ is nonempty and

$$\Phi(f) = \int^* f d\mu = \sup_{s\in\Omega_f} \int^* s d\mu = \sup_{s\in\Omega_f} \left(\sum_{i=1}^n x_i \mu(E_i)\right).$$

3. If $f \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ and $B \in \mathcal{F}$, then by definition

$$\int_B^* f d\mu = \int^* I_B \cdot f d\mu.$$

**Definition A.31.** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu$ a positive measure on $\mathcal{F}$. An $\mathcal{F}$-measurable function $f$ is *$\mu$-integrable* if

$$\int^* f^+ d\mu < +\infty \text{ and } \int^* f^- d\mu < +\infty,$$

where $f^+$ and $f^-$ denote the positive and negative parts of $f$, respectively. The real number

$$\int^* f^+ d\mu - \int^* f^- d\mu$$

is therefore the *Lebesgue integral* of $f$ with respect to $\mu$, denoted by

$$\int f d\mu \text{ or } \int f(x) d\mu(x) \text{ or } \int f(x)\mu(dx).$$

**Proposition A.32.** *Let $(\Omega, \mathcal{F})$ be a measurable space, endowed with measure $\mu$ and $f \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$. Then*

1. $\int^* f d\mu = 0 \Leftrightarrow f = 0$ *almost surely with respect to$\mu$,*
2. *for every $A \in \mathcal{F}, \mu(A) = 0$ we have*

$$\int_A^* f d\mu = 0;$$

3. *for every $g \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ such that $f = g$, almost surely with respect to $\mu$, we have*

$$\int^* f d\mu = \int^* g d\mu.$$

**Theorem A.33.** (Monotone convergence). *Let $(\Omega, \mathcal{F})$ be a measurable space endowed with measure $\mu$, $(f_n)_{n\in\mathbb{N}}$ an increasing sequence of elements of $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$, and $f : \Omega \to \bar{\mathbb{R}}_+$ such that*

$$\forall \omega \in \Omega\colon\ f(\omega) = \lim_{n\to\infty} f_n(\omega) = \sup_{n\in\mathbb{N}} f_n(\omega).$$

*Then $f \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ and*

$$\int^* f d\mu = \lim_{n\to\infty} \int^* f_n d\mu.$$

**Theorem A.34.** (Lebesgue's dominated convergence). *Let $(\Omega, \mathcal{F})$ be a measurable space endowed with measure $\mu$, $(f_n)_{n\in\mathbb{N}}$ a sequence of $\mu$-integrable functions defined on $\Omega$, and $g : \Omega \to \bar{\mathbb{R}}_+$ a $\mu$-integrable function, such that*

*$|f_n| \le g$, for all $n \in \mathbb{N}$. If we suppose that $\lim_{n\to\infty} f_n = f$ exists almost surely in $\Omega$, then $f$ is $\mu$-integrable and we have*

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

**Lemma A.35.** (Fatou). *Let $f_n \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$. Then*

$$\liminf_n \int^* f_n d\mu \ge \int^* \liminf_n f_n d\mu.$$

**Theorem A.36.** (Fatou–Lebesgue).

1. *Let $|f_n| \le g \in \mathcal{L}^1$. Then*

$$\limsup_n \int f_n d\mu \le \int \limsup_n f_n d\mu.$$

2. *Let $|f_n| \le g \in \mathcal{L}^1$. Then*

$$\liminf_n \int f_n d\mu \ge \int \liminf_n f_n d\mu.$$

3. *Let $|f_n| \le g$ and $f = \lim_n f_n$, almost surely with respect to $\mu$. Then*

$$\lim_n \int f_n d\mu = \int f d\mu.$$

**Definition A.37.** Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{B})$ be a measurable space, endowed with measure $\mu$, and let $h : (\Omega, \mathcal{F}) \to (E, \mathcal{B})$ be a measurable function. The mapping $\mu_h : \mathcal{B} \to \bar{\mathbb{R}}_+$, such that $\mu_h(B) = \mu(h^{-1}(B))$ for all $B \in \mathcal{B}$ is a measure on $E$, called *induced measure* $h$ on $\mu$, denoted $h(\mu)$.

**Proposition A.38.** *Given the assumptions of Definition A.37 the function $g : (E, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ is integrable with respect to $\mu_h$ if and only if $g \circ h$ is integrable with respect to $\mu$ and*

$$\int g \circ g d\mu = \int g d\mu_h.$$

**Theorem A.39.** (Product measure). *Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces and the former be endowed with $\sigma$-finite measure $\mu_1$ on $\mathcal{F}_1$. Further suppose that for all $\omega_1 \in \Omega_1$ a measure $\mu(\omega_1, \cdot)$ is assigned on $\mathcal{F}_1$ and that for all $B \in \mathcal{F}_2$, $\mu(\cdot, B) : \Omega_1 \to \mathbb{R}$ is a Borel-measurable function. If $\mu(\omega_1, \cdot)$ is uniformly $\sigma$-finite, then there exists a $(B_n)_{n\in\mathbb{N}} \in \mathcal{F}_2^{\mathbb{N}}$ such that $\Omega_2 = \bigcup_{n=1}^{\infty} B_n$ and, for all $n \in \mathbb{N}$ there exists a $K_n \in \mathbb{R}$ such that $\mu(\omega_1, B_n) \le K_n$ for all $\omega_1 \in \Omega_1$. Then there exists a unique measure $\mu$ on the product $\sigma$-algebra $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ such that*

$$\forall A \in \mathcal{F}_1, B \in \mathcal{F}_2: \qquad \mu(A \times B) = \int_A \mu(\omega_1, B)\mu_1(d\omega_1),$$

*and*

$$\forall F \in \mathcal{F}: \qquad \mu(F) = \int_{\Omega_1} \mu(\omega_1, F(\omega_1))\mu_1(d\omega_1).$$

**Definition A.40.** Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces, endowed with $\sigma$-finite measures $\mu_1$, $\mu_2$ on $\mathcal{F}_1$ and $\mathcal{F}_2$, respectively. Defining $\Omega = \Omega_1 \times \Omega_2$ and $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, the function $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ with

$$\forall F \in \mathcal{F}: \qquad \mu(F) = \int_{\Omega_1} \mu_2(F(\omega_1))d\mu_1(\omega_1) = \int_{\Omega_2} \mu_1(F(\omega_2))d\mu_2(\omega_2),$$

is the unique measure on $\mathcal{F}$ with

$$\forall A \in \mathcal{F}_1, B \in \mathcal{F}_2: \qquad \mu(A \times B) = \mu_1(A) \times \mu_2(B).$$

Moreover, $\mu$ is $\sigma$-finite on $\mathcal{F}$ as well as a probability measure, if so are $\mu_1$ and $\mu_2$. The measure $\mu$ is the *product measure* of $\mu_1$ and $\mu_2$, denoted by $\mu_1 \otimes \mu_2$.

**Theorem A.41.** (Fubini). *Given the assumptions of Definition A.40, let $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ be a Borel-measurable function, such that $\int_\Omega f d\mu$ exists. Then*

$$\int_\Omega f d\mu = \int_{\Omega_1} \int_{\Omega_2} f d\mu_2 d\mu_1 = \int_{\Omega_2} \int_{\Omega_1} f d\mu_1 d\mu_2.$$

**Proposition A.42.** *Let $(\Omega_i, \mathcal{F}_i)_{1 \leq i \leq n}$ be a family of measurable spaces. Further, let $\mu_1 : \mathcal{F}_1 \to \bar{\mathbb{R}}$ be a $\sigma$-finite measure and let*

$$\forall (\omega_1, \ldots, \omega_j) \in \Omega_1 \times \cdots \times \Omega_j: \qquad \mu(\omega_1, \ldots, \omega_j, \cdot) : \mathcal{F}_{j+1} \to \bar{\mathbb{R}}$$

*be a measure on $\mathcal{F}_{j+1}$, $1 \leq j \leq n-1$. If $\mu(\omega_1, \ldots, \omega_j, \cdot)$ is uniformly $\sigma$-finite and for every $c \in \mathcal{F}_{j+1}$*

$$\mu(\ldots, c) : (\Omega_1 \times \cdots \times \Omega_j, \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_j) \to (\bar{\mathbb{R}}, \mathcal{B}_{\bar{\mathbb{R}}}),$$

*such that*

$$\forall (\omega_1, \ldots, \omega_j) \in \Omega_1 \times \cdots \times \Omega_j: \qquad \mu(\ldots, c)(\omega_1, \ldots, \omega_j) = \mu(\omega_1, \ldots, \omega_j, c)$$

*is measurable, then, defining $\Omega = \Omega_1 \times \cdots \times \Omega_n$ and $\mathcal{F} = \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_n$:*

1. *There exists a unique measure $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ such that for every measurable rectangle $A_1 \times \cdots \times A_n \in \mathcal{F}$:*

$$\begin{aligned}&\mu(A_1 \times \cdots \times A_n)\\ &= \int_{A_1} \mu_1(d\omega_1) \int_{A_2} \mu(\omega_1, d\omega_2) \cdots \int_{A_n} \mu(\omega_1, \ldots, \omega_{n-1}, d\omega_n).\end{aligned}$$

   *$\mu$ is $\sigma$-finite on $\mathcal{F}$ and a probability whenever $\mu_1$ and all $\mu(\omega_1, \ldots, \omega_j, \cdot)$ are probability measures.*

2. *If $f : (\Omega, \mathcal{F}) \to (\bar{\mathbb{R}}, \mathcal{B}_{\bar{\mathbb{R}}})$ is measurable and nonnegative, then*

$$\int_{\Omega} f d\mu$$
$$= \int_{\Omega_1} \mu_1(d\omega_1) \int_{\Omega_2} \mu(\omega_1, d\omega_2) \cdots \int_{\Omega_n} f(\omega_1, \ldots, \omega_n)\mu(\omega_1, \ldots, \omega_{n-1}, d\omega_n).$$

**Proposition A.43.** *1. Given the assumptions and the notation of Proposition A.42, if we assume that $f = I_F$, then for every $F \in \mathcal{F}$:*

$$\mu(F)$$
$$= \int_{\Omega_1} \mu_1(d\omega_1) \int_{\Omega_2} \mu(\omega_1, d\omega_2) \cdots \int_{\Omega_n} I_F(\omega_1, \ldots, \omega_n)\mu(\omega_1, \ldots, \omega_{n-1}, d\omega_n).$$

2. *For all $j = 1, \ldots, n-1$, let $\mu_{j+1} = \mu(\omega_1, \ldots, \omega_j, \cdot)$. Then there exists a unique measure $\mu$ on $\mathcal{F}$ such that for every rectangle $A_1 \times \cdots \times A_n \in \mathcal{F}$ we have*
$$\mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n).$$
*If $f : (\Omega, \mathcal{F}) \to (\bar{\mathbb{R}}, \mathcal{B}_{\bar{\mathbb{R}}})$ is measurable and positive, or else if $\int_\Omega f d\mu$ exists, then*
$$\int_{\Omega} f d\mu = \int_{\Omega_1} d\mu_1 \cdots \int_{\Omega_n} f d\mu_n,$$
*and the order of integration is arbitrary. The measure $\mu$ is the product measure of $\mu_1, \ldots, \mu_n$ and is denoted by $\mu_1 \otimes \cdots \otimes \mu_n$.*

**Definition A.44.** Let $(v_i)_{1 \leq i \leq n}$ be a family of measures defined on $\mathcal{B}_{\mathbb{R}}$ and

$$v^{(n)} = v_1 \otimes \cdots \otimes v_n : \mathbb{R}^n \to \mathbb{R}$$

the product measure. The *convolution product* of $v_1, \ldots, v_n$, denoted by $v_1 * \cdots * v_n$, is the induced measure that, for generic functions $f : \mathbb{R}^n \to \mathbb{R}$, associates $(x_i, \ldots, x_n)$ with $\sum_{i=1}^n x_i$ of $v^{(n)}$.

**Proposition A.45.** *Let $v_1$ and $v_2$ be measures on $\mathcal{B}_{\mathbb{R}}$. Then for every $B \in \mathcal{B}_{\mathbb{R}}$ we have*

$$v_1 * v_2(B) = \int_B d(v_1 * v_2) = \int_{\mathbb{R}} I_B(z) d(v_1 * v_2) = \int \int I_B(x_1 + x_2) d(v_1 \otimes v_2).$$

## A.4 Lebesgue–Stieltjes Measure and Distributions

**Definition A.46.** Let $\mu : \mathcal{B}_{\mathbb{R}} \to \bar{\mathbb{R}}$ be a measure. It then represents a *Lebesgue–Stieltjes* measure if for every interval $I$ we have that $\mu(I) < +\infty$.

**Definition A.47.** Every function $F : \mathbb{R} \to \mathbb{R}$ that is right-continuous and increasing is a *distribution function* on $\mathbb{R}$.

It is in fact possible to establish a one-to-one relationship between the set of Lebesgue–Stieltjes measures and the set of distribution functions in the sense that every Lebesgue–Stieltjes measure can be assigned a distribution function and vice versa.

**Proposition A.48.** *Let $\mu$ be a Lebesgue–Stieltjes measure on $\mathcal{B}_{\mathbb{R}}$ and the function $F : \mathbb{R} \to \mathbb{R}$ defined, apart from a constant, as*

$$F(b) - F(a) = \mu(]a, b]) \qquad \forall a, b \in \mathbb{R}, a < b.$$

*Then $F$ is a distribution function, in particular the one assigned to $\mu$.*

**Proposition A.49.** *Let $F$ be a distribution function and*

$$F(b) - F(a) = \mu(]a, b]) \qquad \forall a, b \in \mathbb{R}, a < b.$$

*There exists a unique extension of $\mu$, which is a Lebesgue–Stieltjes measure on $\mathcal{B}_{\mathbb{R}}$. This measure is the Lebesgue–Stieltjes measure canonically associated with $F$.*

**Definition A.50.** Every measure $\mu : \mathcal{B}_{\mathbb{R}^n} \to \bar{\mathbb{R}}$ that for every bounded interval $I$ of $\mathbb{R}^n$ has $\mu(I) < +\infty$ is a Lebesgue–Stieltjes measure on $\mathbb{R}^n$

**Definition A.51.** Let $f : \mathbb{R} \to \mathbb{R}$ be of constant value 1 and we consider the function $F : \mathbb{R} \to \mathbb{R}$ with

$$F(x) - F(0) = \int_0^x f(t)dt \qquad \forall x > 0,$$
$$F(0) - F(x) = \int_x^0 f(t)dt \qquad \forall x < 0,$$

where $F(0)$ is fixed and arbitrary. This function $F$ is a distribution function and its associated Lebesgue–Stieltjes measure is called *Lebesgue measure* on $\mathbb{R}$.

**Definition A.52.** Let $(\Omega, \mathcal{F}, \mu)$ be a space with $\sigma$-finite measure $\mu$ and consider another measure $\lambda : \mathcal{F} \to \bar{\mathbb{R}}_+$. $\lambda$ is said to be defined through its *density* with respect to $\mu$ if there exists a Borel-measurable function $g : \Omega \to \bar{\mathbb{R}}_+$ with

$$\lambda(A) = \int_A g d\mu \qquad \forall A \in \mathcal{F}.$$

This function $g$ is the density of $\lambda$ with respect to $\mu$. In this case $\lambda$ is absolutely continuous with respect to $\mu$ ($\lambda \ll \mu$). If $\mu$ is a Lebesgue measure on $\mathbb{R}$, then $g$ is the density of $\mu$. A measure $\nu$ is called *$\mu$-singular* if there exists $N \in \mathcal{F}$ such that $\mu(N) = 0$ and $\nu(N \setminus \mathcal{F}) = 0$. Conversely, if also $\mu(N) = 0$ whenever $\nu(N) = 0$, then the two measures are *equivalent* (denoted $\lambda \sim \mu$).

**Theorem A.53.** (Radon–Nikodym). *Let $(\Omega, \mathcal{F})$ be a measurable space, $\mu$ a $\sigma$-finite measure on $\mathcal{F}$, and $\lambda$ an absolutely continuous measure with respect to $\mu$. Then $\lambda$ is endowed with density with respect to $\mu$. Hence there exists a Borel-measurable function $g : \Omega \to \bar{\mathbb{R}}_+$ such that*

$$\lambda(A) = \int_A g d\mu, \qquad A \in \mathcal{B}.$$

*A necessary and sufficient condition for $g$ to be $\mu$-integrable is that $\lambda$ is bounded. Moreover, if $h : \Omega \to \bar{\mathbb{R}}_+$ is another density of $\lambda$, then $g = h$, almost surely with respect to $\mu$.*

**Theorem A.54.** (Lebesgue–Nikodym). *Let $\nu$ and $\mu$ be a measure and a $\sigma$-finite measure on $(E, \mathcal{B})$, respectively. There exist a $\mathcal{B}$-measurable function $f : E \to \bar{\mathbb{R}}_+$ and a $\mu$-singular measure $\nu'$ on $(E, \mathcal{B})$ so that*

$$\nu(B) = \int_B f d\mu + \nu'(B) \qquad \forall B \in \mathcal{B}.$$

*Furthermore,*

1. *$\nu'$ is unique.*
2. *If $h : E \to \bar{\mathbb{R}}_+$ is a $\mathcal{B}$-measurable function with*

$$\nu(B) = \int_B h d\mu + \nu'(B) \qquad \forall B \in \mathcal{B},$$

*then $f = h$ almost surely with respect to $\mu$.*

**Definition A.55.** A function $F : \mathbb{R} \to \mathbb{R}$ is *absolutely continuous* if, for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $]a_i, b_i[ \subset \mathbb{R}$ for $1 \leq i \leq n$ with $]a_i, b_i[ \cap ]a_j, b_j[ = \emptyset$, $i \neq j$,

$$b_i - a_i < \delta \Rightarrow \sum_{i=1}^{n} |F(b_i) - F(a_i)| < \epsilon.$$

**Proposition A.56.** *Let $F$ be a distribution function. Then the following two propositions are equivalent:*

1. *$F$ is absolutely continuous.*
2. *The Lebesgue measure canonically associated with $F$ is absolutely continuous.*

**Proposition A.57.** *Let $f : [a, b] \to \mathbb{R}$ be a mapping. The following two statements are equivalent:*

1. *$f$ is absolutely continuous.*

2. *There exists a Borel-measurable function* $g : [a, b] \to \mathbb{R}$ *that is integrable with respect to Lebesgue measure and*

$$f(x) - f(a) = \int_a^x g(t)dt \qquad \forall x \in [a, b].$$

*This function* $g$ *is the density of* $f$.

**Proposition A.58.** *If* $f : [a, b] \to \mathbb{R}$ *is absolutely continuous, then*

1. $f$ *is differentiable almost everywhere in* $[a, b]$,
2. $f'$, *the first derivative of* $f$, *is integrable in* $[a, b]$ *and we have that*

$$f(x) - f(a) = \int_a^x f'(t)dt.$$

**Theorem A.59.** (Fundamental theorem of calculus). *If* $f : [a, b] \to \mathbb{R}$ *is integrable in* $[a, b]$ *and*

$$F(x) = \int_a^x f(t)dt \qquad \forall x \in [a, b],$$

*then*

1. $F$ *is absolutely continuous in* $[a, b]$,
2. $F' = f$ *almost everywhere in* $[a, b]$.

*Vice versa, if we consider a function* $F : [a, b] \to \mathbb{R}$ *that satisfies 1 and 2, then*

$$\int_a^b f(x)dx = F(b) - F(a).$$

**Proposition A.60.** *If* $f : [a, b] \to \mathbb{R}$ *is differentiable in* $[a, b]$ *and has integrable derivatives, then*

1. $f$ *is absolutely continuous in* $[a, b]$,
2. $f(x) = \int_a^x f'(t)dt$.

**Definition A.61.** Let $(\Omega, \mathcal{F}, \mu)$ be a space endowed with measure and $p > 0$. The set of Borel-measurable functions defined on $\Omega$, such that $\int_\Omega |f|^p d\mu < +\infty$ is a vector space on $\mathbb{R}$ and is denoted with the symbols $\mathcal{L}^p(\mu)$ or $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$. Its elements are called functions integrable to the exponent $p$. In particular, elements of $\mathcal{L}^2(\mu)$ are said to be square-integrable functions. Finally, $\mathcal{L}^1(\mu)$ coincides with the space of functions integrable with respect to $\mu$.

## A.5 Stochastic Stieltjes Integration

Suppose $(\Omega, \mathcal{F}, P)$ is a given probability space with $(X_t)_{t\in\mathbb{R}_+}$ a measurable stochastic process whose sample paths $(X_t(\omega))_{t\in\mathbb{R}_+}$ are of locally bounded variation for any $\omega \in \Omega$. Now let $(H_s)_{s\in\mathbb{R}_+}$ be a measurable process, whose sample paths are locally bounded for any $\omega \in \Omega$. Then the process $H \bullet X$ defined by

$$(H \bullet X)_t(\omega) = \int_0^t H(s,\omega) dX_s(\omega), \qquad \omega \in \Omega, t \in \mathbb{R}_+$$

is called the *stochastic Stieltjes integral* of $H$ with respect to $X$. Clearly, $((H * X)_t)_{t\in\mathbb{R}_+}$ is itself a stochastic process.

If we assume further that $X$ is progressively measurable and $H$ is $\mathcal{F}_t$-predictable with respect to the $\sigma$-algebra generated by $X$, then $H * X$ is progressively measurable. In particular, if $N = \sum_{n\in\mathbb{N}^*} \epsilon_{\tau_n}$ is a point process on $\mathbb{R}_+$, then for any nonnegative process $H$ on $\mathbb{R}_+$, the stochastic integral $H * N$ exists and is given by

$$(H \bullet N)_t = \sum_{n\in\mathbb{N}^*} I_{[\tau_n \le t]}(t) H(\tau_n).$$

**Theorem A.62.** *Let $M$ be a martingale of locally integrable variation, i.e., such that*

$$E\left[\int_0^t d|M_s|\right] < \infty \qquad \textit{for any } t > 0,$$

*and let $C$ be a predictable process satisfying*

$$E\left[\int_0^t |C_s| d|M_s|\right] < \infty \qquad \textit{for any } t > 0.$$

*Then the stochastic integral $C * M$ is a martingale.*

# B

# Convergence of Probability Measures on Metric Spaces

## B.1 Metric Spaces

For more details on the following and further results refer to Loève (1963), Dieudonné (1960), and Aubin (1977).

**Definition B.1.** Consider a set $R$. A *distance (metric)* on $R$ is a mapping $\rho : R \times R \to \mathbb{R}_+$, which satisfies the following properties.

D1.For any $x, y \in R$, $\rho(x, y) = 0 \Leftrightarrow x = y$.
D2.For any $x, y \in R$, $\rho(x, y) = \rho(y, x)$.
D3.For any $x, y, z \in R$, $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ (triangle inequality).

**Definition B.2.** A *metric space* is a set $R$ endowed with a metric $\rho$; we shall write $(R, \rho)$. Elements of a metric space will be called *points*.

**Definition B.3.** Given a metric space $(R, \rho)$, a point $a \in R$, and a real number $r > 0$, the *open ball* (or the *closed ball*) of center $a$ and radius $r$ is the set $B(a, r) := \{x \in R | \rho(a, x) < r\}$ (or $B'(a, r) := \{x \in R | \rho(a, x) \leq r\}$).

**Definition B.4.** In a metric space $(R, \rho)$, an *open set* is any subset $A$ of $R$ such that for any $x \in A$ there exists an $r > 0$ such that $B(a, r) \subset A$.

The empty set is open, and so is the entire space $R$.

**Proposition B.5.** *The union of any family of open sets is an open set. The intersection of a finite family of open sets is an open set.*

**Definition B.6.** The family $\mathcal{T}$ of all open sets in a metric space is called its *topology*. In this respect the couple $(R, \mathcal{T})$ is a *topological space*.

**Definition B.7.** The *interior* of a set $A$ is the largest open subset of $A$.

**Definition B.8.** In a metric space $(R, \rho)$, a *closed set* is any subset of $R$ which is a complement of an open set.

The empty set is closed, and so is the entire space $R$.

**Proposition B.9.** *The intersection of any family of closed sets is a closed set. The union of a finite family of closed sets is a closed set.*

**Definition B.10.** In a metric space $(R, \rho)$, the *closure* of a set $A$ is the smallest subset of $R$ containing $A$. It is denoted by $\bar{A}$. Any element of the closure of $A$ is called a *point of closure* of $A$.

**Proposition B.11.** *A closed set is the intersection of a decreasing sequence of open sets. An open set is the union of an increasing sequence of closed sets.*

**Definition B.12.** *A topological space is called a* Hausdorff topological space *if it satisfies the following property:*

*(HT) For any two distinct points $x$ and $y$ there exist two disjoint open sets $A$ and $B$ such that $x \in A$ and $y \in B$.*

**Proposition B.13.** *A metric space is a Hausdorff topological space.*

**Definition B.14.** In a metric space $(R, \rho)$, the *boundary* of a set $A$ is the set $\partial A = \bar{A} \cap (R \setminus A)$. Here $R \setminus A$ is the complement of $A$.

**Definition B.15.** Given two metric spaces $(R, \rho)$ and $(R', \rho')$, a function $f : R \to R'$ is *continuous*, if for any open set $A'$ in $(R', \rho')$, the set $f^{-1}(A')$ is an open set in $(R, \rho)$.

**Definition B.16.** Two metric spaces $(R, \rho)$ and $(R', \rho')$ are said to be *homeomorphic* if a function $f : R \to R'$ exists satisfying the following two properties:

1. $f$ is a bijection (an invertible function);
2. $f$ is bicontinuous; i.e., both $f$ and its inverse $f^{-1}$ are continuous.

The function $f$ above is called a *homeomorphism.*

**Definition B.17.** Given two distances $\rho$ and $\rho'$ on the same set $R$, we say that they are *equivalent distances* if the identity $i_R : x \in R \mapsto x \in R$ is a homeomorphism between the metric spaces $(R, \rho)$ and $(R', \rho')$.

*Remark B.18.* We may remark here that the notions of open set, closed set, closure, boundary, and continuous function are *topological notions.* They depend only on the topology induced by the metric. The topological properties of a metric space are invariant with respect to a homeomorphism.

**Definition B.19.** Given a subset $A$ of a metric space $(R, \rho)$ its *diameter* is given by $\delta(A) = \sup_{x\in A, y\in A} d(x, y)$. $A$ is *bounded* if its diameter is finite.

**Definition B.20.** Given two metric spaces $(R, \rho)$ and $(R', \rho')$, a function $f : R \to E'$ is *uniformly continuous* if for any $\epsilon > 0$, a $\delta > 0$ exists such that $x, y \in R$, $\rho(x, y) < \delta$ implies $\rho'(f(x), f(y)) < \epsilon$.

**Proposition B.21.** *A uniformly continuous function is continuous. (The converse is not true in general.)*

*Remark B.22.* The notions of diameter of a set and of uniform continuity of a function are *metric notions.*

**Definition B.23.** Let $A, B$ be two subsets of a metric space $R$. $A$ is said to be *dense* in $B$ if $B \subseteq \bar{A}$. $A$ is said to be *everywhere dense* in $R$ if $\bar{A} = R$.

**Definition B.24.** A metric space $R$ is said to be *separable* if it contains an everywhere dense countable subset.

Here are some examples of separable spaces with the corresponding everywhere countable subset.

- The space $\mathbb{R}$ of real numbers with distance function $\rho(x, y) = |x - y|$, with the set $\mathbb{Q}$.
- The space $\mathbb{R}^n$ of ordered $n$-tuples of real numbers $x = (x_1, x_2, \dots, x_n)$ with distance function $\rho(x, y) = \{\sum_{k=1}^{n}(y_k - x_k)^2\}^{\frac{1}{2}}$, with the set of all vectors with rational coordinates.
- The space $\mathbb{R}_0^n$ of ordered $n$-tuples of real numbers $x = (x_1, x_2, \dots, x_n)$ with distance function $\rho_0(x, y) = \max\{|y_k - x_k|; 1 \leq k \leq n\}$ with the set of all vectors with rational coordinates.
- $C^2([a, b])$, the totality of all continuous functions on the segment $[a, b]$ with distance function $\rho(x, y) = \int_a^b [x(t) - y(t)]^2 dt$ with the set of all polynomials with rational coefficients.

**Definition B.25.** A family $\{G_\alpha\}$ of open sets in the metric space $R$ is called a *basis* of $R$ if every open set in $R$ can be represented as the union of a (finite or infinite) number of sets belonging to this family.

**Definition B.26.** $R$ is said to be a space with countable basis if there is at least one basis in $R$ consisting of a countable number of elements.

**Theorem B.27.** *A necessary and sufficient condition for $R$ to be a space with countable basis is that there exists in $R$ an everywhere dense countable set.*

**Corollary B.28.** A metric space $R$ is separable if and only if it has a countable basis.

**Definition B.29.** A *covering* of a set is a family of sets, whose union contains the set. If the number of elements of the family is countable, then we have a *countable covering.* If the sets of the family are open, we have an open *covering.*

**Theorem B.30.** *If $R$ is a separable space, then we can select a countable covering from each of its open coverings.*

**Theorem B.31.** *Every separable metric space $R$ is homeomorphic to a subset of $\mathbb{R}^\infty$.*

**Definition B.32.** In a metric space $(R, \rho)$, a sequence $(x_n)_{n\in\mathbb{N}}$ is any function from $\mathbb{N}$ to $R$.

**Definition B.33.** We say that a sequence $(x_n)_{n\in\mathbb{N}}$ admits a *limit* $b \in R$ (is convergent to $b$), if $b$ is such that for any open set $V$, with $x \in V$, there exists an $n_V \in \mathbb{N}$ such that for any $n > n_V$ we have $x_n \in V$. We write $\lim_{n\to\infty} x_n = b$.

**Definition B.34.** A subsequence of a sequence $(x_n)_{n\in\mathbb{N}}$ is any sequence $k \in \mathbb{N} \mapsto x_{n_k} \in R$ such that $(n_k)_{k\in\mathbb{N}}$ is strictly increasing.

**Proposition B.35.** *If* $\lim_{n\to\infty} x_n = b$*, then* $\lim_{k\to\infty} x_{n_k} = b$ *for any subsequence of* $(x_n)_{n\in\mathbb{N}}$.

**Definition B.36.** $b$ is called a *cluster point* of a sequence $(x_n)_{n\in\mathbb{N}}$ if a subsequence exists having $b$ as a limit.

**Proposition B.37.** *Given a subset $A$ of a metric space $(R, \rho)$, for any $a \in \bar{A}$ there exists a sequence of elements of $A$ converging to $a$.*

**Proposition B.38.** *If $x$ is the limit of a sequence $(x_n)_{n\in\mathbb{N}}$, then $x$ is the unique cluster point of $(x_n)_{n\in\mathbb{N}}$. Conversely, $(x_n)_{n\in\mathbb{N}}$ may have a unique cluster point $x$ and still this does not imply that $x$ is the limit of $(x_n)_{n\in\mathbb{N}}$ (see Aubin (1977), page 67, for a counterexample).*

**Definition B.39.** In a metric space $(R, \rho)$, a *Cauchy sequence* is a sequence $(x_n)_{n\in\mathbb{N}}$ such that for any $\epsilon > 0$ an integer $n_0 \in \mathbb{N}$ exists such that $m, n \in \mathbb{N}$, $m, n > n_0$ implies $\rho(x_m, x_n) < \epsilon$.

**Proposition B.40.** *In a metric space, any convergent sequence is a Cauchy sequence. The converse is not true in general.*

**Proposition B.41.** *In a metric space, if a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ has a cluster point $x$, then $x$ is the limit of $(x_n)_{n\in\mathbb{N}}$.*

**Definition B.42.** A metric space $R$ is called *complete* if any Cauchy sequence in $R$ is convergent to a point of $R$.

**Definition B.43.** A subspace of a metric space $(R, \rho)$ is any nonempty subset $F$ of $R$ endowed with the restriction of $\rho$ to $F \times F$.

**Proposition B.44.** *If a subspace of a metric space $R$ is complete, then it is closed in $R$. In a complete metric space any closed subspace is complete.*

**Definition B.45.** A metric space $R$ is said to be *compact* if any arbitrary open covering $\{O_\alpha\}$ of the space $R$ contains a finite subcovering.

**Definition B.46.** A metric space $R$ is called *precompact* if, for all $\epsilon > 0$, there is a finite covering of $R$ by sets of diameter $< \epsilon$.

*Remark B.47.* The notion of compactness is a topological one, while the notion of precompactness is a metric one.

**Theorem B.48.** *For a metric space $R$, the following three conditions are equivalent:*

1. *$R$ is compact.*
2. *Any infinite sequence in $R$ has at least a limit point.*
3. *$R$ is precompact and complete.*

**Proposition B.49.** *Every precompact metric space is separable.*

**Proposition B.50.** *In a compact metric space any sequence which has only one cluster value, $a$ converges to $a$.*

**Proposition B.51.** *Any continuous mapping of a compact metric space into another metric space is uniformly continuous.*

**Definition B.52.** A *compact set* (or *precompact set*) in a metric space $R$ is any subset of $R$ that is compact (or precompact) as a subspace of $R$.

**Proposition B.53.** *Any precompact set is bounded.*

**Proposition B.54.** *Any compact set in a metric space is closed. In a compact metric space, any closed subset is compact.*

**Proposition B.55.** *Any compact set in a metric space is complete.*

**Definition B.56.** A set $M$ in the metric space $R$ is said to be *relatively compact* if $M = \bar{M}$.

**Theorem B.57.** *A relatively compact set is precompact. In a complete metric space a precompact set is relatively compact.*

**Proposition B.58.** *A necessary and sufficient condition that a subset $M$ of a metric space $R$ be relatively compact is that every sequence of points of $M$ has a cluster point in $R$.*

**Definition B.59.** A metric space $R$ is said to be *locally compact*, if for every point $x \in R$ there exists a compact neighborhood of $x$ in $R$.

**Theorem B.60.** *Let $R$ be a locally compact metric space. The following properties are equivalent:*

1. *there exists an increasing sequence $(U_n)$ of open relatively compact sets in $R$, such that $\bar{U}_n \subset U_{n+1}$ for every $n$, and $R = \cup_n U_n$;*
2. *$R$ is the countable union of compact subsets;*
3. *$R$ is separable.*

**Convergence of Probability Measures**

Let $(S, \rho)$ be a metric space and let $\mathcal{S}$ be the $\sigma$-algebra of Borel subsets generated by the topology induced by $\rho$. Let $P, P_1, P_2, \ldots$ be probability measures on $(S, \mathcal{S})$.

**Definition B.61.** A sequence of probability measures $\{P_n\}_{n\in\mathbb{N}}$ converges weakly to the probability measure $P$ (notation $P_n \xrightarrow{\mathcal{W}} P$) if

$$\int_E f dP_n \to \int_E f dP$$

for every function $f \in C_b(S)$, the class of continuous bounded functions on $S$.

**Definition B.62.** A set $A$ in $\mathcal{S}$ such that $P(\partial A) = 0$ is called a $P$-*continuity set.*

**Theorem B.63.** *Let $P_n$ and $P$ be probability measures on $(S, \mathcal{S})$. These five conditions are equivalent:*

1. $P_n \xrightarrow{\mathcal{W}} P$,
2. $\lim_n \int f dP_n = \int f dP$ *for all bounded, uniformly continuous real functions* $f$,
3. $\limsup_n P_n(F) \leq P(F)$ *for all closed* $F$,
4. $\liminf_n P_n(G) \geq P(G)$ *for all open* $G$,
5. $\lim_n P_n(A) = P(A)$ *for all* $P$-continuity sets $A$.

On the set of probability measures on $(S, \mathcal{S})$, we may refer to the topology of weak convergence.

**Definition B.64.** Let $\Pi$ be a family of probability measures on $(S, \mathcal{S})$. $\Pi$ is said to be *relatively compact* if every sequence of elements of $\Pi$ contains a weakly convergent subsequence; i.e., for every sequence $\{P_n\}$ in $\Pi$ there exists a subsequence $\{P_{n_k}\}$ and a probability measure $P$ (defined on $(S, \mathcal{S})$, but not necessarily an element of $\Pi$) such that $P_{n_k} \xrightarrow{\mathcal{W}} P$.

**Definition B.65.** A family $\Pi$ of probability measures on the general metric space $(S, \mathcal{S})$ is said to be *tight* if, for all $\epsilon > 0$, there exists a compact set $K$ such that

$$P(K) > 1 - \epsilon \qquad \forall P \in \Pi.$$

Consider sequences of random variables $(X_n)$ and $(Y_n)$ valued in a metric separable space $(S, \rho)$ having common domain; it makes sense to speak of the distance $\rho(X_n, Y_n)$, the function with value $\rho(X_n(\omega), Y_n(\omega))$ at $\omega$. Since $S$ is separable, $\rho(X_n, Y_n)$ is a random variable (see Billingsley (1968), page 225), and we have the following theorem.

**Theorem B.66.** *If* $X_n \xrightarrow{\mathcal{D}} X$ *and* $\rho(X_n, Y_n) \xrightarrow{P} 0$, *then* $Y_n \xrightarrow{\mathcal{D}} X$.

Let $h$ be a measurable mapping of the metric space $S$ into another metric space $S'$. Denote by $h(P)$ the probability measure induced by $h$ on $(S'\mathcal{S}')$, defined by $h(P)(A) = P(h^{-1}(A))$ for any $A \in \mathcal{S}'$. Let $D_h$ be the set of discontinuities of $h$.

**Theorem B.67.** *If* $P_n \xrightarrow{\mathcal{W}} P$ *and* $P(D_h) = 0$*, then* $h(P_n) \xrightarrow{\mathcal{W}} h(P)$*.*

For a random element $X$ of $S$, $h(X)$ is a random element of $S'$ (since $h$ is measurable), and we have the following corollary.

**Corollary B.68.** If $X_n \xrightarrow{\mathcal{D}} X$ and $P(X \in D_h) = 0$, then $h(X_n) \xrightarrow{\mathcal{D}} h(X)$.

We recall now one of the most frequently used results in analysis.

**Theorem B.69.** (Helly). *For every sequence* $(F_n)$ *of distribution functions there exists a subsequence* $(F_{n_k})$ *and a nondecreasing, right-continuous function* $F$ *(a generalized distribution function) such that* $0 \leq F \leq 1$ *and* $\lim_k F_{n_k}(x) = F(x)$ *at continuity points* $x$ *of* $F$*.*

Consider a probability measure $P$ on $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$ and let $\pi_k$ be the projection from $\mathbb{R}^\infty$ to $\mathbb{R}^k$, defined by $\pi_{i_1,\ldots,i_k}(x) = (x_{i_1}, \ldots, x_{i_k})$. The functions $\pi_k(P) : \mathbb{R}^k \to [0,1]$ are called *finite-dimensional distributions* corresponding to $P$. It is possible to show that probability measures on $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$ converge weakly if and only if all the corresponding finite-dimensional distributions converge weakly.

Let $C := C([0,1])$ be the space of continuous functions on $[0,1]$ with the uniform topology, i.e., the topology obtained by defining the distance between two points $x, y \in C$ as $\rho(x,y) = \sup_t |x(t) - y(t)|$. We shall denote with $(C, \mathcal{C})$ the space $C$ with the topology induced by this metric $\rho$.

For $t_1, \ldots, t_k$ in $[0,1]$, let $\pi_{t_1 \ldots t_k}$ be the mapping that carries the point $x$ of $C$ to the point $(x(t_1), \ldots, x(t_k))$ of $\mathbb{R}^k$. The finite-dimensional distributions of a probability measure $P$ on $(C, \mathcal{C})$ are defined as the measures $\pi_{t_1 \ldots t_k}(P)$. Since these projections are continuous, the weak convergence of probability measures on $(C, \mathcal{C})$ implies the weak convergence of the corresponding finite-dimensional distributions, but the converse fails (perhaps in the presence of singular measures).

**Definition B.70.** A sequence $(X_n)$ of random variables with values in a common measurable space $(S, \mathcal{S})$ converges in distribution to the random variable $X$ (notation $X_n \xrightarrow{\mathcal{D}} X$), if the probability laws $P_n$ of the $X_n$ converge weakly to the probability law $P$ of $X$:

$$P_n \xrightarrow{\mathcal{W}} P.$$

## B.2 Prohorov's Theorem

Prohorov's theorem, gives, under suitable hypotheses, equivalence among relative compactness and tightness of families of probability measures.

**Theorem B.71.** (Prohorov). *Let $\Pi$ be a family of probability measures on the probability space $(S, \mathcal{S})$. Then*

1. *if $\Pi$ is tight, then it is relatively compact;*
2. *suppose $S$ is separable and complete; if $\Pi$ is relatively compact, then it is tight.*

*Proof:* See, e.g., Billingsley (1968). □

## B.3 Donsker's Theorem

### Weak Convergence and Tightness in $C([0,1])$

Consider the space $C := C([0,1])$ of continuous functions on $[0,1]$. Weak convergence of finite-dimensional distributions of a sequence of probability measures on $C$ is not a sufficient condition for weak convergence of the sequence itself in $C$. One can prove (see, e.g., Billingsley (1968)) that an additional condition is needed, i.e., relative compactness of the sequence. Since $C$ is a Polish space, i.e., a separable and complete metric space, by Prohorov's theorem we have the following result.

**Theorem B.72.** *Let $(P_n)$ and $P$ be probability measures on $(C, \mathcal{C})$. If the finite-dimensional distributions of $P_n$ converge weakly to those of $P$, and if $\{P_n\}$ is tight, then $P_n \xrightarrow{\mathcal{W}} P$.*

To use this theorem we provide here some characterization of tightness. Given a $\delta \in ]0,1]$, a $\delta$-continuity modulus of an element $x$ of $C$ is defined by

$$w_x(\delta) = w(x, \delta) = \sup_{|s-t|<\delta} |x(s) - x(t)|, \qquad 0 < \delta \leq 1.$$

Let $(P_n)$ be a sequence of probability measures on $(C, \mathcal{C})$.

**Theorem B.73.** *The sequence $(P_n)$ is tight if and only if these two conditions hold:*

1. *For each positive $\eta$, there exists an $a$ such that*
$$P_n(x||x(0)| > a) \leq \eta, \qquad n \geq 1.$$
2. *For each positive $\epsilon$ and $\eta$, there exists a $\delta$, with $0 < \delta < 1$, and an integer $n_0$ such that*
$$P_n(x|w_x(\delta) \geq \epsilon) \leq \eta, \quad n \geq n_0.$$

The following theorem gives a sufficient condition for compactness.

**Theorem B.74.** *If the following two conditions are satisfied:*

1. *For each positive $\eta$, there exists an $a$ such that*

$$P_n(x||x(0)| > a) \leq \eta \qquad n \geq 1.$$

2. *For each positive $\epsilon$ and $\eta$, there exists a $\delta$, with $0 < \delta < 1$, and an integer $n_0$ such that*

$$\frac{1}{\delta} P_n \left( x \,\middle|\, \sup_{t \leq s \leq t+\delta} |x(s) - x(t)| \geq \epsilon \right) \leq \eta, \qquad n \geq n_0,$$

*for all $t$, then the sequence $(P_n)$ is tight.*

Let $X$ be a mapping from $(\Omega, \mathcal{F}, P)$ into $(C, \mathcal{C})$. For all $\omega \in \Omega$, $X(\omega)$ is an element of $C$, i.e., a continuous function on $[0,1]$, whose value at $t$ we denote by $X(t,\omega)$. For fixed $t$, let $X(t)$ denote the real function on $\Omega$ with value $X(t,\omega)$ at $\omega$. Then $X(t)$ is the projection $\pi_t X$.

Similarly, let $(X(t_1), X(t_2), \ldots, X(t_k))$ denote the mapping from $\Omega$ into $\mathbb{R}^k$ with values $(X(t_1,\omega), X(t_2,\omega), \ldots, X(t_k,\omega))$ at $\omega$. If each $X(t)$ is a random variable, $X$ is said to be a random function. Suppose now that $(X_n)$ is a sequence of random functions. According to Theorem B.73, $(X_n)$ is tight if and only if the sequence $(X_n(0))$ is tight, and for any positive real numbers $\epsilon$ and $\eta$ there exists $\delta$, $(0 < \delta < 1)$ and an integer $n_0$ such that

$$P\left(w_{X_n}(\delta) \geq \epsilon\right) \leq \eta, \qquad n \geq n_0.$$

This condition states that the random functions $X_n$ do not oscillate too much. Theorem B.74 can be restated in the same way: $(X_n)$ is tight if $(X_n(0))$ is tight and if for any positive $\epsilon$ and $\eta$ there exists a $\delta$, $0 < \delta < 1$, and an integer $n_0$ such that

$$\frac{1}{\delta} P\left( \sup_{t \leq s \leq t+\delta} |X_n(s) - X_n(t)| \geq \epsilon \right) \leq \eta \tag{B.1}$$

for $n \geq n_0$ and $0 \leq t \leq 1$. Let $\xi_1, \xi_2, \ldots$ be independent identically distributed random variables on $(\Omega, \mathcal{F}, P)$ with mean 0 and variance $\sigma^2$. We define the sequence of partial sums $S_n = \xi_1 + \cdots + \xi_n$, with $S_0 = 0$. Let us construct the sequence of random variables $X_n$ from the sequence $(S_n)$ by means of rescaling and linear interpolation, as follows:

$$X_n\left(\frac{i}{n}, \omega\right) = \frac{1}{\sigma\sqrt{n}} S_i(\omega) \qquad \text{for} \qquad \frac{i}{n} \in [0,1[; \tag{B.2}$$

$$\frac{X_n(t) - X_n\left(\frac{i-1}{n}\right)}{X_n\left(\frac{i}{n}\right) - X_n\left(\frac{i-1}{n}\right)} - \frac{t - \frac{i-1}{n}}{\frac{1}{n}} = 0 \qquad \text{for} \qquad t \in \left[\frac{(i-1)}{n}, \frac{i}{n}\right]. \tag{B.3}$$

With a little algebra, we obtain

$$\begin{aligned}
X_n(t) &= X_n\left(\frac{i-1}{n}\right) + \frac{t-\frac{i-1}{n}}{\frac{1}{n}}\left(X_n\left(\frac{i}{n}\right) - X_n\left(\frac{i-1}{n}\right)\right) \\
&= \frac{t-\frac{i-1}{n}}{\frac{1}{n}} X_n\left(\frac{i}{n}\right) + \left(\frac{\frac{i}{n}-t}{\frac{1}{n}}\right) \\
&= \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega)\frac{\frac{i}{n}-t}{\frac{1}{n}} + \frac{t-\frac{(i-1)}{n}}{\frac{1}{n}}\frac{1}{\sigma\sqrt{n}} S_i(\omega) \\
&= \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega)\left(\frac{\frac{i}{n}-t}{\frac{1}{n}} + \frac{t-\frac{i}{n}+\frac{1}{n}}{\frac{1}{n}}\right) + \frac{1}{\sigma\sqrt{n}}\frac{t-\frac{(i-1)}{n}}{\frac{1}{n}}\xi_i(\omega) \\
&= \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega) + n\left(t - \frac{i-1}{n}\right)\frac{1}{\sigma\sqrt{n}}\xi_i(\omega).
\end{aligned}$$

Since $i-1=[nt]$, if $t \in [\frac{(i-1)}{n}, \frac{i}{n}]$, we may rewrite equation (B.3) as follows:

$$X_n(t,\omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt-[nt])\frac{1}{\sigma\sqrt{n}}\xi_{[nt]+1}(\omega). \tag{B.4}$$

For any fixed $\omega$, $X_n(.,\omega)$ is a piecewise linear function whose pieces' amplitude decreases as $n$ increases. Since the $\xi_i$ and hence the $S_i$ are random variables it follows by (B.4) that $X_n(t)$ is a random variable for each $t$. Therefore, the $X_n$ are random functions.

The following theorem provides a sufficient condition for $(X_n)$ to be a tight sequence.

**Theorem B.75.** *Suppose $(X_n)$ is defined by (B.4). The sequence $(X_n)$ is tight if for each positive $\epsilon$ there exists a $\lambda$, with $\lambda > 1$, and an integer $n_0$ such that, if $n \geq n_0$, then*

$$P\left(\max_{i\leq n}|S_{k+i} - S_k| \geq \lambda\sigma\sqrt{n}\right) \leq \frac{\epsilon}{\lambda^2} \tag{B.5}$$

*holds for all $k$.*

Let us denote by $P_W$ the probability measure of the Wiener process as defined in Definition 2.134 and whose existence is a consequence of Theorem 2.54. We will refer here to its restriction to $t \in [0,1]$, so that its trajectories are almost surely elements of $C([0,1])$.

**Lemma B.76.** *Let $\xi_1, \ldots, \xi_m$ be independent random variables with mean 0 and finite variance $\sigma_i^2$; put $S_i = \xi_1 + \cdots + \xi_i$ and $s_i^2 = \sigma_1^2 + \cdots + \sigma_i^2$. Then*

$$P\left(\max_{i\leq m}|S_i| \geq \lambda s_m\right) \leq 2P\left(|S_m| \geq (\lambda - \sqrt{2})s_m\right). \tag{B.6}$$

**Theorem B.77.** (Donsker). *Let $\xi_1, \xi_2, \dots, \xi_n$ be independent identically distributed random variables defined on $(\Omega, \mathcal{F}, P)$ with mean 0 and finite, positive variance $\sigma^2$:*

$$E[\xi_n] = 0, \qquad E[\xi_n^2] = \sigma^2.$$

*Let $S_n = \xi_1 + \xi_2 + \cdots + \xi_n$. Then the random functions*

$$X_n(t,\omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega)$$

*satisfy $X_n \xrightarrow{\mathcal{D}} W$.*

*Proof:* We first show that the finite-dimensional distributions of $\{X_n\}$ converge to those of $W$. Consider first a single time point $s$; we need to prove that

$$X_n(s) \xrightarrow{\mathcal{W}} W_s.$$

Since

$$\left| X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]} \right| = (ns - [ns]) \left| \frac{1}{\sigma\sqrt{n}} \xi_{[ns+1]} \right|$$

and since, by Chebyshev's inequality,

$$\begin{aligned} P\left( \left| \frac{1}{\sigma\sqrt{n}} \xi_{[ns]+1} \right| \geq 1 \right) &\leq \frac{E\left[ \left| \frac{1}{\sigma\sqrt{n}} \xi_{[ns]+1} \right|^2 \right]}{\epsilon^2} \\ &= \frac{1}{\sigma n \epsilon^2} E\left[ \xi_{[ns]+1}^2 \right] = \frac{1}{\sigma n \epsilon} \sigma^2 \\ &= \frac{\sigma}{n\epsilon} \to 0, \qquad n \to \infty, \end{aligned}$$

we obtain

$$\left| X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]} \right| \xrightarrow{P} 0. \tag{B.7}$$

Since $\lim_{n\to\infty} \frac{[ns]}{ns} = 1$, by the Lindeberg Theorem 1.92

$$\frac{1}{\sigma\sqrt{ns}} \sum_{k=1}^{[ns]} \xi_k \xrightarrow{\mathcal{D}} N(0,1),$$

so that

$$\frac{1}{\sigma\sqrt{n}} S_{[ns]} \xrightarrow{\mathcal{D}} W_s.$$

Therefore, by Theorem B.66, $X_n(s) \xrightarrow{\mathcal{D}} W_s$. Consider now two time points $s$ and $t$ with $s < t$. We must prove

$$(X_n(s), X_n(t)) \xrightarrow{\mathcal{D}} (W_s, W_t).$$

Since

$$\left|X_n(t) - \frac{1}{\sigma\sqrt{n}} S_{[nt]}\right| \xrightarrow{P} 0 \qquad \text{and} \qquad \left|X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]}\right| \xrightarrow{P} 0$$

by Chebyshev's inequality, so that

$$\left\| (X_n(s), X_n(t)) - \left( \frac{1}{\sigma\sqrt{n}} S_{[ns]}, \frac{1}{\sigma\sqrt{n}} S_{[nt]} \right) \right\|_{\mathbb{R}^2} \xrightarrow{P} 0,$$

and by Theorem B.66, it is sufficient to prove that

$$\frac{1}{\sigma\sqrt{n}} \left( S_{[ns]}, S_{[nt]} \right) \xrightarrow{\mathcal{D}} (W_s, W_t).$$

By Corollary B.68 of Theorem B.67 this is equivalent to proving

$$\frac{1}{\sigma\sqrt{n}} \left( S_{[ns]}, S_{[nt]} - S_{[ns]} \right) \xrightarrow{\mathcal{D}} (W_s, W_t - W_s).$$

For independence of the random variables $\xi_i$, $i = 1, 2, \ldots, n$, the random variables $S_{[ns]}$ and $S_{[nt]} - S_{[ns]}$ are independent, so that

$$\begin{aligned}
&\lim_{n\to\infty} E\left[ e^{\frac{\mathrm{i}u}{\sigma\sqrt{n}} \sum_{j=1}^{[ns]} \xi_j + \frac{\mathrm{i}v}{\sigma\sqrt{n}} \sum_{j=[ns]+1}^{[nt]} \xi_j} \right] \\
&= \lim_{n\to\infty} E\left[ e^{\frac{\mathrm{i}u}{\sigma\sqrt{n}} \sum_{j=1}^{[ns]} \xi_j} \right] \cdot \lim_{n\to\infty} E\left[ e^{\frac{\mathrm{i}v}{\sigma\sqrt{n}} \sum_{j=[ns]+1}^{[nt]} \xi_j} \right]. \qquad \text{(B.8)}
\end{aligned}$$

Since $\lim_{n\to\infty} \frac{[ns]}{ns} = 1$, by the Lindeberg Theorem 1.92

$$\frac{1}{\sigma\sqrt{n}} S_{[ns]} \xrightarrow{\mathcal{D}} N(0, s)$$

and for the same reason

$$\frac{1}{\sigma\sqrt{n}} (S_{[nt]} - S_{[ns]}) \xrightarrow{\mathcal{D}} N(0, t - s),$$

so that

$$\lim_{n\to\infty} E\left[ e^{\frac{\mathrm{i}u}{\sigma\sqrt{n}} S_{[ns]}} \right] = e^{-\frac{u^2 s}{2}}$$

and

$$\lim_{n\to\infty} E\left[ e^{\frac{\mathrm{i}v}{\sigma\sqrt{n}} S_{[nt]} - S_{[ns]}} \right] = e^{-\frac{u^2 s}{2}}.$$

Substitution of these two last equations into (B.8) gives

$$\frac{1}{\sigma\sqrt{n}} \left( S_{[ns]}, S_{[nt]} - S_{[ns]} \right) \xrightarrow{\mathcal{D}} (W_s, W_t - W_s),$$

and consequently

$$(X_n(s), X_n(t)) \xrightarrow{\mathcal{D}} (W_s, W_t).$$

A set of three or more time points can be treated in the same way, and hence the finite-dimensional distributions converge properly. Applying Lemma B.76 to the random variables $\xi_1, \xi_2, \ldots, \xi_n$, we have

$$P\left(\max_{i \le n} |S_i| \ge \lambda\sqrt{n}\sigma\right) \le 2P\left(|S_n| \ge (\lambda - \sqrt{2})\sqrt{n}\sigma\right).$$

For $\frac{\lambda}{2} > \sqrt{2}$ we have

$$P\left(\max_{i \le n} |S_i| \ge \lambda\sqrt{n}\sigma\right) \le 2P\left(|S_n| \ge \frac{\lambda}{2}\sqrt{n}\sigma\right).$$

By the Central Limit Theorem,

$$P\left(|S_n| \ge \frac{1}{2}\lambda\sigma\sqrt{n}\right) \to P\left(|N| \ge \frac{1}{2}\lambda\right) < \frac{8}{\lambda^3} E\left[|N|^3\right],$$

where the last inequality follows by Chebyshev's and $N \sim N(0,1)$. Therefore, if $\epsilon$ is positive, there exists a $\lambda$ such that

$$\limsup_{n \to \infty} P\left(\max_{i \le n} |S_i| \ge \lambda\sigma\sqrt{n}\right) < \frac{\epsilon}{\lambda^2}$$

and then, by Theorem B.75, the family of the distribution functions of $X_n$ is tight. Since $C$ is separable and complete, by Prohorov's theorem this family is relatively compact and then $X_n \xrightarrow{\mathcal{D}} X$. □

### An Application of Donsker's Theorem

Donsker's theorem has the following qualitative interpretation: $X_n \xrightarrow{\mathcal{D}} W$ implies that, if $\tau$ is small, then a particle subject to independent displacements $\xi_1, \xi_2, \ldots$ at successive times $\tau_1, \tau_2, \ldots$ appears to follow approximately a Brownian motion.

More important than this qualitative interpretation is the use of Donsker's theorem to prove limit theorems for various functions of the partial sums $S_n$. By using Donsker's theorem it is possible to use the relation $X_n \xrightarrow{\mathcal{D}} W$ to derive the limiting distribution of $\max_{i \le n} S_i$.

Since $h(x) = \sup_t x(t)$ is a continuous function on $C$, $X_n \xrightarrow{\mathcal{D}} W$ implies, by Corollary B.68, that

$$\sup_{0 \le t \le 1} X_n(t) \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} W_t.$$

The obvious relation

$$\sup_{0\le t\le 1} X_n(t) = \max_{i\le n} \frac{1}{\sigma\sqrt{n}} S_i$$

implies

$$\frac{1}{\sigma\sqrt{n}} \max_{i\le n} S_i \xrightarrow{\mathcal{D}} \sup_{0\le t\le 1} W_t. \tag{B.9}$$

Thus, under the hypotheses of Donsker's theorem, if we knew the distribution of $\sup_t W_t$ we would have the limiting distribution of $\max_{i\le n} S_i$. The technique we shall use to obtain the distribution of $\sup_t W_t$ is to compute the limit distribution of $\max_{i\le n} S_i$ in a simple special case and then using $h(X_n) \xrightarrow{\mathcal{D}} h(W)$, where $h$ is continuous on $C$ or continuous except at points forming a set of Wiener measure 0, we obtain the distribution of $\sup_t W_t$ in the general case.

Suppose that $S_0, S_1, \ldots$ are the random variables for a symmetric random walk starting from the origin; this is equivalent to supposing that $\xi_n$ are independent and satisfy

$$P(\xi_n = 1) = P(\xi_n = -1) = \frac{1}{2}. \tag{B.10}$$

Let us show that if $a$ is a nonnegative integer, then

$$P\left(\max_{0\le i\le n} S_i \ge a\right) = 2P(S_n > a) + P(S_n = a). \tag{B.11}$$

If $a = 0$ the previous relation is obvious; in fact, since $S_0 = 0$,

$$P\left(\max_{0\le i\le n} S_i \ge 0\right) = 1$$

and obviously, by symmetry of $S_n$

$$2P(S_n > 0) + P(S_n = 0) = P(S_n > 0) + P(S_n < 0) + P(S_n = 0) = 1.$$

Suppose now that $a > 0$ and put $M_i = \max_{0\le j\le i} S_j$. Since

$$\{S_n = a\} \subset \{M_n \ge a\}$$

and

$$\{S_n > a\} \subset \{M_n \ge a\},$$

we have

$$P(M_n \ge a) - P(S_n = a) = P(M_n \ge a, S_n < a) + P(M_n \ge a, S_n > a)$$

and

$$P(M_n \ge a, S_n > a) = P(S_n > a).$$

Hence we have to show that

$$P(M_n \geq a, S_n < a) = P(M_n \geq a, S_n > a). \tag{B.12}$$

Because of (B.10), all $2^n$ possible paths $(S_1, S_2, \ldots, S_n)$ have the same probability $2^{-n}$. Therefore, (B.12) will follow, if we show that the number of paths contributing to the left-hand event is the same as the number of paths contributing to the right-hand event. To show this it suffices to find a one-to-one correspondence between the paths contributing to the right-hand event and the paths contributing to the left-hand event.

Given a path $(S_1, S_2, \ldots, S_n)$ contributing to the left-hand event in (B.12), match it with the path obtained by reflecting through $a$ all the partial sums after the first one that achieves the height $a$. Since the correspondence is one-to-one, (B.12) follows. This argument is an example of the reflection principle. See also Lemma 2.144.

Let $\alpha$ be an arbitrary nonnegative number, and let $a_n = -[-\alpha n^{\frac{1}{2}}]$. By (B.12) we have

$$P\left(\max_{i \leq n} \frac{1}{\sqrt{n}} S_i \geq a_n\right) = 2P(S_n > a_n) + P(S_n = a_n).$$

Since $S_i$ can assume only integer values and since $a_n$ is the smallest integer greater than or equal to $\alpha n^{\frac{1}{2}}$,

$$P\left(\max_{i \leq n} \frac{1}{\sqrt{n}} S_i \geq \alpha\right) = 2P(S_n < a_n) + P(S_n = a_n). \tag{B.13}$$

By the central limit theorem

$$P(S_n \geq a_n) \to P(N \geq \alpha),$$

where $N \sim N(0,1)$ and $\sigma^2 = 1$ by (B.10).

Since in the symmetric binomial distribution $S_n \to 0$ almost certainly, the term $P(S_n = a_n)$ is negligible. Thus

$$P\left(\max_{i \leq n} \frac{1}{\sqrt{n}} S_i \geq \alpha\right) \to 2P\left(N \geq \alpha\right), \qquad \alpha \geq 0. \tag{B.14}$$

By (B.14), (B.9), and (B.10), we conclude that

$$P\left(\sup_{0 \leq t \leq 1} W_t \leq \alpha\right) = \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-\frac{1}{2}u^2} du, \qquad \alpha \geq 0. \tag{B.15}$$

If we drop the assumption (B.10) and suppose that the random variables $\xi_n$ are independent and identically distributed and satisfy the hypothesis of Donsker's theorem, then (B.9) holds and from (B.15) we obtain

$$P\left(\frac{1}{\sigma\sqrt{n}} \max_{i \leq n} S_i \leq \alpha\right) \to \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-\frac{1}{2}u^2} du, \qquad \alpha \geq 0. \tag{B.16}$$

Thus we have derived the limiting distribution of $\max_{i \le n} S_i$ by Lindeberg's theorem. Therefore, if the $\xi_n$ are independent and identically distributed with $E[\xi_n] = 0$ and $E[\xi_n^2] = \sigma^2$, then the limit distribution of $h(X_n)$ does not depend on any further properties of the $\xi_n$. For this reason, Donsker's theorem is often called an invariance principle.

# C

# Maximum Principles of Elliptic and Parabolic Operators

The maximum principle is a generalization of the fact that if a function $f : [a, b] \to \mathbb{R}$, endowed with a first and second derivative, has $f'' > 0$ ($f'' < 0$) in $[a, b]$, then it attains its maximum (minimum) at the limits of the interval it is defined on.

In fact, if a function, as the solution of a certain differential equation, attains its maximum on the boundary of the domain $\Omega$ on which it is defined, then it is said to underlie a maximum principle. The latter is a remarkable instrument for the study of partial differential equations (e.g., uniqueness of solutions, comparison of solutions, etc.).

## C.1 Maximum Principles of Elliptic Equations

Let $\Omega \subset \mathbb{R}$ be open bounded and let $a, b, c$, be real-valued functions defined on $\Omega$. We consider the partial differential operator

$$L[u] = \frac{1}{2}a(x)u_{xx} + b(x)u_x + c(x)u. \tag{C.1}$$

$L$ is said to be *elliptic in a point* $x_0 \in \Omega$ if $ax_0 > 0$. If for all $x \in \Omega\colon\ a(x) > 0$, then $L$ is said to be *uniformly elliptic*.

**Lemma C.1.** *For $a(x) > 0$, $c(x) \leq 0$, for all $x \in \Omega$, if*

$$\exists \max_{x\in\Omega} u(x) = u(x_0) > 0, \qquad x_0 \in \Omega,$$

*and $u \in C^2(\Omega)$. Then $L[u](x_0) \leq 0$.*

*Proof:*

$$L[u](x_0) = \frac{1}{2}a(x_0)u_{xx}(x_0) + b(x_0)u_x(x_0) + c(x_0)u(x_0),$$

where $c(x_0)u(x_0) \leq 0$, $b(x_0)u_x(x_0) = 0$ and $a(x_0)u_{xx}(x_0) \leq 0$ with $x_0$ being the maximum point of $u$. □

**Theorem C.2.** *Let $a(x) > 0$, $c(x) \leq 0$ for all $x \in \Omega$ and there exists a $\lambda > 0$ such that $\frac{1}{2}\lambda^2 a(x) + \lambda b(x) > 0$ for all $x \in \Omega$. If $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$, $L[u] \geq 0$ in $\Omega$ and if $\max_{\bar{\Omega}} u(x) > 0$, then $\sup_{\Omega} u(x) \leq \max_{\partial\Omega} u(x)$, where $\partial\Omega$ is the boundary of $\Omega$.*

*Proof:* See, e.g., Friedman (1963). □

**Corollary C.3.** Under the assumptions of the preceding theorem, if $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$, $L[u] \leq 0$ in $\Omega$, and if $\min_{\bar{\Omega}} u(x) < 0$, then $\inf_{\Omega} u \geq \min_{\partial\Omega} u$.

*Proof:* See, e.g., Friedman (1963). □

**Theorem C.4.** (Strong maximum principle). *Let $L$ be a uniformly elliptic operator ($a(x) > 0$ in $\Omega$) with bounded coefficients $a, b, c$ on compact sets of $\Omega$ and let $c(x) \leq 0$ in $\Omega$. If $u \in C^2(\Omega)$, $L[u] \geq 0$ ($L[u] \leq 0$) in $\Omega$, and if $u \neq$ constant, then $u$ cannot attain a positive maximum (negative minimum) in $\Omega$.*

*Proof:* See, e.g., Friedman (1963). □

*Remark C.5.* The boundedness of the coefficients $a, b, c$ is essential, as the following example demonstrates:

$$u_{xx} + b(x)u_x = 0, \tag{C.2}$$

where

$$b(x) = \begin{cases} -\frac{3}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

It is easily verified that $u = 1 - x^4$ is the solution of (C.2), and moreover $\max_{[-1,1]} u(x) = u(0) = 1$. In fact, $b$ is not bounded in compact neighborhoods of zero.

## The First Boundary Value or Dirichlet Problem

The Dirichlet problem consists of finding a solution $u$ of the system

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega, \\ u(x) = \phi(x) & \text{in } \partial\Omega. \end{cases} \tag{C.3}$$

**Theorem C.6.** *Let $a(x) > 0$, $c(x) \leq 0$; $a, b, c, f$ uniformly Hölder continuous with exponent $\alpha$ in $\bar{\Omega}$. Then there exists a unique $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$, solution of the Dirichlet problem.*

*Proof:* See, e.g., Friedman (1963) or (1964). □

## C.2 Maximum Principles of Parabolic Equations

Let $Q \subset \mathbb{R}^2$ be open bounded and $a, b, c$ be real-valued functions defined on $Q$. We consider the partial differential operator

$$M[u] = a(x,t)u_{xx} + b(x,t)u_x + c(x,t)u - u_t. \tag{C.4}$$

$M$ is of *parabolic type* in $(x_0, t_0) \in Q$ if $a(x_0, t_0) > 0$. If $a(x,t) > 0$ in $Q$, then $M$ is said to be *uniformly parabolic.*

We suppose $Q \subset \mathbb{R}\times]0,T[$ and define

$$\begin{aligned} D_T &= \{(x,T)|(x,t-\delta) \in Q, \forall 0 < \delta < \delta_0, \delta_0 \text{ independent of } x\}, \\ Q_0 &= Q \cup D_T, \\ \partial_0 Q &= \partial Q \setminus D_T. \end{aligned}$$

$\partial_0 Q$ is a closed set of $\mathbb{R}^2$ and is called a *parabolic boundary.*

*Example C.7.* $Q = \Omega\times]t_0, T[$.

**Theorem C.8.** *Let* $a(x,t) \geq 0$ *and* $c(x,t) \leq 0$ *in* $Q$. *If* $u \in C^0(\bar{Q})$, $u_x, u_{xx}, u_t$ *belong to* $C^0(Q_0)$ *and if* $M[u] \geq 0$ *in* $Q_0$ *and* $\max_{\bar{Q}} u > 0$, *then* $\sup_{Q_0} u(x,t) \leq \max_{\partial_0 Q} u(x,t)$.

*Proof:* See, e.g., Friedman (1963) or (1964). □

**Definition C.9.** For every $P_0 = (x_0, t_0) \in Q$, let $S(P_0) = \{P \in Q|$ a simple continuous curve $\gamma_{P_0}$ exists that is contained within $Q$ and does not decrease along $t$ passing from $P$ to $P_0$ connecting $P$ to $P_0\}$ and let $C(P_0)$ be the connecting component at $t = t_0$ of $Q \cap \{t = t_0\}$ that contains $P_0$. Clearly $C(P_0) \subset S(P_0)$.

**Theorem C.10.** (Strong maximum principle). *Let* $M$ *be uniformly parabolic in* $Q$ *with bounded coefficients and let* $c(x,t) \leq 0$. *If* $u, u_x, u_{xx}, u_t$ *are continuous in* $Q$ *and* $M[u] \geq 0$ *in* $Q$ *and if* $u$ *attains a positive maximum in the point* $P_0 = (x_0, t_0) \in Q$, *then*

$$u(P) = u(P_0) \qquad \forall P \in S(P_0).$$

*Proof:* See, e.g., Friedman (1963) or (1964). □

### The First Boundary Value Problem

Let $Q$ be a domain bounded in $\mathbb{R}$, $Q \subset \mathbb{R}\times]0,T[$ and define

$$\begin{aligned} \tilde{B}_T &= \bar{Q} \cap \{t = T\}, \\ \tilde{B} &= \tilde{Q} \cap \{t = 0\}, \\ B_T = \tilde{B}_T, &\quad B = \tilde{B}, \\ S_0 &= \{(x,t) \in \partial Q, 0 < t \leq T\}, \\ S &= S_0 \setminus B_T, \\ \partial_0 Q &= B \cup S \text{ parabolic boundary of } Q. \end{aligned}$$

The first boundary value problem consists of finding a solution $u$ of the system

$$\begin{cases} M[u](x,t) = f(x,t) & \text{in } Q \cup B_T, \\ u(x,0) = \phi(x) & \text{in } B \text{ (initial condition)}, \\ u(x,t) = g(x,t) & \text{in } S \text{ (boundary condition)},, \end{cases} \tag{C.5}$$

where $f, \phi, g$ are appropriately chosen functions. If $g = \phi$ in $\bar{B} \cap \bar{S}$, then the solution $u$ is always understood to be continuous in $\bar{Q}$.

**Definition C.11.** $\omega_R(P)$ is a *barrier function* in $R \in \bar{B} \cup S$ if

1. $\omega_R(P)$ is continuous in $\bar{Q}$,
2. $\omega_R(P) > 0$ for $P \in \bar{Q}, P \neq R, \omega_R(R) = 0$,
3. $M[\omega_R] \leq -1$ in $Q \cup B_T$.

*Remark C.12.* If $Q = \Omega \times (0,T)$, then there always exists a barrier function in any point $P_0 = (x_0, t_0)$ of $S$ $(0 < t_0 \leq T)$ that is given by

$$\omega_{P_0} = K e^{\gamma t} \left( \frac{1}{R_0^p} - \frac{1}{R} \right),$$

where $K$ and $p$ are positive constants, $\gamma \geq c(x,t), R_0 = |x - x_0|, R = (|x - \bar{x}|^2 + (t - t_0)^2)^{\frac{1}{2}}$ with $\bar{x} > x_0$.

**Theorem C.13.** *Let $M$ be uniformly parabolic in $Q$; the functions $a, b, c$, and $f$ uniformly Hölder continuous in $\bar{Q}$; $\phi$ continuous in $\bar{B}$; and $g$ continuous in $\bar{S}$ with $\phi = g$ in $\bar{B} \cap \bar{S}$. If, for all $R \in S$, there exists $\omega_R$ a barrier function in $R$, then there exists a unique $u$, solution of (C.5), with $u_x, u_{xx}, u_t$ Hölder continuous.*

*Proof:* See, e.g., Friedman (1963) or (1964). □

## The Cauchy Problem

Let $L[u] = a u_{xx} + b u_x + c u$ be an elliptic operator in $\mathbb{R}$ for all $t \in [0,T]$ and let $f : \mathbb{R} \times [0,T] \to \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$ be two appropriately assigned functions. The Cauchy problem consists of finding a solution $u$ of

$$\begin{cases} M[u] \equiv L[u] - u_t = f(x,t) & \text{in } \mathbb{R} \times ]0,T], \\ u(x,0) = \phi(x) & \text{in } \mathbb{R}. \end{cases} \tag{C.6}$$

The solution is understood to be continuous in $\mathbb{R} \times [0,T]$, with the derivatives $u_x, u_{xx}, u_t$ continuous in $\mathbb{R} \times ]0,T]$.

**Theorem C.14.** *Let*

$$0 \leq a(x,t) \leq C, |b(x,t)| \leq C(|x| + 1), c(x,t) \leq C(|x|^2 + 1), \tag{C.7}$$

*where $C$ is a constant. If $M[u] \leq 0$ in $\mathbb{R} \times ]0,T]$, $u(x,t) \geq -B \exp\{\beta |x|^2\}$ in $\mathbb{R} \times [0,T]$ $(B, \beta$ positive constants), $u(x,0) \geq 0$ in $\mathbb{R}$, then $u(x,t) \geq 0$ in $\mathbb{R} \times [0,T]$.*

*Proof:* See, e.g., Friedman (1963) or (1964). □

**Corollary C.15.** If $a(x,t) \geq 0$, satisfying (C.7), then there exists at least one solution $u$ of the Cauchy problem with

$$|u(x,t)| \leq Be^{\beta|x|^2},$$

where $b, \beta$ are positive constants.

*Proof:* See, e.g., Friedman (1963) or (1964). □

**Theorem C.16.** *Let*

$$a(x,t) \geq 0, |a(x,t)| \leq C(|x|^2+1), |b(x,t)| \leq C(|x|+1), c \leq C, \tag{C.8}$$

*where $C$ is a constant. If $M[u] \leq 0$ in $\mathbb{R}\times]0,T]$, $u(x,t) \geq -N(|x|^q+1)$ in $\mathbb{R} \times [0,T]$ ($N, q$ positive constants), $u(x,0) \geq 0$ in $\mathbb{R}$, then $u(x,t) \geq 0$ in $\mathbb{R} \times [0,T]$.*

*Proof:* See, e.g., Friedman (1964). □

**Corollary C.17.** If $a(x,t) \geq 0$, satisfying (C.8), then there exists at least one solution $u$ of the Cauchy problem with

$$|u(x,t)| \leq N(1+|x|^q),$$

where $N, q$ are positive constants.

*Proof:* See, e.g., Friedman (1964). □

**Definition C.18.** A *fundamental solution of the parabolic operator* $L - \frac{\partial}{\partial t}$ in $\mathbb{R} \times [0,T]$ is a function $\Gamma(x,t;\xi,r)$, defined, for all $(x,t) \in \mathbb{R} \times [0,T]$ and all $(\xi,t) \in \mathbb{R} \times [0,T]$, $t > r$, such that, for all $f$ with compact support[14], the function

$$u(x,t) = \int_{\mathbb{R}} \Gamma(x,t;\xi,r) f(\xi) d\xi$$

satisfies

1. $L[u] - u_t = 0$ if $x \in \mathbb{R}, r < t \leq T$,
2. $u(x,t) \to f(x)$ if $t \downarrow r$.

We impose the following conditions:
$(A_1)$ there exists a $\mu > 0$ such that $a(x,t) \geq \mu$ for all $(x,t) \in \mathbb{R} \times [0,T]$;
$(A_2)$ the coefficients of $L$ are continuous functions, bounded in $\mathbb{R} \times [0,T]$, and the coefficient $a(x,t)$ is continuous in $t$ uniformly with respect to $(x,t) \in \mathbb{R} \times [0,T]$;
$(A_3)$ the coefficients of $L$ are Hölder continuous functions (with exponent $\alpha$) in $x$, uniformly with respect to the variables $(x,t)$ in compacts of $\mathbb{R} \times [0,T]$, and the coefficient $a(x,t)$ is Hölder continuous (with exponent $\alpha$) in $x$, uniformly with respect to $(x,t) \in \mathbb{R} \times [0,T]$.

[14] The support of a function $f : \mathbb{R} \to \mathbb{R}$ is the set $\{x \in \mathbb{R} | f(x) \neq 0\}$.

**Theorem C.19.** *If* $(A_1)$*,* $(A_2)$*, and* $(A_3)$ *are satisfied, then there exists* $\Gamma(x,t;\xi,r)$*, a fundamental solution of* $L - \frac{\partial}{\partial t}$*, with*

$$|D_x^m \Gamma(x,t;\xi,r)| \leq c_1(t-r)^{-\frac{m+1}{2}} \exp\left\{-c_2 \frac{(x-\xi)^2}{t-r}\right\}, \qquad m = 0,1,$$

*where* $c_1$ *and* $c_2$ *are positive constants. The functions* $D_x^m \Gamma$*,* $m = 0,1,2$*, and* $D_t \Gamma$ *are continuous in* $(x,t;\xi,r) \in \mathbb{R} \times [0,T] \times \mathbb{R} \times [0,T]$*,* $t > r$*, and* $L[\Gamma] - \Gamma_t = 0$*, as function of* $(x,t)$*. Finally, for all* $f$ *bounded continuous, we have*

$$\int_{\mathbb{R}} \Gamma(x,t;\xi,r) f(x) dx \to f(\xi) \text{ for } t \downarrow r.$$

*Proof:* See, e.g., Friedman (1963). □

**Theorem C.20.** *Let* $(A_1)$*,* $(A_2)$*,* $(A_3)$ *be satisfied,* $f(x,t)$ *be a continuous function in* $\mathbb{R} \times [0,T]$*, Hölder continuous in* $x$*, uniformly with respect to* $(x,t)$ *in compacts of* $\mathbb{R} \times [0,T]$*, and let* $\phi$ *be a continuous function in* $\mathbb{R}$*. Moreover, we suppose that*

$$|f(x,t)| \leq Ae^{a_1|x|^2} \text{ in } \mathbb{R} \times [0,T],$$
$$|\phi(x,t)| \leq Ae^{a_1|x|^2} \text{ in } \mathbb{R},$$

*where* $A, a_1$ *are positive constants. There exists a solution of the Cauchy problem in* $0 \leq t \leq T^*$*, where* $T^* = \min\{T, \frac{\bar{c}}{a_1}\}$ *and* $\bar{c}$ *is a constant, that only depends on the coefficients of* $L$ *and*

$$|u(x,t)| \leq A' e^{a_1'|x|^2} \text{ in } \mathbb{R} \times [0,T^*],$$

*with positive constants* $A', a_1'$*. The solution is given by*

$$u(x,t) = \int_{\mathbb{R}} \Gamma(x,t;\xi,0)\phi(\xi) d\xi - \int_0^t \int_{\mathbb{R}} \Gamma(x,t;\xi,r) f(\xi,r) d\xi dr.$$

The operator $M^*$, as a supplement to $M = L - \frac{\partial}{\partial t}$, is given by

$$M^*[v] = L^*[v] + \frac{\partial v}{\partial t},$$
$$L^*[v] = \frac{1}{2} a v_{xx} + b^* v_x + c^* v,$$

where $b^* = -b + a_x$, $c^* = c - b_x + \frac{1}{2} a_{xx}$, assuming that $a_x$, $a_{xx}$, $a_t$ exist and are bounded.

*Remark C.21.*

$$\begin{aligned} M^*[v] &= \frac{1}{2} a v_{xx} + a_x v_x - b v_x + \frac{1}{2} a_{xx} v - b_x v + c v + v_t \\ &= \frac{1}{2}(av)_{xx} - (bv)_x + cv + v_t, \end{aligned}$$

from which follows *Green's formula*:

$$\begin{aligned}
&vM[u] - uM^*[v] \\
&= v\left(\frac{1}{2}au_{xx} + bu_x + cu - u_t\right) - u\left(\frac{1}{2}(av)_{xx} - (bv)_x + cv + v_t\right) \\
&= \frac{1}{2}(va)u_{xx} - \frac{1}{2}(av)_{xx} + vbu_x - u(bv)_x + vcu - vcu - vu_t - uv_t \\
&= \frac{1}{2}((va)u_x - (ua)v_x - va_x)_x + (vbu)_x - (uv)_t.
\end{aligned}$$

Therefore, if $u$ and $v$ have compact support in a domain $G$, we have that

$$\int\int_G (vMu - uM^*v)dxdt = 0.$$

**Definition C.22.** A *fundamental solution of the operator* $L^* + \frac{\partial}{\partial t}$ in $\mathbb{R} \times [0, T]$ is a function $\Gamma^*(x, t; \xi, r)$, defined, for all $(x, t) \in \mathbb{R} \times [0, T]$ and all $(\xi, t) \in \mathbb{R} \times [0, T]$, $t > r$, such that, for all $g$ continuous with compact support, the function

$$v(x, t) = \int_{\mathbb{R}} \Gamma^*(x, t; \xi, r)g(\xi)d\xi$$

satisfies

1. $L^*[v] + v_t = 0$ if $x \in \mathbb{R}, 0 \leq t \leq r$;
2. $v(x, t) \to g(x)$ if $t \uparrow r$.

We consider the following additional condition.
$(A_4)$ The functions $a$, $a_x$, $a_{xx}$, $b$, $b_x$, $c$ are bounded and the coefficients of $L^*$ satisfy the conditions $(A_2)$ and $(A_3)$.

**Theorem C.23.** *If* $(A_1)$, $(A_1)$, $(A_3)$, *and* $(A_4)$ *are satisfied, then there exists a fundamental solution* $\Gamma^*(x, t; \xi, r)$ *of* $L^* + \frac{\partial}{\partial t}$ *such that*

$$\Gamma(x, t; \xi, r) = \Gamma^*(\xi, r; x, t), \qquad t > r.$$

*Proof:* See, e.g., Friedman (1963). □

# D

# Stability of Ordinary Differential Equations

We consider the system of ordinary differential equations

$$\begin{cases} \frac{d}{dt}\mathbf{u}(t) = \mathbf{f}(t, \mathbf{u}(t)),\ t > t_0, \\ \mathbf{u}(t_0) = \mathbf{c} \end{cases} \tag{D.1}$$

in $\mathbb{R}^d$ and we suppose that, for all $\mathbf{c} \in \mathbb{R}^d$, there exists a unique general solution $\mathbf{u}(t, t_0, \mathbf{c})$ in $[t_0, +\infty[$. We further suppose that $\mathbf{f}$ is continuous in $[t_0, +\infty[\times\mathbb{R}^d$ and that $\mathbf{0}$ is the equilibrium solution of $\mathbf{f}$. Thus $\mathbf{f}(t, \mathbf{0}) = \mathbf{0}$ for all $t \geq t_0$.

**Definition D.1.** The equilibrium solution $\mathbf{0}$ is *stable* if, for all $\epsilon > 0$:

$$\exists\delta = \delta(\epsilon, t_0) > 0 \text{ such that } \forall\mathbf{c} \in \mathbb{R}^d, |\mathbf{c}| < \delta \Rightarrow \sup_{t_0 \leq t \leq +\infty} |\mathbf{u}(t, t_0, \mathbf{c})| < \epsilon. \tag{D.2}$$

If the condition (D.2) is not verified, then the equilibrium solution is *unstable*. The position of the equilibrium is said to be *asymptotically stable* if it is stable and *attractive*, namely, if along with (D.2), it can also be verified that

$$\lim_{t\to+\infty} \mathbf{u}(t, t_0, \mathbf{c}) = \mathbf{0} \qquad \forall\mathbf{c} \in \mathbb{R}^d, |\mathbf{c}| < \delta \text{ (chosen suitably)}. \tag{D.3}$$

*Remark D.2.* There may be attraction without stability.

*Remark D.3.* If $\mathbf{x}^* \in \mathbb{R}^d$ is the equilibrium solution of $\mathbf{f}$, then the position $\mathbf{y}(t) = \mathbf{u}(t) - \mathbf{x}^*$ tends towards $\mathbf{0}$.

**Definition D.4.** We consider the ball $B_h \equiv \bar{B}_h(0) = \{\mathbf{x} \in \mathbb{R}^d | |\mathbf{x}| \leq h\}, h > 0$, which contains the origin. The continuous function $v : B_h \to \mathbb{R}_+$ is *positive definite* (in the Lyapunov sense) if

$$\begin{cases} v(\mathbf{0}) = 0, \\ v(\mathbf{x}) > 0\ \ \forall\mathbf{x} \in B_h \setminus \{\mathbf{0}\}. \end{cases} \tag{D.4}$$

The continuous function $v : [t_0, +\infty[\times B_h \to \mathbb{R}_+$ is *positive definite* if

$$\begin{cases} v(t, \mathbf{0}) = 0 & \forall t \in [t_0, +\infty[, \\ \exists \omega : B_h \to \mathbb{R}_+ \text{ positive definite such that } v(t, \mathbf{x}) \geq \omega(\mathbf{x}) & \forall t \in [t_0, +\infty[. \end{cases} \tag{D.5}$$

$v$ is negative definite if $-v$ is positive definite.

Now let $v : [t_0, +\infty[\times B_h \to \mathbb{R}_+$ be a positive definite function endowed with continuous first partial derivatives with respect to $t$ and $x_i$, $i = 1, \ldots, d$. We consider the function

$$V(t) = v(t, \mathbf{u}(t, t_0, \mathbf{c})) : [t_0, +\infty[\to \mathbb{R}_+,$$

where $\mathbf{u}(t, t_0, \mathbf{c})$ is the solution of (D.1). $V$ is differentiable with respect to $t$ and we have

$$\frac{d}{dt} V(t) = \frac{\partial v}{\partial t} + \sum_{i=1}^{d} \frac{\partial v}{\partial x_i} \frac{du_i}{dt}.$$

But $\frac{du_i}{dt} = f_i(t, \mathbf{u}(t, t_0, \mathbf{c}))$, therefore

$$\dot{v} \equiv \frac{d}{dt} V(t) = \frac{\partial v}{\partial t} + \sum_{i=1}^{d} \frac{\partial v}{\partial x_i} f_i(t, \mathbf{u}(t, t_0, \mathbf{c})),$$

and this is the derivative of $v$ with respect to time "along the trajectory" of the system. If $\frac{d}{dt} V(t) \leq 0$ for all $t \in (t_0, +\infty[$, then $\mathbf{u}(t, t_0, \mathbf{c})$ does not increase the value $v$, which measures by how much $\mathbf{u}$ moves away from $\mathbf{0}$. Through this observation, the required stability of the Lyapunov criterion for the stability of $\mathbf{0}$ has been formulated.

**Definition D.5.** Let $v : [t_0, +\infty[\times B_h \to \mathbb{R}_+$ be a positive definite function. $v$ is said to be a *Lyapunov function for the system (D.1) relative to the equilibrium position* $\mathbf{0}$, if

1. $v$ is endowed with first partial derivatives with respect to $t$ and $x_i, i = 1, \ldots, d$;
2. for all $t \in ]t_0, +\infty[$: $\dot{v}(t) \leq 0$ for all $c \in B_h$.

**Theorem D.6.** (Lyapunov).

1. *If there exists $v(t, \mathbf{x})$ a Lyapunov function for the system (D.1) relative to the equilibrium position $\mathbf{0}$, then $\mathbf{0}$ is stable;*
2. *if moreover the Lyapunov function $v(t, \mathbf{x})$ is such that, for all $t \in [t_0, +\infty[$: $v(t, \mathbf{x}) \leq \omega(\mathbf{x})$ with $\mathbf{u}$ being a positive definite function and $\dot{v}$ negative definite along the trajectory, then $\mathbf{0}$ is asymptotically stable.*

*Example D.7.* We consider the autonomous linear system

$$\begin{cases} \frac{d}{dt} \mathbf{u}(t) = A \mathbf{u}(t), \; t > t_0, \\ \mathbf{u}(t_0) = \mathbf{c}, \end{cases}$$

where $A$ is a matrix that does not depend on time. A matrix $P$ is said to be positive definite if, for all $\mathbf{x} \in \mathbb{R}^d, \mathbf{x} \neq \mathbf{0} : \mathbf{x}'P\mathbf{x} > 0$. Considering the function $v(\mathbf{x}) = \mathbf{x}'P\mathbf{x}$, we have

$$\dot{v} = \frac{d}{dt}v(\mathbf{u}(t)) = \sum_{i=1}^{d} \frac{\partial v}{\partial x_i}(A\mathbf{u}(t))_i = \mathbf{u}'(t)PA\mathbf{u}(t) + \mathbf{u}'(t)A'P\mathbf{u}(t).$$

Therefore, if $P$ is such that $PA + A'P = -Q$, with $Q$ being positive definite, then $\dot{v} = -\mathbf{u}'Q\mathbf{u} < 0$ and, by 2 of Lyapunov's theorem, $\mathbf{0}$ is asymptotically stable.

# References


1. Aalen O. O., *Nonparametric inference for a family of counting processes*, Annals of Statistics, 701–726, 6; 1978.
2. Aletti G. & Capasso V., *Profitability in a multiple strategy market*, Decis. Econ. Finance, 145–152, 26; 2003.
3. Andersen P. K., Borgan Ø., Gill R.D. & Keiding N., *Statistical Models Based on Counting Processes*, Springer-Verlag, Heidelberg; 1993.
4. Anderson W. J., *Continuous-Time Markov Chains. An Application-Oriented Approach*, Springer–Verlag, New York; 1991.
5. Arnold L., *Stochastic Differential Equations*, Wiley, New York; 1974.
6. Ash R. B., *Real Analysis and Probability*, Academic Press, London; 1972.
7. Ash R. B. & Gardner M. F., *Topics in Stochastic Processes*, Academic Press, London; 1975.
8. Aubin J.-P., *Applied Abstract Analysis*, Wiley, New York; 1977.
9. Bachelier L., *Théorie de la spéculation*, Ann. Sci. École Norm. Sup., 21–86, 17; 1900.
10. Bailey N. T. J., *The Mathematical Theory of Infectious Diseases*, Griffin, London; 1975.
11. Baldi P., *Equazioni differenziali stocastiche*, UMI, Bologna; 1984.
12. Bartholomew D. J., *Continuous time diffusion models with random duration of interest*, J. of Math. Sociology, 187–199, 4; 1976.
13. Bauer H., *Probability Theory and Elements of Measure Theory*, Academic Press, London; 1981.
14. Becker N., *Analysis of Infectious Disease Data*, Chapman and Hall, London; 1989.
15. Bertoin J., *Lévy Processes*, Cambridge University Press, Cambridge, UK; 1996.
16. Bhattacharya R.N. & Waymire E.C., *Stochastic Processes with Applications*, Wiley, New York; 1990.
17. Billingsley P., *Convergence of Probability Measures*, Wiley, New York; 1968.
18. Black F. & Scholes M., *The pricing of options and corporate liabilities*, J. of Pol. Econ., 637–654, 81; 1973.
19. Borodin A. & Salminen P., *Handbook of Brownian Motion: Facts and Formulae*, Birkhä user, Boston; 1996.
20. Boyle P. & Tian Y., *Pricing lookback and barrier options under the CEV process*, J. of Fin. and Quant. Analysis, 241–264, 34; 1999.

21. Brace A., Gatarek D., & Musiela M., *The market model of interest rate dynamics*, Math. Fin., 127–154, 7; 1997.
22. Breiman L., *Probability*, Addison–Wesley, Reading, MA; 1968.
23. Bremaud P., *Point Processes and Queues: Martingale Dynamics*, Springer-Verlag, Heidelberg; 1981.
24. Burger M., Capasso V., & Morale D., *On an aggregation model with long and short range interaction*, MIRIAM reports, Milan; 2003.
25. Capasso V., *A counting process approach for stochastic age-dependent population dynamics*, in *Biomathematics and Related Computational Problems* (L. M. Ricciardi, Ed.), pp. 255–269, Kluwer, Dordrecht; 1988.
26. Capasso V., *A counting process approach for age-dependent epidemic systems*, in *Stochastic Processes in Epidemic Theory* (J. P. Gabriel et al., Eds.) pp. 118–128, Lecture Notes in Biomathematics, Vol. 86, Springer-Verlag, Heidelberg; 1990.
27. Capasso V., *Mathematical Structures of Epidemic Systems*, Springer-Verlag, Heidelberg; 1993.
28. Capasso V., Di Liddo A. & Maddalena L., *Asymptotic behaviour of a nonlinear model for the geographical diffusion of innovations*, Dynamic Systems and Applications, 207–220, 3; 1994.
29. Capasso V., Morale D. & Sioli F., *An agent-based model for "price herding", applied to the automobile market*, MIRIAM reports, Milan; 2003.
30. Carrillo J., *Entropy solutions for nonlinear degenerate problems*, Arch. Rat. Mech. Anal., 269–361, 147; 1999.
31. Cherny A. S. & Shiryaev A. N., *Vector stochastic integrals and the fundamental theorems of asset pricing*, French–Russian A. M. Liapunov Institute for Applied Mathematics and Computer Science Transactions, 5–37, 3; 2001.
32. Chiang C.L., *Introduction to Stochastic Processes in Biostatistics*, Wiley, New York; 1968.
33. Chung K. L., *A Course in Probability Theory*,Second Edition, Academic Press, New York; 1974.
34. Cox J. C., *The constant elasticity of variance option pricing model*, J. of Portfolio Management, 15–17, 22; 1996.
35. Cox J. C., Ross S. A., & Rubinstein M., *Option pricing: A simplified approach*, J. Fin. Econ., 229–263, 7; 1979.
36. Dalang R. C., Morton A., & Willinger W., *Equivalent martingale measures and non-arbitrage in stochastic securities market models*, Stochastics and Stochastic Reports, 185–201, 29; 1990.
37. Darling D. A. D. & Siegert A. J. F., *The first passage time problem for a continuum Markov process*, Ann. Math. Stat., 624–639, 24; 1953.
38. Davis M. H. A., *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, Journal of the Royal Statistical Society, Series B, 353–388, 46; 1984.
39. Delbaen F. & Schachermeyer W., *A general version of the fundamental theorem of asset pricing*, Mathematische Annalen, 463–520, 300; 1994.
40. De Masi A. & Presutti E., *Mathematical Methods for Hydrodynamical Limits*, Springer-Verlag, Heidelberg; 1991.
41. Dieudonné J., *Foundations of Modern Analysis*, Academic Press, New York; 1960.
42. Donsker M. D. & Varadhan S. R. S., *Large deviations from a hydrodynamical scaling limit*, Comm. Pure Appl. Math., 243–270, 42; 1989.

43. Doob J. L., *Stochastic Processes*, Wiley, New York; 1953.
44. Duffie D., *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, NJ; 1996.
45. Dupire B., *Pricing with a smile*, RISK, 18–20, 1; 1994.
46. Durrett R. & Levin S. A., *The importance of being discrete (and spatial)*, Theor. Pop. Biol., 363–394, 46; 1994.
47. Dynkin E. B., *Markov Processes*, Springer-Verlag, Berlin, Vols. 1–2; 1965.
48. Embrechts P., Klüppelberg C., & Mikosch T., *Modelling Extreme Events for Insurance and Finance*, Springer-Verlag, Berlin; 1997.
49. Epstein J. & Axtell R., *Growing Artificial Societies–Social Sciences from the Bottom Up*, Brooking Institution Press & MIT Press, Cambridge, MA; 1996.
50. Ethier S. N. & Kurtz T. G., *Markov Processes, Characterization and Convergence*, Wiley, New York; 1986.
51. Feller W., *An Introduction to Probability Theory and Its Applications*, Wiley, New York; 1971.
52. Flierl G., Grünbaum D., Levin S. A., & Olson D., *From individuals to agrregations: The interplay between behavior and Physics*, J. Theor. Biol., 397–454, 196; 1999.
53. Fournier N. & Méléard S., *A microscopic probabilistic description of a locally regulated population and macroscopic approximations*, Prépublications du Laboratoire de Probabilités et Modèles Aléatoires, Universités Paris 6/7, PMA-798; 2003.
54. Friedman A., *Partial Differential Equations*, Krieger, New York; 1963.
55. Friedman A., *Partial Differential Equations of Parabolic Type*, Prentice–Hall, London; 1964.
56. Friedman A., *Stochastic Differential Equations and Applications*, Academic Press, London; 1975.
57. Fristedt B. & Gray L., *A Modern Approach to Probability Theory*, Birkhäuser, Boston; 1997.
58. Gard T. C., *Introduction to Stochastic Differential Equations*, Marcel Dekker, New York; 1988.
59. Gihman I. I. & Skorohod A. V., *Stochastic Differential Equations*, Springer-Verlag, Berlin; 1972.
60. Gihman I. I. & Skorohod A. V., *The Theory of Random Processes*, Springer-Verlag, Berlin; 1974.
61. Gihman I. I. & Skorohod A. V., *Introduction à la théorie des processes aléatoires*, MIR, Moscow; 1980.
62. Gnedenko B. V., *The Theory of Probability*, Chelsea, New York; 1963.
63. Grigoriu M., *Stochastic Calculus. Applications to Science and Engineering*, Birkhäuser, Boston; 2002.
64. Grünbaum D. & Okubo A., *Modelling social animal aggregations*, in *Frontiers of Theoretical Biology* (S. A. Levin, Ed.) pp. 296–325, Lectures Notes in Biomathematics, Vol. 100, Springer-Verlag, New York; 1994.
65. Gueron S., Levin S. A., & Rubenstein D. I., *The dynamics of herds: From individuals to aggregations*, J. Theor. Biol., 85–98, 182; 1996.
66. Harrison J. M. & Kreps D. M., *Martingales and arbitrage in multiperiod securities markets*, J. Econ. Th., 381–408, 20; 1979.
67. Harrison J. M. & Pliska S. R., *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Processes Appl., 215–260, 11; 1981.

68. Has'minskii R. Z., *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff, The Netherlands; 1980.
69. Heath D., Jarrow R., & Morton A., *Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation*, Econometrica, 77–105, 1; 1992.
70. Hunt P. J. & Kennedy J. E., *Financial Derivatives in Theory and Practice*, Wiley, New York; 2000.
71. Hull J. & White A., *Pricing interest rate derivative securities*, Rev. Fin. Stud., 573–592, 4; 1990.
72. Ikeda N. & Watanabe S., *Stochastic Differential Equations and Diffusion Processes*, North–Holland, Kodansha; 1989.
73. Jacobsen M., *Statistical Analysis of Counting Processes*, Springer-Verlag, Heidelberg; 1982.
74. Jacod J. & Shiryaev A. N., *Limit Theorems for Stochastic Processes*, Springer Lecture Notes in Math., Springer-Verlag, Berlin; 1987.
75. Jamshidian F., *LIBOR and swap market models and measures*, Finance and Stochastics, 43–67, 1; 1997.
76. Jelinski Z. & Moranda P., *Software reliability research*, in *Statistical Computer Performance Evaluation*, pp. 466–484, Academic Press, New York; 1972.
77. Kallenberg O., *Foundations of Modern Probability*, Springer-Verlag, Berlin; 1997.
78. Karatzas I. & Shreve S. E., *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York; 1991.
79. Karlin S. & Taylor H. M., *A First Course in Stochastic Processes*, Academic Press, New York; 1975.
80. Karlin S. & Taylor H. M., *A Second Course in Stochastic Processes*, Academic Press, New York; 1981.
81. Karr A. F., *Point Processes and Their Statistical Inference*, Marcel Dekker, New York; 1986.
82. Kolmogorov A. N., *Foundations of the Theory of Probability*, Chelsea, New York; 1956.
83. Kolmogorov A. N. & Fomin E. S. V., *Elements of Theory of Functions and Functional Analysis*, Groylock, Moscow; 1961.
84. Kou S., *A jump diffusion model for option pricing*, Man. Sc., 1086–1101, 48; 2002.
85. Lamberton D. & Lapeyre B., *Introduction to Stochastic Calculus applied to Finance*, Chapman & Hall, London; 1996.
86. Lamperti J., *Probability*, Benjamin, New York; 1966.
87. Last G. & Brandt A., *Marked Point Processes on the Real Line. The Dynamic Approach*, Springer-Verlag, Heidelberg; 1995.
88. Loève M., *Probability Theory*, Van Nostrand–Reinhold, Princeton, NJ; 1963.
89. Mahajan V. & Wind Y., *Innovation Diffusion Models of New Product Acceptance*, Ballinger, Cambridge, MA; 1986.
90. Méléard S., *Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models*, in *Probabilistic Models for Nonlinear Partial Differential Equations* (D. Talay, L. Tubaro, Eds.) pp. 42–95, Lecture Notes in Mathematics, Vol. 1627, CIME Subseries, Springer-Verlag, Heidelberg; 1996.
91. Merton R. C., *Theory of rational option pricing*, Bell J. Econ. Man. Sc., 141–183, 4; 1973.

92. Merton R. C., *Option pricing when underlying stock returns are discontinuous*, J. Fin. Econ., 125–144, 3; 1976.
93. Métivier M., *Notions fondamentales de la théorie des probabilités*, Dunod, Paris; 1968.
94. Meyer P. A., *Probabilités et Potentiel*, Ilermann, Paris; 1966.
95. Miltersen K. R., Sandmann K., & Sondermann D., *Closed form solutions for term structure derivatives with log-normal interest rates*, J. of Finance, 409–430, 52; 1997.
96. Morale D., Capasso V., & Oelschläger K., *An interacting particle system modelling aggregation behaviour: From individuals to populations*, J. Math. Biology; 2004.
97. Morale D., Capasso V., & Oelschläger K., *A rigorous derivation of the mean-field nonlinear integro-differential equation for a population of aggregating individuals subject to stochastic fluctuations*, Preprint 98–38 (SFB 359), IWR, Universität Heidelberg, Juni; 1998.
98. Musiela M. & Rutkowski M., *Martingale Methods in Financial Modelling*, Springer-Verlag, Berlin; 1998.
99. Nagai T. & Mimura M., *Some nonlinear degenerate diffusion equations related to population dynamics*, J. Math. Soc. Japan, 539–561, 35; 1983.
100. Neveu J., *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco; 1965.
101. Oelschläger K., *A law of large numbers for moderately interacting diffusion processes*, Z. Wahrsch. Verw. Geb., 279–322, 69; 1985.
102. Oelschläger K., *Large systems of interacting particles and the porous medium equation*, J. Differential Equations, 294–346, 88; 1990.
103. Øksendal B., *Stochastic Differential Equations*, Springer-Verlag, Berlin; 1998.
104. Okubo A., *Dynamical aspects of animal grouping: Swarms, school, flocks and herds*, Adv. BioPhys., 1–94, 22; 1986.
105. Parzen E., *Stochastic Processes*, Holden-Day, San Francisco; 1962.
106. Pliska S. R., *Introduction to Mathematical Finance: Discrete-Time Models*, Blackwell Publishers, Oxford; 1997.
107. Protter P., *Stochastic Integration and Differential Equations*, Springer-Verlag, Berlin; 1990.
108. Rebolledo R., *Central limit theorems for local martingales*, Z. Wahrsch. verw. Geb., 269–286, 51; 1980.
109. Robert P., *Stochastic Networks and Queues*, Springer-Verlag, Heidelberg; 2003.
110. Rogers L. C. G. & Williams D., *Diffusions, Markov Processes and Martingales*, Vol. 1, Wiley, New York; 1994.
111. Rolski T., Schmidli H., Schmidt V., & Teugels J., *Stochastic Processes for Insurance and Finance*, Wiley, New York; 1999.
112. Roozen H., *Equilibrium and extinction in stochastic population dynamics*, Bull. Math. Biology, 671–696, 49; 1987.
113. Samorodnitsky G. & Taqqu M. S., *Stable Non-Gaussian Random Processes*, Chapman & Hall/CRC Press, Boca Ration, FL; 1994.
114. Sato K. I., *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK; 1999.
115. Schuss Z., *Theory and Applications of Stochastic Differential Equations*, Wiley, New York; 1980.
116. Shiryaev A. N., *Probability*, Springer-Verlag, New York; 1995.

117. Skellam J. G., *Random dispersal in theoretical populations*,Biometrika, 196–218, 38; 1951.
118. Sobczyk K., *Stochastic Differential Equations. With Applications to Physics and Engineering*, Kluwer, Dordrecht; 1991.
119. Stein R. B., *A theoretical analysis of neuronal variability*, Biophys. J., 173–194, 5; 1965.
120. Stein R. B., *Some models of neuronal variability*, Biophys. J., 37–68, 7; 1967.
121. Taira K., *Diffusion Processes and Partial Differential Equations*, Academic Press, New York; 1988.
122. Tan W. Y., *Stochastic Models with Applications to Genetics, Cancers, AIDS and other Biomedical Systems*, World Scientific, Singapore; 2002.
123. Tuckwell H. C., *On the first exit time problem for temporarily homogeneous Markov process*, J. Appl. Prob., 39–48, 13; 1976.
124. Tuckwell H. C., *Stochastic Processes in the Neurosciences*, SIAM, Philadelphia; 1989.
125. Vasicek O., *An equilibrium characterisation of the term structure*, J. Fin. Econ., 177–188, 5; 1977.
126. Wang F. J. S., *Gaussian approximation of some closed stochastic epidemic models*, J. Appl. Prob., 221–231, 14; 1977.
127. Warburton K. & Lazarus J., *Tendency-distance models of social cohesion in animal groups*, J. Theor. Biol., 473–488, 150; 1991.
128. Wax N., *Selected Papers on Noise and Stochastic Processes*, Dover, New York; 1954.
129. Williams D., *Probability with Martingales*, Cambridge University Press, Cambridge, UK; 1991.
130. Wilmott P., Howison S. D., & Dewynne J. N., *Option Pricing: Mathematical Models and Computation*, Oxford Fin. Press, Oxford; 1993.
131. Yang G., *Stochastic epidemics as point processes*, in *Mathematics in Biology and Medicine* (V. Capasso, E. Grosso, S. L. Paveri-Fontana, Eds.), pp. 135–144, Lectures Notes in Biomathematics, Vol. 57, Springer-Verlag, Heidelberg; 1985.

# Nomenclature

| | |
|---|---|
| | "increasing" is used with the same meaning as "nondecreasing"; "decreasing" is used with the same meaning as "nonincreasing." In the strict cases "strictly increasing/strictly decreasing" is used. |
| $(\Omega, \mathcal{F}, P)$ | probability space with $\Omega$ a set, $\mathcal{F}$ a $\sigma$-algebra of parts of $\Omega$, and $P$ a probability measure on $\mathcal{F}$ |
| $(E, \mathcal{B}_E)$ | measurable space with $E$ a set and $\mathcal{B}_E$ a $\sigma$-algebra of parts of $E$ |
| $:=$ | equal by definition |
| $< f, g >$ | scalar product of two elements $f$ and $g$ in an Hilbert space |
| $< M, N >$ | predictable covariation of the martingales $M$ and $N$ |
| $< M >, < M, M >$ | predictable variation of the martingale $M$ |
| $[a, b[$ | semiopen interval closed at extreme $a$ and open at extreme $b$ |
| $[a, b]$ | closed interval of extremes $a$ and $b$ |
| $\bar{\mathbb{R}}$ | extended set of real numbers; i.e., $\mathbb{R} \cup \{-\infty, +\infty\}$ |
| $\bar{A}$ | closure of a set $A$ depending upon the context |
| $\bar{C}$ | the complement of the set $C$ depending upon the context |
| $\Delta$ | Laplace operator |
| $\delta_x$ | Dirac delta-function localized at $x$ |
| $\delta_{ij}$ | Kronecker delta; i.e., $= 1$ for $i = j$, $= 0$ for $i \neq j$ |
| $\emptyset$ | the empty set |
| $\epsilon_x$ | Dirac delta-measure localized at $x$ |
| $\equiv$ | coincide |
| $\exp\{x\}$ | exponential function $e^x$ |
| $\int^*$ | integral of a nonnegative measurable function, finite or not |
| $\lim_{s \downarrow t}$ | limit for $s$ decreasing while tending to $t$ |
| $\lim_{s \uparrow t}$ | limit for $s$ increasing while tending to $t$ |
| $\mathbb{C}$ | the complex plane |
| $\mathbb{N}$ | the set of natural nonnegative integers |
| $\mathbb{N}^*$ | the set of natural (strictly) positive integers |

| | |
|---|---|
| $\mathbb{Q}$ | the set of rational numbers |
| $\mathbb{R}^n$ | $n$-dimensional Euclidean space |
| $\mathbb{R}_+$ | the set of positive (nonnegative) real numbers |
| $\mathbb{R}_+^*$ | the set of (strictly) positive real numbers |
| $\mathbb{Z}$ | the set of all integers |
| $\mathcal{A}$ | infinitesimal generator of a semigroup |
| $\mathcal{B}_{\mathbb{R}^n}$ | $\sigma$-algebra of Borel sets on $\mathbb{R}^n$ |
| $\mathcal{B}_E$ | $\sigma$-algebra of Borel sets generated by the topology of $E$ |
| $\mathcal{D}_A$ | domain of definition of an operator $A$ |
| $\mathcal{F}_t$ or $\mathcal{F}_t^X$ | history of a process $(X_t)_{t\in\mathbb{R}_+}$ up to time $t$; i.e., the $\sigma$-algebra generated by $\{X_s, s \leq t\}$ |
| $\mathcal{F}_{t+}$ | $\bigcap_{s>t} \mathcal{F}_t$ |
| $\mathcal{F}_{t-}$ | $\sigma$-algebra generated by $\sigma(X_s, s < t)$ |
| $\mathcal{F}_X$ | $\sigma$-algebra generated by the random variable $X$ |
| $\mathcal{L}(X)$ | probability law of $X$ |
| $\mathcal{L}^p(P)$ | set of integrable functions with respect to the measure $P$ |
| $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ | set of all $\mathcal{F}$-measurable functions with values in $\bar{\mathbb{R}}_+$ |
| $\mathcal{M}(E)$ | set of all measures on $E$ |
| $\mathfrak{P}(\Omega)$ | the set of all parts of a set $\Omega$ |
| $\xrightarrow[n]{P}$ or $P$-lim | convergence in probability |
| $\xrightarrow[n]{\mathcal{W}}$ | weak convergence |
| $\xrightarrow[n]{a.s.}$ | almost sure convergence |
| $\xrightarrow[n]{d}$ | convergence in distribution |
| $\xrightarrow[n]{P}$ | convergence in probability |
| $\nabla$ | gradient |
| $\Omega$ | the underlying sample space |
| $\omega$ | an element of the underlying sample space |
| $\otimes$ | product of $\sigma$-algebras or product of measures |
| $\partial A$ | boundary of a set $A$ |
| $\Phi$ | cumulative distribution function of a standard normal probability law |
| $\mathrm{sgn}\{x\}$ | sign function; 1, if $x > 0$; 0, if $x = 0$; $-1$, if $x < 0$ |
| $\sigma(\mathcal{R})$ | $\sigma$-algebra generated by the family of events $\mathcal{R}$ |
| $\square$ | end of a proof |
| $\|a\|$ | absolute value of a number $a$; or modulus of a complex number $a$ |
| $\|A\|$ or $\sharp(A)$ | cardinal number (number of elements) of a finite set $A$ |
| $\|\|x\|\|$ | the norm of a point $x$ |
| $]a, b[$ | open interval of extremes $a$, $b$ |
| $]a, b]$ | semiopen interval open at extreme $a$ and closed at extreme $b$ |
| $a \vee b$ | maximum of two numbers |

| | |
|---|---|
| $A'$ | transpose of a matrix $A$ |
| $A \setminus B$ | the set of elements of $A$ that do not belong to $B$ |
| $a \wedge b$ | minimum of two numbers |
| $B(x,r)$ or $B_r(x)$ | the open ball centered at $x$ and having radius $r$ |
| $C(A)$ | set of continuous functions from $A$ to $\mathbb{R}$ |
| $C(A,B)$ | set of continuous functions from $A$ to $B$ |
| $C^k(A)$ | set of functions from $A$ to $\mathbb{R}$ with continuous derivatives up to order $k$ |
| $C^{k+\alpha}(A)$ | set of functions from $A$ to $\mathbb{R}$ whose $k$-th derivatives are Lipschitz continuous with exponent $\alpha$ |
| $C_0(A)$ | set continuous functions on $A$ with compact support |
| $C_b(A)$ or $BC(A)$ | set of bounded continuous functions on $A$ |
| $Cov[X,Y]$ | the covariance of two random variables $X$ and $Y$ |
| $E[\cdot]$ | expected value with respect to an underlying probability law clearly identifiable from the context |
| $E[Y\|\mathcal{F}]$ | conditional expectation of a random variable $Y$ with respect to the $\sigma$-algebra $\mathcal{F}$ |
| $E_P[\cdot]$ | expected value with respect to the probability law $P$ |
| $E_x[\cdot]$ | expected value conditional upon a given initial state $x$ in a stochastic process |
| $f * g$ | convolution of functions $f$ and $g$ |
| $f \circ X$ or $f(X)$ | a function $f$ composed with a function $X$ |
| $f\|_A$ | the restriction of a function f to the set $A$ |
| $f^-$, $f^+$ | negative (positive) part of $f$; i.e., $f^- = \max\{-f, 0\}$ ($f^+ = \max\{f, 0\}$) |
| $f^{-1}(B)$ | the preimage of the set $B$ by the function $f$ |
| $F_X$ | cumulative distribution function of a random variable $X$ |
| $H \bullet X$ | stochastic Stieltjes integral of the process $H$ with respect to the stochastic process $X$ |
| $I_A$ | indicator function associated with a set $A$; i.e., $I_A(x) = 1$, if $x \in A$ otherwise $I_A(x) = 0$ |
| $L^p(P)$ | set of equivalence classes of a.e. equal integrable functions with respect to the measure $P$ |
| $N(\mu, \sigma^2)$ | normal (Gaussian) random variable with mean $\mu$ and variance $\sigma^2$ |
| $O(\Delta)$ | of the same order as $\Delta$ |
| $o(\delta)$ | of higher order with respect to $\delta$ |
| $P$-a.s. | almost surely with respect to the measure P |
| $P(A\|B)$ | conditional probability of an event $A$ with respect to an event $B$ |
| $P * Q$ | convolution of measures $P$ and $Q$ |
| $P \ll Q$ | the measure $P$ is absolutely continuous with respect to the measure $Q$ |
| $P \sim Q$ | the measure $P$ is equivalent to the measure $Q$ |
| $P_X$ | probability law of a random variable $X$ |

| | |
|---|---|
| $P_x$ | probability law conditional upon a given initial state $x$ in a stochastic process |
| $Var[X]$ | the variance of a random variable $X$ |
| $W_t$ | standard Brownian motion, Wiener process |
| $X \sim P$ | the random variable $X$ has probability law $P$ |
| a.e. | almost everywhere |
| a.s. | almost surely |

# Index