A

Measure and Integration

A.1 Rings and σ-Algebras

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

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

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 Ω is called a σ*-ring* if it satisfies the following additional condition:

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

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

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

1. $\emptyset \in \mathcal{F}$.
2. $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$.
3. $A, B \in \mathcal{F}, A \subset B \Rightarrow \exists (A_j)_{i\leq j\leq m} \in \mathcal{F}^{\{1,\ldots,m\}}$ 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 Ω has the following properties:*

1. *If $\mathcal{F}$ is a σ-algebra of subsets of Ω, then it is an algebra.*
2. *If $\mathcal{F}$ is a σ-algebra of subsets of Ω, then*
   - $E_1, \ldots, E_n \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{n} E_i \in \mathcal{F}$
   - $E_1, \ldots, E_n, \ldots \in \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} E_n \in \mathcal{F}$

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

- $B \in \mathcal{F} \Rightarrow \Omega \setminus B \in \mathcal{F}$

3. If $\mathcal{F}$ is a ring on Ω, then it is also a semiring.

Definition A.5. Every pair $(\Omega, \mathcal{F})$ consisting of a set Ω and a σ-ring $\mathcal{F}$ of the subsets of Ω is a *measurable space*. Furthermore, if $\mathcal{F}$ is a σ-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 σ-algebra on the set E and $X : \Omega \to E$ a generic mapping, then the set
$$X^{-1}(\mathcal{B}) = \left\{A \subset \Omega | \exists B \in \mathcal{B} \text{ such that } A = X^{-1}(B)\right\}$$
is a σ-algebra on Ω.
2. *Generated σ-algebra*. If $\mathcal{A}$ is a set of the elements of a set Ω, then there exists the smallest σ-algebra of subsets of Ω that contains $\mathcal{A}$. This is the σ-algebra *generated* by $\mathcal{A}$, denoted $\sigma(\mathcal{A})$. If, now, $\mathcal{G}$ is the set of all σ-algebras of subsets of Ω 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 σ-algebra.* Let Ω be a topological space. Then the *Borel σ-algebra* on Ω, denoted by $\mathcal{B}_\Omega$, is the σ-algebra generated by the set of all open subsets of Ω. Its elements are called Borelian or Borel-measurable.
4. The set of all left-open, right-closed bounded intervals of $\mathbb{R}$, defined as $(a, b] := \{x \in \mathbb{R} \,|\, a < x \leq b\}$, for $a, b \in \mathbb{R}$, is a semiring but not a ring.
5. The set of all bounded and unbounded intervals of $\mathbb{R}$ is a semialgebra.
6. If $\mathcal{B}_1$ and $\mathcal{B}_2$ are algebras on Ω_1 and Ω_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.
7. *Product σ-algebra.* Let $(\Omega_i, \mathcal{F}_i)_{1 \leq i \leq n}$ 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 elements of Ω. The σ-algebra generated by $\mathcal{R}$ is called the *product σ-algebra* of the σ-algebras $(\mathcal{F}_i)_{1 \leq i \leq n}$.

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

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 σ-algebras $\mathcal{F}$ and $\mathcal{B}_{\mathbb{R}^n}$, the latter being the Borel σ-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 and $\mathcal{U}$ 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. *Let $(\Omega, \mathcal{F})$ be a measurable space and $X_n : \Omega \to \mathbb{R}$, $n \in \mathbb{N}$, a sequence of measurable functions converging pointwise to a function $X : \Omega \to \mathbb{R}$; then X is itself measurable.*

Proposition A.16. *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.17. *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.18. *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 $1 \leq i \leq n$ is the ith projection. Then the product σ-algebra $\bigotimes_{i=1}^n \mathcal{F}_i$ of the family of σ-algebras $(\mathcal{F}_i)_{1 \leq i \leq n}$ is the smallest σ-algebra on Ω for which every projection π_i is measurable.*

Proposition A.19. *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 $i = 1, \ldots n$, $h_i = \pi_i \circ h$ is measurable.*

Proof. 1 ⇒ 2 follows from Proposition A.18 and Lemma A.17. To prove that 2 ⇒ 1, it is sufficient to see that given $\mathcal{R}$, the set of rectangles on Ω, it follows that, for all $B \in \mathcal{R}:\ h^{-1}(B) \in \mathcal{B}$. Let $B \in \mathcal{R}$. Then for all $i = 1, \ldots, n$, there exists a $B_i \in \mathcal{F}_i$ such that $B = \prod_{i=1}^n B_i$. Therefore, by recalling that due to point 2 every h_i is measurable, we have that

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

Corollary A.20. *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 $i = 1, \ldots, n$, h_i is Borel-measurable.*

Definition A.21. Let $(\Omega, \mathcal{F})$ be a measurable space. Every function $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ such that

1. For all $E \in \mathcal{F}$: $\mu(E) \geq 0$.
2. For all $E_1, \ldots, E_n, \ldots \in \mathcal{F}$ 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

$$\mu(\Omega) = 1,$$

then μ is a *probability measure* or a *probability*. Furthermore, a measure μ is *finite* if

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

and σ-*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 $n \in \mathbb{N}\colon\ \mu(A_n) < +\infty$.

Definition A.22. The ordered triple $(\Omega, \mathcal{F}, \mu)$, where Ω denotes a set, $\mathcal{F}$ a σ-ring on Ω, and $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ a measure on $\mathcal{F}$, is a *measure space*. If μ is a probability measure, then $(\Omega, \mathcal{F}, \mu)$ is a *probability space*.[12]

Definition A.23. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $\lambda : \mathcal{F} \to \bar{\mathbb{R}}$ a measure on Ω. Then λ is said to be *absolutely continuous* with respect to μ, denoted $\lambda \ll \mu$, if

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

Proposition A.24 (Characterization of a measure). *Let μ 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. *μ 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 μ is finite, then 1 and 2 are equivalent to the following statements.

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*

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

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

Proposition A.25 (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. *μ is (finitely) additive on $\mathcal{G}$.*
2. *μ 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 σ-ring generated by $\mathcal{G}$.[13] *Moreover, if $\mathcal{G}$ is a semialgebra and $\mu(E) = 1$, then $\bar{\mu}$ is a probability measure.*

Proposition A.26. *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 σ-ring generated by $\mathcal{U}$,*

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

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

Lemma A.27. (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.28. *Under the assumptions of Fatou's Lemma A.27, 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.29. *Let $(\Omega, \mathcal{F})$ be a measurable space and μ a positive measure on $\mathcal{F}$. Then there exists a unique mapping Φ from $\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$ to $\bar{\mathbb{R}}_+$, such that:*

1. *For every $\alpha \in \mathbb{R}_+$, $f, g \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$,*
   $\Phi(\alpha f) = \alpha \Phi(f)$,
   $\Phi(f + g) = \Phi(f) + \Phi(g)$,
   $f \leq g \Rightarrow \Phi(f) \leq \Phi(g)$.

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

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.30. If Φ is the unique functional associated with μ, a 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 μ.

Remark A.31. Let $(\Omega, \mathcal{F})$ be a measurable space, and let Φ be the functional canonically associated with μ measure on $\mathcal{F}$.

1. If $s : \Omega \to \bar{\mathbb{R}}_+$ is an elementary function, and 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 $\Omega_f = \left\{s : \Omega \to \bar{\mathbb{R}}_+ | s \text{ elementary }, s \leq f\right\}$, then Ω_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.32. Let $(\Omega, \mathcal{F})$ be a measurable space and μ a positive measure on $\mathcal{F}$. An $\mathcal{F}$-measurable function f is μ-*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 μ, denoted by

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

Proposition A.33. *Let* $(\Omega, \mathcal{F})$ *be a measurable space endowed with measure* μ *and* $f \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$. *Then*

1. $\int^* f d\mu = 0 \Leftrightarrow f = 0$ *a.s. with respect to* μ.
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* $f = g$, *a.s. with respect to* μ, *we have*

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

Theorem A.34 (Monotone convergence). *Let* $(\Omega, \mathcal{F})$ *be a measurable space endowed with measure* μ, $(f_n)_{n\in\mathbb{N}}$ *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.35 (Lebesgue's dominated convergence). *Let* $(\Omega, \mathcal{F})$ *be a measurable space endowed with measure* μ, $(f_n)_{n\in\mathbb{N}}$ *a sequence of* μ*-integrable functions defined on* Ω, *and* $g : \Omega \to \bar{\mathbb{R}}_+$ *a* μ*-integrable function such that* $|f_n| \leq g$ *for all* $n \in \mathbb{N}$. *If we suppose that* $\lim_{n\to\infty} f_n = f$ *exists almost surely in* Ω, *then* f *is* μ*-integrable and we have*

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

Lemma A.36. (Fatou). *Let* $f_n \in \mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$. *Then*

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

Theorem A.37 (Fatou–Lebesgue).

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

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

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

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

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

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

Definition A.38. Let $(\Omega, \mathcal{F})$ be a measurable space endowed with measure μ, and let $(E, \mathcal{B})$ be an additional measurable space; 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 the *induced* or *image measure* of μ via h, and denoted $h(\mu)$.

Proposition A.39. *Given the assumptions of Definition A.38, the function* $g : (E, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *is integrable with respect to* μ_h *if and only if* $g \circ h$ *is integrable with respect to* μ *and*

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

Theorem A.40 (Product measure). *Let* $(\Omega_1, \mathcal{F}_1)$ *and* $(\Omega_2, \mathcal{F}_2)$ *be measurable spaces, and let the former be endowed with* σ*-finite measure* μ_1 *on* $\mathcal{F}_1$*. Further suppose that for all* $\omega_1 \in \Omega_1$ *a measure* $\mu(\omega_1, \cdot)$ *is assigned on* $\mathcal{F}_2$*, 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* σ*-finite, then there exists a sequence* $(B_n)_{n \in \mathbb{N}} \in \mathcal{F}_2^{\mathbb{N}}$ *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* $\mu(\omega_1, B_n) \leq K_n$ *for all* $\omega_1 \in \Omega_1$*. Then there exists a unique measure* μ *on the product* σ*-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.41. Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces endowed with σ-finite measures μ_1 and μ_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, μ is σ-finite on $\mathcal{F}$ as well as a probability measure if μ_1 and μ_2 are as well. The measure μ is the *product measure* of μ_1 and μ_2, denoted by $\mu_1 \otimes \mu_2$.

Theorem A.42 (Fubini). *Given the assumptions of Definition A.41, 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.43. *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 σ-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 σ-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 $A_1 \times \cdots \times A_n \in \mathcal{F}$:*

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

   μ is σ-finite on $\mathcal{F}$ and a probability whenever μ_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.44.

1. *Given the assumptions and the notation of Proposition A.43, 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* μ *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 μ *is the product measure of* $\mu_1, \ldots, \mu_n$ *and is denoted by* $\mu_1 \otimes \cdots \otimes \mu_n$*.*

Definition A.45. Let $(v_i)_{1 \leq i \leq n}$ be a family of measures defined on $\mathcal{B}_{\mathbb{R}}$, and

$$v^{(n)} = v_1 \otimes \cdots \otimes v_n$$

their product measure on $\mathcal{B}_{\mathbb{R}^n}$. The *convolution product* of $v_1, \ldots, v_n$, denoted by $v_1 * \cdots * v_n$, is the induced measure of $v^{(n)}$ on $\mathcal{B}_{\mathbb{R}}$ via the function $f : (x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto \sum_{i=1}^n x_i \in \mathbb{R}$.

Proposition A.46. *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.47. 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.48. Every function $F : \mathbb{R} \to \mathbb{R}$ that is right-continuous and increasing is a *(generalized) 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 to every Lebesgue–Stieltjes measure can be assigned a distribution function and vice versa.

Proposition A.49. *Let μ 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 μ.

Conversely, the following holds.

Proposition A.50. *Let F be a distribution function, and let μ be defined on bounded intervals of $\mathbb{R}$ by*

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

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

Definition A.51. 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.52. 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 a *Lebesgue measure* on $\mathbb{R}$. It is such that

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

Definition A.53. Let $(\Omega, \mathcal{F}, \mu)$ be a space with σ-finite measure μ, and consider another measure $\lambda : \mathcal{F} \to \bar{\mathbb{R}}_+$. λ is said to be defined through its *density* with respect to μ 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 λ with respect to μ. In this case λ is absolutely continuous with respect to μ ($\lambda \ll \mu$). If μ is a Lebesgue measure on $\mathbb{R}$, then g is the density of μ. A measure ν is called *μ-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.54 (Radon–Nikodym). *Let $(\Omega, \mathcal{F})$ be a measurable space, μ a σ-finite measure on $\mathcal{F}$, and λ an absolutely continuous measure with respect to μ. Then λ is endowed with density with respect to μ. 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 μ-integrable is that λ is bounded. Moreover, if $h : \Omega \to \bar{\mathbb{R}}_+$ is another density of λ, then $g = h$, almost surely with respect to μ.

Theorem A.55 (Lebesgue–Nikodym). *Let ν and μ be a measure and a σ-finite measure on $(E, \mathcal{B})$, respectively. There exists a $\mathcal{B}$-measurable function $f : E \to \bar{\mathbb{R}}_+$ and a μ-singular measure ν' on $(E, \mathcal{B})$ such that*

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

Furthermore,

1. *ν' 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 μ.

Definition A.56. 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.57. *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.58. *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 the 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.59. *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.60 (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 $[a, b]$.*

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

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

Proposition A.61. *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.62. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and $p > 0$. The set of Borel-measurable functions defined on Ω, such that $\int_\Omega |f|^p d\mu < +\infty$, is a vector space on $\mathbb{R}$; it is denoted with the symbols $\mathcal{L}^p(\mu)$ or $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$. Its elements are called integrable functions, 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 μ.

A.5 Radon Measures

Consider a complete metric space E endowed with its Borel σ-algebra $\mathcal{B}_E$.

Definition A.63. A σ-finite measure μ on $\mathcal{B}_E$ is called

(i) *locally finite* if, for any point $x \in E$, there exists an open neighborhood U of x such that $\mu(U) < +\infty$.

(ii) *inner regular* if

$$\mu(A) = \sup\{\mu(K) \,|\, K \quad \text{compact}, \quad K \subset A\} \qquad \forall A \in \mathcal{B}_E.$$

(iii) *outer regular* if

$$\mu(A) = \sup\{\mu(U) \,|\, U \quad \text{open}, \quad A \subset U\} \qquad \forall A \in \mathcal{B}_E.$$

(iv) *regular* if it is both inner and outer regular.

(v) a *Radon measure* if it is an inner regular and locally finite measure.

Proposition A.64. *The usual Lebesgue measure on $\mathbb{R}^d$ is a regular Radon measure. However, not all σ-finite measures on $\mathbb{R}^d$ are regular.*

Proof. See, e.g., Klenke (2008, p. 247). □

Proposition A.65. *If μ is a Radon measure on a locally compact and complete metric space E endowed with its Borel σ-algebra, then*

$$\mu(K) < +\infty, \qquad \forall K \quad \text{compact subset of } E.$$

$$\left| \int_E f \, d\mu \right| < +\infty$$

for any real-valued continuous function f with compact support.

Proof. See, e.g., Karr (1991, p. 411). □

Let us now stick to a locally compact and complete metric space E endowed with its Borel σ-algebra $\mathcal{B}_E$.

Definition A.66. A Radon measure μ on $\mathcal{B}_E$ is

(i) A *point* or (*counting*) measure if $\mu(A) \in \mathbb{N}$, for any $A \in \mathcal{B}_E$.

(ii) A *simple point* measure if μ is a point measure and $\mu(\{x\}) \leq 1$ for any $x \in E$.

(iii) A *diffuse* measure if $\mu(\{x\}) = 0$ for any $x \in E$.

The fundamental point measure is the Dirac measure ϵ_x associated with a point $x \in E$; it is defined by

$$\epsilon_x(A) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

A point $x \in E$ is called an *atom* if $\mu(\{x\}) > 0$.

Proposition A.67. *A Radon measure μ on a locally compact and complete metric space E endowed with its Borel σ-algebra has an at most countable set of atoms. It can be decomposed as*

$$\mu = \mu_d + \sum_{i=1}^{K} a_i \epsilon_{x_i},$$

where μ_d is a diffuse measure, $K \in \mathbb{N} \cup \{\infty\}$, $a_i \in \mathbb{R}_+^*$, $x_i \in E$. The decomposition is unique up to reordering.

Proof. See, e.g., Karr (1991, p. 412). □

A Radon measure is *purely atomic* if its diffuse component is zero.

Remark A.68. A purely atomic measure is a point measure if and only if $a_i \in \mathbb{N}$ for each i, and in this case the family $\{x_i,\ i = 1, \ldots, K\}$ can have no accumulation points in E.

A.6 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 σ-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

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

Theorem A.69. *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 $\rho : R \times R \to \mathbb{R}_+$ that 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 ρ; we shall write (R, ρ). Elements of a metric space will be called *points*.

Definition B.3. Given a metric space (R, ρ), 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, ρ), 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.

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

Definition B.8. In a metric space (R, ρ), a *closed set* is any subset of R that is the 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, ρ), 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, ρ), 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, ρ) and (R', ρ'), a function $f : R \to R'$ is *continuous* if for any open set A' in (R', ρ'), the set $f^{-1}(A')$ is an open set in (R, ρ).

Definition B.16. Two metric spaces (R, ρ) and (R', ρ') 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 ρ and ρ' 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, ρ) and (R', ρ').

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, ρ), 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, ρ) and (R', ρ'), a function $f : R \to R'$ 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 their corresponding everywhere dense countable subsets.

- 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 $x = (x_1, x_2, \ldots, x_n)$ with distance function $\rho(x, y) = \left\{\sum_{k=1}^{n} (y_k - x_k)^2\right\}^{\frac{1}{2}}$, with the set of all vectors with rational coordinates.
- The space $\mathbb{R}_0^n$ of ordered n-tuples of real numbers $x = (x_1, x_2, \ldots, x_n)$ with distance function $\rho_0(x, y) = \max\{|y_k - x_k|; 1 \le k \le n\}$ 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 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, then 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, ρ), 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, ρ), 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, p. 67 for a counterexample).*

Definition B.39. In a metric space (R, ρ), 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, ρ) is any nonempty subset F of R endowed with the restriction of ρ 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, whereas 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 that 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 a 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 now (S, ρ) be a separable metric space endowed with the σ-algebra $\mathcal{S}$ of Borel subsets generated by the topology induced by ρ. As usual, given a probability space $(\Omega, \mathcal{F}, P)$, an S-valued random variable X is an $\mathcal{F} - \mathcal{S}$-measurable function $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$.

Definition B.61. A sequence $(X_n)_{n \in \mathbb{N}}$ of random variables, with values in the common measurable space $(S, \mathcal{S})$, converges almost surely to the random variable X (notation $X_n \stackrel{\text{a.s.}}{\to} X$) if for almost all $\omega \in \Omega$, $X_n(\omega)$ converges to $X(\omega)$ with respect to the metric ρ.

In a metric space, in the foregoing definition only the elements of $(X_n)_{n \in \mathbb{N}}$ are required to be measurable, i.e., random variables, since in any case the limit function will automatically be itself measurable, i.e., a random variable (e.g., Dudley 2005, p. 125). We further remark that, since (S, ρ) is a separable metric space, for any two S-valued random variables X and Y, the distance $\rho(X, Y)$ is a real-valued random variable, so that the following definition makes sense.

Definition B.62. A sequence $(X_n)_{n \in \mathbb{N}}$ of random variables with values in the common measurable space $(S, \mathcal{S})$ converges (*in probability*) to the random variable X (notation $X_n \stackrel{P}{\to} X$) if for any $\varepsilon > 0$,

$$P(\rho(X_n, X) > \varepsilon) \to 0,$$

as $n \to \infty$.

Theorem B.63. *For random variables valued in a separable metric space, almost sure convergence implies convergence in probability.*

The converse of this theorem does not hold in general, though the following theorem holds.

Theorem B.64. *For random variables $(X_n)_{n\in\mathbb{N}}$ and X, valued in a separable metric space, $X_n \xrightarrow{P} X$ if and only if for every subsequence of $(X_n)_{n\in\mathbb{N}}$ there exists a subsubsequence that converges to X a.s.*

Proof. See, e.g., Dudley (2005, p. 288). □

Within the foregoing framework, let $\mathcal{L}^0(\Omega, \mathcal{F}, S, \mathcal{S})$ or simply $\mathcal{L}^0(S, \mathcal{S})$ denote the set of all $\mathcal{F} - \mathcal{S}$-measurable functions (i.e., S-valued random variables); we will then denote by $L^0(S, \mathcal{S})$ the set of equivalence classes of elements of $\mathcal{L}^0(S, \mathcal{S})$ with respect to the usual P-a.s. equality. Given two elements $X, Y \in \mathcal{L}^0(S, \mathcal{S})$, define

$$\alpha(X, Y) := \inf \{\varepsilon \geq 0 \mid P(\rho(X, Y) > \varepsilon) \leq \varepsilon\}.$$

Theorem B.65. *On $L^0(S, \mathcal{S})$, α is a metric that metrizes convergence in probability, so that for random variables $(X_n)_{n\in\mathbb{N}}$ and X, valued in the separable metric space S, $X_n \xrightarrow{P} X$ if and only if $\alpha(X_n, X) \to 0$.*

Proof. See, e.g., Dudley (2005, p. 289). □

The metric α is called the *Ky Fan* metric.

Theorem B.66. *If (S, ρ) is a complete separable metric space, then $L^0(S, \mathcal{S})$, endowed with the Ky Fan metric α, is complete.*

Proof. See, e.g., Dudley (2005, p. 290). □

Let (S, ρ) be a metric space endowed with its Borel σ-algebra $\mathcal{S}$ as above. Let $P, P_1, P_2, \ldots$ be probability measures on $(S, \mathcal{S})$, and let $C_b(S)$ be the class of all continuous bounded real-valued functions on S.

Definition B.67. A sequence of probability measures $(P_n)_{n\in\mathbb{N}}$ on $(S, \mathcal{S})$ converges weakly to a probability measure P (notation $P_n \xrightarrow{\mathcal{W}} P$) if

$$\int_S f dP_n \to \int_S f dP$$

for every function $f \in C_b(S)$.

Proposition B.68. *If (S, ρ) is a metric space, then P and Q are two probability laws on $\mathcal{S}$, and, for any $f \in C_b(S)$, $\int_S f dP = \int_S f dQ$, then $P = Q$.*

An important consequence of the previous proposition is uniqueness of the weak limit of a sequence of probability laws.

Definition B.69. A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables with values in a common measurable space $(S, \mathcal{S})$ converges in distribution to the random variable X (notation $X_n \stackrel{\mathcal{D}}{\to} X$) if the probability laws P_n of the X_n converge weakly to the probability law P of X:

$$P_n \stackrel{\mathcal{W}}{\to} P.$$

If we denote by $\mathcal{L}(X)$ the probability law of a random variable X, then the foregoing convergence can be equivalently written as

$$\mathcal{L}(X_n) \stackrel{\mathcal{W}}{\to} \mathcal{L}(X).$$

Proposition B.70. *If (S, ρ) is a separable metric space, for random variables $(X_n)_{n\in\mathbb{N}}$ and X, valued in S,*

$$X_n \stackrel{P}{\to} X \Rightarrow X_n \stackrel{\mathcal{D}}{\to} X.$$

Recall that if for some $x \in S$, $\mathcal{L}(X) = \epsilon_x$, i.e., X is a degenerate random variable, then

$$X_n \stackrel{P}{\to} X \Longleftrightarrow X_n \stackrel{\mathcal{D}}{\to} X.$$

Theorem B.71 (Skorohod representation theorem). *Consider a sequence $(P_n)_{n\in\mathbb{N}}$ of probability measures and a probability measure P on a separable metric space $(S, \mathcal{S})$ such that $P_n \xrightarrow[n\to\infty]{\mathcal{W}} P$. Then there exists a sequence of S-valued random variables $(Y_n)_{n\in\mathbb{N}}$ and a random variable Y defined on a common (suitably extended) probability space such that Y_n has probability law P_n, Y has probability law P, and*

$$Y_n \xrightarrow[n\to\infty]{\text{a.s.}} Y.$$

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

Consider sequences of random variables $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ valued in a metric separable space (S, ρ) having a common domain; it makes sense to speak of the distance $\rho(X_n, Y_n)$, i.e., the function with value $\rho(X_n(\omega), Y_n(\omega))$ at ω. Since S is separable, $\rho(X_n, Y_n)$ is a random variable (Billingsley 1968, p. 225), and we have the following theorem.

Theorem B.72. *If $X_n \stackrel{\mathcal{D}}{\to} X$ and $\rho(X_n, Y_n) \stackrel{P}{\to} 0$, then $Y_n \stackrel{\mathcal{D}}{\to} X$.*

Let h be a measurable mapping of the metric space S into another metric space S'. If P is a probability measure on $(S, \mathcal{S})$, then we 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.73. *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.74. *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.75. (Helly). *For every sequence* $(F_n)_{n\in\mathbb{N}}$ *of distribution functions there exists a subsequence* $(F_{n_k})_{k\in\mathbb{N}}$ *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.

Definition B.76. A set A in $\mathcal{S}$ such that $P(\partial A) = 0$ is called a P-*continuity set.*

Theorem B.77 (Portmanteau theorem). *Let* $(P_n)_{n\in\mathbb{N}}$ *and* P *be probability measures on a metric space* (S, ρ) *endowed with its Borel* σ*-algebra. 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.

Consider a metric space (S, ρ). Given a bounded real-valued function f on S, we may consider its Lipschitz seminorm defined as

$$\|f\|_L := \sup_{x\neq y} \frac{| f(x) - f(y) |}{\rho(x, y)}$$

and its supremum norm $\|f\|_\infty := \sup_x | f(x) |$. Let

$$\|f\|_{BL} := \|f\|_L + \|f\|_\infty,$$

and consider the set $BL(S, \rho)$ of all bounded real-valued Lipschitz functions on S, i.e.,

$$BL(S, \rho) := \{f : S \to \mathbb{R} | \ \|f\|_{BL} < \infty\} .$$

Theorem B.78. *Let* (S, ρ) *be a metric space.*

1. $BL(S, \rho)$ *is a vector space.*
2. $\|\cdot\|_{BL}$ *is a norm.*
3. $(BL(S, \rho), \|\cdot\|_{BL})$ *is a Banach space.*

For any two probability laws P and Q on the Borel σ-algebra of (S, ρ) we may define

$$\beta(P, Q) := \sup \left\{ | \int f dP - \int f dQ | \mid \|f\|_{BL} \leq 1 \right\}.$$

Theorem B.79. *Let (S, ρ) be a metric space endowed with its Borel σ-algebra $\mathcal{S}$. β is a metric on the set of all probability laws on $\mathcal{S}$.*

Now on a metric space (S, ρ) consider any subset $A \subset S$ and for any $\varepsilon > 0$ let

$$A^\varepsilon := \{ y \in S \mid \rho(x, y) < \varepsilon \text{ for some } x \in A \} .$$

For any two probability laws P and Q on the Borel σ-algebra $\mathcal{S}$ we may define

$$\gamma(P, Q) := \inf \{ \varepsilon > 0 \mid P(A) \leq Q(A^\varepsilon) + \varepsilon, \text{ for all } A \in \mathcal{S} \} .$$

Theorem B.80. *Let (S, ρ) be a metric space endowed with its Borel σ-algebra $\mathcal{S}$. γ is a metric on the set of all probability laws on $\mathcal{S}$.*

The metric γ is known as the *Prohorov metric*, or sometimes the *Lévy–Prohorov metric.*

Theorem B.81. *Let (S, ρ) be a separable metric space endowed with its Borel σ-algebra $\mathcal{S}$; consider a sequence $(P_n)_{n \in \mathbb{N}}$ and a P probability measure on $\mathcal{S}$. These four statements are equivalent.*

(a) $P_n \xrightarrow{\mathcal{W}} P$
(b) $\lim_n \int f dP_n = \int f dP$ *for all functions* $f \in BL(S, \rho)$
(c) $\lim_n \beta(P_n, P) = 0$
(d) $\lim_n \gamma(P_n, P) = 0$

Proof. See, e.g., Dudley (2005, p. 395). □

The fact that convergence in probability implies convergence in law can be expressed in terms of the Prohorov and the Ky Fan metrics as follows.

Theorem B.82. *Let (S, ρ) be a separable metric space endowed with its Borel σ-algebra $\mathcal{S}$, and let X, Y be two S-valued random variables defined on the same probability space. Then*

$$\gamma(\mathcal{L}(X), \mathcal{L}(Y)) \leq \alpha(X, Y).$$

Convergence of Empirical Measures

Consider a metric space (S, ρ) endowed with its Borel σ-algebra $\mathcal{S}$, and let $(X_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. S-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$. The sequence $(P_n)_{n\in\mathbb{N}^*}$ of *empirical measures* associated with $(X_n)_{n\in\mathbb{N}^*}$ is defined by

$$P_n(B)(\omega) := \frac{1}{n}\sum_{j=1}^{n} \epsilon_{X_j(\omega)}(B), \qquad B \in \mathcal{S}, \quad \omega \in \Omega,$$

where ϵ_x is the usual Dirac measure associated with a point $x \in S$.

The following theorem is a generalization of the Glivenko–Cantelli theorem, also known as the Fundamental Theorem of Statistics.

Theorem B.83 (Varadarajan). *Let (S, ρ) be a separable metric space endowed with its Borel σ-algebra $\mathcal{S}$; let $(X_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. S-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$; and let P_X denote their common probability law on $\mathcal{S}$. Then the sequence of empirical measures $(P_n)_{n\in\mathbb{N}^*}$ associated with $(X_n)_{n\in\mathbb{N}^*}$ converges to P_X almost surely, i.e.,*

$$P(\{\omega \in \Omega \mid P_n(\cdot)(\omega) \to P_X\}) = 1.$$

Proof. See, e.g., Dudley (2005, p. 399). □

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

Definition B.84. Let Π be a family of probability measures on $(S, \mathcal{S})$. Π is said to be *relatively compact* if every sequence of elements of Π contains a weakly convergent subsequence, i.e., for every sequence $(P_n)_{n\in\mathbb{N}}$ in Π there exists a subsequence $(P_{n_k})_{k\in\mathbb{N}}$ and a probability measure P [defined on $(S, \mathcal{S})$, but not necessarily an element of Π] such that $P_{n_k} \xrightarrow{\mathcal{W}} P$.

Theorem B.85. *Let $(P_n)_{n\in\mathbb{N}}$ be a relatively compact sequence of probability measures and P an additional probability measure on $(S, \mathcal{S})$. Then the following propositions are equivalent:*

(a) $P_n \xrightarrow{\mathcal{W}} P$.
(b) All weakly converging subsequences of $(P_n)_{n\in\mathbb{N}}$ weakly converge to P.

Definition B.86. A family Π 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_ε such that

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

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.87 (Prohorov). *Let Π be a family of probability measures on the measurable space $(S, \mathcal{S})$. Then*

1. *If Π is tight, then it is relatively compact.*
2. *Suppose S is separable and complete; if Π is relatively compact, then it is tight.*

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

Corollary B.88. *Let (S, ρ) be a Polish space endowed with its Borel σ-algebra $\mathcal{S}$; then the metric space of all probability measures on $\mathcal{S}$ is complete with either metric β or γ.*

Proof. See, e.g., Dudley (2005, p. 405). □

B.3 Donsker's Theorem

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

Consider a probability measure P on $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$, and let π_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 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 ρ. For $t_1, \ldots, t_k$ in $[0,1]$, let $\pi_{t_1 \ldots t_k}$ be the mapping that carries point x of C to 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), i.e., 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 (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.89. *Let $(P_n)_{n\in\mathbb{N}}$ and P be probability measures on $(C,\mathcal{C})$. If the sequence of the finite-dimensional distributions of P_n, $n \in \mathbb{N}$ converge weakly to those of P, and if $(P_n)_{n\in\mathbb{N}}$ 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 δ-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 $(P_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $(C,\mathcal{C})$.

Theorem B.90. *The sequence $(P_n)_{n\in\mathbb{N}}$ is tight if and only if these two conditions hold:*

1. *For each positive η there exists an a_η such that*
$$P_n(x||x(0)| > a_\eta) \leq \eta, \qquad n \geq 1.$$
2. *For each positive ϵ and η there exists a δ, 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.91. *If the following two conditions are satisfied:*

1. *For each positive η, there exists an a such that*
$$P_n(x||x(0)| > a) \leq \eta \qquad n \geq 1.$$
2. *For each positive ϵ and η, there exists a δ, 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 $t \in [0,1]$, then the sequence $(P_n)_{n\in\mathbb{N}}$ 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 Ω with value $X(t,\omega)$ at ω. 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 Ω into $\mathbb{R}^k$ with values $(X(t_1,\omega), X(t_2,\omega), \ldots, X(t_k,\omega))$ at ω. If each $X(t)$ is a random variable, X is said to be a random function. Suppose now that $(X_n)_{n\in\mathbb{N}}$ is a sequence of random functions. According to Theorem B.90, $(X_n)_{n\in\mathbb{N}}$ is tight if and only if the sequence $(X_n(0))_{n\in\mathbb{N}}$ is tight, and for any positive real numbers ϵ and η there exists δ, $(0 < \delta < 1)$ 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 $X_n,\ n \in \mathbb{N}$, do not oscillate too much. Theorem B.91 can be restated in the same way: $(X_n)_{n\in\mathbb{N}}$ is tight if $(X_n(0))_{n\in\mathbb{N}}$ is tight, and if for any positive ϵ and η there exists a δ, $0 < \delta < 1$, and an integer n_0 such that

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

for $n \geq n_0$ and $0 \leq t \leq 1$. Let $(\xi_n)_{n\in\mathbb{N}\setminus\{0\}}$ be a sequence of i.i.d. random variables on $(\Omega, \mathcal{F}, P)$ with mean 0 and variance σ^2. We define the sequence of partial sums $S_n = \xi_1 + \cdots + \xi_n,\ n \in \mathbb{N}$, with $S_0 = 0$. Let us construct the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ from the sequence $(S_n)_{n\in\mathbb{N}}$ by means of rescaling and linear interpolation, as follows:

$$\begin{aligned} X_n\left(\frac{i}{n}, \omega\right) &= \frac{1}{\sigma\sqrt{n}} S_i(\omega) \qquad &&\text{for} \qquad \frac{i}{n} \in [0, 1[; \\ \frac{X_n(t) - X_n\left(\frac{i-1}{n}\right)}{X_n\left(\frac{i}{n}\right) - X_n\left(\frac{i-1}{n}\right)} - \frac{t - \frac{i-1}{n}}{\frac{1}{n}} &= 0 \qquad &&\text{for} \qquad t \in \left[\frac{(i-1)}{n}, \frac{i}{n}\right]. \end{aligned} \tag{B.1}$$

With a little algebra, we obtain

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

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

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega). \tag{B.2}$$

For any fixed ω, $X_n(\cdot, \omega)$ is a piecewise linear function whose pieces' amplitude decreases as n increases. Since the ξ_i and hence the S_i are random variables it follows by (B.2) 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 $(X_n)_{n\in\mathbb{N}}$ to be a tight sequence.

Theorem B.92. *Suppose X_n, $n \in \mathbb{N}$ is defined by (B.2). The sequence $(X_n)_{n\in\mathbb{N}}$ is tight if for each positive ϵ there exists a λ, with $\lambda > 1$, and an integer n_0 such that, if $n \geq n_0$, then*

$$P\left(\max_{i\leq n}|S_{k+i} - S_k| \geq \lambda\sigma\sqrt{n}\right) \leq \frac{\epsilon}{\lambda^2} \tag{B.3}$$

holds for all k.

If the sequence $(\xi_n)_{n\in\mathbb{N}\setminus\{0\}}$ is made of i.i.d. random variables, then condition (B.3) reduces to

$$P\left(\max_{i\leq n}|S_i| \geq \lambda\sigma\sqrt{n}\right) \leq \frac{\epsilon}{\lambda^2}. \tag{B.4}$$

Let us denote by P_W the probability measure of the Wiener process as defined in Definition 2.157 and whose existence is a consequence of Theorem 2.55. We will refer here to its restriction to $t \in [0,1]$, so that its trajectories are almost sure elements of $C([0,1])$.

Theorem B.93 (Donsker). *Let $(\xi_n)_{n\in\mathbb{N}\setminus\{0\}}$ be a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, P)$ with mean 0 and finite, positive variance σ^2:*

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

Let $S_n = \xi_1 + \xi_2 + \cdots + \xi_n$, $n \in \mathbb{N}$. 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 wish to apply Theorem B.89; we first show that the sequence of the finite-dimensional distributions of X_n, $n \in \mathbb{N}$ 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,

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

we obtain

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

Since $\lim_{n\to\infty} \frac{[ns]}{ns} = 1$, by the Central Limit Theorem for i.i.d. variables

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

so that

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

Therefore, by Theorem B.72, $X_n(s) \xrightarrow{\mathcal{D}} W_s$. 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.72, 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.74 of Theorem B.73 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 ξ_i, $i = 1, 2, \ldots, n$, the random variables $S_{[ns]}$ and $S_{[nt]} - S_{[ns]}$ are independent, so that

$$\lim_{n\to\infty} E\left[e^{\frac{iu}{\sigma\sqrt{n}} \sum_{j=1}^{[ns]} \xi_j + \frac{iv}{\sigma\sqrt{n}} \sum_{j=[ns]+1}^{[nt]} \xi_j} \right] = \lim_{n\to\infty} E\left[e^{\frac{iu}{\sigma\sqrt{n}} \sum_{j=1}^{[ns]} \xi_j} \right] \cdot \lim_{n\to\infty} E\left[e^{\frac{iv}{\sigma\sqrt{n}} \sum_{j=[ns]+1}^{[nt]} \xi_j} \right]. \tag{B.5}$$

Since $\lim_{n\to\infty} \frac{[ns]}{ns} = 1$, by the Lindeberg Theorem 1.190

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

and for the same reason

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

so that

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

and

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

Substitution of these two last equations into (B.5) 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. To prove tightness we apply Theorem B.92; under the assumptions of the present theorem, it can be shown (Billingsley 1968, p. 69) that

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

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 inequality, and $N \sim N(0,1)$. Therefore, if ϵ is positive, there exists a λ 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.92, the sequence of the distribution functions of $(X_n)_{n\in\mathbb{N}}$ is tight. □

An Application of Donsker's Theorem

Donsker's theorem has the following qualitative interpretation: $X_n \overset{\mathcal{D}}{\to} W$ implies that, if τ 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. Using Donsker's theorem it is possible to use the relation $X_n \overset{\mathcal{D}}{\to} W$ 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, $X_n \overset{\mathcal{D}}{\to} W$ implies, by Corollary B.74, that

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

The obvious relation

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

implies

$$\frac{1}{\sigma\sqrt{n}} \max_{i\le n} S_i \overset{\mathcal{D}}{\to} \sup_{0\le t\le 1} W_t. \tag{B.6}$$

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 $h(X_n) \overset{\mathcal{D}}{\to} h(W)$, 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 ξ_n are independent and satisfy

$$P(\xi_n = 1) = P(\xi_n = -1) = \frac{1}{2}. \tag{B.7}$$

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

$$P\left(\max_{0\le i\le n} S_i \ge a\right) = 2P(S_n > a) + P(S_n = a). \tag{B.8}$$

If $a = 0$, then 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

$$P(M_n \ge a, S_n < a) = P(M_n \ge a, S_n > a). \tag{B.9}$$

Because of (B.7), all 2^n possible paths $(S_1, S_2, \ldots, S_n)$ have the same probability 2^{-n}. Therefore, (B.9) 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.9), 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.9) follows. This argument is an example of the reflection principle. See also Lemma 2.170.

Let α be an arbitrary nonnegative number, and let $a_n = -[-\alpha n^{\frac{1}{2}}]$. By (B.9), we have

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

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

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

By the central limit theorem,

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

where $N \sim N(0,1)$ and $\sigma^2 = 1$ by (B.7).

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

$$P\left(\max_{i \le n} \frac{1}{\sqrt{n}} S_i \ge \alpha\right) \to 2P\left(N \ge \alpha\right), \qquad \alpha \ge 0. \tag{B.10}$$

By (B.10), (B.6), and (B.7), we conclude that

$$P\left(\sup_{0 \le t \le 1} W_t \le \alpha\right) = \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-\frac{1}{2}u^2} du, \qquad \alpha \ge 0. \tag{B.11}$$

If we drop assumption (B.7) and suppose that the random variables ξ_n are i.i.d. and satisfy the hypothesis of Donsker's theorem, then (B.6) holds and from (B.11) we obtain

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

Thus we have derived the limiting distribution of $\max_{i \le n} S_i$ by Lindeberg's theorem. Therefore, if the ξ_n are i.i.d. 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 ξ_n. For this reason, Donsker's theorem is often called an invariance principle.

C

Elliptic and Parabolic Equations

We recall here basic facts about the existence and uniqueness of elliptic and parabolic equations; for further details, the interested reader may refer to Friedman (1963, 1964).

C.1 Elliptic Equations

Consider an open bounded $\Omega \subset \mathbb{R}^n$, (for $n \geq 1$). We are given $a_{ij},\ b_i,$ and $c,\ i,j = 1,\ldots,n$, real-valued functions defined on Ω. Consider the partial differential operator

$$M \equiv \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial}{\partial x_i} + c(x). \tag{C.1}$$

The operator M is said to be *elliptic at a point* $x_0 \in \Omega$ if the matrix $(a_{i,j}(x_0))_{i,j=1,\ldots,n}$ is positive-definite, i.e., for any real vector $\xi \neq 0$, $\sum_{i,j=1}^{n} a_{ij}(x_0)\xi_i\xi_j > 0$.

If there is a positive constant μ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \mu \mid \xi \mid^2$$

for all $x \in \Omega$, and all $\xi \in \mathbb{R}^n$, then M is said to be *uniformly elliptic* in Ω.

Definition C.1. A *barrier* for M at a point $y \in \partial\Omega$ is a continuous nonnegative function w_y defined on $\overline{\Omega}$ that vanishes only at the point y and such that $M[w_y](x) \leq -1$, for any $x \in \Omega$.

Proposition C.2. *Let $y \in \partial\Omega$. If there exists a closed ball K such that $K \cap \Omega = \emptyset$, and $K \cap \bar{\Omega} = \{y\}$, then y has a barrier for M.*

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

The First Boundary Value or Dirichlet Problem

Given a real-valued function f defined on Ω and a real-valued function ϕ defined on $\partial\Omega$, the Dirichlet problem consists of finding a solution u of the system

$$\begin{cases} M[u](x) = f(x) & \text{in} \quad \Omega, \\ u(x) \quad\quad = \phi(x) & \text{in} \quad \partial\Omega. \end{cases} \tag{C.2}$$

Theorem C.3. *Assume that M is uniformly elliptic in Ω, that $c(x) \leq 0$, and that a_{ij}, b_i, $(i, j = 1, \ldots, n)$, c, f are uniformly Hölder continuous with exponent α in $\bar{\Omega}$. If every point of $\partial\Omega$ has a barrier, and ϕ is continuous on $\partial\Omega$, then there exists a unique $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$ solution of the Dirichlet problem (C.2).*

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

C.2 The Cauchy Problem and Fundamental Solutions for Parabolic Equations

Let

$$L_0 \equiv \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x,t) \frac{\partial}{\partial x_i} + c(x,t) \tag{C.3}$$

be an elliptic operator in $\mathbb{R}^n$, for all $t \in [0, T]$, and let $f : \mathbb{R}^n \times [0, T] \to \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$ be two appropriately assigned functions.

The Cauchy problem consists in finding a solution $u(x, t)$ of

$$\begin{cases} L[u] \equiv L_0[u] - u_t = f(x,t) & \text{in } \mathbb{R}^n \times]0, T], \\ u(x, 0) = \phi(x) & \text{in } \mathbb{R}^n. \end{cases} \tag{C.4}$$

The solution is understood to be a continuous function defined for $(x, t) \in \mathbb{R}^n \times [0, T]$, with its derivatives $u_{x_i}, u_{x_i x_j}, u_t$ continuous in $\mathbb{R}^n \times]0, T]$.

Theorem C.4. *Let the matrix $(a_{ij}(x,t))_{i,j=1,\ldots,n}$ be a nonnegative definite real matrix, and let*

$$| a_{ij}(x,t) | \leq C, \quad |b_i(x,t)| \leq C(|x| + 1), \quad c(x,t) \leq C(|x|^2 + 1), \tag{C.5}$$

for a suitable constant C. If $L[u] \leq 0$ in $\mathbb{R}^n \times]0, T]$, and if $u(x,t) \geq -B \exp\left\{\beta |x|^2\right\}$ in $\mathbb{R}^n \times [0, T]$ (for some B, β positive constants), and if $u(x, 0) \geq 0$ in $\mathbb{R}^n$, then $u(x, t) \geq 0$ in $\mathbb{R}^n \times [0, T]$.

Proof. See, e.g., Friedman (2004, p. 139). □

Corollary C.5. *Let the matrix $(a_{ij}(x,t))_{i,j=1,\ldots,n}$ be a nonnegative definite real matrix, and let (C.5) hold. Then there exists at most one solution of the Cauchy problem (C.4) satisfying*

$$|u(x,t)| \leq -B \exp\left\{\beta |x|^2\right\}$$

in $\mathbb{R}^n \times [0,T]$ *(for some* B, β *positive constants).*

The next theorem, and consequent corollary, considers different growth conditions on the coefficients of the operator L_0.

Theorem C.6. *Let the matrix* $(a_{ij}(x,t))_{i,j=1,\dots,n}$ *be a nonnegative definite real matrix, and let*

$$|a_{ij}(x,t)| \leq C(|x|^2+1), \quad |b_i(x,t)| \leq C(|x|+1), \quad c(x,t) \leq C, \tag{C.6}$$

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

Proof. See, e.g., Friedman (2004, p. 140). □

Corollary C.7. *Let the matrix* $(a_{ij}(x,t))_{i,j=1,\dots,n}$ *be a nonnegative definite real matrix, and let conditions (C.6) be satisfied; then there exists at most one solution* u *of the Cauchy problem with*

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

where N, q *are positive constants.*

Later the following conditions will be required.

(A_1) There exists a $\mu > 0$ such that $\sum_{i,j=1}^n a_{ij}(x,t)\xi_i\xi_j \geq \mu\xi^2$ for all $(x,t) \in \mathbb{R}^n \times [0,T]$.

(A_2) The coefficients of L_0 are bounded continuous functions in $\mathbb{R}^n \times [0,T]$, and the coefficients $a_{ij}(x,t)$ are continuous in t, uniformly with respect to $(x,t) \in \mathbb{R}^n \times [0,T]$.

(A_3) The coefficients of L_0 are Hölder continuous functions (with exponent α) in x, uniformly with respect to the variables (x,t) in compacts of $\mathbb{R}^n \times [0,T]$, and the coefficients $a_{ij}(x,t)$ are Hölder continuous (with exponent α) in x, uniformly with respect to $(x,t) \in \mathbb{R}^n \times [0,T]$.

Definition C.8. A *fundamental solution of the parabolic operator* $L_0 - \frac{\partial}{\partial t}$ in $\mathbb{R}^n \times [0,T]$ is a function $\Gamma(x,t;\xi,r)$, defined, for all $(x,t) \in \mathbb{R}^n \times [0,T]$ and all $(\xi,t) \in \mathbb{R}^n \times [0,T]$, $t > r$, such that, for all ϕ with compact support,[14] the function

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

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

satisfies

(i) $L[u](x,t) - u_t(x,t) = 0$ for $x \in \mathbb{R}^n$, $r < t \leq T$
(ii) $u(x,t) \to \phi(x)$ as $t \downarrow r$, for $x \in \mathbb{R}^n$

Theorem C.9. *If conditions* (A_1), (A_2), *and* (A_3) *hold, then there exists a fundamental solution* $\Gamma(x,t;\xi,r)$, *for* $L_0 - \frac{\partial}{\partial t}$, *satisfying the inequalities*

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

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* $(x,t;\xi,r) \in \mathbb{R}^n \times [0,T] \times \mathbb{R}^n \times [0,T]$, $t > r$, *and* $L_0[\Gamma] - \Gamma_t = 0$, *as a function of* (x,t).

Finally, for any bounded continuous function ϕ *we have*

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

Proof. See, e.g., Friedman (2004, p. 141). □

Theorem C.10. *Let* (A_1), (A_2), (A_3) *hold, let* $f(x,t)$ *be a continuous function in* $\mathbb{R}^n \times [0,T]$, *Hölder continuous in* x, *uniformly with respect to* (x,t) *in compacts of* $\mathbb{R}^n \times [0,T]$, *and let* ϕ *be a continuous function in* $\mathbb{R}^n$. *Moreover, assume that*

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

where A, a_1 *are positive constants. Then there exists a solution of the Cauchy problem (C.4) in* $0 \leq t \leq T^*$, *where* $T^* = \min\left\{T, \frac{\bar{c}}{a_1}\right\}$ *and* $\bar{c}$ *is a constant, which depends only on the coefficients of* L_0, *and*

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

with positive constants A', a_1'.

The solution is given by

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

The adjoint operator L^* of $L = L_0 - \frac{\partial}{\partial t}$ is given by

$$L^*[v] = L_0^*[v] + \frac{\partial v}{\partial t},$$
$$L_0^*[v](x,t) = \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(x,t)v(x,t)) - \sum_{i=1}^n \frac{\partial}{\partial x_i}(b_i(x,t)v(x,t)) + c(x,t),$$

by assuming that all quoted derivatives of the coefficients exist and are bounded functions.

Definition C.11. A *fundamental solution of the operator* $L_0^* + \frac{\partial}{\partial t}$ in $\mathbb{R}^n \times [0, T]$ is a function $\Gamma^*(x, t; \xi, r)$, defined, for all $(x, t) \in \mathbb{R}^n \times [0, T]$ and all $(\xi, r) \in \mathbb{R}^n \times [0, T]$, $t < r$, such that, for all g continuous with compact support, the function

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

satisfies

1. $L^*[v] + v_t = 0 \quad$ for $\quad x \in \mathbb{R}^n, \ 0 \le t \le r$
2. $v(x, t) \to g(x) \quad$ as $\quad t \uparrow r, \ $ for $\ x \in \mathbb{R}^n$

We consider the following additional condition.

(A_4) The functions a_{ij}, $\frac{\partial a_{ij}}{\partial x_i}$, $\frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}$, b_i, $\frac{\partial b_i}{\partial x_i}$, c are bounded and the coefficients of L_0^* satisfy conditions (A_2) and (A_3).

Theorem C.12. *If* (A_1)–(A_4) *are satisfied, then there exists a fundamental solution* $\Gamma^*(x, t; \xi, r)$ *of* $L_0^* + \frac{\partial}{\partial t}$*; it is such that*

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

Proof. See, e.g., Friedman (2004, p. 143). □

D

Semigroups of Linear Operators

In this appendix we will report the main results concerning the structure of contraction semigroups of linear operators on Banach spaces, as they are closely related to evolution semigroups of Markov processes. For the present treatment we refer to the now classic books by Lamperti (1977), Pazy (1983), and Belleni-Morante and McBride (1998).

Throughout this appendix, E will denote a Banach space.

Definition D.1. A one-parameter family $(T_t)_{t\in\mathbb{R}_+}$ of linear operators on E is a strongly continuous semigroup of bounded linear operators or, simply, a C_0 *semigroup* if

(i) $T_0 = I$ (the identity operator)
(ii) $T_{s+t} = T_s T_t$, for all $s, t \in \mathbb{R}_+$
(iii) $\lim_{t\to 0+} \parallel T_t x - x \parallel = 0$, for all $x \in E$

Theorem D.2. *Let $(T_t)_{t\in\mathbb{R}_+}$ be a C_0 semigroup. There exist constants $\omega \geq 0$ and $M \geq 1$ such that*

$$\parallel T_t \parallel \leq M e^{\omega t}, \qquad \text{for} \quad t \in \mathbb{R}_+. \tag{D.1}$$

Corollary D.3. *If $(T_t)_{t\in\mathbb{R}_+}$ is a C_0 semigroup, then, for any $x \in E$, the map $t \in \mathbb{R}_+ \mapsto T_t x \in E$ is a continuous function.*

Definition D.4. Let $(T_t)_{t\in\mathbb{R}_+}$ be a semigroup of bounded linear operators. The linear operator $\mathcal{A}$ defined by

$$\mathcal{D}_{\mathcal{A}} = \left\{ x \in E \,|\, \lim_{t\to 0+} \frac{T_t x - x}{t} \quad \text{exists} \right\}$$

$$\mathcal{A}x = \lim_{t\to 0+} \frac{T_t x - x}{t}, \qquad \text{for} \quad x \in \mathcal{D}_{\mathcal{A}},$$

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

where the limit is taken in the topology of the norm of E.

Theorem D.5. *Let $(T_t)_{t\in\mathbb{R}_+}$ be a C_0 semigroup, and let $\mathcal{A}$ be its infinitesimal generator. Then*

(a) For $x \in E$,

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} T_s x\, ds = T_t x.$$

(b) For $x \in E$, $\int_0^t T_s x ds \in \mathcal{D}_{\mathcal{A}}$, and

$$\mathcal{A}\Big(\int_0^t T_s x\, ds\Big) = T_t x - x.$$

(c) For $x \in \mathcal{D}_{\mathcal{A}}$, $T_t x \in \mathcal{D}_{\mathcal{A}}$, and

$$\frac{d}{dt}T_t x = \mathcal{A}T_t x = T_t\mathcal{A}x$$

(the derivative is taken in the topology of the norm of E).

(d) For $x \in \mathcal{D}_{\mathcal{A}}$,

$$T_t x - T_s x = \int_s^t T_\tau \mathcal{A}x\, d\tau = \int_s^t \mathcal{A}T_\tau x\, d\tau.$$

Corollary D.6. *If $\mathcal{A}$ is the infinitesimal generator of a C_0 semigroup, then its domain $\mathcal{D}_{\mathcal{A}}$ is dense in E.*

Corollary D.7. *Let $(T_t)_{t\in\mathbb{R}_+}$, $(S_t)_{t\in\mathbb{R}_+}$ be C_0 semigroups with infinitesimal generators $\mathcal{A}$, and $\mathcal{B}$, respectively. If $\mathcal{A} = \mathcal{B}$, then $T_t = S_t$, for $t \in \mathbb{R}_+$.*

Definition D.8. Let $(T_t)_{t\in\mathbb{R}_+}$ be a C_0 semigroup. If in (D.1) $\omega = 0$, we say that $(T_t)_{t\in\mathbb{R}_+}$ is uniformly bounded; if, moreover, $M = 1$, we say that $(T_t)_{t\in\mathbb{R}_+}$ is a C_0 *semigroup of contractions.*

The resolvent set $\rho(A)$ of a linear operator A on E (bounded or not) is the set of all complex numbers λ for which the operator $\lambda I - A$ is invertible, and its inverse is a bounded operator on E. The family

$$\big\{R(\lambda : A) = (\lambda I - A)^{-1}, \quad \lambda \in \rho(A)\big\}$$

is called the *resolvent* of A.

Definition D.9. A linear operator $A : \mathcal{D}_A \subset E \to E$ is *closed* if and only if for any sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{D}_A$ such that $u_n \to u$ and $Au_n \to v$ in E we have that $u \in \mathcal{D}_A$ and $v = Au$.

Theorem D.10. *Let $(T_t)_{t\in\mathbb{R}_+}$ be a C_0 semigroup of contractions. $\mathcal{A}$ is its infinitesimal generator if and only if*

(i) $\mathcal{A}$ *is a closed linear operator, and* $\overline{\mathcal{D}_\mathcal{A}} = E$.
(ii) *The resolvent set* $\rho(\mathcal{A})$ *of* $\mathcal{A}$ *contains* $\mathbb{R}_+^*$, *and for any* $\lambda > 0$,

$$\|R(\lambda : \mathcal{A})\| \leq \frac{1}{\lambda}. \tag{D.2}$$

Further, for any $\lambda > 0$ *and any* $x \in E$,

$$R(\lambda : \mathcal{A})x = \int_0^{+\infty} e^{-\lambda t}\, T_t\, x\, dt.$$

For any $\lambda > 0$, *and any* $x \in E$, $R(\lambda : \mathcal{A})x \in \mathcal{D}_\mathcal{A}$.

Proof. See, e.g., Pazy (1983, p. 8). □

Note that, since the map $t \to T_t\, x$ is continuous and uniformly bounded, the integral exists as an improper Riemann integral and defines indeed a bounded linear operator satisfying (D.2).

Theorem D.11. *Let* $(T_t)_{t\in\mathbb{R}_+}$ *be a* C_0 *semigroup of contractions, and let* $\mathcal{A}$ *be its infinitesimal generator. Then, for any* $t \in \mathbb{R}_+$ *and any* $x \in E$,

$$T_t\, x = \lim_{n\to\infty} \left(I - \frac{t}{n}\mathcal{A}\right)^{-n} x = \lim_{n\to\infty} \left[\frac{n}{t} R(\frac{n}{t} : \mathcal{A})\right]^n x.$$

The foregoing theorem induces the notation $T_t = e^{t\mathcal{A}}$.

Finally, based on all the foregoing treatment we may further notice that if $x \in \mathcal{D}_\mathcal{A}$, then we know that $T_t\, x \in \mathcal{D}_\mathcal{A}$, for any $t \in \mathbb{R}_+$, and it is the unique solution of the initial value problem

$$\frac{d}{dt} u(t) = \mathcal{A}\, u(t), \quad t > 0,$$

subject to the initial condition

$$u(0) = x.$$

E

Stability of Ordinary Differential Equations

We consider the system of ordinary differential equations

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

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

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

If condition (E.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 (E.2), it can also be verified that

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

Remark E.2. There may be attraction without stability.

Remark E.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 toward $\mathbf{0}$.

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

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

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

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

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

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 (E.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 E.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 (E.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 E.6 (Lyapunov).

1. *If there exists $v(t, \mathbf{x})$ a Lyapunov function for system (E.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 E.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

Aalen O., *Nonparametric inference for a family of counting processes*, Annals of Statistics, 701–726, 6; 1978.

Aletti G. and Capasso V., *Profitability in a multiple strategy market*, Decis. Econ. Finance, 145–152, 26; 2003.

Andersen P. K., Borgan Ø., Gill R.D., and Keiding N., *Statistical Models Based on Counting Processes*, Springer, Heidelberg; 1993.

Anderson W. J., *Continuous-Time Markov Chains: An Application-Oriented Approach*, Springer, New York; 1991.

Applebaum D., *Levy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge; 2004.

Arnold L., *Stochastic Differential Equations: Theory and Applications*, Wiley, New York; 1974.

Ash R. B., *Real Analysis and Probability*, Academic, London; 1972.

Ash R. B. and Gardner M. F., *Topics in Stochastic Processes*, Academic, London; 1975.

Aubin J.-P., *Applied Abstract Analysis*, Wiley, New York; 1977.

Bachelier L., *Théorie de la spéculation*, Ann. Sci. École Norm. Sup., 21–86, 17; 1900.

Bailey N. T. J., *The Mathematical Theory of Infectious Diseases*, Griffin, London; 1975.

Baldi P., *Equazioni differenziali stocastiche*, UMI, Bologna; 1984.

Bartholomew D. J., *Continuous time diffusion models with random duration of interest*, J. Math. Sociol., 187–199, 4; 1976.

Bauer H., *Probability Theory and Elements of Measure Theory*, Academic, London; 1981.

Becker N., *Analysis of Infectious Disease Data*, Chapman & Hall, London; 1989.

Belleni-Morante A. and McBride A. C., *Applied Nonlinear Semigroups*, Wiley, Chichester; 1998.

Bertoin J., *Lévy Processes*, Cambridge University Press, Cambridge, UK; 1996.

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

Bhattacharya R. N. and Waymire E. C., *Stochastic Processes with Applications*, Wiley, New York; 1990.

Bianchi A., Capasso V., and Morale D., *Estimation and prediction of a nonlinear model for price herding* in *Complex Models and Intensive Computational Methods for Estimation and Prediction* (C. Provasi, Ed.), CLUEP, Padova, 2005, pp. 365–370.

Billingsley P., *Convergence of Probability Measures*, Wiley, New York; 1968.

Billingsley P., *Probability and Measure*, Wiley, New York; 1986.

Black F. and Scholes M., *The pricing of options and corporate liabilities*, J. Pol. Econ., 637–654, 81; 1973.

Bohr H., *Almost Periodic Functions*, Chelsea, New York; 1947.

Borodin A. and Salminen P., *Handbook of Brownian Motion: Facts and Formulae*, Birkhä user, Boston; 1996.

Boyle P. and Tian Y., *Pricing lookback and barrier options under the CEV process*, J. Fin. Quant. Anal., 241–264, 34; 1999.

Brace A., Gatarek D., & Musiela M., *The market model of interest rate dynamics*, Math. Fin., 127–154, 7; 1997.

Breiman L., *Probability*, Addison-Wesley, Reading, MA; 1968.

Bremaud P., *Point Processes and Queues: Martingale Dynamics*, Springer, Heidelberg; 1981.

Burger M., Capasso V. and Morale D., *On an aggregation model with long and short range interaction*, Nonlinear Anal. Real World Appl., 939–958, 3; 2007.

Cai G. Q. and Lin Y. K., *Stochastic analysis of the Lotka-Volterra model for ecosystems*, Phys. Rev. E, 041910, 70; 2004.

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.

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, Heidelberg; 1990.

Capasso V., *Mathematical Structures of Epidemic Systems*, Springer, Heidelberg; 1993. 2nd corrected printing; 2008.

Capasso V., Di Liddo A., and Maddalena L., *Asymptotic behaviour of a nonlinear model for the geographical diffusion of innovations*, Dynamic Systems and Applications, 207–220, 3; 1994.

Capasso V. and Morale D., *Asymptotic behavior of a system of stochastic particles subject to nonlocal interactions.* Stoch. Anal. Appl. 574–603, 27; 2009.

Capasso V., Morale D., and Sioli F., *An agent-based model for "price herding", applied to the automobile market*, MIRIAM reports, Milan; 2003.

Carrillo J., *Entropy solutions for nonlinear degenerate problems*, Arch. Rat. Mech. Anal., 269–361, 147; 1999.

Carrillo J. A., McCann R. J. and Villani C., *Kinetic equilibration rates for granular nedia and related equations: entropy dissipation and mass transportation estimates*, Rev. Mat. Iberoam., 971–1018, 19; 2003.

Champagnat N., Ferriére R., and Méléard S. *Unifying evolutionary dynamics: From individula stochastic processes to macroscopic models*, Theor. Pop. Biol., 297–321, 69; 2006.

Chan K. C., Karolyi G.A., Longstaff F.A. and Sanders A. B., *An empirical comparison of alternative models of the short-term interest rate*, J. Fin., 1209–1227, 47; 1992.

Chiang C.L., *Introduction to Stochastic Processes in Biostatistics*, Wiley, New York; 1968.

Chow Y. S. and Teicher H., *Probability Theory: Independence, Interchangeability, Martingales*, Springer, New York; 1988.

Chung K. L., *A Course in Probability Theory*, Second Edition, Academic, New York; 1974.

Cox J. C., *The constant elasticity of variance option pricing model*, J. Portfolio Manage., 15–17, 22; 1996.

Cox J. C., Ross S. A. and Rubinstein M., *Option pricing: A simplified approach*, J. Fin. Econ., 229–263, 7; 1979.

Çynlar E., *Introduction to Stochastic Processes*, Prentice Hall, Englewood Cliffs, NJ; 1975.

Dalang R. C., Morton A. and Willinger W., *Equivalent martingale measures and non-arbitrage in stochastic securities market models*, Stochast. Stochast. Rep., 185–201, 29; 1990.

Daley D. and Vere-Jones D., *An Introduction to the Theory of Point Processes*, Springer, Berlin; 1988.

Daley D. and Vere-Jones D., *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*, Springer, Heidelberg; 2008.

Darling D. A. D. and Siegert A. J. F., *The first passage time problem for a continuum Markov process*, Ann. Math. Stat., 624–639, 24; 1953.

Davis M. H. A., *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, J. R. Stat. Soc. Ser. B, 353–388, 46; 1984.

Delbaen F. and Schachermeyer W., *A general version of the fundamental theorem of asset pricing*, Math. Annal., 463–520, 300; 1994.

De Masi A. and Presutti E., *Mathematical Methods for Hydrodynamical Limits*, Springer, Heidelberg; 1991.

Dieudonné J., *Foundations of Modern Analysis*, Academic, New York; 1960.

Di Nunno G., Øksendal B. and Proske F., *Malliavin Calculus for Levy Processes with Applications to Finance*, Springer, Berlin Heidelberg; 2009.

Donsker M. D. and Varadhan S. R. S., *Large deviations from a hydrodynamical scaling limit*, Comm. Pure Appl. Math., 243–270, 42; 1989.

Doob J. L., *Stochastic Processes*, Wiley, New York; 1953.

Dudley R. M., *Real Analysis and Probability*, Cambridge University Press, Cambridge; 2005.

Duffie D., *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, NJ; 1996.

Dupire B., *Pricing with a smile*, RISK, 18–20, 1; 1994.

Durrett R. and Levin S. A., *The importance of being discrete (and spatial)*, Theor. Pop. Biol., 363–394, 46; 1994.

Dynkin E. B., *Markov Processes*, Springer, Berlin, Vols. 1–2; 1965.

Einstein A., *Über die von der molekularkinetischen Theorie der Wärme geforderten Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Annal. Phys., 549–560, 17; 1905.

Embrechts P., Klüppelberg C. and Mikosch T., *Modelling Extreme Events for Insurance and Finance*, Springer, Berlin; 1997.

Epstein J. and Axtell R., *Growing Artificial Societies–Social Sciences from the Bottom Up*, Brookings Institution Press and MIT Press, Cambridge, MA; 1996.

Ethier S. N. and Kurtz T. G., *Markov Processes, Characterization and Convergence*, Wiley, New York; 1986.

Feller W., *An Introduction to Probability Theory and Its Applications*, Wiley, New York; 1971.

Flierl G., Grünbaum D., Levin S. A. and Olson D., *From individuals to aggregations: The interplay between behavior and physics*, J. Theor. Biol., 397–454, 196; 1999.

Fournier N. and Méléard S., *A microscopic probabilistic description of a locally regulated population and macroscopic approximations*, Ann. Appl. Prob., 1880–1919, 14; 2004.

Franke J., Härdle W. K. and Hafner C. M., *Statistics of Financial Markets: An Introduction*, Second Edition, Springer, Heidelberg; 2011.

Friedman A., *Partial Differential Equations*, Krieger, New York; 1963.

Friedman A., *Partial Differential Equations of Parabolic Type*, Prentice-Hall, London; 1964.

Friedman A., *Stochastic Differential Equations and Applications*, Academic, London; 1975. Two volumes bounded as one, Dover, Mineola, NY; 2004.

Fristedt B., and Gray L., *A Modern Approach to Probability Theory*, Birkhäuser, Boston; 1997.

Gard T. C., *Introduction to Stochastic Differential Equations*, Marcel Dekker, New York; 1988.

Gihman I. I. and Skorohod A. V., *Stochastic Differential Equations*, Springer, Berlin; 1972.

Gihman I. I. and Skorohod A. V., *The Theory of Random Processes*, Springer, Berlin; 1974.

Gnedenko B. V., *The Theory of Probability*, Chelsea, New York; 1963.

Grigoriu M., *Stochastic Calculus: Applications to Science and Engineering*, Birkhäuser, Boston; 2002.

Grünbaum D. and 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, New York; 1994.

Gueron S., Levin S. A. and Rubenstein D. I., *The dynamics of herds: From individuals to aggregations*, J. Theor. Biol., 85–98, 182; 1996.

Hagan P. S., Kumar D., Lesniewski A. and Woodward D. E., *Managing smile risk*, Wilmott Mag., 84–102, 1; 2002

Harrison J. M. and Kreps D. M., *Martingales and arbitrage in multiperiod securities markets*, J. Econ. Theory, 381–408, 20; 1979.

Harrison J. M. and Pliska S. R., *Martingales and stochastic integrals in the theory of continuous trading*, Stochast. Process. Appl., 215–260, 11; 1981.

Has'minskii R. Z., *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff, The Netherlands; 1980.

Heath D., Jarrow R. and Morton A., *Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation*, Econometrica, 77–105, 1; 1992.

Heston S. L., *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, Rev. Fin. Stud., 327–343, 6; 1993.

Hunt P. J. and Kennedy J. E., *Financial Derivatives in Theory and Practice*, Wiley, New York; 2000.

Hull J. and White A., *Pricing interest rate derivative securities*, Rev. Fin. Stud., 573–592, 4; 1990.

Ikeda N. and Watanabe S., *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Kodansha; 1989.

Itô K. and McKean H. P., *Diffusion processes and their sample paths*, Springer, Berlin; 1965.

Jacobsen M., *Statistical Analysis of Counting Processes*, Springer, Heidelberg; 1982.

Jacod J. and Protter P., *Probability Essentials*, Springer, Heidelberg; 2000.

Jacod J. and Shiryaev A. N., *Limit Theorems for Stochastic Processes*, Springer Lecture Notes in Mathematics, Springer, Berlin; 1987.

Jamshidian F., *LIBOR and swap market models and measures*, Fin. Stochast., 43–67, 1; 1997.

Jeanblanc M., Yor M. and Chesney M., *Mathematical Methods for Financial Markets*, Springer, London; 2009.

Jelinski Z. and Moranda P., *Software reliability research*, in *Statistical Computer Performance Evaluation*, pp. 466–484, Academic, New York; 1972.

Kallenberg O., *Foundations of Modern Probability*, Springer, Berlin; 1997.

Karatzas I. and Shreve S. E., *Brownian Motion and Stochastic Calculus*, Springer, New York; 1991.

Karlin S. and Taylor H. M., *A First Course in Stochastic Processes*, Academic, New York; 1975.

Karlin S. and Taylor H. M., *A Second Course in Stochastic Processes*, Academic, New York; 1981.

Karr A. F., *Point Processes and Their Statistical Inference*, Marcel Dekker, New York; 1986.

Karr A. F., *Point Processes and Their Statistical Inference*, Second Edition Revised and expanded. Marcel Dekker, New York; 1991.

Klenke A., *Probability Theory*, Springer, Heidelberg; 2008.

Kloeden P. E. and Platen E., *Numerical Solution of Stochastic Differential Equations*, Springer, Heidelberg; 1999.

Kolmogorov A. N., *Foundations of the Theory of Probability*, Chelsea, New York; 1956.

Kolmogorov A. N. and Fomin E. S. V., *Elements of Theory of Functions and Functional Analysis*, Groylock, Moscow; 1961.

Kou S., *A jump-diffusion model for option pricing*, Manage. Sci., 1086–1101, 48; 2002.

Ladyzenskaja O. A., Solonnikov V. A. and Ural'ceva N.N., *Linear and Quasi-linear Equations of Parabolic Type*, AMS, Providence, RI, 1968.

Lamperti J., *Stochastic Processes: A Survey of the Mathematical Theory*, Springer, New York; 1977.

Lapeyre B., Pardoux E. and Sentis R. *Introduction to Monte-Carlo methods for transport and diffusion equations*, Oxford University Press, Oxford; 2003.

Last G. and Brandt A., *Marked Point Processes on the Real Line: The Dynamic Approach*, Springer, Heidelberg; 1995.

Lewis A. L., *Option Valuation Under Stochastic Volatility*, Finance Press, Newport Beach; 2000.

Lipster R. and Shiryaev A. N., *Statistics of Random Processes, I: General Theory*, Springer, Heidelberg; 1977.

Lipster R. and Shiryaev A. N., *Statistics of Random Processes, II: Applications*, Springer, Heidelberg; Second Edition, 2010.

Loève M., *Probability Theory*, Van Nostrand-Reinhold, Princeton, NJ; 1963.

Lu L., *Optimal control of input rates of Stein's models*, Math. Med. Biol., 31–46, 28; 2011.

Ludwig D., *Stochastic Population Theories*, Lecture Notes in Biomathematics, Vol. 3, Springer, Heidelberg; 1974.

Lukacs E., *Characteristic Functions*, Griffin, London; 1970.

Mahajan V. and Wind Y., *Innovation Diffusion Models of New Product Acceptance*, Ballinger, Cambridge, MA; 1986.

Malrieu F., *Convergence to equilibrium for granular media equations and their Euler scheme*, Ann. Appl. Prob., 540–560, 13; 2003.

Mandelbrot K. and van Ness J., *Fractional Brownian motions, fractional noises and applications*, SIAM Rev., 422–437, 10; 1968.

Mao X., *Stochastic Differential Equations and Applications*, Horwood, Chichester; 1997.

Mao X., Marion G. and Renshaw E., *Environmental Brownian noise suppresses explosions in population dynamics*, Stoch. Proc. Appl., 95–110, 97; 2002.

Mao X, et al. *Stochastic differential delay equations of population dynamics*, J. Math. Anal. Appl., 296–320, 304;2005.

Markowich P. A. and Villani C., *On the trend to equilibrium for the Fokker-Planck equation: An interplay between physics and functional analysis*, Math. Contemp., 1–31, 19; 2000.

Medvegyev P., *Stochastic Integration Theory*, Oxford University Press, Oxford; 2007.

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, Heidelberg; 1996.

Merton R. C., *Theory of rational option pricing*, Bell J. Econ. Manage. Sci., 141–183, 4; 1973.

Merton R. C., *Option pricing when underlying stock returns are discontinuous*, J. Fin. Econ., 125–144, 3; 1976.

Métivier M., *Notions fondamentales de la théorie des probabilités*, Dunod, Paris; 1968.

Meyer P. A., *Probabilités et Potentiel*, Ilermann, Paris; 1966.

Mikosch T., *Non-Life Insurance Mathematics*, Springer, Berlin Heidelberg; Second Edition, 2009.

Miltersen K. R., Sandmann K. and Sondermann D., *Closed form solutions for term structure derivatives with log-normal interest rates*, J. Fin., 409–430, 52; 1997.

Morale D., Capasso V. and Oelschläger K., *An interacting particle system modelling aggregation behaviour: From individuals to populations*, J. Math. Biol.; 2004.

Musiela M. and Rutkowski M., *Martingale Methods in Financial Modelling*, Springer, Berlin; 1998.

Nagai T. and Mimura M., *Some nonlinear degenerate diffusion equations related to population dynamics*, J. Math. Soc. Japan, 539–561, 35; 1983.

Neveu J., *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco; 1965.

Nowman K. B., *Gaussian estimation of single-factor continuous time models of the term structure of interest rate*, J. Fin., 1695–1706, 52; 1997.

Oelschläger K., *A law of large numbers for moderately interacting diffusion processes*, Z. Wahrsch. verw. Geb., 279–322, 69; 1985.

Oelschläger K., *Large systems of interacting particles and the porous medium equation*, J. Differential Equations, 294–346, 88; 1990.

Øksendal B., *Stochastic Differential Equations*, Springer, Berlin; 1998.

Okubo A., *Dynamical aspects of animal grouping: Swarms, school, flocks and herds*, Adv. BioPhys., 1–94, 22; 1986.

Parzen E., *Stochastic Processes*, Holden-Day, San Francisco; 1962.

Pascucci A., *Calcolo Stocastico per la Finanza*, Springer Italia, Milano; 2008.

Pazy A., *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York; 1983.

Pliska S. R., *Introduction to Mathematical Finance: Discrete-Time Models*, Blackwell, Oxford; 1997.

Protter P., *Stochastic Integration and Differential Equations*, Springer, Berlin; 1990. Second Edition 2004.

Rebolledo R., *Central limit theorems for local martingales*, Z. Wahrsch. verw. Geb., 269–286, 51; 1980.

Revuz D. and Yor M., *Continuous Martingales and Brownian Motion*, Springer, Heidelberg; 1991.

Robert P., *Stochastic Networks and Queues*, Springer, Heidelberg; 2003.

Rogers L. C. G. and Williams D., *Diffusions, Markov Processes and Martingales*, Vol. 1, Wiley, New York; 1994.

Rolski T., Schmidli H., Schmidt V. and Teugels J., *Stochastic Processes for Insurance and Finance*, Wiley, New York; 1999.

Roozen H., *Equilibrium and extinction in stochastic population dynamics*, Bull. Math. Biol., 671–696, 49; 1987.

Samorodnitsky G. and Taqqu M. S., *Stable Non-Gaussian Random Processes*, Chapman & Hall/CRC Press, Boca Ration, FL; 1994.

Sato K. I., *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK; 1999.

Schuss Z., *Theory and Applications of Stochastic Differential Equations*, Wiley, New York; 1980.

Schuss Z., *Theory and Applications of Stochastic Processes: An Analytical Approach*, Springer, New York; 2010.

Shiryaev A. N., *Probability*, Springer, New York; 1995.

Shiryaev A. N. and A. S. Cherny *Vector stochastic integrals and the fundamental theorems of asset pricing*, Tr. Mat. Inst., Steklova, 12–56, 237; 2002.

Skellam J. G., *Random dispersal in theoretical populations*, Biometrika, 196–218, 38; 1951.

Skorohod A. V., *Studies in the Theory of Random Processes*, Dover, New York; 1982.

Skorohod A. V., *Asymptotic Methods in the Theory of Stochastic Differential Equations*, AMS, Providence, RI; 1989.

Sobczyk K., *Stochastic Differential Equations: With Applications to Physics and Engineering*, Kluwer, Dordrecht; 1991.

Stein R. B., *A theoretical analysis of neuronal variability*, Biophys. J., 173–194, 5; 1965.

Stein R. B., *Some models of neuronal variability*, Biophys. J., 37–68, 7; 1967.

Taira K., *Diffusion Processes and Partial Differential Equations*, Academic, New York; 1988.

Tan W. Y., *Stochastic Models with Applications to Genetics, Cancers, AIDS and Other Biomedical Systems*, World Scientific, Singapore; 2002.

Tucker H.G., *A Graduate Course in Probability*, Academic Press, New York; 1967.

Tuckwell H. C., *On the first exit time problem for temporarily homogeneous Markov process*, J. Appl. Prob., 39–48, 13; 1976.

Tuckwell H. C., *Stochastic Processes in the Neurosciences*, SIAM, Philadelphia; 1989.

Vasicek O., *An equilibrium characterisation of the term structure*, J. Fin. Econ., 177–188, 5; 1977.

Ventcel' A. D., *A Course in the Theory of Stochastic Processes*, Nauka, Moscow (in Russian); 1975. Second Edition 1996.

Veretennikov A. Y., *On subexponential mixing rate for Markov processes*, Theory Prob. Appl., 110–122, 49; 2005.

Wang F. J. S., *Gaussian approximation of some closed stochastic epidemic models*, J. Appl. Prob., 221–231, 14; 1977.

Warburton K. and Lazarus J., *Tendency-distance models of social cohesion in animal groups*, J. Theor. Biol., 473–488, 150; 1991.

Wax N., *Selected Papers on Noise and Stochastic Processes*, Dover, New York; 1954.

Williams D., *Probability with Martingales*, Cambridge University Press, Cambridge, UK; 1991.

Wilmott P., Dewynne J. N. and Howison S. D.,*Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, Oxford; 1993.

Wu F., Mao X. and Chen K., *A highly sensitive mean-reverting process in finance and the Euler–Maruyama approximations*, J. Math. Anal. Appl., 540–554, 348; 2008.

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, Heidelberg; 1985.

Nomenclature

Notation	**Description**
$(E, \mathcal{B}_E)$	Measurable space with E a set and $\mathcal{B}_E$ a σ-algebra of parts of E
$(\Omega, \mathcal{F}, P)$	Probability space with Ω a set, $\mathcal{F}$ a σ-algebra of parts of Ω, and P a probability measure on $\mathcal{F}$
$:=$	Equal by definition
$< M, N >$	Predictable covariation of martingales M and N
$< M >$, $< M, M >$	Predictable variation of martingale M
$< f, g >$	Scalar product of two elements f and g in a Hilbert space
A'	Transpose of matrix A
$A \setminus B$	Set of elements of A that do not belong to B
$B(x, r)$ **or** $B_r(x)$	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 kth derivatives are Lipschitz continuous with exponent α
$C_0(A)$	Set of continuous functions on A with compact support
$C_b(A)$ **or** $BC(A)$	Set of bounded continuous functions on A
$Cov[X, Y]$	Covariance of two random variables X and Y
$E[Y \| \mathcal{F}]$	Conditional expectation of random variable Y with respect to σ-algebra $\mathcal{F}$
$E[\cdot]$	Expected value with respect to an underlying probability law clearly identifiable from context
$E_P[\cdot]$	Expected value with respect to probability law P
$E_x[\cdot]$	Expected value conditional upon a given initial state x in a stochastic process

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

Notation	Description
F_X	Cumulative distribution function of a random variable X
$H \bullet X$	Stochastic Stieltjes integral of process H with respect to 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 measure P
$N(\mu, \sigma^2)$	Normal (Gaussian) random variable with mean μ and variance σ^2
$O(\Delta)$	Of the same order as Δ
P**-a.s.**	Almost surely with respect to measure P
$P(A\|B)$	Conditional probability of event A with respect to event B
$P * Q$	Convolution of measures P and Q
$P \ll Q$	Measure P is absolutely continuous with respect to measure Q
$P \sim Q$	Measure P is equivalent to 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
$Var[X]$	Variance of a random variable X
W_t	Standard Brownian motion, Wiener process
$X \sim P$	Random variable X has probability law P
$[a, b[$	Semiopen interval closed at extreme a and open at extreme b
$[a, b]$	Closed interval of extremes a and b
Δ	Laplace operator
Ω	Underlying sample space
Φ	Cumulative distribution function of a standard normal probability law
$\bar{A}$	Closure of a set A depending on context
$\bar{C}$	Complement of set C depending on context
$\bar{\mathbb{R}}$	Extended set of real numbers, i.e., $\mathbb{R} \cup \{-\infty, +\infty\}$
δ_x	Dirac delta function localized at x
δ_{ij}	Kronecker delta, i.e., $= 1$ for $i = j$, $= 0$ for $i \neq j$
$\emptyset$	Empty set
ϵ_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

Notation	Description
$\mathbb{C}$	Complex plane
$\mathbb{N}$	Set of natural nonnegative integers
$\mathbb{N}^*$	Set of natural (strictly) positive integers
$\mathbb{Q}$	Set of rational numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{R}_+$	Set of positive (nonnegative) real numbers
$\mathbb{R}_+^*$	Set of (strictly) positive real numbers
$\mathbb{Z}$	Set of all integers
$\mathcal{A}$	Infinitesimal generator of a semigroup
$\mathcal{B}_E$	σ-algebra of Borel sets generated by the topology of E
$\mathcal{B}_{\mathbb{R}^n}$	σ-algebra of Borel sets on $\mathbb{R}^n$
$\mathcal{D}_A$	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., σ-algebra generated by $\{X_s, s \leq t\}$
$\mathcal{F}_X$	σ-algebra generated by random variable X
$\mathcal{F}_{t+}$	$\bigcap_{s>t} \mathcal{F}_t$
$\mathcal{F}_{t-}$	σ-algebra generated by $\sigma(X_s, s < t)$
$\mathcal{L}(X)$	Probability law of X
$\mathcal{L}^p(P)$	Set of integrable functions with respect to measure P
$\mathcal{M}(E)$	Set of all measures on E
$\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$	Set of all $\mathcal{F}$-measurable functions with values in $\bar{\mathbb{R}}_+$
$\mathfrak{P}(\Omega)$	Set of all parts of a set Ω
$\xrightarrow[n]{P}$	Convergence in probability
$\xrightarrow[n]{P}$ **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
∇	Gradient
ω	Element of underlying sample space
$\otimes$	Product of σ-algebras or product of measures
∂A	Boundary of a set A
$\operatorname{sgn}\{x\}$	Sign function; 1 if $x > 0$, 0 if $x = 0$, -1 if $x < 0$
$\sigma(\mathcal{R})$	σ-algebra generated by family of events $\mathcal{R}$
$\square$	End of a proof
$\|A\|$ **or** $\sharp(A)$	Cardinal number (number of elements) of a finite set A
$\|\|x\|\|$	Norm of a point x
$\|a\|$	Absolute value of a number a; or modulus of a complex number a
$]a, b[$	Open interval of extremes a and b
$]a, b]$	Semiopen interval open at extreme a and closed at extreme b

Notation	**Description**
$a \vee b$	Maximum of two numbers
$a \wedge b$	Minimum of two numbers
$f * g$	Convolution of functions f and g
$f \circ X$ **or** $f(X)$	A function f composed with a function X
$f\|_A$	Restriction of a function f to set A
f^-, f^+	Negative (positive) part of f, i.e., $f^- = \max\{-f, 0\}$ ($f^+ = \max\{f, 0\}$)
$f^{-1}(B)$	Preimage of set B by function f
$o(\delta)$	Of higher order with respect to δ
a.e.	Almost everywhere
a.s.	Almost surely

Index

A
absolutely continuous, 363, 371
absorbing state, 123
adapted, 91, 141
affine
  rate model, 310
algebra, 359
  σ-, 359
  Borel, 15, 360
  generated, 9, 360
  product, 15, 77, 360
  semi, 359, 360
  smallest, 78, 360, 362
  tail, 48
annuity, 310
arithmetic Brownian motion, 270
asset
  riskless, 278
  risky, 278
attainable, 280
attractive, 407
autonomous, 234

B
Bachelier model, 296
ball, 377
  closed, 377
  open, 377
barrier, 397
basis, 379
bicontinuous, 378
bijection, 378
binomial
  distribution, 14
  variable, 22, 29
Black–Scholes
  equation, 285
  formula, 287
  model, 283
Borel
  σ-algebra, 360
  –Cantelli lemma, 48
  algebra, 15
  measurable, 360, 361, 367
boundary, 378
boundary point
  accessible, 264
  regular, 264
bounded, 378
bounded Lipschitz norm, 385
Brownian bridge, 88, 269
Brownian motion, 125
  absorbed , 138
  arithmetic, 224, 270
  first passage time, 168
  fractional, 188
  geometric, 224, 270
  Hölder continuity, 136
  Lévy characterization, 130
  recurrence , 168
  reflected, 138

C
càdlàg, 81
Cantor function, 11

V. Capasso and D. Bakstein, *An Introduction to Continuous-Time Stochastic Processes*, MSSET, DOI 10.1007/978-0-8176-8346-7,

Cauchy
  –Schwarz inequality, 50
  distribution, 13
  problem, 247, 398
  sequence, 380
  variable, 23, 71
cemetery, 306
Chapman–Kolmogorov, 103
Characteristic functions
  vectors, 27
Chernoff, 75
class, 7, 17, 37
  D, 98
  DL, 98
  equivalence, 34
closed
  ball, 377
  set, 378
closure, 378
  point, 378
compact, 381
  locally, 382
  relatively, 381, 387
  set, 381
compacts, 196
compatible system, 79, 88
compensator, 99, 144
complement, 378
complete, 380
  market, 280
composite
  projection, 78
condition
  Feller, 55
conditional
  density, 44
  distribution, 41
  expectation, 33
  expectation, on σ-algebra, 36
  probability, 4
  probability, regular version, 35, 41, 43
conservative, 122
contingent claim, 280
  linear, 286
continuous
  absolutely, 363, 371
  function, 361, 378
  Hölder, 247
  in probability, 81
  left-, 81
  random variable, 11
  right-, 10, 81, 83, 369
  uniformly, 379
convergence
  almost sure, 49
  dominated, 366
  in distribution, 51
  in mean, 49
  in probability, 49
  monotone, 366
  pointwise, 52
  uniform, 52
  weak, 51, 383
convex, 38
convolution, 20, 90, 339, 369
  semigroup, 90
corollary
  Lévy's continuity, 52
correlation, 24
countable, 81
  additive, 364
  base, 360
  basis, 379
  family, 13
  partition, 8
  set, 33
covariance, 23, 25
covariation, 100
covering, 379
cylinder, 15, 77

D
decomposable, 98
decomposition
  Doob, 69
  Doob–Meyer, 97
  Lévy-Itô, 158
definite
  negative, 408
deflator, 279, 282
delta, 308
  Dirac, 259
  Kronecker, 191
dense, 81, 379
  everywhere, 379
density, 80, 370
  conditional, 44
  Gaussian, 12

lognormal, 12
normal, 12
probability, 11
risk-neutral transition, 297
transition, 248
uniform, 12, 22
depolarization, 349
derivative, 280
diameter, 378
diffusion
coefficient, 118, 235
matrix, 212
operator, 249
process, 118, 235
Dirac
delta, 259
measure, 146
Dirichlet problem, 398
distance, 377
equivalent, 378
distribution
binomial, 14
Cauchy, 13
conditional, 41
cumulative, 10
discrete, 13
empirical, 327
exponential, 12
finite-dimensional, 388
function, 80, 369
Gamma, 13
Gaussian, 87
initial, 104
joint, 19, 41
marginal, 15, 19, 33
Poisson, 14
uniform, 13
divisible
infinitely, 57
Doléans exponential, 199
Doob
–Meyer decomposition, 97
decomposition, 69
inequality, 94
drift, 118, 235, 283
change of, 237
vector, 212
dynamic system, 233
Dynkin's formula, 114, 253

E
elementary
event, 3
function, 35, 361, 365
random variable, 35
elliptic, 253
equation, 397
operator, 254
entropy, 71
equation
Black–Scholes, 285
Chapman–Kolmogorov, 103, 250
elliptic, 397
evolution, 327
Fokker–Planck, 250
Kolmogorov, 211
Kolmogorov backward, 122
Kolmogorov forward, 250, 298
parabolic, 397
equiprobable, 3
equivalent, 239, 371
distance, 378
process, 80
essential
infimum, 48
supremum, 48
event, 3, 148
T-preceding, 85
complementary, 7
elementary, 3
independent, 6
mutually exclusive, 6
tail, 48
excited, 349
expectation, 20
conditional, 33
conditional on σ-algebra, 36
expected value, 20

F
Feynman–Kac formula, 246
filtration, 83, 91
generated, 83
natural, 83
right-continuous, 107
finite, 3
σ-, 363
additive, 364
base, 78

finite (*cont.*)
dimensional distribution, 388
horizon, 279
measure, 363, 367
product, 79
space, 3
stopping time, 85
uniformly, 367
Fokker–Planck, 250
formula
Black–Scholes, 287
Dynkin, 114, 253
Feynman–Kac, 246
Itô, 197
Kolmogorov, 253
forward, 285
measure, 295
function
Cantor, 11
characteristic, 52, 142
characteristic of a random vector, 27
continuous, 361, 378
convex, 38
cumulative distribution, 10
distribution, 80, 369
elementary, 35, 361, 365
equivalence class of, 34
Gamma, 13
indicator, 361
Lyapunov, 255, 408
Markov transition, 103
matrix transition, 122
measurable, 361
moment generating, 75
partition, 10, 51
piecewise, 175
random, 78
space, 78
test, 353
fundamental solution, 398

G
gamma, 308
Gaussian
bivariate, 45
density, 12
distribution, 87
process, 87
variable, 22, 30, 58
vectors, 30
generating triplet, 72
geometric Brownian motion, 270
Girsanov, 239
Glivenko-Cantelli theorem, 54
gradient
generalized, 338
Greeks, 308

H
Hausdorff topological space, 378
heat operator, 249
Heston
model, 300
Hille-Yosida theorem, 404
history, 150
Hölder
continuous, 247
inequality, 50
holding, 278
homeomorphic, 378
homeomorphism, 378
horizon
finite, 279
infinite, 279

I
independent
class, 7
classes, 17
event, 6
increments, 88
marking, 302
mutually, 7
variable, 25, 42
index
indistinguishable, 80
modifications, 80
indicator, 34
function, 361
indistinguishable, 80
induced
measure, 367
probability, 87, 367
inequality
Cauchy–Schwarz, 50
Chebyshev, 22
Doob, 94

Hölder, 50
Jensen, 38, 92
Markov, 22
Minkowski, 50
infinitely
divisible, 57
infinitesimal generator, 107
extended, 116
inhibited, 349
initial reserve, 301
instantaneous state, 123
integrable
P-, 20
$P_{(X,Y)}$-, 47
μ-, 365
square-, 21, 372
uniformly, 40, 97
integral
Itô, 173, 181
Lebesgue, 184, 365
Lebesgue–Stieltjes, 173
Riemann–Stieltjes, 174
stochastic Stieltjes, 374
Stratonovich, 185
upper, 365
intensity, 142
cumulative stochastic, 151
matrix, 122
multiplicative, 313
stochastic, 148
interior, 377
Itô
formula, 197
integral, 173, 181
isometry, 178, 226
Itô-Lévy integral, 204
representation theorem, 199

J
jumps
bounded, 155
fixed, 141

K
killing, 246
Kolmogorov
–Chapman equation, 103, 250
backward equation, 122, 245
continuity theorem, 128
equation, 211, 241
formula, 253
forward equation, 250, 298
theorem, 60
zero-one law, 48

L
Lagrangian, 326
Laplacian, 249
large deviations, 75
law
of iterated logarithms, 75, 137
of large numbers, 136
tower, 37
law of large numbers
strong, 54
weak , 53
Lebesgue
–Stieltjes integral, 173
–Stieltjes measure, 11, 369
dominated convergence theorem, 366
integral, 184, 365
integration, 364
measure, 11, 370
Nikodym theorem, 371
lemma
Borel–Cantelli, 48
Fatou, 222, 364, 366
Gronwall, 214
Lévy
–Khintchine decomposition, 72
–Khintchine formula, 159
characterization of Brownian motion, 130
continuity theorem, 52
measure, 156
process, 154
Lévy-Itô
decomposition, 158
limit, 380
McKean–Vlasov, 341
moderate, 341
projective, 79, 87
Lindeberg
theorem, 56
Lindeberg condition, 55, 64
Lipschitz, 75, 247
local
volatility, 297

locally compact, 382
Lyapunov
  criterion, 408
  function, 255, 408
  theorem, 408

M
mark, 149
market, 278
  complete, 280
  discount-bond, 290
Markov
  chain, 166
  inequality, 22
  process, 101
    Feller process, 113
    Feller property, 106
  property, 231, 267
  property, strong, 108, 237
  sequence, 102
  stopping time, 85
  transition function, 103
Markov process
  holding time, 117
martingale, 68, 91, 185
  central limit theorem, 70
  convergence theorem, 70
  innovation, 147
  local, 99
  orthogonal, 100
  problem, 116
  purely discontinuous, 100
  representation theorem, 198, 199
  reversed, 305
  semi-, 170
  sub-, 69, 91
  super-, 91
mean reverting process, 270
measurable, 35, 83
  $(\mathcal{F} - \mathcal{B}^T)$-, 77
  $\mathcal{F}$-, 360, 365
  $\mathcal{F}_t$-, 91
  Borel-, 360, 361, 367
  function, 361
  jointly, 83
  mapping, 361
  progressively, 83, 84
  projection, 362
  rectangle, 368
  set, 360
  space, 360, 361
measure, 362
  characterization of, 363
  compatible system of, 79
  Dirac, 146
  empirical, 327
  equivalent, 239, 371
  finite, 363, 367
  forward, 295
  generalization of, 364
  induced, 367
  jump, 212
  Lebesgue, 11, 370
  Lebesgue–Stieltjes, 11, 369
  Lévy, 156
  physical, 291
  probability, 3, 34, 362
  product, 367–369
  Radon, 373
  random, 145
  regular, 373
  space, 363
metric, 377
  Ky Fan, 383
  notion, 379
  space, 377
modifications
  progressively measurable, 83
  separable, 82, 83
moment
  centered, 21
  generating function, 75

N
nonexplosive, 141
norm, 105
  Euclidean, 257
  semi-, 50
  sup, 113
normal
  bivariate, 45
  density, 12
Novikov condition, 239
numeraire, 279, 282

O
open
  ball, 377
  set, 377
operator
  closed, 404
  diffusion, 249
  elliptic, 254
  expectation, 23
  heat, 249
  parabolic, 399
option, 285
  American, 289
  barrier, 288
  binary, 287
  Call, 285
  digital, 287
  European, 289
  Put, 285
  vanilla, 285
Ornstein–Uhlenbeck process, 270
orthogonal, 100

P
parabolic
  differential equation, 246
  operator, 399
partition, 8, 174
  function, 10
path, 77, 82
  space, 345
payoff, 285
point, 377
  cluster, 380
  of closure, 378
Poisson
  compound process, 165, 303
  compound variable, 59
  distribution, 14
  generalized process, 166
  intensity, 14
  marked process, 152, 302
  process, 142
  variable, 22, 29, 58
polynomial, 196
portfolio, 278
positive
  definite, 408
precede, 85
precompact, 381
predictable, 83
  covariation, 100
premia, 301
probability, 3, 362
  axioms, 4
  conditional, 4
  conditional probability measure, 5
  conditional, regular version, 35, 41, 43
  density, 11
  generating function, 142
  induced, 87
  joint, 15, 46, 87
  law, 9, 87, 88
  law of a process, 79
  measure, 3, 34, 362
  one-point transition, 121
  product, 79
  ruin, 302
  space, 3, 77, 363
  survival, 302
  total law of, 8
  transition, 103, 234
process
  adapted, 91
  affine, 171
  canonical form, 87
  claims surplus, 302
  compound Poisson, 165, 303
  counting, 141
  Cox, 145
  cumulative claims, 302
  diffusion, 118, 235
  equivalent, 80
  Feller, 107
  Gaussian, 87
  generalized Poisson, 166
  holding, 278
  homogeneous, 112, 234
  Itô–Lévy, 207
  Lévy
    stable, 163
  Lévy, 154
  marked point, 149, 302
  marked Poisson, 152, 302
  Markov, 101
  mean reverting, 270
  orderly, 145
  Ornstein–Uhlenbeck, 224, 270

process (*cont.*)
  piecewise deterministic, 306
  point, 145
  Poisson, 142
  portfolio, 278
  self-similar, 189
  separable, 81
  simple, 145
  stochastic, 77
  time-homogeneous, 90
  Wiener, 125
    Gaussian, 126
  with independent increments, 88
product
  convolution, 369
  measure, 367–369
  scalar, 178, 203
projection, 15, 33, 362
  canonical, 77
  composite, 78
  orthogonal, 39
projective system, 79
property
  Beppo–Levi, 35, 364
  Feller, 106, 237
  scaling, 163

R
Radon measures, 373
Radon–Nikodym derivative, 239
random
  function, 78
  variable, 9
  variables, family of, 77
  variables, family of,
    reproducible, 30
  vector, 14, 44, 87
rate
  forward, 293
  interest, 290
  riskless, 283
  short, 291
  swap, 310
  zero, 290
RCLL, 81
rectangle, 19, 80, 360
  measurable, 368
reflection principle, 132, 395
regression, 46
ring, 78, 359
  σ-, 359
  semi-, 359
risk
  insurance, 301
  reserve, 302
riskless
  account, 283
  asset, 278
  rate, 283
risky
  asset, 278
ruin
  probability, 302
  time of, 302

S
SABR model, 300
sample, 170
scaling property, 136, 163
self-similar, 170
semicircle, 19
semigroup, 105
  contraction, 116, 404
  convolution, 90
  infinitesimal generator, 403
  property, 233
  transition, 113
semigroup of linear operators, 403
separable, 379
sequence, 35
  Cauchy, 180, 380
  Markovian, 102
set
  closed, 378
  compact, 381
  countable, 33
  empty, 3
  negligeable, 82
  open, 377
  separating, 82, 129
singular, 370
space
  complete, 44
  function, 78
  Hilbert, 39, 178
  mark, 149
  measurable, 360, 361
  measure, 363

metric, 44, 377
normed, 14
on which a probability measure can be built, 360
path, 345
phase, 77
Polish, 79
probability, 3, 77, 363
product, 78
separable, 44
state, 77
topological, 360, 377
uniform, 3
spaces
Borel, 43
isomorphic, 43
stable, 407
asymptotically, 407
Lévy process, 163
law
absolute continuity, 68
domain of attraction, 66
stability index, 66
state, 123
un-, 407
standard deviation, 21
stationary
strictly, 163
weakly, 164
stochastic
differential, 193, 201
differential equation, 213
autonomous, 223
existence and uniqueness, 214
global existence, 223
Itô-Lévy , 265
process, 77
stability, 257
stopping time, 70, 85, 289
subordinator, 169
support, 13, 399
swap rate, 310

T

Taylor's formula, 119
term structure, 290
test
function, 353
theorem approximation of measurable functions through elementary functions, 361
Bayes, 8
central limit, 54
Cramér-Wold, 30
dominated convergence, 366
Donsker, 391
Doob–Meyer, 98
Fatou–Lebesgue, 366
first fundamental of asset pricing, 280
Fubini, 46, 83, 90, 368
fundmental theorem of calculus, 372
Girsanov, 239
inversion, 26
Itô representation theorem, 199
Jirina, 44
Kolmogorov zero-one law, 48
Kolmogorov's continuity, 128
Kolmogorov–Bochner, 79
Lagrange, 245
law of iterated logarithms, 75, 137
Lebesgue–Nikodym, 371
Lévy's continuity, 52
Lyapunov, 408
martingale representation, 152, 198, 199
mean value, 237
measure extension, 102
monotone convergence, 366
numeraire invariance, 282
Polya, 52
portmanteau, 385
product measure, 367
Prohorov, 388
Radon–Nikodym, 36, 371
second fundamental of asset pricing, 281
Skorohod representation, 53, 384
strong law of large numbers, 136
total law of probability, 8
Weierstrass, 196
theta, 308
threshold, 350
tight, 387
time
exit, 108, 252, 288
explosion, 141
hitting, 108, 252

time (*cont.*)
  homogeneous, 90, 156
  of ruin, 302
  stopping, 85, 107, 189, 289
  value of money, 290
topological
  notions, 378
topological space, 377
topology, 377
tower law, 37
trajectory, 77, 82
translation invariance, 161

U
uniformly
  continuous, 379
  integrable, 97
usual hypotheses, 83

V
variable
  binomial, 22, 29
  Cauchy, 23, 71
  centered, 21
  compound Poisson, 59
  elementary, 35
  extensive, 9
  Gaussian, 22, 30, 58
  independent, 25, 42
  Poisson, 22, 29, 58
  random, 9
  sums of, 19
variance, 21, 25
  constant elasticity of, 310
variation
  bounded, 174
  predictable, 100
  quadratic, 100
    Wiener, 184
  total, 174
vega, 308
volatility, 283
  implied, 296
  local, 297
  skew, 297, 300
  smile, 300
  stochastic, 300